



T.C.

BURSA ULUDAĞ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

İŞLETME ANABİLİM DALI

MUHASEBE VE FİNANSMAN BİLİM DALI

**SOSYAL MEDYANIN FİNANSAL PİYASALARA
ETKİSİ VE HİSSE SENEDİ FİYAT
ÖNGÖRÜLERİNDE KULLANILMASI: BORSA
İSTANBUL ÖRNEĞİ**

(DOKTORA TEZİ)

Yunus Emre AKDOĞAN

BURSA – 2023



T.C.

BURSA ULUDAĞ ÜNİVERSİTESİ

SOSYAL BİLİMLER ENSTİTÜSÜ

İŞLETME ANABİLİM DALI

MUHASEBE VE FİNANSMAN BİLİM DALI

**SOSYAL MEDYANIN FİNANSAL PİYASALARA ETKİSİ VE
HİSSE SENEDİ FİYAT ÖNGÖRÜLERİNDE
KULLANILMASI: BORSA İSTANBUL ÖRNEĞİ**

(DOKTORA TEZİ)

Yunus Emre AKDOĞAN

**Danışman:
Prof. Dr. Adem ANBAR
BURSA – 2023**

Yemin Metni

Doktora tezi olarak sunduđum “Sosyal Medyanın Finansal Piyasalara Etkisi ve Hisse Senedi Fiyat Öngörülerinde Kullanılması: Borsa İstanbul Örneđi” başlıklı çalışmanın bilimsel araştırma, yazma ve etik kurallarına uygun olarak tarafımdan yazıldığına ve tezde yapılan bütün alıntıların kaynaklarının usulüne uygun olarak gösterildiđine, tezimde intihal ürünü cümle veya paragraflar bulunmadığına şerefim üzerine yemin ederim.

Tarih ve İmza

Adı Soyadı: Yunus Emre AKDOĞAN

Öğrenci No: 711314017

Anabilim Dalı: İşletme

Programı: Doktora

Tezin Türü: Yüksek Lisans / ☐ Doktora / ☒



SOSYAL BİLİMLER ENSTİTÜSÜ
DOKTORA İNTİHAL YAZILIM RAPORU
BURSA ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
İŞLETME ANABİLİM DALI BAŞKANLIĞI'NA

Tarih: 04/07/2023

Tez Başlığı / Konusu: Sosyal Medyanın Finansal Piyasalara Etkisi ve Hisse Senedi Fiyat Öngörülerinde Kullanılması: Borsa İstanbul Örneği

Yukarıda başlığı gösterilen tez çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 334 sayfalık kısmına ilişkin, 04/07/2023 tarihinde şahsım tarafından *Turnitin* adlı intihal tespit programından (Turnitin)* aşağıda belirtilen filtrelemeler uygulanarak alınmış olan özgünlük raporuna göre, tezimin benzerlik oranı %6'dır.

Uygulanan filtrelemeler:

- 1- Kaynakça hariç
- 2- Alıntılar hariç/dahil
- 3- 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Bursa Uludağ Üniversitesi Sosyal Bilimler Enstitüsü Tez Çalışması Özgünlük Raporu Alınması ve Kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Gereğini saygılarımla arz ederim.

Tarih ve İmza

Adı Soyadı: Yunus Emre AKDOĞAN

Öğrenci No: 711314017

Anabilim Dalı: İşletme

Programı: İşletme Doktora

Statüsü: ☐ Y.Lisans ☒ Doktora

Danışman

(Prof. Dr. Adem ANBAR, 04.07.2023)

T. C.
BURSA ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İşletme Anabilim / Ana sanat Dalı, Bilim Dalı'nda 711314017 numaralı Yunus Emre AKDOĞAN'ın hazırladığı “Sosyal Medyanın Finansal Piyasalara Etkisi ve Hisse Senedi Fiyat Öngörülerinde Kullanılması: Borsa İstanbul Örneği” başlıklı doktora tezi ile ilgili savunma sınavı, .../...../20.... günü - saatleri arasında yapılmıştır. Alınan cevaplar sonunda adayın (başarılı / başarısız) olduğuna

..... (oybirliği / oy çokluğu) ile karar verilmiştir.

Üye (Tez Danışmanı ve Sınav Komisyonu
Başkanı)
Prof. Dr. Adem ANBAR
Bursa Uludağ Üniversitesi

Üye
Prof. Dr. Mehmet ÇINAR
Bursa Uludağ Üniversitesi

Üye
Prof. Dr. İbrahim Halil EKŞİ
Gaziantep Üniversitesi

Üye
Prof. Dr. Abdulkadir KAYA
Bursa Teknik Üniversitesi

Üye
Prof. Dr. Sinan AYTEKİN
Balıkesir Üniversitesi

.../.../ 2023

ÖZET

Yazar Adı Soyadı	Yunus Emre AKDOĞAN
Üniversite	Bursa Uludağ Üniversitesi
Enstitü	Sosyal Bilimler Enstitüsü
Anabilim Dalı	İşletme ABD
Bilim Dalı	
Tezin Niteliği	Doktora Tezi
Mezuniyet Tarihi	/...../2023
Tez Danışmanı	Prof. Dr. Adem ANBAR

SOSYAL MEDYANIN FİNANSAL PİYASALARA ETKİSİ VE HİSSE SENEDİ FİYAT ÖNGÖRÜLERİNDE KULLANILMASI: BORSA İSTANBUL ÖRNEĞİ

Dijital dönüşüm, sosyal yaşamın geçmişte verileştirilmesinin en zor alanlarıyla ilgili deneysel ve gözlemsel verilerine, makroskobik ve mikroskobik ölçekte ulaşmak için benzeri görülmemiş fırsatlar sunmaktadır. Bu bağlamda sosyal bilimler açısından dijital çağın en önemli veri kaynaklarından biri sosyal medya platformlarıdır. Bu çalışmada, geleneksel finansın genel kabul görmüş “tam bilgi”, “rasyonel insan” ve “izole birey” varsayımları yerine, eksik bilgisini tamamlamaya çalışan, sınırlı ve ekolojik rasyonaliteye sahip, sosyal çevresinin ve duygularının karar süreçlerinde pay sahibi olabildiği bir insan modelini esas alarak, bireysel yatırımcıların davranışlarını Twitter’den toplanan büyük veri ile yapay zekâ ve makine öğrenmesi yöntemlerini kullanarak çözümlemek amaçlanmaktadır. Bu amaçla Twitter kullanıcılarının 01.01.2012-28.02.2020 döneminde paylaştığı tweetler toplanmıştır. Tweetlerin Borsa İstanbul (BIST) pay piyasası ile ilgili olup olmadığını belirlemek için hem anahtar kelime tabanlı hem makine öğrenmesi tabanlı iki farklı bağlam sınıflandırıcı geliştirilmiştir. Makine öğrenmesi tabanlı bağlam sınıflandırıcı derin öğrenme yaklaşımlarından Gated Recurrent Units (GRU) yöntemi ile eğitilmiş ve %98 sınıflandırma başarısı elde edilmiştir. Tweetlerin fikri ve duygusal yöneliminin pozitif, negatif ya da nötr olarak sınıflandırılabilmesi için fikir ve duygu [sentiment] sınıflandırıcısı, ön eğitilmiş Bidirectional Encoder Representations from Transformers (BERT) yöntemiyle eğitilmiş ve pozitif ve negatif sınıflar için %91, nötr sınıf için %89 sınıflandırma başarısı elde edilmiştir. Twitter verilerinden elde edilen öznitelikler ile Borsa İstanbul pay piyasası endekslerinden BIST30, BIST100, BISTTUM, BIST SİNAİ, BIST TEKNOLOJİ, BIST HİZMETLER, BIST MALİ arasındaki ilişki makine öğrenmesi yöntemlerinden Lineer Regresyon, Lasso Regresyon, Rassal Orman ve XGBoost ile analiz edilmiştir. Analiz sonucunda BIST 100 endeksinin açılış değerindeki değişimin %91’inin, işlem hacmindeki değişimin %63’ünün ve volatilitedeki değişimin %67’sinin tweetlerden elde edilen bilişsel, davranışsal ve sosyal öznitelikler ile açıklanabildiği bulgusuna ulaşılmıştır. Benzer bulgular diğer endeksler için de geçerlidir.

Anahtar kelimeler: Davranışsal Finans, Borsa İstanbul, Yapay Zekâ, Makine Öğrenmesi, Büyük Veri, Duygu ve Fikir Analizi, Sosyal Medya Analizi, Algoritmik Finansal İşlemler

ABSTRACT

Name & Surname	Yunus Emre AKDOGAN
University	Bursa Uludag University
Institute	Institute of Social Sciences
Field	Business Administration
Subfield	
Degree Awarded	Doctorate
Date of Degree Awarded	/...../2023
Supervisor	Prof. Dr. Adem ANBAR

THE IMPACT OF SOCIAL MEDIA ON FINANCIAL MARKETS AND USING STOCK PRICE PREDICTION: CASE OF BORSA ISTANBUL

Digital transformation offers unprecedented opportunities to access empirical and observational data on the macroscopic and microscobic scales of social life in areas that were most difficult to datafication in the past. In this context, one of the most important data sources of the digital age in terms of social sciences is social media platforms. In this study, instead of the generally accepted "perfect information", "rational human" and "isolated individual" assumptions of traditional finance, a human model that tries to complete the missing information, has bounded and ecological rationality, and whose social network and emotions can have an effect in the decision processes, is based on an individual model. It is aimed to analyze the behavior of investors by using big data collected from Twitter, artificial intelligence and machine learning methods. For this purpose, tweets shared by Twitter users between 01.01.2012-28.02.2020 were collected. Two different context classifiers, both keyword-based and machine learning-based, have been developed to determine whether the tweets are related to the Borsa İstanbul Equity Market. The machine learning-based context classifier was trained with the Gated Recurrent Units (GRU) method, one of the deep learning approaches, and 98% classification success was achieved. In order to classify the sentiment of tweets as positive, negative or neutral, the sentiment classifier was trained with the pre-trained BERT method, and an accuracy of 91% for positive and negative classes and 89% for neutral class was achieved. The relationship between the features obtained from Twitter data and the BIST30, BIST100, BISTTUM, BIST SERVICES, BIST FINANCIAL, BIST INDUSTRY, BIST TECHNOLOGY indices was analyzed by machine learning methods Linear Regression, Lasso Regression, Random Forest and XGBoost methods. As a result of the analysis, it was found that 91% of the change in the opening values of the BIST 100 index, 63% of the change in the trading volume and 67% of the change in volatility can be explained by cognitive, behavioral and social features obtained from the tweets. Similar findings are also valid for other indices.

Keywords: *Behavioral Finance, Borsa Istanbul, Artificial Intelligence, Machine Learning, Big Data, Data Analytics, Sentiment Analysis, Social Media Analysis, Algorithmic Trading*

TEŞEKKÜR

Bu araştırmada yer alan nümerik hesaplamaların büyük bir kısmı TÜBİTAK ULAKBİM, Yüksek Başarım ve Grid Hesaplama Merkezi'nde (TRUBA kaynaklarında) gerçekleştirilmiştir.

Bu çalışmada kullanılan hesaplama kaynaklarının bir kısmı Ulusal Yüksek Başarımlı Hesaplama Merkezi'nin (UHeM), #4007652020# numaralı desteğiyle sağlanmıştır.

Bu çalışmada sırasında üç ay misafir araştırmacı olarak bulunduğum Sabancı Üniversitesi'ne ve desteğinden dolayı Prof. Dr. Yücel SAYGIN'a teşekkür ederim.

Bu araştırmada sırasında bilgisayar laboratuvarlarını kullanma imkânı veren Yozgat Bozok Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dekanlığı'na ve İşletme Bölüm Başkanlığına teşekkür ederim. Ayrıca destekleri için Prof. Dr. Hikmet ULUSAN, Prof. Dr. Mustafa SAÇMACI, Prof. Dr. Tansel HACIHASANOĞLU ve Prof. Dr. Mahmut AKIN'a teşekkür ederim.

Desteğini ve rehberliğini her zaman yanıbaşımnda bulduğum ve kıymetli katkılarıyla bu çalışmanın son halini almasını sağlayan danışmanım Prof. Dr. Adem ANBAR'a teşekkür ederim. Ayrıca çalışmanın son halini almasında yine çok değerli katkıları olan Prof. Dr. İbrahim Halil EKŞİ, Prof. Dr. Mehmet ÇINAR, Prof. Dr. Abdulkadir KAYA, Prof. Dr. Sinan AYTEKİN ve Prof. Dr. Fuat OĞUZ'a teşekkür ederim.

Son olarak bu çalışmada desteği ve katkısı olan herkese teşekkür ederim.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
TEŞEKKÜR	iii
İÇİNDEKİLER	iv
TABLOLAR	ix
ŞEKİLLER	xi
GRAFİKLER	xii
KISALTMALAR	xiv
GİRİŞ	1

BİRİNCİ BÖLÜM

RASYONALİTE, GELENEKSEL FİNANS VE DAVRANIŞSAL FİNANS

1.1 İNSAN EYLEMLERİNİN VE DAVRANIŞLARININ RASYONALİTESİ	8
1.1.1 Normatif Karar Teorisi ve Rasyonalite	9
1.1.2 Betimsel Karar Teorileri ve Rasyonalite	14
1.1.2.1 Beklenti Teorisi	14
1.1.2.2 Piyasa Süreci Teorisi	18
1.1.2.3 Karar Vermenin Psikolojik ve Nörobiyolojik Temelleri	21
1.1.2.3.1 Akıl Yürütme ve Karar Verme	24
1.1.2.3.2 Sezgi	28
1.1.2.4 Davranışsal Yaklaşımların Rasyonalite Varsayımına Yönelik Eleştirileri	29
1.2 GELENEKSEL FİNANS KURAMLARI VE RASYONALİTE	35
1.2.1 Belirsizlik Koşullarında Rasyonel Seçim: Beklenen Fayda Teorisi	35
1.2.2 Modern Portföy Teorisi	38
1.2.3 Sermaye Varlıklarını Fiyatlama Modeli	39
1.2.4 Etkin Piyasalar Hipotezi	41
1.2.5 Rassal Yürüyüş Hipotezi	43
1.3 DAVRANIŞSAL FİNANSIN RASYONALİTE ELEŞTİRİLERİ	44
1.4 DAVRANIŞSAL FİNANSTA YATIRIMCI EĞİLİMLERİ	47
1.4.1 Doğrulama Eğilimi	47
1.4.2 Örneklem Eğilimi	48

1.4.3 Bilgi İllüzyonu.....	49
1.4.4 Kontrol İllüzyonu	50
1.4.5 Hazırda Bulunma Etkisi.....	52
1.4.6 Aşinalık Eğilimi	53
1.4.7 Belirsizlikten Kaçınma	54
1.4.8 Pişmanlıktan Kaçınma	54
1.4.9 Aşırı Tepki Hipotezi	55
1.4.10 Sosyal Etkileşim.....	57
1.4.10.1 Sürü Davranışı.....	58
1.4.10.2 Sosyal Medya ve Duygusal Durum	59
1.4.11 Gürültü Tüccarı Riski	60
1.4.12 Anlatı İktisadı.....	62
1.4.13 Geleneksel Medya ve Yatırımcı Eğilimleri.....	64
1.4.14 İnternet ve Yatırımcı Eğilimleri	65
1.5 DAVRANIŞSAL FİNANSTA BİLİŞSEL KESTİRMELER	67
1.5.1 Kahneman’ın Bilişsel Kestirmeleri ve Yanlılıkları.....	67
1.5.2 Gigerenzer’in Hızlı, Yalın ve Adaptif Sezgileri	70

İKİNCİ BÖLÜM

BÜYÜK VERİ, SOSYAL MEDYA VE FİNANS

2.1 DİJİTAL ÇAĞDA VERİLEŞTİRME VE SOSYAL BİLİMLER	74
2.1.1 Büyük Verinin Tanımı.....	77
2.1.2 Yapay Zekâ, Makine Öğrenmesi ve Finansal Sektör	82
2.2 SOSYAL MEDYA VE YATIRIMCILAR	87
2.2.1 Sosyal Medyanın Tanımı ve Kullanım İstatistikleri	88
2.2.2 Sosyal Medya ve Kurumsal Yatırımcılar	93
2.2.3 Sosyal Medya ve Bireysel Yatırımcılar.....	96
2.2.4 Sosyal Medyada Finansal Etkili Kişiler	101
2.2.5 Sosyal Medya ve Sermaye Piyasası Manipülasyonu	104
2.2.6 Sosyal Medya ve Sermaye Piyasası Regülasyonları.....	107

ÜÇÜNCÜ BÖLÜM

LİTERATÜR TARAMASI

3.1 FİNANSAL METİNLERİN YATIRIMCI DUYARLILIĞI VE FİNANSAL VARLIK FİYATLARI ÜZERİNDEKİ ETKİSİ	110
3.2 FİNANSAL HABERLER VE HİSSE SENEDİ FİYATLARI İLİŞKİSİNE İLİŞKİN LİTERATÜR.....	111
3.3 SOSYAL MEDYADA FİKİR VE DUYGU ANALİZİ: TWITTER VE HİSSE SENEDİ FİYATLARI İLİŞKİSİNE İLİŞKİN LİTERATÜR	115
3.3.1 Twitter’da Paylaşılan Mesajların (Tweetlerin) Kullanıldığı Çalışmalar	116
3.3.1.1 ABD Hisse Senedi Piyasaları İçin Yapılan Çalışmalar	117
3.3.1.2 Türkiye Hisse Senedi Piyasası İçin Yapılan Çalışmalar.....	145
3.3.1.3 Diğer Ülke Hisse Senedi Piyasaları İçin Yapılan Çalışmalar	148
3.3.1.4 Diğer Piyasalar İçin Yapılan Çalışmalar	153
3.3.2 Twitter Dışındaki Diğer Sosyal Medya Mesajlarının Kullanıldığı Çalışmalar....	155
3.4 LİTERATÜR DEĞERLENDİRMESİ	157

DÖRDÜNCÜ BÖLÜM

SOSYAL MEDYANIN HİSSE SENEDİ FİYAT ÖNGÖRÜLERİNDE KULLANILABİLİRLİĞİNE İLİŞKİN BORSA İSTANBUL ÜZERİNE BİR UYGULAMA

4.1 ARAŞTIRMANIN AMACI.....	161
4.2 ARAŞTIRMANIN KAPSAMI.....	164
4.3 ARAŞTIRMADA KULLANILAN VERİ SETLERİ.....	165
4.3.1 Biçimsel Dilin Kullanıldığı Kaynakların Derlenmesi	167
4.3.1.1 Yatırım Kuruluşlarının Raporları	168
4.3.1.2 Ekonomi ve Finans Web Siteleri.....	169
4.3.1.3 Diğer Günlük Ekonomi ve Finans Haber Kaynakları.....	170
4.3.2 Twitter Verilerinin Derlenmesi	171
4.3.3 Finansal Verilerin Derlenmesi	177
4.4 ARAŞTIRMANIN KISITLARI.....	177
4.5 ARAŞTIRMADA KULLANILAN YÖNTEMLER VE ANALİZ AŞAMALARI.....	180
4.5.1 Borsa Endeksleri ile İlgili Değişkenler	181
4.5.2 Bilişsel Katman: Doğal Dil İşleme ve Derin Öğrenme Yöntemleri ile Finansal Bağlam, Fikir ve Duygu Sınıflandırması	183
4.5.2.1 Doğal Dil İşleme ve Sözcük Temsilleri.....	183

4.5.2.2 Derin Öğrenme ve Yinelemeli Sinir Ağları.....	200
4.5.2.3 Finansal Bağlam Sınıflandırıcı.....	215
4.5.2.4 Finansal Fikir ve Duygu Sınıflandırıcı	226
4.5.2.5 Bilişsel Özniteliklere İlişkinin Analizi	238
4.5.3 Davranışsal Katman: Etki Tepki Analizi.....	242
4.5.4 Sosyal Katman: Sosyal Değişkenlerin Finansal Kararlar ile İlişkisi	243
4.5.5 Bütünleşik Karar Modeli ve Borsa İstanbul Pay Piyasası Endeksleri ile İlişkisi.....	244
4.5.5.1 Tweet Öznitelikleri ile BIST Pay Endekslerinin Genel İlişkisi	244
4.5.5.1.1 Açılış Değeri.....	246
4.5.5.1.2 İşlem Hacmi.....	250
4.5.5.1.3 Volatilité	255
4.5.5.1.4 Getiri	259
4.5.5.2 Öznitelik Seçimi.....	260
4.5.5.3 Makine Öğrenmesi Yöntemleri ile Bütünleşik Karar Analizi.....	263
4.5.5.3.1 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini	263
4.5.5.3.1.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini.....	267
4.5.5.3.1.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini.....	270
4.5.5.3.1.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini.....	273
4.5.5.3.1.4 XGBoost Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini.....	275
4.5.5.3.1.5 Açılış Değerlerinin Tahmininde Özniteliklerin Önem Düzeyleri	277
4.5.5.3.2 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini	285
4.5.5.3.2.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini	286
4.5.5.3.2.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini	288
4.5.5.3.2.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini	289
4.5.5.3.2.4 XGBoost Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini.....	291
4.5.5.3.2.5 İşlem Hacimlerinin Tahmininde Öznitelik Önem Düzeyleri	292

4.5.5.3.3 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Getirilerinin Tahmini.....	300
4.5.5.3.4 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Volatilitelerinin Tahmini.....	301
4.5.5.3.4.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini	303
4.5.5.3.4.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini	304
4.5.5.3.4.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini	305
4.5.5.3.4.4 XGBoost Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini	307
4.5.5.3.4.5 Volatilitelerin Tahmininde Öznitelik Önem Düzeyleri	308
4.6 BULGULARIN DEĞERLENDİRİLMESİ VE ÖNERİLER	315
SONUÇ	332
YARARLANILAN KAYNAKLAR.....	335
ÖZGEÇMİŞ.....	374

TABLolar

Tablo 1.1: Varlığın Faydası	15
Tablo 1.2: İki Bilişsel Sistem	68
Tablo 2.1: Geleneksel Medya Sosyal Medya Karşılaştırması.....	89
Tablo 4.1: Biçimsel Dil Modeli İçin Finansal Derlem Kaynakları	167
Tablo 4.2: Finansal Derlem Kaynağına Göre Belge Sayısı	171
Tablo 4.3: Twitter Kullanıcısı Profil Bilgileri Örneği	176
Tablo 4.4: Birlikte Bulunma Matrisi	189
Tablo 4.5: GloVe Örneği	195
Tablo 4.6: Ayırt Etme Matrisi	214
Tablo 4.7: Sınıflandırma Performansı Değerlendirme Metrikleri.....	214
Tablo 4.8: RNN-GRU Bağlam Sınıflandırıcı Model Özeti	217
Tablo 4.9: RNN-LSTM ile Bağlam Sınıflandırıcı Model Özeti.....	218
Tablo 4.10: Tahminlerin Doğruluk Durumuna Göre Dağılımı	220
Tablo 4.11: Bağlam Sınıflandırıcısının Eğitim Performans Sonuçları	222
Tablo 4.12: Yatırım Kuruluşları Uzmanlarınca Etiketlemiş Veri Örneği.....	228
Tablo 4.13: Duygu-Fikir Etiketli Tweet Örnekleri	230
Tablo 4.14: BERT Fikir ve Duygu Sınıflandırıcı Model Özeti.....	235
Tablo 4.15: Ön Eğitimli BERT Mimarisi ile Üç Kategorili Sınıflandırma Performansı....	236
Tablo 4.16: Üç Sınıflı Fikir ve Duygu Sınıflandırıcısının Eğitim Performans Sonuçları ..	237
Tablo 4.17: Üç Sınıflı Fikir ve Duygu Sınıflandırıcısının Kategorilere Göre Eğitim Performans Sonuçları	238
Tablo 4.18: Yükseliş Potansiyeli İndeksleri	239
Tablo 4.19: Sözbirliği İndeksleri.....	240
Tablo 4.20: Nötr Tweet Öznitelikleri.....	241
Tablo 4.21: Pozitif-Negatif Tweet Öznitelikleri	241
Tablo 4.22: Davranışsal Öznitelikler	242
Tablo 4.23: Sosyal Öznitelikler	243
Tablo 4.24: Pearson Korelasyon Katsayısının Nitelendirilmesi	245
Tablo 4.25: Korelasyon Tablosu (Açılış Değeri)	248
Tablo 4.26: Korelasyon Tablosu (İşlem Hacmi)	252
Tablo 4.27: Korelasyon Tablosu (Volatilite).....	256
Tablo 4.28: Korelasyon Tablosu (Getiri).....	260
Tablo 4.29: BIST Pay Endekslerinin Açılış Değerleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları	267
Tablo 4.30: BIST Pay Endekslerinin İşlem Hacimleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları	286
Tablo 4.31: BIST Pay Endekslerinin Getirileri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları	301

Tablo 4.32: BIST Pay Endekslerinin Volatiliteleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları	302



ŞEKİLLER

Şekil 1.1: Parlaklık Algısının Referansa Bağlılığı	17
Şekil 2.1: Büyük Verinin 3V'si	78
Şekil 2.2: Büyük Verinin Ayırt Edici Özellikleri	81
Şekil 2.3: Yapay Zekânın Kapsamı	82
Şekil 2.4: Yapay Zekâ ve Finansal İşlemlerin Tarihsel Gelişimi	86
Şekil 4.1: Araştırma Kapsamında Toplanan Twitter Kullanıcı Veri Ağı	175
Şekil 4.2: Bireysel Yatırımcı Karar Modeli	180
Şekil 4.3: One-Hot Encoding Örneği	186
Şekil 4.4: Skip-Gram Yaklaşımı	191
Şekil 4.5: Skip-Gram Örneği	193
Şekil 4.6: CBOW Yaklaşımı	194
Şekil 4.7: Doğal Dil İşleme Akışı	199
Şekil 4.8: Biyolojik Sinir Hücresi-Yapay Sinir Hücresi Karşılaştırması	201
Şekil 4.9: Yapay Sinir Ağı	202
Şekil 4.10: Kayıp Fonksiyonu ile Ağıın Çıktısı Arasındaki İlişki	203
Şekil 4.11: Yapay Sinir Ağı Öğrenme Süreci	204
Şekil 4.12: Aktivasyon Fonksiyonları	205
Şekil 4.13: Yinelemeli Ağ Mimarisi	207
Şekil 4.14: LSTM Durum Hücresi	209
Şekil 4.15: LSTM Sinir Hücresini Düzenleyen Kapılar	210
Şekil 4.16: LSTM Durum Hücresin Yeni Durumu	210
Şekil 4.17: GRU Sinir Hücresini Düzenleyen Kapılar	211
Şekil 4.18: RNN Mimarilerinin Karşılaştırması	212
Şekil 4.19: Öğrenme Başarımı	213
Şekil 4.20: Alıcı İşletim Karakteristik Eğrisi [ROC]	223
Şekil 4.21: Bireysel Yatırımcı Duygu ve Fikir Analizi Yaklaşımı Özeti	319

GRAFİKLER

Grafik 1.1: Beklenti Teorisi	18
Grafik 2.1: Küresel Verinin Boyutu.....	79
Grafik 2.2: Dünyada Genelinde En Çok Kullanılan Sosyal Medya Platformları	91
Grafik 2.3: Türkiye’de En Çok Kullanılan Sosyal Medya Platformları	92
Grafik 2.4: Kurumsal Yatırımcıların Veri Topladıkları Kaynaklar	95
Grafik 2.5: Yatırım Fikirlerini Araştırmak İçin Başvurulan Kaynaklar	97
Grafik 2.6: Yatırım Bilgisi Kaynağı.....	98
Grafik 2.7: Yatırımla İlgili Bilgi Almak İçin Kullanılan Sosyal Medya Kanalları.....	99
Grafik 2.8: Yaş Gruplarına Göre Yatırım Bilgisi Almak İçin Başvurulan Sosyal Medya Platformları.....	100
Grafik 4.1: Öğrenme Oranının Belirlenmesi	206
Grafik 4.2: Borsa Derlemi	216
Grafik 4.3: Test Veri Kümesi Doğruluk-Hata Matrisi	220
Grafik 4.4: RNN GRU Mimarisi ile Eğitim ve Doğrulama Başarımı	221
Grafik 4.5: RNN GRU Kesinlik, Duyarlılık ve AUC Oranları	222
Grafik 4.6: Tweetlerin Borsa İle İlgisine Göre Bağlam Sınıflandırıcı Sonuçları.....	224
Grafik 4.7: 2012-2020 Tarihleri Arasında Borsa Bağlamında Günlük Tweetler	225
Grafik 4.8: 2012-2018 Tarihleri Arasında Borsa Bağlamında Günlük Tweetler ve Dikkat Çekici Artışlar.....	225
Grafik 4.9: Üç Sınıflı Finansal Fikir ve Duygu Derleminin Duyguya ve Edinildiği Kaynağa Göre Dağılımı	232
Grafik 4.10: Ön Eğitimli BERT Mimarisi ile Üç Sınıflı Eğitim ve Doğrulama Başarımı	236
Grafik 4.11: Lineer Regresyon Çapraz Doğrulama Skoru (Açılış Değeri)	269
Grafik 4.12: Lasso Regresyon Çapraz Doğrulama Skoru (Açılış Değeri)	272
Grafik 4.13: Rassal Orman Çapraz Doğrulama Skoru (Açılış Değeri).....	274
Grafik 4.14: XGBoost Çapraz Doğrulama Skoru (Açılış Değeri)	276
Grafik 4.15: XU030 Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	279
Grafik 4.16: XU100 Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	280
Grafik 4.17: XUTUM Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	280
Grafik 4.18: XUSIN Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	281
Grafik 4.19: XUMAL Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	282
Grafik 4.20: XUTEK Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	283

Grafik 4.21: XUHIZ Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	284
Grafik 4.22: Lineer Regresyon Çapraz Doğrulama Skoru (İşlem Hacmi)	287
Grafik 4.23: Lasso Regresyon Çapraz Doğrulama Skoru (İşlem Hacmi)	289
Grafik 4.24: Rassal Orman Çapraz Doğrulama Skoru (İşlem Hacmi).....	290
Grafik 4.25: XGBoost Çapraz Doğrulama Skoru (İşlem Hacmi).....	292
Grafik 4.26: XU030 Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	293
Grafik 4.27: XU100 Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	294
Grafik 4.28: XUTUM Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	295
Grafik 4.29: XUSIN Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	296
Grafik 4.30: XUMAL Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	297
Grafik 4.31: XUTEK Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	298
Grafik 4.32: XUHIZ Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	299
Grafik 4.33: Lineer Regresyon Çapraz Doğrulama Skoru (Parkinson Volatilitesi)	303
Grafik 4.34: Lasso Regresyon Çapraz Doğrulama Skoru (Parkinson Volatilitesi).....	305
Grafik 4.35: Rassal Orman Çapraz Doğrulama Skoru (Parkinson Volatilitesi).....	306
Grafik 4.36: XGBoost Çapraz Doğrulama Skoru (Parkinson Volatilitesi)	308
Grafik 4.37: XU030 Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	309
Grafik 4.38: XU100 Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	310
Grafik 4.39: XUTUM Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	311
Grafik 4.40: XUSIN Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	312
Grafik 4.41: XUMAL Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	313
Grafik 4.42: XUTEK Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	314
Grafik 4.43: XUHIZ Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi.....	315
Grafik 4.44: Makine Öğrenmesi Yöntemlerinin Performanslarının Karşılaştırılması.....	331

KISALTMALAR

ABD	: Amerika Birleşik Devletleri
AMEX	: The American Stock Exchange (Amerikan Borsası)
ANN	: Artificial Neural Networks (Yapay Sinir Ağları)
API	: Application Programming Interface (Uygulama Programlama Arayüzü)
Ar-Ge	: Araştırma ve Geliştirme
ASE	: Amman Stock Exchange (Amman Menkul Kıymetler Borsası)
AUC	: Area Under Curves (Eğri Altında Kalan Alan)
B2B	: Business to Business (İşletmeden İşletmeye)
B2C	: Business to Consumer (İşletmeden Tüketicieye)
BERT	: Bidirectional Encoder Representations from Transformers (Transformerstan Çift Yönlü Kodlayıcı Temsilleri)
BIST	: Borsa İstanbul
bkz.	: Bakınız
BOW	: Bag of Words (Kelime Çantası)
CAPM	: Capital Asset Pricing Model (Sermaye Varlıklarını Fiyatlama Modeli)
CBDDO	: Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı
CBOE	: Chicago Board Options Exchange (Chicago Opsiyon Borsası)
CBOW	: Continuous Bag of Words (Kesintisiz Kelime Çantası)
CDS	: Credit Default Swap (Kredi Temerrüt Swapı)
CNN	: Convolutional Neural Network (Evrişimsel Sinir Ağları)
csv	: Comma-Separated Values (Virgülle Ayrılan Değerler)
Çev.	: Çeviren
DJIA	: Dow Jones Industrial Average (Dow Jones Borsası Endüstri Endeksi)
ed.	: Edition (Baskı)
Ed.	: Editör
Eds.	: Editors (Editörler)
ELMo	: Embeddings from Language Models (Dil Modellerinden Gömülmeler)
eWOM	: Electronic Word of Mouth (Elektronik Ağızdan Ağıza)
FAV	: Beğeni Tweet
FGNHI	: Facebook's Gross National Happiness Index (Facebook'un Gayri Safi Milli Mutluluk Endeksi)
FN	: False Negative (Yanlış Negatif)
FP	: False Positive (Yanlış Pozitif)
FTSE 100	: Financial Times Stock Exchange Index (Financial Times Borsa Endeksi)

GloVe	: Global Vectors for Word Representation (Kelime Temsili için Küresel Vektörler)
GRU	: Gated Recurrent Units (Geçitli Tekrarlayan Birimler)
GSYİH	: Gayri Safi Yurt İçi Hâsıla
HAN	: Hierarchical Attention Network (Hiyerarşik Dikkat Ağı)
HPC	: High Performance Computing (Yüksek Başarımlı Hesaplama)
IoT	: Internet of Things (Nesnelerin İnterneti)
json	: JavaScript Object Notation (JavaScript Nesnesi Gösterimi)
KAP	: Kamuyu Aydınlatma Platformu
kNN	: K-Nearest Neighbors (K-En Yakın Komşu)
KOSPI	: Korean Composite Stock Price Indexes (Kore Bileşik Hisse Senedi Fiyat Endeksi)
LDA	: Latent Dirichlet Allocation (Gizli Dirichlet Ayırımı)
LibSVM	: Library for Support Vector Machines (Destek Vektör Makineleri için Kitaplık)
LSA	: Latent Semantic Analysis (Gizli Semantik Analiz)
LSTM	: Long Short-Term Memory (Uzun Kısa Dönem Bellek)
LTAM	: Local Trend Association Measure (Yerel Eğilim İlişkilendirme Ölçüsü)
MACD	: Moving Average Convergence Divergence (Hareketli Ortalama Yakınsama İraksama)
MAP	: Moving Approximation (Hareketli Yaklaşım)
MCC	: Matthew's Correlation Coefficient (Matthews Korelasyon Katsayısı)
MLM	: Masked Language Models (Maskelenmiş Dil Modeli)
MLP	: Multilayer Perceptron (Çok Katmanlı Algılayıcı)
MSH	: Morgan Stanley Technology Index (Morgan Stanley Teknoloji Endeksi)
NASAA	: The North American Securities Administrators Association (Kuzey Amerika Menkul Kıymet Yöneticileri Derneği)
NB	: Naive Bayes
NIFTY 50	: Indian Stock Market Index (Nifty 50 Endeksi)
NIKKEI 225	: Prime Market of the Tokyo Stock Exchange (Tokyo Menkul Kıymetler Borsası)
NLP	: Natural Language Processing (Doğal Dil İşleme)
NNet	: Neural Network (Yapay Sinir Ağı)
NoSQL	: Not only SQL (SQL ve Daha Fazlası)
NSE	: National Stock Exchange of India (Hindistan Ulusal Borsası)
NSP	: Next Sentence Prediction (Sıradaki Cümlemin Tahmini)
NYSE	: The New York Stock Exchange (New York Menkul Kıymetler Borsası)

OECD	: The Organization for Economic Cooperation and Development (İktisadi İşbirliği ve Gelişme Teşkilatı)
RNN	: Recurrent Neural Networks (Yinelemeli Sinir Ağları)
ROC	: Receiver Operating Curve (Alıcı İşletim Karakteristik Eğrisi)
RSS	: Really Simple Syndication (Gerçekten Basit Dağıtım)
RT	: Retweet
S&P 100	: Standard and Poor's 100 Index (Standard & Poor's 100 Endeksi)
S&P 500	: Standard and Poor's 500 Index (Standard & Poor's 500 Endeksi)
SEC	: Securities and Exchange Commission (Amerika Birleşik Devletleri Menkul Kıymetler ve Borsa Komisyonu)
SHOMP	: Shanghai Stock Exchange (Şanghay Kompozit Endeksi)
SPK	: Sermaye Piyasası Kurulu
SQL	: Structured Query Language (Yapılandırılmış Sorgu Dili)
STB	: Sanayi ve Teknoloji Bakanlığı
SVM	: Support Vector Machine (Destek Vektör Makinesi)
TF-CDF	: Category Frequency x Category Discrimination (Kategori Sıklığı x Kategori Ayrımı)
TF-IDF	: Term Frequency - Inverse Document Frequency (Terim Frekansı - Ters Belge Frekansı)
TN	: True Negative (Doğru Negatif)
TP	: True Positive (Doğru Pozitif)
Trans.	: Translator (Çeviren)
TW	: Tweet
ULMFiT	: Universal Language Model Fine-Tuning (Evrensel Dil Modeli İnce Ayarı)
VADER	: Valence Aware Dictionary and Sentiment Reasoner (Değerlik Farkındalık Sözlüğü ve Duygu Akıl Yürütücüsü)
VAR	: Vector Autoregression (Vektör Otoregresyon)
vb.	: ve benzeri
vd.	: ve diğerleri
VIX	: Chicago Board Options Exchange Volatility Index (Korku Endeksi)
VSM	: Vector Space Model (Vektör Uzay Modeli)
WOM	: Word of Mouth (Ağızdan Ağıza)
Word2Vec	: Words to Vectors (Kelimelerden Vektörlere)
XU030	: Borsa İstanbul 30 Endeksi
XU100	: Borsa İstanbul 100 endeksi
XUHIZ	: Borsa İstanbul Hizmetler Endeksi
XUMAL	: Borsa İstanbul Mali Endeksi

XUSIN : Borsa İstanbul Sınai Endeksi
XUTEK : Borsa İstanbul Teknoloji Endeksi
XUTUM : Borsa İstanbul Tüm Endeksi



GİRİŞ

Geleneksel finans kuramları temelde insanların rasyonel olduđu varsayımı üzerine inşa edilmiştir. Bu kuramlardaki karar vericilerin rasyonalitesi ise, neoklasik iktisadın *beklenen fayda teorisindeki* rasyonel insan kurgusuna dayanmaktadır. Bu soyut modele göre insanlar, yatırım alternatifleri hakkında noksansız bilgiye sahip, belirsizlik koşullarında beklenen faydasının maksimizasyonu için kusursuz hesaplar yapabilen, yeni bir bilgi geldiğinde beklentilerini anında ve mükemmel bir şekilde Bayes kuralına uygun olarak güncelleyebilen, bilişsel kısıtları olmayan, duygularından arınmış ve sosyal çevresinin etkisinden azade karar vericilerdir. Geleneksel finans kuramlarına göre insanlar rasyonel olduđu için, yani hem tam bilgiye hem de bu bilgiyi kusursuz bir şekilde işleyerek fiyatlara yansıtabilecek bilişsel yeteneklere sahip olduđu için, piyasalar da etkindir. Dolayısıyla piyasada fiyatlar rassal bir yürüyüşle ve mevcut tüm bilgiyi yansıtacak şekilde rasyonel insanlar tarafından şekillendirilmektedir. Bu çerçevede geleneksel finans teorisinin yapı taşlarından biri olan etkin piyasa, yeniden değerlemenin mümkün olmadığı, bireysel yatırımcıların piyasadan daha iyi performans gösteremediğı, fiyat hareketlerini önceden tahmin etmenin mümkün olmadığı, özetle her türlü bilginin fiyatlara yansımış olduđu piyasayı ifade etmektedir. Rasyonalite varsayımının ruhuna sirayet ettiğı geleneksel finansın bir diğerk yapı taşı, modern portföy teorisidir. Portföyün beklenen getirisine, standart sapmasına ve portföy içerisindeki korelasyonlara dayanan bu teoride rasyonalite anlayışı, riskten kaçman yani riskini minimize ederken beklenen getirileri maksimize etmeyi amaçlayan birey olarak tezahür etmiştir. Modern portföy teorisi üzerine inşa edilmiş olan sermaye varlıklarını fiyatlama modeli de benzer bir rasyonalite anlayışı ile beklenen getirilerin sadece riskin bir fonksiyonu olduđu görüşünü savunmuştur.

Geleneksel finansın yapı taşı niteliğindeki bu kuramlar, fizik ve matematiğe öykünen iktisat biliminin pozitivist felsefeden beslenen rasyonalite anlayışı çerçevesinde şekillendiğı için, psikolojik faktörleri çözümlenmenin dışında bırakmıştır. Özellikle psikoloji alanında kaydedilen ilerlemeler ve geleneksel yaklaşımların soyut modellerinde dışlanmış olan

insanın gerçekliğine ilişkin elde edilen bulgular geleneksel finans varsayımlarını ve bu varsayımlar üzerine inşa edilen kuramları tartışılır hale getirmiştir. Böylece geleneksel finans kuramlarının, insanların nasıl karar vermesi gerektiğini tarif eden normatif ve soyut modellerinin yerini, özellikle psikoloji alanından elde edilen bulgulardan beslenerek insanların gerçekte nasıl karar verdiklerini açıklamaya çalışan modeller almaya başlamıştır. Bu modeller üzerine inşa edilen davranışsal finans, psikolojik faktörlerin insanların finansal kararları ve dolayısıyla piyasalar üzerindeki etkilerini açıklamaya çalışan bir alan olarak ortaya çıkmıştır.

Geleneksel finansın temel hipotezlerine yapılan ilk itirazlardan biri, sınırsız rasyonalite varsayımıyla ilgilidir. Literatürde *sınırlı rasyonalite* yaklaşımı olarak kendisine yer edinen bu itiraz, insanların hem tam bilgi sahibi olmasının hem de tüm hesaplamaları kusursuz bir şekilde yapabilmesinin gerçeklikle bağdaşmadığını savunmaktadır. Başka bir ifadeyle fiili durumda insan, alternatifler hakkında *tam bilgi* sahibi olmayan, isabetli bir karar verebilmek için gerekli olan tüm hesaplamaların yapılmasında *bilişsel kısıtlara* sahip bir model ortaya koymaktadır. Esasen bu görüş, tümüyle etkin bir piyasanın mümkün olmadığının ilk sinyallerini vermiştir. Nitekim daha sonra yapılan pek çok çalışma etkin bir piyasanın teorik bir idealden öteye geçemediğini çeşitli piyasa anomalilerinin tespitiyle ortaya koymuş, böylece geleneksel finansın yapı taşlarından biri olan etkin piyasalar hipotezi sorgulanır hale gelmiştir. Davranışsal finans, piyasaların etkin olmamasının temel nedenini insanın doğasından kaynaklanan bilişsel kestirmeleri, duyguları ve eğilimleri ile açıklamaktadır. Ancak ilk kuşak davranışsal finans araştırmaları, insanların seçim yaparken başvurduğu bilişsel kestirmelere ve karar verirken yaptığı hatalara odaklanmıştır. Başka bir ifadeyle 1980'lerin başından itibaren ilk nesil davranışsal finans çalışmaları, geleneksel finansı standart kabul etmiş, insanları genellikle bilişsel ve duygusal kusurlarına yenik düşen ve rasyonaliteye giden yolda yanlış yönlendirilen, yani irrasyonel modeller ile tanımlamışlardır. Benzer şekilde piyasaların etkinliğini test eden ve etkin olmadığı sonucuna ulaşan çalışmaların da bulgularını normalden sapma anlamında anomali olarak ifade etmelerinin temelinde yatan neden de budur.

Davranışsal finansın ikinci kuşağı normal insanların daha gerçekçi bir portresini sunmaktadır. Başka bir ifadeyle davranışsal finansın ikinci kuşağı, insan davranışlarını anlamak için daha gerçekçi hipotezler ortaya koymuştur. Bu nedenle ikinci kuşak davranışsal finans, insanların isteklerini ve bunların faydalarını normal kabul ederek analiz yapmaktadır. Normal istekleri kusurlardan ayırmakta ve kusurlu davranışlardan kaçınma konusunda rehberlik sunmaktadır.

Bu noktada davranışsal finansa yöneltilen en temel eleştiri ise, geleneksel finans kuramlarındaki izole birey varsayımının (sürü psikolojisi çalışmaları dışında) büyük oranda her iki kuşakta da korunmuş olmasıdır. İnsanın sosyal yönünün göz önünde bulundurulması halinde, davranışın şekillenmesinde sosyal ağların, sosyal kimliklerin, sosyal statünün ve insanlar arasındaki etkileşim modellerinin pay sahibi olduğu açıktır. Dolayısıyla insanların karar verme biçimlerinin ve davranışlarının analizinde insanın sosyalliğinin de çözümlemeye dâhil edilmesi gerekmektedir. Bu bağlamda üçüncü kuşak davranışsal finans hem bireyin iradeli eylemlerinin hem de sosyolojik etkinin çözümlemeye dâhil edildiği bir duruma evrilmektedir.

Davranışsal finansın bu evriminde özellikle sosyal medya platformlarından elde edilen etkileşim verilerinin önemli bir payı vardır. Sosyal medya, insanların sosyal çevreleri ile dijital ortamlarda etkileşim kurarak duygu ve düşüncelerini ifade edebilmesine ve bunları sosyal ağları üzerinden yayabilmesine imkân sağlayan platformlardır. İletişimde dijital platformların her geçen gün daha fazla tercih edilmesi, sosyal hayata ilişkin elde edilmesi en zor verilerin pasif ve gerçek zamanlı olarak makroskobik ve mikroskobik ölçekte toplanmasını sağlamaktadır. Böylece çevrimiçi platformlardan elde edilen etkileşim verileri davranışsal finansın ilk iki kuşağında göz ardı edilen *sosyal bireyin* modellemesine olanak tanımaktadır. Başka bir ifadeyle gerçek dünyada karar vermenin daha önce gözlemlenmesi mümkün olmayan sosyal süreçleri hakkında daha fazla fikir vermekte ve gelecekteki eylem ve davranışların daha başarılı bir şekilde tahmin edilmesine imkân sunmaktadır.

Sosyal medya sadece insan davranışlarına değil, fakat aynı zamanda insan eylemlerine odaklanma imkânı da sunmaktadır. Zira insanı anlamaya dönük çalışmalara bakıldığında bir kısmının *davranış kuramları* çerçevesinde, diğer bir kısmının ise *eylem kuramları* çerçevesinde öbeklendiği görülmektedir. Davranışsal yaklaşım çerçevesinde öbeklenenler kuramsal kabullerini bir uyarıcı ve tepkileri bağlamıyla sınırlandırırken, eylemsel yaklaşımı benimseyenler önermelerini yönelimsel eylem üzerine yapmaktadırlar. Bu bağlamda davranış, özel bir anlamı ya da yönelmişliği olmayan, tümüyle mekanik bedensel bir hareket, ya da belirli bir tür uyarıcıya veya dürtüye cevap niteliğinde otomatik bir reaksiyon iken; eylem, ilgili kişiler için öznel bir anlamı veya amacı olan bilinçli bir faaliyettir. Eylem öğrenmeyle, öğrenme ise öğrenenin zihnindeki, inanç, amaç, beklenti gibi öğelerin değişmesiyle ilgilidir. Bireysel yatırımcıların sosyal medyada bilgi arayışları bir yönüyle inanç ve beklenti inşa etmekle, yani öğrenmeyle ilgilidir. Bu anlamda sosyal medya paylaşımlarının içeriğinin fikri ve duygusal yöneliminin analizine imkân vererek bilişsel boyutta, etkileşim örüntülerini (tweet, retweet, beğeni gibi) analize imkân vererek davranışsal boyutta çözümlemelerle yeni bir bakış açısı kazandırabilir.

Davranışsal finansın ilk iki kuşağında ihmal edilen konulardan bir diğeri eksik bilgiye sahip insandır. Bu bağlamda sosyal medya, geleneksel finansın varsaydığı gibi tam bilgi sahibi olmayan insanın, bilgi arayışlarını modelleme imkânı da vermektedir. Başka bir deyişle bilginin bireyler arasında dağılmış olması durumunda (-ki kişisel farklılıklardan dolayı genelde böyledir) nasıl elde edildiği ya da bireyler arasında nasıl aktarıldığı, nasıl kullanıldığı, farklı zihinlerde var olan bu bilginin nasıl birbiriyle uyumlu hale gelebildiğinin araştırılmasına imkân verecek veriler sunmaktadır. Dolayısıyla sosyal medyadan elde edilecek veriler kısmi bilgiye sahip olan insanların etkileşimlerinin ve bilginin iletim yollarının finansal piyasalarla ilişkisini aydınlatma potansiyeline sahiptir. Bilginin insanlar arasında homojen dağılmaması, bir başka deyişle asimetrik olması aynı zamanda eksik bilgiye sahip insanın yönlendirilebileceğine de işaret etmektedir. Dolayısıyla insanların beklenti ve inançlarının inşasında, sosyal etkileşimlerinin ya da kısmen sosyal medyanın rolü de analiz edilmesi gereken bir başka boyutu temsil etmektedir.

Bütün bunlara ek olarak sosyal medya bir taraftan mikro düzeyde bireylerin kararlarında sosyal çevresinin etkisini çözümlene imkânı sunarken diğer taraftan makro düzeyde kolektif yönelimi gözlemleme fırsatı da vermektedir. Bu gözlem insanların özellikle ekonomi ve finansla ilgili konularda paylaştıkları düşünce, yorum ve görüşlerinin makineler tarafından anlamlandırılıp ve buradan elde edilen bilginin karar süreçlerine dâhil edilmesinde oldukça önemlidir. Özellikle algoritmik finansal işlemler açısından düşünüldüğünde, ekonomi ve finansla ilgili sayısal ve yapılandırılmış verilerin modern finans çerçevesinde analizlerine ek olarak sosyal medya ve haber siteleri gibi dijital platformlarından elde edilecek yöneltimsel verilerle desteklenmesi yatırım kararlarının başarı performansını artırma potansiyeline sahiptir.

Sosyal medyadan elde edilen verilerin hızı, büyüklüğü ve çeşitliliği analiz yöntemlerini de temelden etkilemektedir. Zira büyük verinin insan gücüyle anlamlandırılmasının mümkün olmaması nedeniyle eğilimlerin ve örüntülerin ortaya çıkartılmasında disiplinlerarası yaklaşımlara ihtiyaç duyulmaktadır. Bu nedenle bilgisayar bilimleri, istatistik, matematik ve alan bilgisini kullanan yapay zekâ, makine öğrenmesi, veri analitiği gibi yaklaşımlar davranışsal finans çalışmalarına yeni bir soluk getirmektedir. Analiz araçları, veri seti ve bakış açısındaki bu köklü değişim esasında bir paradigma değişikliğinin habercisi olma potansiyeline de sahiptir.

Bu çalışma üçüncü kuşak davranışsal finans bakış açısına sahiptir. Başka bir deyişle bu çalışmada geleneksel finansın genel kabul görmüş *tam bilgi, rasyonel insan* ve *izole birey* varsayımları yerine, eksik bilgisini tamamlamaya çalışan, sınırlı ve ekolojik rasyonaliteye sahip, sosyal çevresinin ve duygularının karar süreçlerinde pay sahibi olduğu bir insan modeli esas alınmaktadır. Diğer bir ifadeyle bu çalışmanın amacı, eksik bilgiye ve bilişsel kısıtlara sahip bireysel yatırımcıların hem iradeli eylemlerinin altında yatan beklenti inşa etme (öğrenme) sürecinin hem davranışsal eğilimlerinin hem de sosyal çevresinin üç katmanlı bir model çerçevesinde kararları üzerindeki bütünleşik etkisini araştırmaktır. Bu insan modeline ilişkin veriler duygu, düşünce ve fikirlerini dijital bir ortamda toplumsal iletişime ve etkileşime izin veren sosyal medya platformlarından toplanmaktadır. Sosyal medya

platformlarından toplanan veriler miktar, hız ve çeşitlilik açısından büyük veri olarak nitelendirilmektedir ve genellikle yapılandırılmamış verilerdir. Bu nedenle bu çalışma kapsamında çerçevesi çizilen insan varsayımına göre bireysel yatırımcıların eylem ve davranışları küçük bir örneklem yerine anakütleyle yakın bir veri seti üzerinden yapay zekâ ve makine öğrenmesi yöntemleri yardımıyla çözümlemeye çalışılmaktadır. Bu amaçla bireysel yatırımcılar için bilişsel, davranışsal ve sosyal boyutları göz önünde bulunduran bütünsel bir karar modeli elde edilmekte ve bu karar modelinin Borsa İstanbul (BIST) pay endeksleri ile ilişkisi araştırılmaktadır.

Çalışma dört bölümden oluşmaktadır. Birinci bölümde geleneksel finans teorisine hâkim olan rasyonalite yaklaşımı kuramsal çerçevede ortaya konularak, normatif ve betimsel karar teorileri bağlamında ele alınmaktadır. Bu bölümde ayrıca; iktisat alanında doğup farklı alanlara nüfuz eden rasyonalite anlayışının geleneksel finans teorisinin inşasında oynadığı role; geleneksel finansın soyut insan modellerine itiraz ederken, psikolojik ve sosyolojik faktörlerin yatırımcıların finansal kararları ve dolayısıyla finansal piyasalar üzerinde etkili olduğunu savunan davranışsal finans teorisine, yatırımcı eğilimlerine ve davranışsal finans teorisinin argümanlarının temel dayanağı niteliğindeki bilişsel kestirmelere değinilmektedir.

Çalışmanın ikinci bölümünde dijitalleşme, büyük veri, sosyal medya ve finans konusu ele alınmakta, büyük veri ve yapay zekâ alanlardaki gelişmelerin finans ile ilişkisi ortaya konulmaktadır. Bu bölümde ayrıca, sosyal medyanın tanımı, yatırımcılar açısından önemi ve değişen finansal olaylara etkisine değinildikten sonra sosyal medya aracılığıyla sermaye piyasasında manipülasyon riskinden ve buna ilişkin düzenlemelerden bahsedilmektedir.

Üçüncü bölümde internet ya da sosyal medyada, bireyler, yatırım kuruluşları, aracı kurumlar ve çeşitli haber kaynakları tarafından paylaşılan finans içerikli metinlerin yapay zekâ algoritmaları ile fikir ve duygu analizi yapılarak elde edilen duygusal yönelimlerinin borsa endeksleri veya hisse senedi fiyatları ile ilişkisini araştırarak çalışmalara değinilmektedir. Başka bir ifadeyle literatür özeti bu bölümde sunulmaktadır.

Dördüncü bölümde Borsa İstanbul örneği üzerinden sosyal medyanın finansal piyasalara etkisi araştırılmaktadır. Bu bölümde araştırmanın amacı, kapsamı, veri setleri, araştırmanın kısıtları ve araştırmada kullanılan yöntemler açıklanmaktadır. Borsa İstanbul’da işlem gören hisse senetleri ile ilgili Twitter’den elde edilen metinlerdeki fikir ve duygunun nicelleştirilmesine, fikir ve duygunun nicelleştirilmesinde kullanılan derin öğrenme yöntemlerine, yine Twitter’den elde edilen sosyal etkileşim verilerinin analizlerinden elde edilen bulgulara değinilmektedir. Bu bölümde ayrıca fikir ve duygu eğilimleri ile sosyal etkileşim değişkenlerinin, BIST Pay Endeksleri (BIST30, BIST100, BISTTUM, BIST SINAİ, BIST TEKNOLOJİ, BIST HİZMETLER, BIST MALİ) ile ilişkisi (açılış fiyatı, işlem hacmi, volatilité ve getiri açısından) makine öğrenmesi yöntemleri ile ampirik olarak test edilmekte ve sonuçlar karşılaştırılmaktadır. Başka bir ifadeyle bu bölümde bireysel yatırımcıların karar süreçlerinde sosyal, bilişsel ve davranışsal boyutların etkisi ampirik olarak ortaya konulmaktadır. Bu bölümde son olarak, elde edilen bulgulara ilişkin genel bir değerlendirme yapılmakta ve bulgular davranışsal finans bağlamında tartışılmaktadır. Ayrıca elde edilen bulgular ışığında borsalara, yatırımcılara, aracı kurumlara, düzenleyici ve denetleyici kurumlara önerilerde bulunmaktadır.

Son olarak, çalışmada ulaşılan bulguların genel bir değerlendirmesi yapılmakta ve gelecek çalışmalar için öneriler sunulmaktadır.

BİRİNCİ BÖLÜM

RASYONALİTE, GELENEKSEL FİNANS VE DAVRANIŞSAL FİNANS

1.1 İNSAN EYLEMLERİNİN VE DAVRANIŞLARININ RASYONALİTESİ

Finans biliminin temel amaçlarından biri, pek çok bilim dalında olduğu gibi, geleceği tahmin edebilmektir. Bu anlamda iktisadi ve finansal seyri tahmin edebilmek ve verilen kararların ne tür sonuçlar doğuracağını kestirebilmek açısından insanların hangi durumlarda hangi eylem ve davranışları sergileyebileceklerini öngörebilmek önem kazanmaktadır. Dolayısıyla insanların iktisadi ve finansal bağlamda tezahür eden eylem ve davranışlarının öngörülebilmesi için karar verme fiilinin, yani aralarında bir tercih yapma zorunluluğu olan çeşitli alternatiflerden birini seçme süreci ve ediminin kuramsal bir çerçevede izah edilmesi gerekmektedir. Bu amaçla ortaya atılan karar teorileri, ya mevcut durumun izahı sadedinde ortalama bir bireyin belirsizlik koşullarında nasıl karar verdiğini veya seçimlerini nasıl yaptığını *tasvir eden* betimsel karar teorisi ya da kural koyucu bir şekilde insanların nasıl karar vermesi gerektiğini soyut modellerle *tarif eden* normatif karar teorisi etrafında toplanmaktadırlar. Literatürde karar teorileri olarak kendisine yer bulan bu karar modellerini rasyonalite kavramı şekillendirmektedir. Bu bağlamda gerek normatif karar teorilerinin, gerekse betimleyici karar teorilerinin doğru bir şekilde anlaşılabilmesi için, rasyonalite kavramının iktisadi açıdan ihtiva ettiği anlamın bilinmesi zaruridir. Bu nedenle geleneksel finans kuramındaki karar teorilerinin şekillenmesinde oynadığı rol kapsamında iktisadi açıdan rasyonalite kavramının tarihsel seyrine kısaca değinilmekte, ardından da bu teoriler izah edilerek davranışsal finans ile ilişkisi ortaya konulmaktadır.

1.1.1 Normatif Karar Teorisi ve Rasyonalite

Klasik iktisat yazınında rasyonalite kavramsal olarak kullanılmamış olsa da anlayışın başlangıcı, kişisel çıkarla ilişkilendirilerek, iktisadın bağımsız bir disiplin olmasında önemli rolü olan Adam Smith'e atfedilmektedir. Adam Smith'in (1776/2016) *Milletlerin Zenginliği* [Wealth of Nations] adlı eserinde ortaya koyduğu insan modelinin kişisel çıkarını düşünen bencil birey olduğu düşüncesiyle modern terimlerle olgunlaşmış bir araçsal rasyonalite ortaya koyduğu yönünde¹ yaygın bir görüş bulunmaktadır (Yılmaz, 2009). Aynı dönemde Ricardo ise Smith'in *kendiliğinden doğan düzen* ve *bencil insan* modelini veri alan soyut bir model geliştirmiştir (O'Brien, 2004). Yine klasik iktisat döneminde insana ilişkin bir diğer tanımlama da faydacı felsefenin etkisi ile Mill (2008) tarafından, *insanın en az emekle en çok hazzın peşinde* olan ve amacına ulaşmak için etkin araçları kullanabilen bir varlık olduğunu varsaymak suretiyle yapılmıştır. Mill bütün apriorici ahlak kuramcıları için faydacı argümanların elzem olduğunu dile getirmiştir. Bu bağlamda Mill'in insana ilişkin apriori ilkesi *azami fayda* ilkesidir. Zira ona göre insan, var olan bilgi durumundan elde edebildikleri asgari emek ve asgari fiziksel özveri miktarıyla, gereksinimlerini azami biçimde elde edebilmenin peşindedir. Dolayısıyla Smith'te çoğu aza tercih etmek şeklinde tezahür eden apriori ilke Mill'de azamileştirme şeklinde ortaya çıkar. Bir diğer deyişle klasik iktisatta bireyin kendi çıkarını takip etmesi rasyonalitenin merkezinde olmakla birlikte, bu eylemin maksimizasyon formunda değil, *çoğu aza tercih etme* ve *azamileştirme* forumunda olduğunun kabulü hâkimdir (Arrow, 1986).

İktisadın neoklasik dönemini temsil eden 1870'lere gelindiğinde Menger, Jevons ve Walras'ın birbirinden bağımsız katkıları ile iktisatta marjinal bir devrim yaşanmıştır. Modern iktisadın temellerinin atıldığı bu dönemde emek, değer kaynağı olmaktan çıkartılırken bunun yerini *marjinal fayda* kavramı almıştır. Jevons'la (1871) birlikte (bireyden ziyade bireylerin toplamı bağlamında (Yay, 2004)) iktisadın fayda kavramı matematiksel bir dilde optimizasyon sorunu olarak ifadesini bulmuştur. Zira ona göre iktisat teorisi tamamen

¹ Ancak Smith'in Ahlaki Duygular Kuramı (1759/2018) ve Ulusların Zenginliği (1776/2016) adlı eserleri birlikte değerlendirildiğinde Smith'in aslında rasyonalist değil, empirisit olduğu görülmektedir.

matematiksel karakterde bir bilimdir. Diğer bir deyişle Jevons insanı haz maksimizasyonu peşinde koşan bir varlık olarak tanımlarken hazzın maksimizasyonunu matematiksel bir hesaplamayla ifade etmiştir. Jevons’a benzer şekilde iktisadın deneysel bir yöntemden ziyade rasyonel bir yöntem olduğunu söylerken rasyonaliteyi fiziğin (Mirowski ve Cook, 1990) ve matematiğin dünyasında arayan Walras (1954), iktisat biliminin öykündüğü mekaniğe uygun olarak (Sarfati, 2011) statik genel denge yaklaşımını ortaya koymuştur. Walras (1954), Smith’in düştüğü yanıltan imtina edebilmek için etiği bilimden ayırmayı önermiştir. Bu durum Smithçi yaklaşımı reddetmenin yolunu açarken iktisat biliminin bugünkü ontolojisini temellendirmiştir. Walras’la birlikte nitel olanın rasyonel kabul edilmeyip, sadece nicel olanın makbul olduğu bir dönemin kapıları aralanmıştır. Teoremler gerçek tiplerden ziyade ideal tiplerden yola çıkılarak ve akıl yürütülerek inşa edilmiştir (Walras, 1954). Böylece modelin ürettiği prototip ile rasyonalitenin gerçekliği örtüştürülmüştür (Sarfati, 2011). Smith’tin modelinde çoğu aza tercih eden insan, Mill’de azami faydayı kovalarken, Jevons ve Walras’ta hazzın matematiksel olarak maksimizasyonunun peşindedir. Bu suretle psikolojik ve sübjektif olan fayda nesnel bir hesaplama birimine dönüştürülmüştür (Kırmızıaltın, 2017).

Marjinalist devrimin bir diğer yüzü olan Menger rasyonalite anlayışını izah ederken, insan eylemini ihtiyacın en mükemmel tatminiyle ilişkilendirmiş, ancak Jevons ve Walras’tan farklı olarak matematiğin diline başvurmamıştır. Menger, iktisadı insan doğasıyla ihtiyaçları arasındaki ilişkinin bilimi olarak; faydayı ise şeyin insan ihtiyaçlarının tatminine hizmet etme kapasitesi olarak tanımlamak suretiyle aslında araçsal rasyonalitenin çerçevesini ihtiyacın en mükemmel şekilde tatmini olarak çizmiştir. Ayrıca Menger insanın amaçların ve araçların bilgisine sahip olduğunu, bu nedenle iktisadın bireyden hareketle anlaşılacağı varsayımını *metodolojik bireycilik* olarak iktisat kuramına yerleştirmiştir (Yılmaz, 2009). Fayda sağlamaya indirgenen bireyin eylemi böylece sadece *kendisi* ile sınırlandırılmıştır (Sarfati, 2011). Diğer bir deyişle Menger, ekonomik faaliyeti bireysel kararların belirlediğini ve insan davranışlarının amaçlı davranışlar olduğunu ifade ederek, esasen araçsal rasyonalite tanımındaki bireyciliğin olgunlaşmasına katkı sağlamıştır (Buğra, 2011).

Walras ve Menger’le birlikte iktisat biliminin temel postülları bütünüyle dönüşüme uğramıştır. Bu dönüşüm aslında iktisat biliminde yeni bir dönemin başlangıcıdır. Zira iktisat teorisinin bugünkü epistemolojik temelleri bu devrimsel dönüşümle atılmıştır. Nitekim finansı temel aksı yapan post modern kapitalizmin iktisadi ve sosyal kurgusu bu teorisinin postüllarıyla ve temel epistemolojisiyle büyük oranda örtüşmektedir (Sarfati, 2011).

Walras ve Jevons’ın iktisat teorisini matematikselleştirme çabasına en büyük desteği Edgeworth, Fisher ve Pareto vermiştir (Yılmaz, 2009). Bu çerçevede insanı kendi çıkarları peşinde koşan bir haz makinesi olarak gören Edgeworth (1881) marjinal faydayla rasyonaliteyi izah ederken bütün hazların aynı birimlerle ölçülebilir olduğu iddiasını ortaya atmıştır. Bu yaklaşım onu rasyonel seçim teorisine ulaşmada önemli bir araç olan farksızlık eğrilerine ulaştırmıştır (Kırmızıaltın, 2017). Her şey nicel olanın dünyasına indirgenmeye çalışılırken duygu artık bilimin sınırlarının dışında bırakılmıştır. Aslında duygunun bilgi kaynağı olmaktan çıkması, Humecu yaklaşımın tümüyle iktisat biliminin dışına atılması demektir. İnsanın **anlama yeteneği**, *hesaplama* ve *rasyo* ile ikame edilerek evrensel bireyin eylem nedeni salt çıkar olarak ilan edilmiştir (Sarfati, 2011).

Fisher da Edgeworth’a benzer şekilde iktisadın fizik gibi nesnel bir bilim olabilmesi için fayda gibi psikolojik faktörlere başvurmadan, matematiksel olarak izah edilmesi gerektiğini savunmuştur. İktisadı doğa bilimlerindeki benzer bir kesinliğe ulaştırmak isteyen Fisher, termodinamik ile iktisat arasında analojiler kurarak fayda, haz gibi psikolojik faktörlerden arındırmaya çalışmıştır (Yılmaz, 2009). İktisadın doğa bilimlerini taklit etmesi gerektiğini düşünen bir diğer isim Pareto’dur. Ancak Pareto rasyonel yöntem ile ampirik yöntemin bir bileşimi olduğunu düşündüğü deneysel [experimental] yöntemi ön plana çıkarmıştır (Marchionatti ve Gambino, 1997). Pareto, Edgeworth’ün farksızlık eğrilerinden yola çıkarak, teorisinin merkezine *seçimi* koymuştur. Pür iktisadi denklemlerin zevk ve acı kavramından bağımsız olarak seçim gerçeğini en sade haliyle ifade edebileceğini, bir başka deyişle farksızlık eğrilerinden türetilebileceğini düşünmüştür (Marchionatti ve Gambino, 1997). Zira ona göre (pür) iktisadın konusu sadece ekonomik güçlerin bir sonucu olarak hareket eden ve “mükemmel hedonist” olan *homo economicus*un soyut kavramıdır (Marchionatti ve

Gambino, 1997). *Homo economicus*, gerçekliği açıklamada bazı noksanlıklar sergilese de iktisadın mekanik bir kesinlikle işleyişini tesis etmek için gerekli olan soyut insan modelidir. Böylece Pareto'un rasyonel *homo economicus*u, iktisattaki baskın rasyonalite anlayışının eylemde bulunanın seçimine bağlı olarak tanımlandığı rasyonel seçime indirgenmesinin öncülüğünü yapmıştır (Kırmızıaltın, 2017). Başka bir deyişle marjinaslitlerin haz ve fayda maksimizasyonu olarak tanımladıkları rasyonalite Pareto'yla birlikte *rasyonel seçime* dönüşmüştür (Kırmızıaltın, 2017).

1930'lu yıllar pozitivistin etkisiyle yüksek bir metodolojik hassasiyete ulaşılan yıllardır. Bu döneme damgasını vuran gelişmelerden biri *ordinalist devrim*, bir diğeri ise *açıklanmış tercihler teorisidir* (Yılmaz, 2009). Ordinalist devrimle birlikte faydanın sayısal olarak ölçülme gerekliliği ortadan kalkarken, bunun yerini kıyaslamacı (ya da sıralamacı) bir yaklaşım almıştır. Hicks ve Allen'in (1934) bu yaklaşımında tercihler skalası kardinal fayda teorisindeki gibi matematiksel olarak hesaplanmamaktadır. Onun yerine bir kayıtsızlık eğrisindeki mal bileşiminin öteki kayıtsızlık eğrisindeki mal bileşimine tercih edilip edilmemesiyle kıyaslanmakta ve bu kıyasa bağlı olarak sıralanmaktadır (Kırmızıaltın, 2017). Böylece ölçümü esas alan *kardinal fayda* teorisinin yerini sıralamayla esas alan *ordinal fayda* teorisi alırken rasyonalitenin tanımı da ölçümden ziyade sıralamayla ilişkilendirilmiş olmaktadır.

Açıklanmış tercihler [revealed preference] teorisi ise belirli bir mal bileşiminin tercih edilmesi durumunda aksiyomatik olarak bireyin faydasının maksimize edildiğini ifade etmektedir. Kuramını Hicks ve Allen'in kardinal fayda teorisi üzerine inşa eden Samuelson (1938), alternatifler arasında seçim yapan tüketicilerin tercihlerinde *tutarlı* olacağını varsayarak esasen *rasyonel seçimin* kapılarını aralayan bir yaklaşımı da ortaya koymuştur. Zira bu varsayımlara göre A'yı B'ye tercih eden bir tüketici, artık B'yi A'ya tercih ettiğini ifade etmeyecektir. Benzer şekilde A'yı B'ye, B'yi C'ye tercih eden biri, A'yı C'ye tercih ediyor demektir. Dolayısıyla bu yaklaşımda tercihlerin derecesi önemini yitirmiştir. Başka bir deyişle faydanın kardinal olarak ölçülmesine gerek kalmadan ordinal bir ölçekte sıraya konulması analiz için yeterlidir. Tüketici talebinin tutarlılık varsayımı tüketicinin amacının

en yüksek tatmini sağlamak olduđu varsayımıyla birlikte ele alındığında tüketicinin mal demetlerinden birini tercih etmesi, o demetin tüketiciye en yüksek faydayı sağlıyor olması anlamı da taşıyacaktı. Böylece Samuelson’da rasyonalite anlayışı hem *tercihlerin sıralanmasına* hem de *insanın tutarlılığına* bağlanmış olmaktadır.

Bu dönemde, pozitivist bakış açısına muhalif görüşler de vardır. Örneğin Knight (1921/2016) insan davranışlarına yön veren dürtülerin sadece iktisadi boyutu olmadığını, sosyal ilişkilerin belirlediği sembolik değerlerin de etkili olabileceğini ifade etmiştir. Bu bağlamda bilgiyi üç türe ayıran Knight (1940), birinci tür bilginin dış dünyaya ait fiziksel gerçekliğin bilgisi; ikinci tür bilginin matematik ve mantığın ortaya koyduğu gerçekliğin bilgisi; üçüncü tür bilginin ise insanın kendi değerler sistemi ile diğer insanların değerler sistemine ulaşarak onların davranışlarını anlayıp yorumlama yeteneği ile temin edilen bilgi olduğunu ifade etmiştir. İnsan davranışları ile ilgili benzer bir kanaate sahip olan Keynes (1937) ise, *belirsizlik koşullarında*² *nesnellikten söz edilemeyeceğini vurgularken, bireylerin tüm bilgilere ulaşmak suretiyle maksimizasyona dönük bir karar alma modeli yerine; eldeki bilgiyi kullanarak sezgiye, alışkanlıklara ve geçmiş deneyimlere dayanarak karar aldıklarını ifade etmiştir.* Buna ilave olarak insanî dürtüler, beklentiler ve psikolojik belirsizliklerle ilgili bir ahlâk bilimi olan iktisadı, bir doğa bilimine çevirmeye kalkışmanın tehlikelerine dikkat çekmiştir (Buğra, 2011).

Pozitivist rasyonel insan modelinin, Knight ve Keynes’le sarsılan itibarını 1950’li yıllara gelindiğinde Friedman (1953) iade etmiştir. Friedman bir teörinin varsayımlarının gerçekçi olması gerekmediğini, önemli olanın varsayımlarla oluşturan yaklaşımların amaca ulaşmak için yeterince iyi tahminler yapıp yapmadığıyla ilgili olduğunu savunmuştur. Basit teorilerin karmaşık olayların yerine tercih edilebileceğini de ifade eden Friedman, insan doğasına

² Sınırlı rasyonalite bağlamında Keynes ile Simon’nun görüşleri mukayese edildiğinde, her ikisinde insanların *niyetsel olarak rasyonel* ancak gerçekte sınırlı olduklarını düşündükleri görülür. Ancak Keynes bunu temel belirsizliğe bağlarken, Simon karmaşıklık ve hesaplama kabiliyeti eksikliğine bağlamaktadır (Yılmaz, 2009).

ilişkin eleştirilerle ilgili olarak ise, insanların doğaları gereği değil, aksi takdirde varlıklarını sürdüremeyecekleri için belirli koşulların onları rasyonel davranmaya ittiğini ifade etmiştir.

Artık genel denge modeli ile iktisada nüfuz etmeye başlayan matematiksel dil 1950’li yıllarda teoriye yerleşmiş ve rasyonalite kavramının kazandığı araçsal form, rasyonel seçim teorisi ile birlikte sağlam bir yapıya bürünmüştür (Yılmaz, 2009). Bir diğer ifadeyle Humecu araçsal rasyonalitenin iktisat teorisindeki araçsal ve deterministik karakteri *belirsizlik koşullarında rasyonel seçim* ya da *beklenen fayda teorisi* olarak ortaya çıkmıştır.

1.1.2 Betimsel Karar Teorileri ve Rasyonalite

İnsanların nasıl karar vermesi gerektiğini tarif eden normatif ve soyut bir model kurmak yerine, ortalama bir bireyin belirsizlik koşullarında nasıl karar aldığını tasvir eden bir takım kuramlar da ortaya konulmuştur. Bu kuramlardan bazıları insanların riskli seçenekler karşısında nasıl bir tutum sergilediğini ya da riskli seçenekler ile kesin şeyler arasında seçimlerini hangi kurallara göre belirlediğini araştırmaya odaklanırken, bazıları piyasayı bir süreç olarak görmek gerektiğini, bazıları ise karar vermenin nöropsikolojik temellerine eğilmek gerektiğini savunmaktadır.

1.1.2.1 Beklenti Teorisi

Karar teorilerinin pek çoğunun arka planında, alternatifler arasında tercih yapmanın ihtiva ettiği belirsizliklerle, şans oyunlarının sonuçlarına ilişkin belirsizliklerin kıyaslanabileceğine dair bir ön kabul vardır. Dolayısıyla karar teorisyenleri kuramlarını genellikle şans oyunları deneyleri üzerinden şekillendirirler. Bunun teorik anlamda ilk tezahürleri, her sonucun olasılığına göre ağırlıklandırıldığı, muhtemel sonuçların ağırlıklı bir ortalamasına tekabül eden beklenen değer [expected value] hesaplamalarında görülmektedir. Örneğin %70 olasılıkla 1.000 TL kazanmak ile %30 olasılıkla 100 TL kazanmanın beklenen değeri 730 TL ($0,7 \times 1.000 + 0,3 \times 100$) olarak hesaplanır ve bu teori insanların bu sonuca bakarak karar verdiğini varsayar (Kahneman, 2011). Ancak bu varsayımın aksine, beklenen değeri 730 TL

olan bir riskli seçenek ile mesela 720 TL’lik kesin bir kazanç arasında seçme hakkı olduğunda, büyük çoğunluğun kesin kazancı tercih ettiği görülmektedir. Buradan hareketle Bernoulli (1954), insanların belirsizlikten ya da riskten kaçınmak için belirli bir düzeye kadar kesin olanın lehine prim ödemeye eğilimli olduklarını, dolayısıyla kararların beklenen değer hesaplamalarından ziyade psikolojik değere göre verildiğini ortaya koymuştur. Bu bağlamda psikolojik değeri de, her biri olasılığına göre ağırlıklandırılmış olan sonuçların *marjinal faydalarının* ortalamasını alarak hesaplamaktadır.

Tablo 1.1: Varlığın Faydası

Varlık (Milyon)	1	2	3	4	5	6	7	8	9	10
Fayda (Değer)	10	30	48	60	70	78	84	90	96	100

Kaynak: Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.

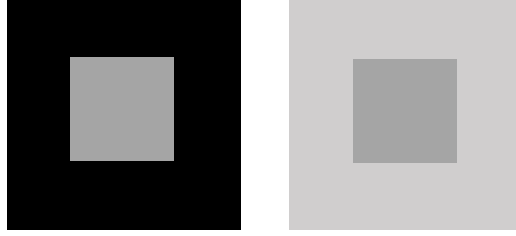
Tablo 1.1 bu şekilde hazırlanan fayda fonksiyonunun bir versiyonunu temsil etmektedir ve 1 milyondan 10 milyona kadar farklı varlık düzeylerine tekabül eden fayda puanlarını göstermektedir. Bu tabloya göre örneğin 2 milyonluk bir varlığa 1 milyon eklemek faydada 18 puanlık bir artışa neden olurken; 9 milyonluk bir varlığa 1 milyon ilave etmek faydada ancak 4 puanlık bir artış sağlamaktadır.

Bernoulli’nin fayda fonksiyonu aynı zamanda, fakir insanların neden sigorta satın aldığını ve zenginlerin neden bunu onlara satmaya istekli olduğunu da izah etmektedir. Zira tabloya göre 9 milyonu olan biri için 1 milyonluk bir kayıp 6 fayda puana, 2 milyonu olan biri için ise 1 milyonluk bir kayıp 20 fayda puana tekabül etmektedir. Bu nedenle daha fakir kişi riski, daha zengin kişiye aktarmak için prim ödemeye eğilimli olmaktadır (Kahneman, 2011). İşte bu şekilde Benoulli’nin temellerini attığı beklenen fayda teorisi, insanların olasılıklara göre ağırlıklandırılmış toplam değere göre değil, ilave bir kazancın ya da kaybın fayda cinsinden değerinin olasılıklarla ağırlıklandırılmış neticesine göre karar verdiğini ortaya koymaktadır. Dolayısıyla teori, insanları riskten kaçınma eğilimini varlığın marjinal faydasına dayandırarak açıklamaktadır.

Örneğin kesin olarak 4 milyona sahip olmak ile eşit olasılıkla 1 milyona veya 7 milyona sahip olmanın beklenen değerleri kıyaslandığında, her ikisinin de beklenen değeri 4 milyon olarak hesaplanır ve bu durumda insanların alternatifler arasında kayıtsız kalması gerekir. Ancak böyle bir seçenekle karşı karşıya kalındığında genellikle kesin olanının tercih edildiği görülmektedir. Zira varlıkların faydaları olasılıklarına göre ağırlıklandırıldığında kesin olan 4 milyonun faydasının 60, riskli durumun faydasının ise 47 olduğu ortaya çıkar. Dolayısıyla beklenen fayda teorisine göre birey kesin olanı faydasını maksimize ettiği için tercih etmektedir. Böylece riskten kaçınma eğiliminin temelde varlığın psikolojik değeriyle (yani beklenen fayda teorisine göre azalan marjinal fayda) ilgili olduğu ortaya konulmaktadır (Kahneman, 2011).

1970’li yılların sonlarına doğru Tversky ve Kahneman, modele fiziksel nicelik değişiklikleri ile öznel deneyimin yoğunluğu ve kalitesinde meydana gelen değişiklikleri dâhil etmişlerdir. Zira Tversky ve Kahneman’a göre psikolojik değerlerin temel belirleyicisi, referans noktasıdır. Algının referansın etkisinde nasıl kaldığını gözler önüne seren Şekil 1.1’de her iki büyük kare içerisine yerleştirilen küçük karelerin esasen renk tonu aynıdır. Ancak büyük karenin, yani referans noktasının etkisi ile örneğin siyah çerçevenin içerisindeki kare, gri çerçevenin içerisindeki kareye kıyasla daha açıkmiş gibi; gri çerçevenin içerisindeki kare ise, siyah çerçevenin içerisindeki kareye kıyasla daha koyuymuş gibi algılanmaktadır (Kahneman, 2011). Referanstan kaynaklanan bu algısal fark beklenti teorisinde, eş değer bir tutarın kayıplar söz konusu iken farklı, kazançlar söz konusu iken farklı bir psikolojik değere sahip olduğunu anlatmaktadır. Buradan hareketle Tversky ve Kahneman, örneğin A ve B’nin bugün 5’er milyona sahip olduğu, ancak A’nın dün 1 milyonu, B’nin ise 9 milyonu olduğu göz önünde bulundurulursa her ikisi için de 5 milyonun psikolojik değerinin aynı şeye tekabül etmediğini savunmaktadırlar.

Şekil 1.1: Parlaklık Algısının Referansa Bağlılığı

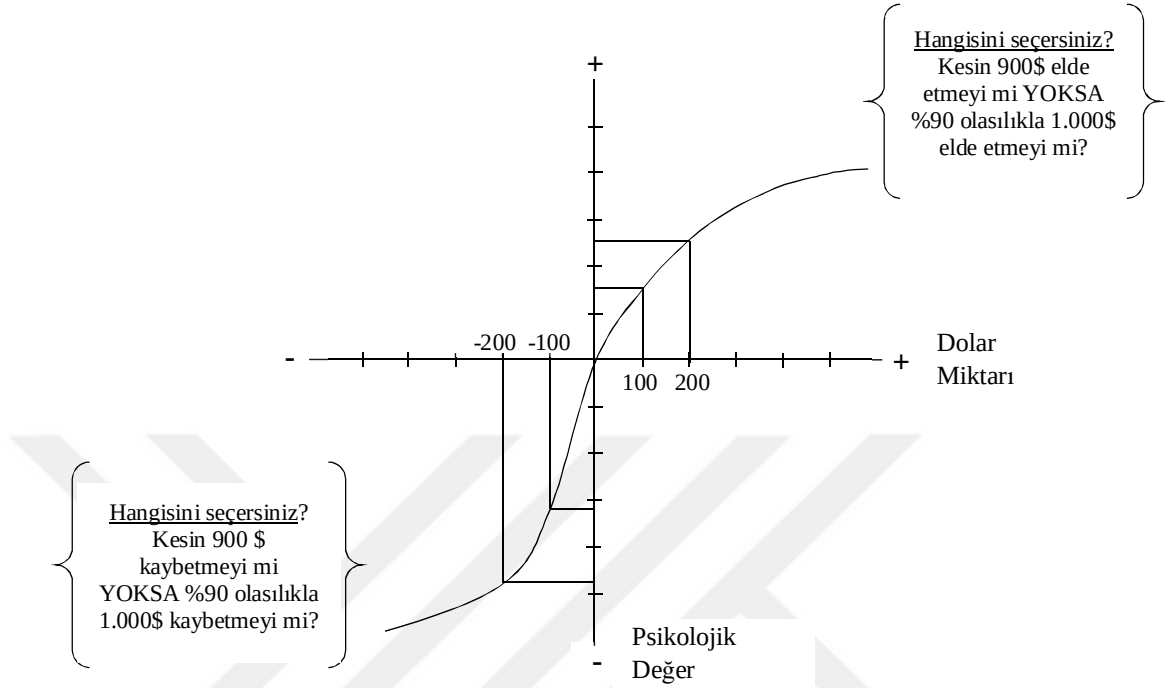


Kaynak: Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5), 1449-1475.

Grafik 1.1 kazanç ve kayıpların psikolojik değeri esas alınarak çizilmiştir. Buna göre grafiğin sağ tarafı kazançları, sol tarafı kayıpları, sıfır noktası ise referans noktasını temsil etmektedir. Eğrinin şekli, referansın iktisadi anlamda bireyin kayıp ya da kazançlara karşı tutumu üzerindeki tesirine işaret etmektedir. Üstelik bu tutum kesinlik etkisiyle³, Grafik 1.1’de de görüldüğü gibi, kazanç durumunda davranışlara *riskten kaçınma* (değer fonksiyonunun kazançlar için içbükey olması) şeklinde yansırken, kayıp durumunda azalan duyarlılık nedeniyle *risk alma* (değer fonksiyonunun kayıplar için dışbükey olması) şeklinde yansımaktadır. Değer fonksiyonunun kayıplarda, kazançlara göre daha dik olması durumu ise esasen, bir kaybın eşdeğer bir kazançtan daha büyük öznel etkiye sahip olduğunu göstermektedir. Bu durum, kayıplara karşı tepkilerin kazançlara karşı tepkilerden daha güçlü olduğuna işaret etmektedir. Literatürde bu durum kayıptan kaçınma kavramıyla anlatılmaktadır. Grafiğin sağ üst tarafında kayıptan kaçınma, kesin bir kazancın ihtimal dâhilindeki daha yüksek bir kazanca tercih edilmesi şeklinde tezahür ederken; kayıplar alanında kayıptan kaçınma, ihtimal içeren daha yüksek bir kaybı, kesin bir kayba tercih etme eğilimine tekabül etmektedir. Bu öznel etki beklenti teorisinde *olasılıklar* yerine *karar ağırlıklarının* kullanılmasıyla ölçülmektedir.

³ *Kesinlik etkisi* bireylerin kesin olduğunu düşündükleri sonuçları ihtimal içeren sonuçlara göre daha fazla ağırlıklandırıdıkları bir durumu ifade eder. *Yansıma etkisi* ise, kesinlik etkisiyle bireylerin kesin kazançlar içeren seçimlerde riskten kaçma, kesin kayıplar içeren seçimlerde ise risk alma eğiliminde olduklarını anlatır.

Grafik 1.1: Beklenti Teorisi



Kaynak: Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.

Sonuç olarak beklenti teorisi beklenen fayda teorisinden farklı olarak ortalama bir bireyin, kayıp ve kazançları öznel bir şekilde belirledikleri bir referans noktasına göre değerlendirdiklerini, kayıpları kazançlardan daha fazla ağırlıklandıklarını, gerçekte beklenen değer üzerinde etkisi olmasa da beklentilerin çerçevesinin tesirinde kaldıklarını ortaya koymaktadır. Bu bulgular küçük olasılıklara çok fazla değer atfedilmesi durumunda riskten kaçınma eğiliminin insanları *gereksiz işlem maliyetlerine katlanmaya* sevk edebildiğine de işaret etmektedir (Kahneman, 2011).

1.1.2.2 Piyasa Süreci Teorisi

Neoklasik iktisadın rasyonel insan kurgusuna yönelik daha kökten bir eleştiri, Avusturya ekolünün ilkeleri üzerine inşa edilen ve çerçevesini Menger, Mises ve Hayek'in çizdiği *piyasa süreci teorisinde* [market process theory] ifadesini bulmaktadır (Oğuz, 2001). Esasen bu karşı duruş, sosyal olguların da doğa bilimlerinin yasaları gibi düzenli bir karaktere sahip

olduğu, dolayısıyla sosyal bilimlere ait olguların nedensel olarak açıklanabileceği ve bu açıklamada da ideal bilimsel yöntemin doğa bilimlerinin yöntemlerine tekabül ettiğini öne süren pozitivist yaklaşıma tepkinin bir ürünüdür (Cevizci, 2014). Zira bu görüşün öncülerinden Dilthey'e göre sosyal bilimlerde doğa bilimlerinde olduğu iddia edilen anlamda yasalar yoktur; ancak toplumsal ve bireysel özgünlüğe ait değerler, normlar ve simgelerden oluşan *düzenlilikler* vardır. Bu bağlamda sosyal bilimlerin olguları bireyden bireye, toplumdan topluma ve devirden devire değişebildiği için doğa bilimlerindeki gibi genel geçerliliği de yoktur, tümevarım yöntemiyle de genelleştirilemezler (Hira, 2000). İşte bu nedenlerden dolayı, doğa bilimlerinin yönteminin doğal olayları *açıklarken* işe yarayacağını, ancak bu yöntemin insan yaşamını *anlamada* kullanılamayacağını, dolayısıyla sosyal bilimlerin kendisine özgü bir yöntemle sahip olması gerektiğini savunmaktadır⁴.

Piyasa süreci teorisinin epistemolojik yaklaşımı, neoklasik iktisat teorisinden oldukça farklı bir yere konumlanırken, bu yol ayrımının belirleyici kavramı *sübjektivizmdir*. Zira Menger ***değerin, insan zihninin bir ürünü*** olduğunu düşünür. Menger'in sübjektivizmle ilgili bu açılımı, Avusturya ekolünün ikinci kuşak temsilcileri olan Mises ve Hayek'in teorik söylemlerinde de yer bulmaktadır. Mises'a göre iktisat, *insanların eylemlerinin* [human action] bir bilimidir. *İnsanın eylemlerinin ise belirli bir amacı* vardır. Bu bağlamda insanların eylemlerinin belirleyici unsuru maksimizasyon değil, her bir bireyin kendine özgü amacına ulaşma gayretidir (Kurzner, 1992). Nitekim Lachmann, neoklasik düşünme tarzının aşırı soyutlamasının, insanın gerçek yaşam dünyasını yok saydığını; hatta insanı mekanik bir yapıya büründürerek *eylem* ile *reaksiyonu* birbirine karıştırdıklarını savunmaktadır (Lavoie, 1994). Bu bağlamda piyasa süreci teorisinin özünde mekanik [mechanistic] bir *davranış*tan ziyade, iradî [voluntaristic] bir *eylem* teorisi olduğu söylenebilir. Ayrıca insanların tüm eylemleri gerçek yaşamda zamana yayılmış olarak, belirsizliklerin hüküm sürdüğü ve araçların bireyin isteklerini karşılamak hususunda kısıtlı olduğu bir âlemde gerçekleşmektedir.

⁴ Hira'ya göre (2000) pozitivist yaklaşım, toplumun temel dinamiği olan ve topluma esas rengini veren değerlerden tecrit edilmiş soyut bir bilim paradigması kurmuştur.

Bu nedenle piyasa süreci teorisine göre *zamana yayılmışlık*, *bilgi eksikliği* ve *kıtlık* esasen iktisadi faaliyetin söz konusu olabilmesi için zarurî olan temel güçlerdir (Yay, 2004).

Avusturya iktisat okulunun ikinci kuşak temsilcilerinden bir diğeri Hayek'tir. Hayek (1945) de *sosyal bilimlerin nesnelerinin, anlamını, bireylerin inançları ve fikirleri doğrultusunda yüklendiği* ⁵, bu nedenle *fizik bilimlerinin kullandığı yöntemlerle açıklanamayacağı, dolayısıyla sosyal bilimlerin metodolojik yaklaşımının sübjektif olması gerektiği kanısındadır*. Hayek için iktisadî problemin odağında bilhassa *bilgi, bilginin sınırları ve ekonomik süreçlerde nasıl elde edildiği* konusu yer almaktadır. Piyasaya yön veren bilginin sadece teorik bilgi olmadığı kanısında olan Hayek, günlük hayatımızı tümüyle kuşatan bir bilgi türüne dikkat çekmektedir (Oğuz, 1999). Teorik ve objektif (yani cümle ya da formel ifade kalıpları içerisinde aktarılabilen) bilgiden farklı olan bu bilgi, maharet kazanmak ya da uzmanlaşmak şeklinde ifade edilen, bir şeyi yapabilmenin sübjektif ve pratik bilgisidir. Öğrenmenin tecrübeye dayalı süreciyle ilgili olan bu bilgi tekâmül ettikçe, bireyin kararlarında daha merkezi bir konuma yerleşmektedir⁶. Esasen bilgiyi bireye özgü hale getiren de, bireyin nevi şahsına münhasır ***ilişkiler ağı*** ve ***deneyimleridir***.

Hayek'in (1955) kuramsal dünyasında bilgi, bireylerin zihinlerindeki farklı farklı bilgilerin, beklentilerin, amaçların, kusurların, noksanlıkların ya da gözden kaçırılanların tüm gerçekliğiyle yansıdığı bir bilgidir. Bu bilgi her bir birey için önlenemeyecek şekilde eksik, başka bir deyişle bireyler arasında dağınık, noksan ve tutarsız olduğundan, "tam bilgi" diye bir şey yoktur. Dolayısıyla bireyin sübjektif dünyasına ait olan bu bilginin bağlamından çıkarılıp evrensel bilgi olarak karar verme süreçlerinde kullanılması mümkün değildir. Bilginin bu şekilde dağınık bir halde bulunması da, rasyonel seçimin iktisadî karar verme süreçlerinde oynadığı rolün, gerçeği varsayıldığı kadar yansıtmadığına işaret etmektedir. *Zira bireylerin genel hakkındaki bilgisizliği arttıkça, her bir bireyin sahip olduğu bilginin nispi önemi artmaktadır* (Oğuz, 1999). Bu bağlamda her birey, diğer bireylere kıyasla eşsiz

⁵ Örneğin *para*, insanların ona yüklediği anlam ile ekonominin merkezine yerleşirken, fiziksel olarak değerlendirildiğinde bir takım simgelerle bezenmiş boyalı bir kâğıttan ibarettir (Yılmaz, 2004).

⁶ Karar Vermenin Psikolojik ve Nörobiyolojik Temelleri başlığı altında öğrenme, bilgi ve karar süreçleri arasındaki ilişki sinirbilimsel açıdan açıklanmaktadır.

bir bilgiye ve dolayısıyla şahsına özgü bir avantaja sahiptir. Ancak bu tür bir bilgi, pozitivist bakış açısına göre “irrasyonel” olarak kabul edilmektedir (Hayek, 1945). Başka bir ifadeye rasyonel teorinin insan zihninin kusuru olarak gördüğü bu durum, aslında Hayek’e göre sosyal bilimlerin tam da başlamak zorunda olduğu şeydir. Bundan dolayı Hayek’e (1945) göre iktisadî meselenin temelinde, bilginin yalnızca tek tek bireyler arasında dağılmış, eksik ve ekseriyetle çelişkili parçacıklardan oluştuğu ve asla toplu olarak hiçbir kimseye verilmediği gerçeği yatmaktadır. Bu nedenle de bilginin nasıl elde edildiği ya da bireyler arasında nasıl aktarıldığı, nasıl kullanıldığı, farklı zihinlerde ve mekânlarda var olan bu bilginin nasıl birbiriyle uyumlu hale gelebildiğinin araştırılması gerekmektedir. Diğer bir ifadeyle, etkin bir iktisadi sistem tasarlayabilmek için, her biri yalnızca kısmî bilgiye sahip olan *insanların etkileşimlerinin*, çözümü nasıl üretebildiği gösterilmelidir. Bunun için de, *bilginin iletildiği yolların aydınlatılması* gerekmektedir (Hayek, 1937).

1.1.2.3 Karar Vermenin Psikolojik ve Nörobiyolojik Temelleri

İnsanların zihinsel süreçlerinin ve eylemlerinin bilimsel olarak incelenmesi, 19. yüzyılın ikinci yarısından sonra başlamıştır. Bir tarafta felsefenin zihin, irade ve bilgi hususunda muhtelif görüşleri; diğer tarafta his, duygu, düşünce ve bilinci, biyoloji ve fizyoloji bağlamında izah etme çabası, psikolojinin bilhassa bu iki alandan yoğun bir şekilde etkilenmesine ve ayrıca epistemolojik bağlamda yönünün belirlenmesinde farklı görüşlerin, farklı yaklaşımların ortaya çıkmasına yol açmıştır.

Bu yaklaşımlardan biri, insan eylemlerini zihinsel süreçlere odaklanmak suretiyle, iç gözlem yoluyla açıklama çabasında olan *bilişsel yaklaşımdır*. 19. yüzyılda bilişsel yaklaşımı benimseyen psikologlar, insanların eylemlerinin algı, bellek, akıl yürütme, karar verme ve problem çözme gibi zihinsel süreçlerle izah edilebileceğini savunmuşlardır. Ancak o dönemde bilişsel yaklaşım, biyolojik bir temelden yoksun olması ve sadece iç gözleme ağırlık vermesi gibi nedenlerle yoğun eleştirilere maruz kalmıştır.

Yine bu yıllarda insan davranışlarını incelemek için deney hayvanlarının kullanılabileceğinin kabulüyle birlikte ilk deneysel psikoloji çalışmaları da meyvelerini vermeye başlamıştır. Rus fizyolog İvan Pavlov bir hayvana iki uyarıyı bağdaştırmayı öğrettiği *klasik koşullamayı*; Amerikalı psikolog Edward Thorndike ise hayvana bir davranış tepkisini, o tepkinin sonuçlarıyla bağdaştırmayı öğrettiği *araçsal koşullamayı* keşfetmiştir. İşte bu ilk deneysel çalışmaların neticesinde ulaşılan uyarı-tepki psikolojisiyle ilgili bulgular davranışsal yaklaşımı doğurmuştur⁷. Dönemin hâkim paradigması olan pozitivist yaklaşımın karakterize ettiği davranışsal yaklaşım, insan psikolojisinin doğa bilimleri gibi yalnızca somut, nesnel ve deneysel yöntemlerle incelenebileceğini (Watson, 1913); dolayısıyla bir organizmanın sadece gözlemlenebilir eylemlerine, yani davranışlarına odaklanması gerektiğini savunmaktadır (Nolen-Hoeksema, 2009). Zira psikolojinin ancak bu suretle nesnel ve deneysel bir doğa bilimi haline gelebileceğine inanmaktadır. Bu bağlamda psikolojinin teorik amacının tahmin ve davranış kontrolü olması gerektiği görüşü hâkimdir (Watson, 1913). Watson, Skinner gibi davranışçılar, hisler, duygular, düşünceler ve değerler gibi insanı insan yapan etkinlikleri, zihinsel süreçlere başvurmaksızın, sırf gözlenebilir davranışlara odaklanmak suretiyle açıklanabileceğine kanaat getirerek insan eylem ve davranışlarıyla ilgili bakış açısını oldukça daraltmışlardır (Kandel, 2011).

Davranışçılık ekolünün Amerika Birleşik Devletlerde geliştiği yıllarda, Viyana’da Freud 19. yüzyıl bilişsel yaklaşımından da izler taşıyan yeni bir kuram ortaya koymuştur (Nolen-Hoeksema, 2009). Psikanalitik yaklaşım olarak kabul gören Freud’un yapısal zihin kuramı, zihnin bilinçli ve bilinçdışı işlevlerinin olduğunu ve insanların eylemlerinin büyük bir kısmının bilinçdışı süreçlerden kaynaklandığını iddia etmektedir⁸ (Schopenhauer, 1966;

⁷ Diğer bir ifadeyle sinir hücreleri özgül patikalar halinde birbirine bağlanır. Dolayısıyla bu patikalar üzerinden sinyaller öngörülebilir bir şablonla ilerler (Kandel, 2011).

⁸ Esasen Sigmund Freud (1856-1939)’un kuramını şekillendirirken kullandığı kavramların altında yatan bazı görüşlerin özellikle Arthur Schopenhauer (1788-1860) ve Friedrich Wilhelm Nietzsche (1844-1900)’den alındığına ilişkin iddialar da söz konusudur. Örneğin Schopenhauer bilinç dışını istencin kararlarının gizli atölyesi olarak tanımlarken (-ki bu Freud’un teorisinde id teriminde ifadesini buluyordu), bilinci ise su tabakasına benzetir. Bilinçli fikirlerin ise bu suyun yüzeyinde şekillendiğini ifade eder. Nietzsche ise insanların bilmeden sürekli düşündüğünü, bilince erişen düşüncenin bunun sadece küçük bir kısmı olduğunu ifade eder. Bu benzerliklerin ayrıntılı bir şekilde ortaya konulduğu bir çalışma için Cybulska (2015)’ya bakılabilir. Ayrıca bu çalışmada bahsi geçen Nietzsche (2008) ve Schopenhauer (1966)’a ait eserlere de bakılabilir.

Nietzsche, 2008; Cybulska, 2015; Freud, 1989). Freud'a göre zihin, ego (benlik), id (alt benlik) ve süper-ego (üst benlik) olmak üzere başlıca üç ruhsal yapıdan oluşmaktadır. İcracı organı, hem bilinçli hem de bilinçsiz bileşene sahip olan ego temsil eder. Egonun bilinçli bileşeni duyu organları vasıtasıyla dış dünya ile temas halindedir ve algı, mantık ve eylem planı gibi işlerle meşgul olur. Bilinçdışı bileşen ise bastırma, yadsıma gibi ruhsal savunmalarla uğraşır (Kandel, 2011). Temel güdülerin mahalli olan id, acıdan kaçınmak, haz peşinde koşmak gibi bir ilkenin, yani haz ilkesinin [pleasure principle] hükmü altındadır ve bütünüyle bilinçdışıdır. Süper-ego ise toplumsal değerler çerçevesinde insana yön veren bir ahlâk mertebesi olarak görülür. Yapısal zihin kuramına göre bu üç ruhsal yapı kompleks ilişkilerle insanın gelişimini ve kişiliğini tesis etmektedir. Bu yaklaşıma göre insanın karar verme süreçlerini anlayabilmek için bu üç ruhsal organın işlevlerinin göz önünde bulundurulması gerekmektedir.

Bilişsel ve davranışsal yaklaşımdan farklı olarak *biyolojik yaklaşım*, bütün psikolojik olayların beyin ve sinir sisteminde bir karşılığının olduğu düşüncesinden yola çıkarak, insanların eylemlerini sinir sisteminde oluşan elektriksel ve kimyasal olaylarla izah etmeye çalışmaktadır (Nolen-Hoeksema, 2009). Biyolojik yaklaşımın özellikle öğrenme ve hafıza çalışmalarında ulaştığı bulgular, bilginin ve deneyimlerin sinir hücreleri ve bu hücrelerin arasındaki bağlantılarda nasıl muhafaza edildiğinin sırlarını aralamaya başlamıştır. Böylece insan, güdülerinin, düşüncelerinin, davranışlarının; belleğin hayata kattığı süreklilik ve eşsiz benlik bilincinin biyolojik bağlamda köklerine yönelik keşifsel bir yolcuğa çıkmaya hazır hale gelmiştir. 1960'lara gelindiğinde zihin felsefesi, davranışçı psikoloji ve bilişsel psikoloji, biyolojik yaklaşımla birlikte *çağdaş bilişsel psikolojinin* temelini atmışlardır. Artık çağdaş bilişsel psikoloji, zihinsel süreçleri, biyolojinin yöntemleriyle incelemeye başlamıştır (Kandel, 2011). Diğer bir deyişle 19. yüzyıl ruhsal içebakış yönteminin yerini, sinir sisteminin içerisine girip zihinsel işlevleri sinir hücreleri düzeyinde nesnel bir şekilde inceleme imkânı sunan bir nevi biyolojik içebakış yöntemi almıştır. Bütün bu gelişmeler eşliğinde zihnin dünyaya ilişkin bilgiyi nasıl elde ettiğine; bunları nasıl muhafaza ettiğine; doğuştan gelen ile sonradan öğrenilen ayırımına; öğrenirken ve hatırlarken merkezi ve çevresel sinir sisteminde ne tür değişikliklerin olduğuna; hislerin ve duyguların zihinsel

süreçlerdeki işlevlerine; karar verme süreçlerinde akıl ve duygular arasındaki ilişkiye; her bir bireyin eşsiz ve benzersiz benlik bilincinin nasıl tezahür ettiğine ve yapısal özelliklerindeki müşterek noktalara ve farklılıklara; benliğin sosyal davranışla ilişkisine ilişkin geçmişte gizemini korumuş pek çok soruya deneysel olarak kısmen cevap verebilir duruma ulaşılmıştır (Kandel, 2011).

1.1.2.3.1 Akıl Yürütme ve Karar Verme

Akıl yürütmenin (muhakemenin) amacı karar vermek, karar vermenin temeli ise muhtelif alternatifler arasından seçim yapmaktır. “Akıl yürütme ve karar verme süreci genellikle karar veren kişinin (1) karar verilmesi gereken durum hakkında, (2) farklı eylem seçenekleri hakkında ve (3) bu seçeneklerin her birinin kısa ve uzun vadeli neticeleri hakkında bilgi ya da fikir sahibi olduğunu ima etmektedir. Diğer taraftan akıl yürütme ve karar verme terimleri, karar veren kişinin akıl yürütürken gerekli olan destek süreçlerinin hazır olduğuna da işaret etmektedir. Bu destek süreçleriyle ilgili olarak da ekseriyetle *dikkat* ve *çalışma hafızası* [işleyen bellek] öne çıkarılırken, *duygu ve hislerden* ve ayrıca bir seçimden bahsedebilmek için gerekli olan *seçenekler setini* (Kahneman, 2011) zihnin nasıl hazırladığından hemen hemen hiç söz edilmemektedir” (Damasio, 1994). Dolayısıyla kararların zihinde nasıl bir etkileşimin neticesinde tezahür ettiğini anlayabilmek için, sinir bilimleri açısından duygu ve hislerin işlevsel olarak nasıl bir vazife üstlendiklerine ve seçenek oluşturma süreçlerinin nasıl işlediğine kısaca değinmek gerekir.

Duygu [emotion], bir uyarının neticesinde, fizyolojik, davranışsal ve bilişsel tepkilerle birlikte tezahür eden kişisel deneyimlerdir. Başka bir ifadeyle duygu; belirli nesne, olay veya bireylerin insanın iç dünyasında uyandırdığı izlenimdir (TDK, 2009). Sözcüğün Latince [motus] ve İngilizce [emotion] köküne bakıldığında ise *hareket*, *devinim* sözcüklerinden türetildiği görülmektedir. Bu tanımlara göre bir duyguya bilişsel reaksiyonlar, fizyolojik reaksiyonlar ve davranışsal reaksiyonlar eşlik etmektedir. Bilişsel etkiler kişinin hafızasına, düşüncesine ve bir olayın algılanışına atıfta bulunmaktadır. Olup biteni anlama ve anlamlandırma çabası olarak da ifade edilebilir. Fizyolojik reaksiyonlar vücudun hormonal

seviyelerindeki, sinir sistemi tepkilerindeki deęişikliklerden kaynaklanmaktadır. Davranış reaksiyonları ise duygunun aktif olarak ifade edilmesini içermektedir. Başka bir deyişle duygunun, vücut veya yüz ifadesinden anlaşılabilir bir hale bürünmesidir. Duygunun olumlu veya olumsuz sahip olduęu içerik ise duyguyu bilinçli ve öznel bir deneyim haline getirmektedir.

Duygu, bireyin nesne, olay ya da dięer insanlarla olan ilişkilerine baęlı olarak iç dünyasında meydana gelen ve çoęu kez fizyolojik temelleri ve kaynakları bilinmeyen korku, öfke, üzüntü, mutluluk, ięrenme ve şaşırma gibi *yönlendirici tepkiler* olarak da tanımlanabilmektedir (Demir ve Acar, 2006; Damasio, 1994). Büyük ölçüde doğuştan gelen ve otomatik eylem programlarına tekabül eden bu kompleks duygular (Damasio, 2010) insanı hayatta tutmayı amaç edinmiş temel mekanizmalar olduğundan *birincil duygular* olarak adlandırılır. Ancak birincil duygular, duygusal davranışlar yelpazesini tümüyle tasvir edememektedir. Bu nokta da ikincil duygular devreye girer. İkincil duygular ise bireyin hem nesne ve durum kategorileri, hem de birincil duygular arasında sistematik bağlantılar kurmaya başladığı andan itibaren ortaya çıkmaktadırlar. Dolayısıyla ikincil duygular bireysel deneyimlerin rehberliğinde, deneyimler ise doğuştan var olan duygusal temsiller rehberliğinde tezahür etmektedirler (Damasio, 1994).

Sinir bilimleri açısından biraz daha farklı bir anlama sahip olan *his* [feeling] kavramı, bireyin zihninde ve bedenindeki deęişikliklerin algılanmasına karşılık gelmektedir. Eylemlerin kendisinden ziyade fotoğrafı hükmünde olan bu hisler, beynin ürettiği bir çeşit algıdır. Bunlardan bazıları duygulardan kaynaklanır. Yani, duygusal tepkiyi oluşturan deęişikliklerin algılanmasıyla ilgilidir (Damasio, 2010). Duygulardan kaynaklanan hislerden farklı olarak, belirli bir duygunun vücuda hakim olmadığı zamanki manzarasına tekabül eden *arka plan hisleri* [background feelings] de söz konusudur. “Ruh hali” kavramını da kapsayan bu durum, aynı zamanda bedenden gelen ve arka planda işleyen kesintisiz veri akışıyla sistemin sürekli açık kalmasını sağlayan süreçleri de ifade etmektedir (Damasio, 1994; Kahneman, 2011; Eagleman, 2016). Bu suretle bireyin hem içsel dünyasının hem de çevresinin bilincinde olması, çevreyle etkileşimlerinde bir tepki esnekliği kazandırmaktadır (Damasio, 1994).

Esasen hissedebilme kabiliyeti de, *hareket edebilme* kabiliyetiyle ilişkilidir. Zira hareket kabiliyeti olan bütün canlılar, çevreden alınacak detaylı bilgilere ihtiyaç duyduğundan sinir sistemine sahiptirler. Diğer bir deyişle sinir sistemi bireye, etrafta neler olup bittiğini *hissedebilme* kabiliyeti kazandırmaktadır (Canan, 2015) ⁹. Hissedebilme kabiliyeti sayesinde beyin, içerden ve dışardan gelen değişiklik bilgilerine göre insanı hayatta tutabilmek için gerekli önlemleri alarak, değişikliklere “uyum” gösterilmesini sağlamaktadır (Canan, 2015). Bu biyolojik bir uyum olduğu kadar, sosyal varlığını sürdürebilmesi açısından *sosyal uyumu* da gerekli kılmaktadır. Dolayısıyla sosyal biliş ve davranışla ilgili donanım da sahip olan insan, eğitim ve kültürel etkilerle akıl yürütme ve karar verme stratejilerine katmanlar eklemektedir. Böylece insanlar hem kendi zihinleri hem de diğer kişilerin zihinleri hakkında *kuramlar* oluşturabilme kabiliyeti kazanırlar (Damasio, 1994).

Akıl yürütme ve karar verme süreciyle ilgili bir diğer mesele de zihnin seçenekler setini *nasıl oluşturduğuyla* ilgilidir. Akıl yürütmek ve karar vermek için gerekli olan verisel bilgiler zihne imgeler şeklinde gelmektedir. İmgeler de *algılanan imgeler* ve *anısal imgeler* olmak üzere iki türe ayrılırlar. Bir şeyi deneyimlerken (örneğin pencereden dışarıyı izlerken) ortaya çıkan imgeler algılanan imgelerdir [perceptual images]. *Anısal imgeler* [recalled images] ise bir şey düşünüldüğünde, bununla ilgili verilerin kayıtlardan toplanarak yönlendirici bir temsiline [dispositional representations] yeniden kurulduğu bir duruma tekabül eder (Damasio, 1994; Eagleman, 2016). Bunlar düşünce için şarttır, ancak düşüncelerin bir içeriği değildir. Bu durum gelecekle ilgili bir plan oluşturulurken de aynı şekilde işlemektedir. Diğer bir ifade ile amaçları geçmişte olanları kaydetmekle sınırlı olmayan bu mekanizmalar, geleceğin kurgusunu, geçmişe ilişkin bu kayıtları yeniden yapılandırmak suretiyle inşa etmektedirler (Eagleman, 2016). Diğer bir deyişle henüz hiç gerçekleşmemiş ve belki de hiç gerçekleşmeyecek olan bir şeyin imgeleri, doğası bakımından, çoktan olmuş bir şeyin akılda kalan imgelerinden farklı değildir. İmgelerin oluşumu, ya beynin dışına yönelik duyuşsal alıcıların ya da beynin içinde, korteks bölgelerindeki yönlendirici temsillerin denetimi altındadır. Yönlendirici temsiller şeklinde bellekte var olan bilgi, neredeyse eşzamanlı

⁹ Belirsizlikten kaçınma sebebi de buna bağlıdır. Belirsizlik hiçbirşey hissedememesi demektir. Hissedemediği anda karar vermek oldukça güçleşecektir,

olarak, hem sözel hem de sözsüz biçimleriyle bilinç tarafından erişilebilir hale getirilebilmektedir (Damasio, 1994). Dolayısıyla yönlendirici temsiller zihnin pek çok bölgesinin (ses, görüntü, koku vb.) işbirliğiyle inşa edilmektedir. Bu suretle hazırlanan seçenekler setinde, daha deneyimler arşivlenirken kodlanan *duygunun izleri* bulunmaktadır¹⁰. Bu bağlamda yönlendirici temsiller hem doğuştan gelen, hem de deneyimlerle kazanılmış ve yine deneyimlerle sürekli değişen bir hazine hükmündedir (Damasio, 1994).

Zihnin seçenekler setini hazırlamasıyla karar verme süreci nihayete ermiş olmaz. Zira bu kez alternatifler arasından seçim yapabilmek için bunların sıraya dizilmesi, sıraya dizilebilmesi için de bir takım kıstasların olması gerekmektedir. *Meyillendiriciler* olarak da ifade edilebilecek olan bu kıstaslar, bireyin birikimleriyle ilgilidir. Bu birikimler, içsel bir tercih sisteminin denetiminde, sosyal kurumlar [institutions] (gelenek, töre, din vb.) ve ahlâk kurallarını kapsayan bir dizi dış koşulun etkisi altında (yani eğitim ve sosyalleşme sürecinde), belirli dürtü türleriyle belirli haller arasında bağlantı kurmak suretiyle, belirli senaryoların gelecekteki tahmini sonuçlarıyla ilgilidirler. İçsel tercih sistemi ise doğuştan acıdan kaçınmaya, zevk içeren olasılıkları aramaya meyillidir (Damasio, 1994).

Karar vermeyeyle ilgili bir diğer husus ise insanın toplumsal boyutuyla ilgilidir. Zira insanlar toplum içerisinde sadece biyolojinin kurallarına göre yaşamamaktadırlar. Sosyal kurumlar, ahlâk kuralları gibi içgüdüsel davranışları biçimlendirerek, karmaşık ve hızla değişen çevreye esnekçe uyum sağlamalarına yardımcı olan ek denetim mekanizmaları da söz konusudur. Bunlar biyolojik düzenleyici süreçlerin sinirsel temsilleriyle birbirinden ayrılmaz biçimde bağlıdırlar. Diğer bir deyişle belirli bir toplumda hayatta kalmayı sağlayacak şekilde tasarlanmış kültürel bir reçeteye göre ayarlanmıştır. Bu kültürel reçete, bireye acı ve zevk veren deneyimlerle birlikte düşünüldüğüne, mekanizmanın harekete geçeceği eşiği ya da şiddetini değiştirmek için araçlar sunmaktadır. Sağlıklı bir kültürde yetişen bir birey için

¹⁰ “Sözel hafıza bileşenleri hipokampus gibi bölgeler tarafından işlenip kaydedilirken, duyguların ve duyuların hafızası farklı yerlerde ve farklı biçimlerde kaydediliyor gibi görünüyor. Hipokampus bölgeleri ameliyatla çıkarılsa bile, duyuların ve duyguların anları beyinde başarıyla kaydedilmeye devam etmektedir. Duyularla ilişkili beyin bölgeleri beynin lisan bölümleriyle doğrudan bağlantılı olmadığı için sözlerle ifade dilemeyen duygular, dile dökülemez fakat sandığımızdan çok daha etkili bir şekilde davranışları yönlendirir.” (Canan, 2015)

sosyal değerlere ve ahlâk kurallarına uymak, insanın biyolojik mekanizmasında “ussal” bir davranış olarak kuruludur (Damasio, 1994).

1.1.2.3.2 Sezgi

Birey bir karar vermek durumu ile karşı karşıya kaldığında, seçenekler setinin hazırlanabilmesi için anısal imgelerin rehberliğine ihtiyaç duyulduğu; bu çerçevede beynin üst katmanlarının, hatırlama işlemi için alttaki duygusal beyni (orta beyin) harekete geçirip o duyguların tekrar yaşanmasını sağladığı ve böylece seçenekler setinin hazırlandığı yukarıda ifade edilmişti (Canan, 2015). Ancak bazı durumlarda bu meyillendiriciler doğrudan ortaya çıkmak yerine, karar süreçlerine perde arkasından müdahil olurlar. İşte bu durum *sezgi* (Damasio, 1994) ya da *önsezi* (Eagleman, 2011) denilen kavrama tekabül etmektedir.

Sezgiyi ortaya koyabilmek amacıyla mümkün olduğunca gerçek yaşama benzer bir deney kurgulanmış ve katılımcıların fizyolojik tepkileri deri iletkenliği ölçülerek tespit edilmiştir (Damasio, 1994). Bir şans oyunu deneyi olarak tasarlanan oyunda, (A, B, C ve D) etiketli iskambil kâğıtlarına benzer dört farklı deste kullanılır. Oyuncuların amacının çok kazanıp az kaybetmek olarak tanımlandığı bu oyun, aynı anda hepsinin sunulduğu kâğıt destelerinden her seferinde bir kart çekerek oynanır. Destelerden ikisi (C, D) daha düşük kayıplar içerirken, geri kalanlar çok büyük kayıplara neden olacak şekilde hazırlanır. Kurallar değiştirilmeden 100 tur devam eden (ve oyunculara oyunun kaç tur olduğuna ilişkin bilgi verilmeyen) bu oyunda, oyuncular önce dört desteden örnekleme yaparak işe başlarlar. Başta ödülün cazibesine kapılarak, A ve B destelerine meylederler. Ancak ilk 30 tur içinde, tercihlerini C ve D destelerine kaydırırlar ve ekseriyetle oyunun sonuna kadar da buna sadık kalırlar. Risk almayı sevenler, A ve B destelerini arada deneseler de, daha güvenli olan C ve D’ye geri dönerler.

Beynin en önemli özelliklerinden biri, deneyimlerinin sonucunda bir *sezgisel tahmin* mekanizması oluşturabilme yeteneğidir. İnsanlar ihtimalleri deneyimleyerek ağırlıklandırmakta ve kararlarını da, geri planda çalışan bu sistemin ürettiği tahminlerin

yönlendirici etkisi altında vermektedirler. Nitekim destelerle ilgili bu deneyde deneklerin riskli desteden bir kart seçmeden hemen önce, yani denekler deneycinin riskli olduğunu bildiği bir desteden kart açmayı düşündüğü veya niyet ettiği sırada, bir deri iletkenliği tepkisi olduğu ve tepkinin oyun ilerledikçe artış gösterdiği bulgusuna ulaşılmıştır (Eagleman, 2011). Benzer bir deneyde de bireylerin 40. hamleye geldiklerinde riski desteyi bilinçli olarak belirleyebildikleri, 20. hamlede ise sezgilerinin yol göstermeye başlamış olduğu bulgusuna ulaşılmıştır (Canan, 2015). Ancak burada dikkat edilmesi gereken en önemli husus, bu kalıpların temelde bilinçdışı süreçlerde, ancak bilinçle işbirliği halinde tesis edildiğidir. Zira bilinçdışının derinliklerine tek tek ayırt edilemeyecek ölçüde işlenen bilginin büyük bir kısmı, beynin bilinçli bölümlerinde hayat bulmakta ve tekrarlarla sinir devrelerine nakşedilerek bilinçdışına gönderilmektedir (Eagleman, 2011).

1.1.2.4 Davranışsal Yaklaşımların Rasyonalite Varsayımına Yönelik Eleştirileri

İktisat biliminin pozitivist felsefeden beslenen rasyonalite anlayışı çerçevesinde doğa bilimlerine öykündüğü yıllarda, esasen klasik iktisat döneminde teorinin içerisinde var olan psikolojik faktörlerin çözümlemenin dışında bırakılmasıyla, psikoloji ile iktisat biliminin arasına mesafe girmeye başlamıştır. Başka bir söyleyişle psikoloji biliminin yeni doğuyor olduğu bu yıllarda pozitivist anlamda bilimsel bir nitelik kazanamaması, iktisatçıların psikolojiye mesafeli yaklaşımlarına neden olmuştur (Ruben ve Dumludağ, 2015). Sonraki yıllarda iktisatta olduğu gibi psikolojide de, pozitivist yaklaşımın kapalı sistem ve mekanik nedensellik anlayışı ile insanların yalnızca duyu deneylerini kaydeden pasif alıcılar olduğu düşüncesi hâkim olmaya başlamıştır (Özel, 2016). Bu anlayış psikoloji alanında *davranışçı* ekolde vücut bulmuştur. İnsan davranışlarında sadece *uyarı* ve *tepkinin* belirleyici olduğunu kabul eden bu ekol, zihin, akıl ve zekânın yönlendirdiği bilişsel süreçleri *kara kutu* nitelendirmesi ile çözümlemenin dışında tutmaktadır. İnsan davranışlarına yönelik analizlerini zihni faktörlere referans vermeden, gözlemlenebilir şeylerle açıklamaya çalışmaktadır. Yirminci yüzyıl ortalarında ortaya çıkan bu yöntemsel yaklaşım birliği, iktisat ve psikolojinin aralarındaki ilişkiyi yeniden canlandırmaya başlamıştır (Ruben ve Dumludağ,

2015). Ancak bu ilişkinin iktisadi ve finansal kararlara yansması davranışçı ekolün benimsediğı uyaran-tepki çerçevesi ile sınırlı kalmıştır.

Aynı dönemlerde Simon, karar almayı merkeze alarak yaptığı çalışmalarda, kompleksite ve sınırlı rasyonellik kavramları ile yapay zekâ ve bilgisayar teknolojilerini bir araya getirerek insan davranışlarının açıklanmasına yeni bir boyut kazandırmıştır. Davranışsal analizlerde yeni tekniklerin kullanılmasına imkân tanıyan bilişsel devrimle birlikte zihinsel süreçlerin kısmen analize dâhil edilebilir ve yapay zekâ teknolojisi ile kısmen modellenebilir hale gelmesi, bilişsel süreçlerin, iktisadi davranışlar ve belirleyicileri üzerine etkilerinin incelenmesine olanak sağlamaktadır.

Neoklasik iktisadın rasyonel insan tanımına ilişkin eleştirilerden biri, bu tanımın *sınırsız rasyonalite* kabulü ile ilgilidir. Bu duruma dikkat çekerek *sınırlı rasyonalite*¹¹ kavramını alanyazınına kazandıran Simon’a (1955; 1956) göre, neoklasik iktisadın rasyonalite tanımının iki önemli eksiğı vardır. Bunlardan birincisi bireylerin tüm alternatifler hakkında *tam bilgi* sahibi olmasının hem oldukça maliyetli olduğu ve hatta çoğu zaman mümkün olmadığıyla ilgilidir. İkinci eksiğı ise bireylerin tüm bilgileri elinde bulundursa bile bu kez isabetli bir karar verebilmek için gerekli olan tüm hesaplamaların yapılmasında *bilişsel kısıtlarla* karşılaşmasıdır. Her bir bireyin bilişsel kısıtları da aynı düzeyde olmadığından, hâlihazırdaki bilgilerin işlenmesinde ortaya çıkan eşitsizlik (ya da çeşitlilik; ya da asimetri) tüm finansal aktörlerin tam bilgiye sahip olmasını da imkânsız hale getirmektedir (Pompian, 2012). İşte bu bağlamda Simon (1956) bireylerin, bu bilişsel kısıtların çizdiği sınırlar içerisinde kalan alternatifler arasından *foyda fonksiyonunu optimize eden* değil, kendilerini *tatmin eden* alternatifi (yani mümkün olanın en iyisi) seçeceğini ifade etmektedir.

Neoklasik iktisadın rasyonalite varsayımına yönelik bir diğer eleştirisi ise, bir davranışın belirli bir amaca ulaşmak açısından tutarlı olup olmadığını ifade eden “özel rasyonalite” kavramına tekabül ettiği, bireyden ve ortamdan kaynaklanan çeşitli etkenleri dikkate alan

¹¹ Ekonomide sınırlı rasyonalitenin rolüne ilişkin kapsamlı bir çalışma için bkz. (Conlisk, 1996).

“yöntemsel rasyonaliteyi” göz ardı ettiği yönündedir (Buğra, 2011). Bu bağlamda rasyonel davranışı, bireyin zihni süreçlerini ima eden “içsel çevre” ile davranışlarının yansıdığı ve etkileşim içerisinde oldukları “dışsal çevre” belirlemektedir (Simon, 2000). Diğer bir deyişle sınırlı rasyonalitede bireyin davranışlarını tahmin edebilmek için etkisi altında bulundukları psikolojik ve sosyolojik faktörlerin bilinmesi gerekir (Simon, 1984). Bu durumda rasyonel davranışı Simon (1990), bir bıçağı bireyin bilişsel kısıtlarına [cognitive limitations] diğer bıçağı ortamın yapısına (ya da bağlama) [structures of the environment] tekabül eden bir makasa benzetir. Buna göre eksik bilgi, dikkat eksikliği, zaman baskısı, sonuçları hesaplamadaki yetersizlikler ve yanlışlıklar gibi bilişsel sınırlar ile ortamın denetlenmez güç ve dinamikleri insanın rasyonelliğini sınırlandırmaktadır (Özcan, 2011). Bu çerçevede Simon’un görüşleri etrafında öbeklenen çalışmalar, neoklasik iktisadın varsayımlarını temelden eleştirerek, alternatif bir model oluşturmaya odaklanmışlardır (Ruben ve Dumludağ, 2016).

Neoklasik iktisadın rasyonel insan tanımına ilişkin eleştirilerden bir diğeri ise, kökleri Adam Smith’e kadar uzandığı iddia edilen, kişisel çıkarının peşinde koşan¹² ve böylece kaynakların optimal bir şekilde dağılmasını temin ederek toplumsal harmoninin oluşmasına katkı sağlayan birey varsayımı ile ilgilidir. Ancak böyle bir varsayım genel olarak iki farklı yönden eleştirilmektedir. Birincisi, kişisel çıkar maksimizasyonunun her zaman bireyin yararına olmayacağı yönündedir. Nitekim bu durum oyun teorisinin literatürde *mahkûmlar açmazı* olarak kendisine yer bulan örneklerinden biriyle izah edilmektedir (Kuhn, 2019). Söz gelişi A ve B bir soruşturma kapsamında gözaltına alınan ve suç ortağı olan iki kişiyi temsil etsin. Tutuklanan bu iki kişi ayrı ayrı hücrelere konulduktan sonra sorgulama sırasında her birine bir anlaşma sunulur. Sunulan bu anlaşmaya göre şüphelilerden birinin, diğeri aleyhinde tanıklık etmesi halinde, diğeri suskun kalırsa tanıklık eden serbest kalacak, susmayı tercih eden ise 10 yıl hapse mahkûm edilecektir. Şayet ikisi de birbirleri aleyhinde tanıklık etmez,

¹² Burada fayda ve kişisel çıkar ayırımına dikkat etmek gerekir. İktisatta fayda mal ve hizmetlerin insan ihtiyaçlarını karşılama özelliğini ifade eder. Esasen iktisadi bakış açısı ile ele alındığında, niteliğine bakmaksızın insan isteklerinin karşılanmasına yönelik yapılan tüm faaliyetler faydalı olarak nitelendirilir. Bireylerin nelerden fayda temin ettikleri de *kişisel çıkar* ya da *kendi yararını gözetme* kavramı ile ortaya konulur. Bu bağlamda kişisel yararını gözetme, mevcut şartlarda en çok yarar umulan seçeneklerin diğerlerine tercih edilmesini ifade eden bir karar verme biçimidir (Demir, 2013).

suskun kalmayı tercih ederlerse, bu kez her ikisi de 1'er yıl hapis cezasına çarptırılacaktır. Ve nihayet her iki şüphelinin de birbirleri aleyhinde tanıklık etmesi durumunda ise, her ikisi de 5'er yıl hapis cezası alacaktır.

		A	
		İnkâr	İtiraf
B	İnkâr	(-1, -1)	(-10,0)
	İtiraf	(0,-10)	(-5,-5)

Kişisel çıkarının maksimizasyonu peşinde koşan (yani rasyonel olan) bireyin, böyle bir senaryo ile karşı karşıya kaldığında, kendi çıkarını düşünerek itiraf etmesi gerekir. Ancak kendi çıkarını maksimize etme gayesiyle yapılan bu itiraf, esasen öteki birey de aynı şekilde rasyonel davrandığında, her biri için 5'er yıllık bir hapis cezası ile sonuçlanmaktadır. Diğer taraftan şüpheliler rasyonel davranmayıp, susmayı tercih ettiklerinde 1 yıllık bir hapis cezası ile kurtulabilmektedirler. Sonuç olarak bu durum bireylerin kişisel çıkarını maksimize etmek için aldığı kararların, her zaman bu amaca hizmet etmeyebileceğini göstermektedir. Nitekim 2001 yılında yaşanan Enron skandalı ve 2008'de yaşanan mortgage krizi bunun yakın tarihteki en bariz örnekleridir¹³.

Enron skandalında şirket, gerçekleşmeyen kârlarını bir takım muhasebe hileleriyle kayıtlara geçirirken, hisse senetlerinin değeri karşılıksız bir şekilde yükselmeye başlar. Enron'un muhasebe işlemlerini yürüten ve bütün bunlardan haberdar olan Arthur Anderson şirketi ise durumu ihbar etmek yerine, oldukça cömert bir şekilde yapılmış danışmanlık anlaşmalarını sürdürmeyi tercih eder. Hissedarlar, kârlarının bir kısmını hileli işlemler nedeniyle fiyatları artmakta olan Enron hisselerinin opsiyonlarından elde ederken şirket çalışanları da yüksek maaşlar ve ödenekler almaktadır. Enron'a verilen kredilerden sorumlu olanlar ise verdikleri krediler ve danışmanlık hizmetleri karşılığında büyük primler elde ederler. Ancak bu uzun sürmez ve kişisel çıkarlar maksimizasyonuna dönük çabalar bir borsa çöküşü ile nihayete erer. Bu durumdan muzdarip olan yatırımcılar bu piyasalardan çekilerek yatırımlarını bu kez gayrimenkul gibi alternatif yatırım araçlarına yöneltirler. 1990'dan 2006 yılına kadarki

¹³ Enron skandalı ve ABD Mortgage krizinin rasyonalite bağlamında analizi için Bkz. (Akerlof & Shiller, 2010)

zaman diliminde, ABD’de konut fiyatları hızla yükselmeye başlar. İpotekli konut kredileri başlangıçta piyasanın %5’ini oluştururken, zamanla bu oran %20’ye ulaşır. Üstelik önceden bu krediler sıkı bir denetim sonunda verilirken, artık borç alanların ödeme koşullarına uygun olmayacak şekilde ve daha sonra yüksek faiz talep edileceği gizlenmek suretiyle kullanılmaya başlanılır. Bu durumun mimarı olarak kredi verenler kendi ürünlerine güvenmediklerinden, çabucak elden çıkarmanın yollarını aramaya başlarlar. Bunun için de bu kredileri dilimleyip muhtelif yöntemlerle farklı dilimleri bir araya getirmek suretiyle yeniden paketlenen bu ürünler, alınan komisyonlar karşılığında derecelendirme kuruluşlarınca yüksek notlarla ambalajlanırlar. Dolayısıyla konut satın alanlardan, ipotekleri menkul kıymetleştirenlere, derecelendirme kuruluşlarına ve bu menkul kıymetlerin alıcılarına kadar uzanan, her birinin çıkarlarının maksimizasyonu üzerine kurulu yeni bir iktisadi denge oluşur. Ancak bu denge 2008 yılına gelindiğinde bir krizle nihayete erer ve “Mortgage Krizi” olarak adlandırılarak iktisat ve finans tarihindeki yerini alır.

Esasen her iki vakada da, iktisadi rasyonalite açısından bakıldığında herkesin kendi çıkarını maksimize ettiği bir denge durumu söz konusudur. Ancak tıpkı *mahkûmlar açmazında* ortaya konulduğu gibi, kişisel çıkarlarının maksimizasyonu güdüsüyle hareket eden taraflar bir krizle yüz yüze gelivermişlerdir. Bu bağlamda rasyonalite kavramının kişisel çıkarın maksimizasyonu şeklinde tanımlanmasına yönelik bir diğer eleştiri ise, bireyin kendi çıkarının maksimizasyonu yönünde kararlar verdiğiğinde rasyonel davrandığını kabul etmenin, ettiği iktisattan dışlayabileceği hususundadır (Sen, 1988).

Rasyonalite tanımının temel unsurlarından *bireyin seçiminin tutarlılığı* kavramına dönük eleştiriler de söz konusudur (Sen, 1993). Buna göre bireyin içsel tutarlılığı, değerler ya da normlar gibi dışsal faktörlerden etkilenebilmekte, böyle bir durumda ise birey birbiriyle çelişen kararlar verebilmektedir¹⁴. Söz gelişi akşam yemeğinde tabakta kalan son elmayı nezaket kuralları çerçevesinde almamayı tercih eden bir bireyin, tabakta iki elma olduğunda,

¹⁴ Üstelik bireyin içerisinde bulunduğu bağlamın doğrudan bir normla ya da nezaket kuralı ile ilişkili olması da gerekmemektedir. Nitekim alternatiflerin sunuş tarzı da, bireyi içsel tutarsızlığa sürükleyebilmektedir. Sunuş tarzının kararlara etkisi, çerçeve yanılıgısı başlığı altında ayrıntılı bir şekilde ele alınmaktadır.

seenekleri bu kez hi elma yememe, elmalardan birini ya da diğeri yeme şeklinde olacaktır. Elma yemeye karar vermesi durumunda ise iki elmadan birini, *nezaket kurallarını* ihlal etmeden yiyebilecektir. Tercihlerin tutarsız olmasından dolayı rasyonalite kuralının ihlal edildiğı bu durumun, bireyin karar sürecindeki bağlam göz önünde bulundurulduğunda, rasyonel olmadığı söylenemez. İşte tam da bu noktada Hayek (2012) ve Sugden (1989), kim tarafından ve hangi rasyonel ilkeler göz önünde bulundurularak tesis edildiğinin bile bilinmediğı normların ya da kurumların, toplumun evrimsel sürecinin planlanmamış sonuçları olduğunu, toplumsal düzenin oluşmasını sağlayan bu kurumların ve altında yatan inançların rasyonalite ile ilişkisini irdeler. Geleneksel kuramın rasyonel tercihler bağlamındaki kabullerine ilişkin temel bir eleştiri sosyoloji alanından gelir. Zira insanların tercih ve beğenilerinin sınıf bağlamıyla açıklanabileceğine dair güçlü bulgular vardır. Bu bulgulara göre insanların beğeni ve tercihleri estetik bir kaygıdan ziyade ait oldukları sınıfsal pozisyonlarla ilgilidir. Zira bu bakış açısına göre tercih ve beğeni insanların yaşamlarını sürdüregeldikleri sınıfların yaşam tarzı pratiklerinin davranışsal eğilimler şeklinde tecessümüdür. Buna göre yemek yeme usulleri, müzik zevkleri ya da sportif zevkler gibi beğeni kalıpları, hem bireylerin ait oldukları sınıfsal konumlarına göre, hem de aynı sınıf içerisindeki diğeri bireylerden kendilerini farklılaştırmasının bir ölçütü olarak tezahür eder (Bourdieu, 2015).

Neoklasik iktisadın varsayımlarına yönelik bu kökten eleştirilerin yanı sıra, noksanlıklarına ya da kusurlarına içeriden bir gözle bakarak revize edici bir üslupla eleştiren yaklaşımlar da vardır. Başka bir söyleyişle davranışsal yaklaşımların diğeri cephesinde ise standart iktisat teorisini referans alıp rasyonalite anlayışını merkezde tutarak, neoklasik yaklaşımın reddine ya da alternatif model arayışına teşebbüs etmeden, bilişsel sapmalara yönelik varsayımları revize ederek aykırılıkları alternatif teorilerle açıklamaya çalışan yaklaşımlar bulunur. Neoklasik yaklaşıma daha yakın duran bu anlayışı benimseyen başta Kahneman olmak üzere pek çok araştırmacı, rasyonel bireyi referans noktası kabul ederek aile içi eleştirilerle psikolojiyi iktisada entegre etmeyi amaçlar (Ruben ve Dumludağ, 2016).

1.2 GELENEKSEL FİNANS KURAMLARI VE RASYONALİTE

Geleneksel finans ekseriyetle Modigliani ve Miller'in (1958) arbitraj ilkeleri ya da kâr payı kayıtsızlığı¹⁵ görüşü (Miller ve Modigliani, 1961); Sharp (1964), Lintner (1965), Black, Jensen ve Scholes (1972) sermaye varlıklarını fiyatlama teorileri; Markowitz'in portföy ilkeleri (Markowitz, 1952; Markowitz, 1959); Black ve Sholes (1973) ve Merton'ın (1976) opsiyon fiyatlama teorileri ile Fama'nın (1970) etkin piyasalar hipotezi çerçevesinde hayat bulur (Statman, 1999). Geleneksel finansın başyapıtları payesine sahip olan bu çalışmalarda teoriler, yatırımcıların ve piyasaların rasyonel olduğu varsayılmak suretiyle geliştirilmiştir. İşte geleneksel finansın bu rasyonalite varsayımı, neoklasik iktisadın *beklenen fayda teorisindeki* rasyonel insan kurgusundan başkası değildir. Diğer bir deyişle geleneksel finans teorileri için de rasyonel insan, alternatifler hakkında kusursuz bilgiye sahip, kişisel çıkarları için maksimizasyon hesapları ve belirsizlik koşullarında mükemmel bir şekilde olasılık tahminleri yapabilen ve nihayetinde duygularından ve duyu deneyimlerinden arınmış bağımsız bir şekilde karar veren soyut bir insan modelidir.

Bu bağlamda geleneksel finansın, neoklasik iktisadın rasyonalite anlayışı üzerine inşa ettiği bütün bu teorilerde, alternatifler hakkında tam bilgiye sahip ve sadece kişisel çıkarlarının peşinde koşan ve bu suretle rasyonalitenin kendisinde toplandığı bir insan modeli varsayılmaktadır (Pompian, 2012). Ancak gerçeğe örtüşmeyen bu soyut rasyonel insan kurgusu, pek çok yönden eleştirilere açık haldedir.

1.2.1 Belirsizlik Koşullarında Rasyonel Seçim: Beklenen Fayda Teorisi

Belirsizlik koşullarında rasyonel seçimin literatürdeki karşılığı *beklenen fayda teorisidir* (Neumann ve Morgenstern, 1953). Beklenen fayda teorisinin temelinde ise, Bernoulli'nin

¹⁵ Miller ve Modigliani (1961), yatırımcıların rasyonel olduğunu, yani çok serveti aza tercih edeceklerini, üstelik servetlerindeki belirli bir artışın nakit temettü dağıtımı ya da sahip oldukları hisse senetlerinin piyasa değerindeki bir artış şeklinde gerçekleşmesi arasında kayıtsız kalacaklarını ifade ederler.

(1954) St. Petersburg Paradoksuna¹⁶ ilişkin çözümü (Plous, 1993) ile Pascal ve Fermat'ın olasılıklı bir oyunda beklenen değeri hesaplamak için ortaya koydukları Eşitlik 1.1'de yer alan (Yılmaz, 2009), beklenen değer formülünün¹⁷ yerini beklenen fayda¹⁸ formülü almıştır.

$$u = \sum U(x_i)p_i \quad (1.1)$$

Beklenen fayda teorisinin temel ilkesi, riskli koşullarda beklenen faydanın maksimize edilmesidir. Diğer bir deyişle John Von-Neumann ve Oscar Morgenstern'in ilk aksiyomatik altyapısını sağladıkları bu teoriye göre *rasyonel bir birey, bir seçimle karşı karşıya kaldığında beklenen faydası en yüksek olan alternatifi tercih eder*. Bu teori ile birlikte iktisattaki ilk yıllarında psikolojik bir mahiyet arz eden fayda kavramı, zaman içerisinde bu anlamların tasfiye olarak seçimini belirli kurallara (aksiyomlara) göre yapan insan davranışını anlatan bir niteliğe bürünmüştür. Bu davranış modeli, bilindiği kadarıyla ilk kez Öklid tarafından kullanılan aksiyomatik yöntemle izah edilmiştir (Hastie ve Dawes, 2010). Bu bağlamda, her ne kadar farklı kaynaklar bunlara eklemeler yapsa da, ideal bir rasyonel seçimin nasıl olacağına ilişkin temel aksiyomlar şu şekilde sıralanabilir (Arrow, 1970;

¹⁶ St. Petersburg Paradoksu: Yazı gelene kadar yazı-tura atmak suretiyle oynanan bir oyunda kazanılacak ödül, madeni paranın yazı gelene kadar kaç kez atıldığına bağlıdır. Bu oyunun beklenen değeri, aşağıda hesaplandığı gibi sonsuz olacaktır.

Beklenen değer = $(1/2)(\$2.00) + (1/4)(\$4.00) + (1/8)(\$8.00) + \dots + (1/2)^K(\$2.00)^K$

Beklenen değer = $(\$1.00) + (\$1.00) + (\$1.00) + \dots + (\$1.00)$

Beklenen değer = ∞

Bu durumun paradoks olarak ifade edilmesinin nedeni ise, oyunun beklenen değerinin sonsuz olmasına rağmen, bu oyunu oynamak için çok az insanın büyük paralar yatırmasıyla ilgilidir. Bernaulli bu paradoksu paranın *azalan marjinal faydası* ile izah etmektedir. Diğer bir deyişle insanlar karar verirken paranın değerinden ziyade faydasını göz önünde bulundururlar. Neo-klasik iktisadın, tüm psikolojik değişkenleri temsilen modele dahil ettiği “faydanın maksimizasyonu” ilkesi, üstü kapalı bir şekilde bireyin *alternatiflerin faydasına ilişkin bilgi sahibi olduğunu* da varsayar; ancak bireyin alternatiflerin faydasına ilişkin *bilgiyi nasıl temin ettiğini* dikkate almaz.

¹⁷ Beklenen değer bir değişkenin alabileceği bütün değerlerin, olasılıklarıyla çarpılması sonucunda elde edilen değerlerin toplamını ifade etmektedir. x olası sonucu, p olasılığı temsil ederken beklenen değer: $\sum x_i p_i$ formülü ile hesaplanmaktadır.

¹⁸ Beklenen fayda eşitliğinde yer alan u , beklenen faydayı, $U(x_i)$, x_i sonucu elde etmenin faydasını, p ise olasılığı temsil etmektedir.

Pindyck ve Rubinfeld, 2013; Nicholson ve Snyder, 2011; Hastie ve Dawes, 2010; Yılmaz, 2009):

- **Bütünlük [Completeness]:** Bireyler herhangi iki alternatifi karşılaştırmak veya sıralamak suretiyle ya bu alternatiflerden birini diğerine tercih ederler ya da ikisi arasında kayıtsız kalırlar. Diğer bir deyişle birey iki alternatif arasında seçim yaparken ya A'yı B'ye veya B'yi A'ya tercih eder ya da A ve B arasında kayıtsız kalırlar.
- **Geçişlilik [Transitivity]:** Bir birey A'yı B'ye, B'yi C'ye tercih ediyorsa, A'yı C'ye tercih ediyor demektir. Bu durum bireyin tercih sırlamasında içsel bir tutarlılık gösterdiğini ifade eder.
- **Süreklilik [Continuity]:** Şayet birey "A'yı B'ye tercih ettiğini" ifade etmişse, A'ya yakın tüm durumları B'ye tercih eder. Diğer bir deyişle olasılıklarda meydana gelen küçük değişiklikler bireyin tercih sırlamasını etkilemez.
- **Bağımsızlık [Independence]:** A ve B gibi iki alternatif arasında tercih yaparken, C gibi üçüncü bir alternatfin sürece dâhil olmasının, bireyin A ve B alternatifleri için yaptığı tercih sırlamasını etkilemeyeceğini ifade eder.
- **Alternatiflerin Olasılık Dağılımlarını Bütünsel Olarak Değerlendirilmesi:** Bu durum birden çok aşamalı kararlarda bireylerin alternatiflerin her bir aşamasındaki olasılıkları dikkate aldıklarını ifade eder.

Rasyonel seçim kavramını bütün psikolojik varsayımlardan arınmış, saf bir forma kavuşturarak nihai noktasına ulaştıran isim ise Savage'tir (1954). Mantığın belirsizliği de kapsayacak şekilde genişletilmesini amaç edinen Savage, olasılıkları tercihler cinsinden tanımlamıştır. Çok paranın daha az olana tercih edileceği varsayımıyla aksiyomlara uyacak şekilde, A olayına bağlı 100 TL'nin, B olayına bağlı 100 TL'ye tercih edilmesini, bireyin A'yı B'den daha olası görmesine bağlamak suretiyle *özel olasılık* kavramını ortaya atmıştır (Sugden, 1991). Rasyonalite kavramının tarihsel seyri içerisinde geldiği bu nokta, büründüğü araçsal formunun en radikal boyutunu temsil ederken, matematiksel ve aksiyomatik bir dil ile ifade edilen şekliyle hâkim iktisadi ve finansal anlayışın temel varsayımlarından biri olmuştur. *İktisadi ya da finansal aktörün seçimleri artık beklenen fayda teorisinin*

aksiyonlarını tatmin ediyorsa rasyonel olarak kabul edilir (Yılmaz, 2009). Diğer bir deyişle neoklasik iktisadın rasyonel insanı artık (Savage, 1954); bütün alternatifler hakkında bilgi sahibi olan, her bir alternatifi beklenen faydasını hesaplayabilen ve seçim yaparken fayda fonksiyonunun beklenen değerini maksimize edecek alternatifi tercih eden bir bireydir. Neoklasik iktisadın ulaştığı bu rasyonel insan tanımı, finansın başlıca teorilerinin şekillenmesinde de üzerine düşen rolü oynamaya hazır hale gelmiştir.

1.2.2 Modern Portföy Teorisi

Modern portföy teorisinin temelleri, Harry Markowitz'in (1952; 1959) portföy seçimi problemine getirdiği, ortalama-varyans modeli olarak da bilinen yaklaşıma dayanmaktadır. Markowitz, portföy riskini azaltmak için basit bir çeşitlendirme mantığı ile sadece menkul kıymet sayısının artırılmasına dayalı geleneksel portföy yaklaşımından farklı olarak, etkin portföy seçiminde getiri ve riski göz önünde bulunduran bir model ortaya koymuştur. Modelde beklenen getiri ve risk objektif kriterlere dayalı olarak matematiksel ve istatistiksel yöntemler yoluyla ölçülmektedir. Bir diğer ifadeyle modern portföy teorisi beklenen getirilerin ortalaması ve varyansının bir seçimidir. Markowitz çeşitlendirmesinde portföye, aralarında tam pozitif ilişki bulunmayan varlıklar dâhil edilmekte, böylece beklenen getiriden vazgeçmeden risk azaltılabilmektedir. Başka bir deyişle Markowitz çeşitlendirmesi, getiriler arasındaki korelasyonları ve getirilerin standart sapması ile ölçülen riski göz önünde bulundurarak belirli bir getiri düzeyi için riskin minimize edilebileceğini ortaya koymaktadır. Üstelik portföyün varyansı olarak ölçülen risk, yalnızca farklı varlıkların getirisinin bireysel varyanslarına değil, aynı zamanda tüm varlıkların ikili kovaryanslarına da bağlıdır. Dolayısıyla, bir varlığın riskine ilişkin temel husus, her varlığın tek başına riski değil, her varlığın toplam portföyün riskine katkısıdır. Bu bağlamda modern portföy teorisi, geleneksel portföy yaklaşımından farklı olarak varlıkların birbirleri ile ilişkisi üzerinden portföyü bir bütün olarak değerlendirmektedir.

Markowitz, portföy seçimine ilişkin çözümünü yatırımcıların rasyonel olduğu, yani seçimlerini portföylerin risk ve getiri düzeylerine göre yaptığı varsayımı üzerine inşa

etmiştir. Dolayısıyla modern portföy teorisinde birey, riskini minimize ederken iskonto edilmiş beklenen getirileri maksimize etme gayretindedir. Beklenen getirileri öznel ya da nesnel olasılıkları kullanarak hesaplayan birey, riski, yani beklenen getiriden sapma ihtimalini hesaplarken de yine bu olasılıklara müracaat etmektedir. Daha yüksek olan beklenen getiriye daha düşük olana, daha düşük olan riski, yani standart sapmayı daha yüksek olana tercih etmektedir. Diğer bir deyişle modern portföy teorisinde rasyonel bireyin daima riskten kaçan bir yapıya sahip olduğu varsayılmakta, risk peşinde koşan bir birey göz önünde bulundurulmamaktadır (Statman, 1999).

Sonuç olarak modern portföy teorisi, her biri değişen özelliklere sahip çok sayıda farklı varlığa ilişkin karmaşık ve çok boyutlu portföy seçimi sorununu, fayda fonksiyonu maksimizasyonu, farklı varlıkların beklenen getirileri, varlıkların varyansı, kovaryansı ve yatırımcının bütçe kısıtlarına indirgemiş olmaktadır.

1.2.3 Sermaye Varlıklarını Fiyatlama Modeli

Sermaye varlıklarını fiyatlama modeli (Treynor, 1962; Sharpe, 1964; Lintner, 1965; Mossin, 1966), Markowitz'in ortaya koyduğu, bir dönemlik ufka sahip, riskten kaçınan yatırımcıların, yalnızca beklenen getirileri ve getirilerin varyansını önemseydiği ortalama-varyans modeli üzerine inşa edilmiştir (Fama ve French, 2004). Markowitz'in, varlığın riskinin portföyün riskine katkısını göz önünde bulunduran portföy seçimi yaklaşımından farklı olarak model, her bir varlığa ilişkin risk ölçütünü ve piyasa dengede iken risk ile getiri arasındaki ilişkileri analiz etmede kullanılmaktadır. Başka bir ifadeyle sermaye varlıklarını fiyatlama modeli sistematik riskin¹⁹ ve bir varlığın değeri üzerindeki etkisinin ölçülmesine odaklanmakta ve her bir varlığın riskini piyasa portföyüne göre tanımlamaktadır. Bu suretle model, varlığın

¹⁹ Piyasa riski olarak da ifade edilen sistematik risk, tüm ekonomiyi ve tüm yatırım varlıklarını etkileyen gelişmelerden kaynaklanan genel risk, sistematik olmayan risk ise her varlığın kendine özgü riskler olarak tanımlanmaktadır.

sistematik risk veya piyasa riski olarak da bilinen ve çeşitlendirme ile ortadan kaldırılamayan riske karşı duyarlılığını ölçmektedir.

Sermaye varlıklarını fiyatlama modelinde, bir menkul kıymetin getirisinin pazar portföyünün getirisi ile ilişkisini, yani riski beta katsayısı göstermektedir. Başka bir deyişle belirli bir varlığın beta değeri, o varlığın riskli menkul kıymetlerden oluşan tüm piyasa portföyünün riskine olan marjinal katkısını açıklamaktadır. Beta katsayısı 1'den büyük olan hisse senetleri, toplam portföy riski üzerinde ortalamanın üzerinde bir etkiye sahipken, beta katsayısı 1'den küçük olan hisse senetleri, toplam portföy riski üzerinde ortalamanın altında bir etkiye sahiptir. Sonuç olarak sermaye varlıklarını fiyatlama modeline göre, etkin bir sermaye piyasasında risk primi ve dolayısıyla bir varlığın beklenen getirisi de beta değeri ile doğru orantılı olarak değişmektedir.

Sermaye varlıklarını fiyatlama modelinden yola çıkarak oluşturulan menkul kıymet pazar doğrusu [security market line, SML] yatırımcının hiç risk almadığı durumda risksiz orandan getiri elde edebileceğini göstermektedir. Sermaye piyasası doğrusu aynı zamanda daha fazla getiri elde etmek isteyen bir yatırımcının belirli bir riske katlanacağına da işaret etmekte, yatırımcının alacağı ilave getirinin, aldığı riskin bir ödülü olduğunu açıklamaktadır (Karan, 2022). Dolayısıyla yatırımcıya bir hisse senedinin piyasa şartlarında istenilen getirisinin ne olması gerektiği konusunda fikir vermektedir.

Sermaye varlıklarını fiyatlama modellerinin varsayımlarına bakıldığında geleneksel finansın rasyonalite anlayışının etkileri varlık fiyatlama modellerinde de göze çarpmaktadır. Başka bir ifadeyle varlık fiyatlama modelleri de, neoklasik iktisadın rasyonalite anlayışını benimsemek suretiyle beklenen değer ve standart sapma hesapları üzerine kurulmuştur (Baker ve Nofsinger, 2010). Buna göre varlık fiyatlama modelleri yatırımcıların riskleri aynı olan iki portföyden getirisi yüksek olanı, getirileri aynı olan iki portföyden standart sapması düşük olanı seçeceğini varsaymaktadır. Başka bir deyişle yatırımcılar riskten kaçarken fayda fonksiyonunu maksimize etmenin peşindedir. Modeller ayrıca piyasanın etkin olduğunu, yani her türlü bilginin fiyatlara yansıdığını varsaymaktadır. Buna göre bütün yatırımcılar

aynı bilgiye, analiz ve yorumlama yeteneğine sahip olduğundan beklentileri de homojendir (Firth, 1977).

1.2.4 Etkin Piyasalar Hipotezi

Etkin piyasalar hipotezi, bir varlığın fiyatının, piyasada mevcut olan tüm bilgileri yansıttığını ifade etmektedir (Fama, 1970). Zira Fama'nın etkin piyasalar hipotezindeki birey, tüm bilgilere haiz ve bu bilgileri fiyatlara rasyonel bir şekilde yansıtan bireydir. Bu bağlamda geleneksel finansa göre piyasaların rasyonelliği, esasen piyasa katılımcılarının rasyonelliğinden beslenmektedir. Başka bir deyişle rasyonel piyasa katılımcıları, sahip olunan bütün bilgiyi ve gelecekteki belirsizliklerle ilgili olasılıkları cari fiyatlara rasyonel bir şekilde yansıtmakta ve böylece Fama'nın iddia ettiği ve normalüstü getiri elde etmenin mümkün olmadığı etkin piyasalara ulaşılmaktadır²⁰. Kısa vadede bu piyasalardaki fiyat değişiklikleri artık yeni bilginin fiyatlara yansımından ibaret iken, uzun vadede ise alınan riskin bedelini yansıtmaktadır (Baker ve Nofsinger, 2010)²¹.

Bu bağlamda teori, piyasalardaki etkinliğin bilginin içeriğine göre üç formda tezahür edeceğini savunmaktadır (Fama, 1970). *Zayıf form etkinliği*, geçmiş fiyat hareketleri kullanılarak normal üstü getiri elde edilemeyeceği varsayımına dayanmaktadır. Buna göre, geçmiş fiyat verilerinin teknik analizinin normalüstü getiri sağlamaya katkısı olmayacaktır.

²⁰ Eugene Fama'nın ortaya koyduğu etkin piyasalar hipotezine göre etkin bir piyasada hiçbir yatırımcı normalüstü bir getiri sağlayamaz. Ancak ampirik bulgular, piyasalarda normaüstü getiri sağlanabildiği yönündedir. Bu anomaliler hakkında ayrıntılı bilgi için bkz. (Karan, 2022).

²¹ Nobel ödül komitesi 2013 yılında iktisat dalındaki ödül sahiplerine karar verirken Eugene Fama ve Robert Shiller arasında çelişki yaşarlar. Bir tarafta *piyasaların etkin olduğunu ve normal üstü getiri elde edilemeyeceğini* savunan Fama, diğer tarafta *piyasaların sıklıkla irrasyonel olduğunu ve bu nedenle normal üstü getiri elde edilebileceğini* savunan Shiller bulunmaktadır. Komite bu çelişkili durumu çözebilmek için değerlendirmeyi "piyasada düzenli bir şekilde normaüstü getiri elde etmenin mümkün olup olmadığı" ilkesine göre yapmaya karar verir. Etkin piyasaları, rasyonel piyasalar [rational markets] ve yenilmesi güç piyasalar [hard-to-beat] olarak alt gruplara ayırırlar. Bu durumda rasyonel piyasalar, menkul kıymetin içsel değerinin piyasa değerine daima eşit olduğu bir piyasayı ifade ederken, yenilmesi güç piyasalar ise menkul kıymetin piyasa fiyatı ve içsel değeri arasındaki farklılıkları kullanarak bazı yatırımcıların her zaman normal üstü getiri elde edebilecekleri, ancak ortalama bir yatırımcının buna genellikle pek muktedir olamadığı piyasaları ifade eder. Böylece piyasanın her zaman rasyonel olmadığı, ortalama bir yatırımcının da normal üstü getiri elde etmesinin oldukça güç olduğu hususunda fikir birliğine varılır. Nihayetinde 2013 Nobel Ekonomi Ödülü "varlık fiyatlarının deneysel analizi" üzerine çalışmaları nedeniyle Eugene Fama, Lars Peter Hansen ve Robert Shiller üçlüsüne verilir.

Zira fiyatlar rassal yürüyüş hipotezine göre oluştuğu için hisse senedi fiyatındaki değişimlerin geçmişteki fiyat değişimleri ile herhangi bir bağlantısı yoktur. *Yarı güçlü form etkinliğinde*, fiyatların geçmiş fiyat verilerinin yanı sıra halka açık bilgileri de yansıttığı varsayılmaktadır. Daha açık bir ifadeyle yarı güçlü formdaki piyasa etkinliğinde firma ile ilgili haberler, gazete yazıları, faaliyet raporları, finansal raporlar, firma duyuruları, ekonomik tahminler gibi herkesin erişimine açık bilgi anında fiyatlara yansıdığı için temel analiz ile kalıcı olarak normalin üstünde getiri sağlanamaz. Başka bir söyleyişle etkin piyasalar hipotezinin bu formu, yatırımcıların kazanç beyanları, medya raporları, temel analiz, mali tablolar analizi gibi bilgileri kullanarak normalin üstünde getiri elde edemeyeceklerini öne sürmektedir. Etkin piyasalar hipotezinin diğer formu ise fiyatların kamuya açık bilgilere ek olarak özel bilgileri de yansıttığını iddia eden en güçlü biçimdir (Malkiel, 1973). *Güçlü formdaki piyasa etkinliği* yöneticiler ve çalışanlar gibi bazı bilgilere tekeli erişimi olan kişilerin, bu bilgileri kullanarak sürekli olarak normalüstü getiri sağlayamadığı piyasa formunu ifade eder. Bu nedenle kuvvetli formda etkin bir piyasada, içsel bilgiye ulaşabilecekler dahi sürekli normalüstü getiri elde etmek için bir strateji geliştiremezler.

Etkin piyasalar hipotezinin taraftarları kendilerine yönelik eleştirileri ya da optimal kararlar veremeyen bireyler ile ilgili izahlarını ise “arbitraj” argümanına dayandırmaktadırlar. Nitekim bu izahın temelleri Friedman’dan (1953) izler taşımaktadır. Buna göre “rasyonel olmayan bireyler, elde ettikleri yetersiz sonuçlarla cezalandırılmış olacaklarından, zaman içerisinde ya daha iyi kararlar vermeyi öğrenecekler ya da piyasayı terk edeceklerdir. Üstelik piyasa katılımcılarının yaptığı herhangi bir hata da birbirleri ile ilişkili olmadığından, bu hatalar piyasa fiyatlarını etkileyecek güce de sahip olmayacaktır” (Baker ve Nofsinger, 2010). Diğer bir deyişle rasyonel yatırımcıların bir arbitraj imkânı gördüğünde bunu değerlendireceklerini ve böylece etkin piyasa koşullarının tekrar sağlanacağını savunmaktadırlar.

1.2.5 Rassal Yürüyüş Hipotezi

Rassal yürüyüş hipotezi, ardışık fiyat değişikliklerinin birbirinden bağımsız ve tamamıyla rastlantısal olduğunu, dolayısıyla hisse senedi fiyatlarını öngörmenin mümkün olmadığını savunmaktadır (Cootner, 1964; Fama, 1965a; Fama, 1965b). Başka bir ifadeyle rassal yürüyüş hipotezi, fiyatları tahmin etmeye yönelik teknik analiz ve temel analiz olmak üzere başlıca iki yaklaşım olduğunu, ancak piyasadaki fiyat oluşumunun rastlantısal karakteri nedeniyle ne teknik analiz, ne de temel analiz kullanılarak hisse senetlerinin gelecekteki fiyat hareketlerini tahmin etmenin mümkün olmadığını iddia etmektedir.

Aslında rassal yürüyüş hipotezi de temelinde etkin piyasa öncülünden yola çıkmaktadır (Fama, 1965b). Bu bağlamda etkin bir piyasa, her birinin bireysel menkul kıymetlerin gelecekteki piyasa değerlerini tahmin etmeye çalıştığı ve önemli güncel bilgilerin tüm katılımcılara neredeyse ücretsiz olarak sunulduğu, çok sayıda rasyonel ve kâr maksimizasyonu peşinde koşan yatırımcının aktif olarak rekabet ettiği bir piyasa olarak tanımlanmaktadır. Dolayısıyla etkin bir piyasada çok sayıda zeki katılımcı arasında ortaya çıkan rekabet, gerçekleşmiş ya da gerçekleşmesi muhtemel olaylara dayalı bilgilerin etkilerinin fiyatlara *anında* yansıdığı bir duruma yol açmaktadır. Fama'ya (1965b) göre etkin bir piyasada gerçek fiyatın, gerçek değere anlık olarak uyum sağlamanın [instantaneous adjustment] iki anlamı vardır. Birincisi gerçek fiyatların başlangıçta gerçek değerdeki değişikliğe ya aşırı uyum sağlayacağı ya da yetersiz uyum sağlayacağı, ikincisi ise gerçek fiyatların gerçek değere tam olarak uyum sağlamaındaki gecikme ya da önceden uyum sağlama durumlarının ortaya çıkmasıdır. Bu nedenle her iki anlamda da etkin bir piyasada anlık uyum sağlama özelliği bireysel hisse senetlerindeki ardışık fiyat değişikliklerinin rassal olacağı anlamına gelmektedir. Dolayısıyla menkul kıymetin etkin bir piyasada herhangi bir zamanda oluşan gerçek fiyatı, o menkul kıymetin temel değerinin iyi bir tahminine eşdeğerdir. Başka bir ifadeyle bir menkul kıymetin fiyatını etkileyen bilanço, yönetim kalitesi, sektör, ekonominin görünümü gibi temel faktörlere ilişkin bilginin etkin bir piyasa da fiyatlara hemen yansımalarının fiyatlarında oluşturacağı rassallık nedeniyle temel analiz ile fiyat öngörüsünün mümkün olamayacağına işaret etmektedir.

Buna ek olarak rassal yürüyüş hipotezi, fiyatları tahmin etmenin teknik analiz ile de mümkün olmadığını savunmaktadır. Zira teknik analizin temel varsayımı bireysel menkul kıymetlerdeki geçmiş fiyat davranışı kalıplarının gelecekte de tekrar etme eğiliminde olacağı yönündedir. Dolayısıyla teknik analiz ile hisse senedi fiyatlarını tahmin ederek potansiyel kazancı artırmanın yolu bireysel menkul kıymetlerdeki ardışık fiyat değişikliklerinin bağımlı olmasına, yani geçmişin tekrerrüt etmesine bağlıdır. Bu bağlamda rassal yürüyüş hipotezi, hisse senedi fiyatlarının rastlantısal olarak oluştuğunu, dolayısıyla bir hisse senedinin geçmiş fiyat hareketlerinin veya eğiliminin, gelecekteki hareketini tahmin etmek için kullanmanın anlamsız bir değerinin olmadığını iddia etmektedir.

Dikkat edileceği üzere rassal yürüyüş hipotezinin temel analiz ile ilgili boyutu etkin piyasalar hipotezinin yarı-güçlü form etkinliği ile ilgili iken teknik analiz ile ilgili boyutu zayıf formda piyasa etkinliği ile ilgilidir. Sonuç olarak rassal yürüyüş hipotezini savunanlar basit bir “al-tut” stratejisinin, alım ve satımların doğru zamanını belirleyebilmek için teknik ve temel analizin karmaşık mekaniğini kullanan bir strateji kadar iyi olacağını ileri sürmektedirler. Dolayısıyla yatırımcıların hisse senedi fiyatlarını tahmin etmek için temel ve teknik analiz kullanmasının bir zaman kaybı olduğunu, bunun yerine bir endeks fonu alıp ellerinde tutmalarının daha iyi olduğunu savunmaktadırlar (Malkiel, 1973).

1.3 DAVRANIŞSAL FİNANSIN RASYONALİTE ELEŞTİRİLERİ

Neoklasik iktisadın pozitif bilimlerde kullanılan matematiksel modelleme araçlarına olan ilgisi, geleneksel finans kuramının teorisyenlerinde de rasyonel yatırımcının finansal davranışlarını matematiksel bir dille ifade etme çabası şeklinde tezahür etmiştir. Rasyonel, iradeli ve bencil olarak tanımlanan insanın risklerini minimize ederken verili faydasını [given utility] maksimize etme isteği yüksek ve soyut matematiksel problemlere dönüşmüştür. Ancak 1960’lardan sonra ana akım finans kuramıyla çelişen ampirik bulgular, kuramın soyut varsayımlarının sorgulanmasına yol açmıştır. Esasen Adam Smith’le (yani Ahlakî Duygular Kuramı (1759/2018) ve Milletlerin Zenginliği (1776/2016) adlı eserleriyle) başlayan iktisadın klasik döneminde psikolojik faktörlerin insanların davranışlarına yön verdiğinin

kabulü, rasyonel tercih teorisiyle ana akım kuramın dışında bırakılırken, 1980'lerden sonra geleneksel akımın temel varsayımlarına dönük eleştiriler getiren *davranışsal finans* yaklaşımları üzerinden alanyazınında tekrar görünür olmaya başlamıştır.

Davranışsal finans, psikolojik ve sosyolojik faktörlerin, yatırımcıların finansal kararları ve dolayısıyla finansal piyasalar üzerindeki etkilerini incelemektedir (Ricciardi ve Simon, 2000). İnsanların rasyonel olduğu varsayımı yerine, normal bir insan kabulüyle yatırımcıların akıl yürütme biçimlerini açıklamaya çalışmaktadır (Statman, 2008). Bu nedenle geleneksel finans teorisinden bazı noktalarda ayrılmaktadır. Ayrıştığı bu noktalardan birisi analiz araçlarıyla ilgilidir. Geleneksel finans teorisi varsayımlar üzerine kurulu soyut modeller üzerinde analizlerini gerçekleştirirken, davranışsal finans teorisyenleri bulgularını deneysel yöntemlerden elde etmektedirler. Geleneksel yaklaşımda öncelikle modeller belirlenip verilere göre modellerin doğruluğu test edilirken, davranışsal finans yaklaşımında insan davranışlarının gözleme dayalı olarak çözümlenmesi söz konusudur (Bayrak, 2012). Davranışsal finans, finansal problemleri matematiksel olarak çözmek için tasarlanmış, ideal ve haddinden fazla basitleştirilmiş varsayımlara dayanan bir anlayış yerine insanların, esasen kararlarını nasıl aldığını anlayabilmenin ve teorilerini bu gerçeklik üzerine inşa edebilmenin peşindedir. Buna göre yatırımcıların ve dolayısıyla piyasaların davranışlarını şekillendiren hususlar doğru bir şekilde belirlenebilirse finansal teoriler, kurgusal finansal davranışlar yerine gözlemlenen finansal davranışlar üzerine inşa edilebilecek ve böylece finansal neticeler daha iyi ve daha etkin bir hale getirilebilecektir (Pompian, 2012).

Geleneksel finans teorisinin temel varsayımlarından bir diğeri *beklenen fayda* teorisidir. Beklenen fayda teorisine göre insanlar seçeneklerin olası sonuçları ve olasılıkları hakkında tam bilgiye sahiptir. Bu verilere göre beklenen faydayı hesaplayarak en üst seviyede fayda sağlaması muhtemel alternatifi tercih etmektedirler. Ancak davranışsal finans teorisyenleri, insanların bir takım davranışsal eğilimlere sahip olmalarından dolayı geleneksel modellerin insanı tam olarak tanımlayamadığını savunmaktadır. Örneğin bulgular, insanların kayıptan kaçınma eğilimleri nedeniyle, aynı miktardaki kâr ve zarar durumları karşısındaki tutumlarının simetrik olmadığını göstermektedir. Benzer şekilde kayıptan kaçınma eğilimi

insanların kesin bir kazanç söz konusu olduğunda riskten kaçınırken, kesin bir kayıp karşısında riskli olanı tercih ettiğine de işaret etmektedir (Ackert, 2014).

Geleneksel kurama yönelik eleştirilerden biri de *etkin piyasalar hipotezi*yle ilgilidir. Etkin piyasalar hipotezine göre fiyatlar piyasadaki tüm bilgileri yansıttığından hiçbir yatırımcının sürekli olarak normalüstü getiri elde etmesi mümkün değildir. Bu bağlamda teori, piyasalardaki etkinliğin bilginin içeriğine göre üç formda tezahür edeceğini savunmaktadır. Zayıf form etkinliği, fiyatların geçmişlerindeki bilgileri yansıtması durumunda gerçekleşmektedir. Buna göre, geçmiş fiyat verilerinin teknik analizinin normalüstü getiri sağlamaya katkısı olmayacaktır. Yarı güçlü form etkinliğinde, fiyatlar geçmiş fiyat verilerinin yanı sıra halka açık bilgileri de yansıtmaktadır. Etkin piyasalar hipotezinin bu formu, yatırımcıların kazanç beyanları veya medya raporları gibi bilgileri kullanarak normal üstü getiri elde edemeyeceklerini öne sürmektedir. Etkin piyasalar hipotezinin diğer formu ise fiyatların özel bilgileri de yansıttığını iddia eden en güçlü biçimdir. Bu nedenle, içsel bilgiye ulaşabilecek şirket çalışanları dahi sürekli normalüstü getiri elde etmek için bir strateji geliştiremez. Ancak araştırmacılar, piyasalarda normalüstü getiri sağlanabildiğine işaret eden pek çok *anomali* ortaya koymuşlardır. Ampirik araştırmalardaki bu akım, davranışsal yaklaşımların güç kazanmasına büyük oranda destek olmuştur (Ackert, 2014).

Geleneksel kurama yönelik eleştirilerden bir diğeri, bireyleri özerk karar vericiler olarak kabul eden *izole birey* yaklaşımıdır. Ancak davranışsal finans yaklaşımları insanın sosyal bir varlık olmasının, kararlarının ve davranışlarının üzerinde etkili olduğunu iddia etmektedir. Başka bir deyişle insanın sosyal yönüne atıfta bulunarak, insan davranışının sosyal tercihlerden, ağlardan, sosyal kimliklerden, sosyal statüden, sosyal normlardan ve insanlar arasındaki işbirliği modellerinden etkilendiğini ifade etmektedir. Çoğu insanın işbirliğine sahip bir ruha, adalet ve karşılıklılık duygusuna sahip olduğuna dikkat çekmektedir. Davranışsal finans bu bağlamda insanın sosyalliğinin, insanların karar verme biçimlerine ve davranışlarının analizine karmaşıklık ve gerçeklik katmanı eklediğini savunur. Bu gerçekliğin geleneksel kuram tarafından ihmal edilmiş olması, politika yapıcılar tarafından sadece maddi teşviklere odaklanılmasına neden olmuştur. Sonuç olarak davranışsal finans

sosyal etkinin kurumların tasarımında, dolayısıyla da toplumun gelişimi üzerinde hem olumlu hem de olumsuz etkileyebilecek bir güce sahip olduğunu savunmaktadır (World Bank , 2015).

Geleneksel finans teorisinin temel kabullerinden bir diğeri tam bilgiye sahip olan insanların, bütün bilgiyi analize dâhil edip ince hesaplar yapmak suretiyle, yaptıkları hesapların ve verdikleri kararların farkında olarak tercihte bulunduklarıdır. Ancak davranışsal finans cephesindeki bulgular, insanların karar verirken kompleks hesaplar yapmak yerine bir takım kestirme [heuristics] yöntemler kullandıklarına işaret etmektedir. Başka bir söyleyişle insanların problem çözerken hem bilişsel hem de ortama bağlı kısıtlarla karşı karşıya kaldıklarında –ki çoğu zaman gerçek durum böyledir- rasyonel hesaplar yapmak yerine başparmak [rule-of-thumb] kurallarına başvurdukları görülür.

1.4 DAVRANIŞSAL FİNANSTA YATIRIMCI EĞİLİMLERİ

Davranışsal finans, geleneksel finansın temel kuramlarının gerçek hayatla bağdaşmayan varsayımlarını, geleneksel finansın standart olarak kabul ettiği durumlardan sapmaları ortaya koymak suretiyle reddetmektedir. Bu bağlamda rasyonaliteden sapmalar genellikle davranışsal eğilimler ya da yanlışlıklar olarak ifade edilmektedir.

1.4.1 Doğrulama Eğilimi

Algıda seçiciliğin bir türü olan doğrulama eğilimi (Pompian, 2006) bireyin inançları ya da hipotezlerini, kasıtsız bir şekilde, bunları tasdik ve teyit eden bilgi ve delillerle çözümlerken, bunların yanlış olduğunu gösteren bilgi ve delilleri görmezden gelme temayülüdür (Colman, 2006).

Yatırımcılar açısından bakıldığında doğrulama eğilimi, hali hazırda yapmış oldukları yatırımlara dair inançlarına aykırı olabilecek bilgileri görmezden gelirken sadece kendi düşünceleri ile uyumlu olan bilgileri araştırmalarına yol açabilmektedir. Nitekim piyasalar

ilişkin bilgilerin ve yatırımcıların yer aldığı sanal topluluklar ve doğrulama eğilimi üzerine yapılan bir çalışmada (Park ve diğerleri, 2013), bilgilerin tek taraflı bir bakış açısı ile değerlendirilmesinin genellikle yatırımcıların aşırı bir öz güvene sahip olarak performansları hakkında daha yüksek beklentiler içerisine girmelerine ve riskli yatırımlar yapmalarına yol açabildiği bulgusuna ulaşılmıştır. Buna ek olarak fiyat oluşumunu da etkileyen bu eğilim (Pouget ve Villeneuve, 2009), hisse senedinin fiyatının düşeceğine ilişkin emarelerin çok güçlü olmasına rağmen yatırımcıların bu belirtileri görmezden gelmesine ve hatta bazı hisse senetlerine aşırı bağlılıklarından dolayı, yetersiz düzeyde çeşitlendirilmiş portföyleri ellerinde tutmaya devam etmelerine de yol açabilmektedir (Pompian, 2006).

1.4.2 Örneklem Eğilimi

İnsanlar doğaları gereği, geçmişte yaşadıkları olaylardan edindikleri tecrübî bilgilerine dayanarak nesneleri ve düşünceleri sınıflandırma eğilimindedirler. Sınıflandırma sürecinde kullanılan bu tecrübî fikirler ya da bilgiler, bir takım örneklerle dayanarak yapılan genellemelerdir. Diğer bir deyişle bütüne ait özellikleri, bu bütünü temsil ettiği düşünülen küçük bir kısmını inceleyerek elde etme sürecidir (Som, 1995). Ancak bu genelleme, olasılık kurallarına göre değil de tecrübelerle dayanarak yapıldığı için öznel bir değerlendirmedir (Tversky ve Kahneman, 1974).

İnsanlar bu şekilde öznel olasılık değerlendirmesi yaparken istatistik biliminin bir takım kurallarını ihlal ederler (Kahneman ve Tversky, 1973, 1972). Bunlardan birincisi *önsel olasılığın ihlalidir*. Önsel olasılık kısaca, gözlemlere atıf yapmadan önce değerlendirilen ya da değerlendirilmesi gereken olasılıktır. Örneğin kişilik özellikleri bilinen bir bireyin mühendis mi yoksa avukat mı olduğunun tahmin edilmeye çalışıldığını düşünelim. Bu tahminin, toplumda yaşayan 100 kişinin 70'inin mühendis 30'unun avukat olduğunu göz önünde bulundurmadan yapılmaya çalışılması, temel-oran hatası ya da önsel olasılığın ihlali anlamına gelmektedir.

İstatistik kurallarının göz önünde bulundurulmasına engel olan sebeplerden üçüncüsü *kumarbaz yanılığıdır (gambler's fallacy)*. Kumarbaz yanılığında, tesadüfi süreçlerin oluşturduğu olaylar dizisinin, bu sürecin temel karakteristiğini hem genel olarak hem de lokal olarak temsil edeceğine inanılmaktadır. Bu duruma rulet çarkında dokuz kez kırmızı gelmesinin ardından, bir sonrakinde siyah gelme ihtimalinin çok yüksek olduğu düşüncesi örnek gösterilebilir. Oysa şans yanılığı (gambler's fallacy) olarak da bilinen bu yanılığa, önceki dokuz denemede olduğu gibi siyah gelme ihtimali yine yüzde ellidir. Kumarbaz yanılığı aynı zamanda, az miktarda bir veri ile genelleme yapma eğilimini ifade eden **küçük sayılar yasası** ile de ilgilidir (Tversky ve Kahneman, 1971) .

Örnekleme eğiliminin finansal piyasalara yansımaları gösteren çalışmalardan birinde (Cooper ve diğerleri, 2001), internet balonunun hüküm sürdüğü dönemde firma isminin sonuna “.com” ifadesini ekleyen firmaların hisse senetleri, başka bir deyişle firma ismine eklenen bu ifadenin firmanın hisse senedi fiyatlarına etkisi incelenmiştir. Firma ismine bu ekin getirilmesine ilişkin duyuruyu takip eden günlerde, bu firmalara ait hisse senetlerinin ortalama olarak %53 normalüstü getiri sağladığı, üstelik bu etkinin firmanın faaliyet alanının gerçekte internetle ne ölçüde ilgili olduğundan bağımsız olduğu bulgusuna ulaşılmıştır. Bu durum yatırımcıların internet balonunun yaşandığı dönemde, diğer internet şirketlerinin (dot-com) hisse senetlerini örnekleyerek, isim değişikliğini buna göre yeniden fiyatlandırıdıklarını göstermektedir.

1.4.3 Bilgi İllüzyonu

Bilgi illüzyonu [illusion of knowledge]; insanların, daha fazla bilgiye sahip olduklarında tahminlerinin daha doğru olacağına inanmaları temayülüdür (Montier, 2007). Bilginin niteliğinden²² ziyade, niceliğini esas almak suretiyle tesis edilen bu inanç (Tsai, 2008), genellikle öz güveni dayanaksız bir şekilde körükleyen (Tsai, 2008) zihin aldatmacalarından birisi olarak görülmektedir.

²² Bilginin niteliğini ne kadar çok olduğu değil, sahip olunan bilgi ile ne yapılabildiği belirler.

Bu durumun yatırımcı davranışlarına yansımaları ekseriyetle bilgi toplama, bu bilgiyi analiz etme ve karar alma süreçlerinde kendisini göstermektedir. Bilgi toplama sürecinde bilginin niteliğine bakmaksızın, niceliğini göz önünde bulundurmasına; analiz sürecinde bilginin yanlış yorumlanmasına ve karar sürecinde ise yatırımcının kendisini olduğundan fazla yetenekli görmesi ve kararlarını bu bakış açısıyla vermesine neden olmaktadır. Bu durum ise genellikle aşırı işlem, ölçsüz risk alma ve nihayetinde de portföy zararları olarak sonuçlanan sağlıklı yatırım kararları anlamına gelmektedir (Nofsinger, 2005).

Bilgi illüzyonunun bilgi toplama ve yorumlama bağlamında yatırımcı üzerindeki tesiri bilhassa günümüzde bilgi kaynağı olarak ekseriyetle interneti kullanan piyasa oyuncuları üzerinde görülmektedir. İnternet sayesinde, özellikle bireysel yatırımcılar, geçmiş fiyatlar, getiriler ve firmanın faaliyetlerine ilişkin performansı gibi tarihsel veriler ve buna ilave olarak gerçek zamanlı haberler, fiyatlar ve işlem hacimlerine ilişkin bilgilere kolaylıkla erişme imkânı elde etmişlerdir. Ancak pek çok bireysel yatırımcı, profesyonel yatırımcıların sahip olduğu deneyim ve eğitime sahip olmadıkları için temin ettikleri bilgileri nasıl yorumlayacakları konusunda yetersiz kalabilmekte ve bu sorunun üstesinden gelebilmek için ise yine internete başvurumaktadırlar (Nofsinger, 2005). Nitekim bu durumu ortaya koyan bir çalışmada (Dewally, 2003) iki internet sitesinin hisse senedi tavsiyelerini içeren mesajları incelenmiş ve satın alınması tavsiye edilen hisse senetlerinin, genel itibarıyla piyasadan önemli bir ölçüde farklı bir performans göstermediği bulgusuna ulaşılmıştır. Bu sonuç bir mânâ da bilginin niceliğinden ziyade niteliğinin önemli olduğu görüşünü desteklerken, çok bilginin yol açabileceği yanılgılara da işaret etmektedir. Bir diğer çalışma da (Tumarkin ve Whitelaw, 2001) ise bir internet sitesinden paylaşılan olumlu mesajların, takip eden gün ya da haftadaki pozitif hisse senedi getirileri ile ilişkili olmadığı, ancak nadiren de olsa yüksek miktardaki gönderilerin, yüksek işlem hacimleri ile ilişkili olabileceği bulgusuna ulaşılmıştır.

1.4.4 Kontrol İllüzyonu

Kontrol illüzyonu [illusion of control] insanların, gerçekte neticelerine hükmedemedikleri olayları kontrol edebileceklerine ya da en azından bunlar üzerinde etkilerinin olabileceğine

inanma temayülüdür (Pompian, 2006). Bu durumu ortaya koyan bir çalışmada (Langer, 1975), kontrol illüzyonuna neden olan etmenler ayrıntılı bir şekilde ele alınmıştır. Buna göre kontrol illüzyonuna neden olan etmenlerden bazıları, *seçim özgürlüğü* ve *aşinalık* ile ilgilidir. Bir deneyde katılımcılara, aşına olunan harfler ya da aşına olunmayan semboller içeren piyango biletleri içerisinde seçme hakkı tanınmış; daha sonra, seçtikleri bu biletleri, kazanma ihtimalinin daha yüksek olduğu başka bir piyango bileti ile değiştirme ya da mevcut biletleri ellerinde tutma hakkı verilmiştir. Deneyin sonucunda, biletlerini kendileri seçenlerin ya da aşına oldukları harfler içeren biletleri tercih edenlerin, bu biletleri yeni bilet ile değiştirmekten ziyade ellerinde tutma eğilimleri daha yüksek olmuş; bu bulgu benzer bir çalışma ile de desteklenmiştir (Wortman, 1975).

Bu bağlamda kontrol illüzyonu yatırımcı davranışları üzerinde de farklı şekillerde tezahür etmektedir. Bunlardan en önemlisi kontrol illüzyonunun yatırımcının aşırı öz güven duymasına neden olmasıdır. Bu durumda aşırı öz güven eğiliminin yol açtığı yatırımcı davranışları kontrol illüzyonunun da kapsamına girmektedir. Nitekim kontrol illüzyonu, yatırımcıların ihtiyatlı davranmaktan ziyade, piyasada çok fazla işlem yapmalarına ve bunun sonucunda da daha düşük getiri elde etmelerine de yol açmaktadır. Ayrıca yatırımcıların piyasaları etkin bir şekilde analiz edebilmesinin önündeki engellerden de biridir (O'Creevy ve diğerleri, 2003). Bunun kaynağında ise yatırımcıların, bilhassa online yatırımcıların yatırımcılarının performansları üzerinde gerçekte olduğundan çok daha fazla kontrol gücüne sahip olduklarına inanmalarıdır (Fellner-Röhling, 2004). Buna ilave olarak kontrol illüzyonu, riskin gerçekte olduğundan daha düşük algılanmasına neden olmakta (Houghton, 2000); bu da yatırım ile ilgili hem risk yönetimi hem de kârlılık boyutunda olumsuz sonuçlara yol açabilmektedir (O'Creevy ve diğerleri, 2003). Ayrıca kontrol illüzyonu, yatırımcıların yeterince çeşitlendirilmemiş bir portföy oluşturup, bu portföyü gerekli değişiklikleri de yapmadan muhafaza etmelerine de neden olmaktadır (Pompian, 2006).

1.4.5 Hazırda Bulunma Etkisi

Hazırda bulunma etkisi, daha hızlı ve kolay hatırlanabilen, daha dikkat çekici ya da görece daha yakın zamanda gerçekleşen olayların akıl yürütme süreçlerine daha yoğun bir şekilde nüfuz etmeleri suretiyle insanların kararlarında oynadıkları baskın rolü ifade eden bir eğilimdir (Baker ve Nofsinger, 2010). Daha kolay hatırlanan ihtimalleri, zihinde canlandırılması ya da kavranması daha zor olanlara kıyasla, daha olası görme eğilimi olarak da tanımlanabilen (Pompian, 2006) bu durum, çoğunlukla hazırda bulunanlarla örnekleme yapıldığından örnekleme eğilimi ile de yakından ilişkilidir (Demir, 2013).

Hazırda bulunma etkisinin düzeyini etkileyen unsurlardan biri dikkat unsurudur. Bu konuda Shiller (1999) yatırımcıların hisse senedi, tahvil ya da gayrimenkul gibi yatırım türlerine olan ilgisinin kamuoyu dikkatinin değişken yapısından etkilendiğini ifade etmekte ve özellikle kriz dönemlerinde bu ilginin piyasalara aşırı bir şekilde yöneldiği gerçeğini de gözler önüne sermektedir. (Shiller, 1992). Satın alınan hisse senetlerinin seçim süreçlerine ışık tutan bir başka çalışmada (Barber ve Odean, 2008) da benzer şekilde yatırımcıların hisse senedi tercihlerinin özellikle son dönemde haberlere konu olmasıyla, yüksek işlem hacmi sergilemesiyle ya da günlük getirilerinin dikkat çekiciliğiyle yakından ilgili olduğu bulgusuna ulaşırken başka çalışmalar da bu bulguyu desteklemektedir (Andrei ve Hasler, 2015; Gadarowski, 2002).

Hatırlama kolaylığının yatırımcı davranışlarına etkisiyle ilgili olarak yapılan bir çalışmada (Kliger ve Kudryavtsev, 2010) ise yatırımcıların, kayıp ya da kazancı değerlendirirken genel piyasa eğilimini esas aldıkları varsayımı ile uzman tavsiyelerindeki güncellemelere tepkileri incelenmiştir. Buna göre pozitif borsa endeksiyle birlikte uzmanların tavsiyelerindeki olumlu güncellemelere ve negatif borsa endeksiyle birlikte uzmanların tavsiyelerindeki olumsuz güncellemelere tepkiler daha güçlü olmaktadır. Hazırda bulunan sonucun etki gücü firmanın piyasa kapitilizasyonu ile negatif ilişkili iken, hisse senedinin betası ve tarihsel getiri oynaklığı ile pozitif ilişkili olduğu bulgusuna ulaşılmıştır. Risk algısının yüksek olduğu, diğer bir deyişle hazırda bulunan şeyin yüksek risk olduğu bir durumda ise hisse senedi fiyat

tepkileri, yukarı yönlü güncellemelere karşı daha zayıfken, aşağı yönlü güncellemelere karşı daha güçlü olmaktadır.

1.4.6 Aşinalık Eğilimi

Aşinalık eğilimi (Tversky ve Kahneman, 1974), genel olarak insanların seçim yaparken daha önceden doğrudan ya da dolaylı bir biçimde tanışıklığı olması ya da bir yönüyle yakınlık bağı kurması nedeniyle, alternatifler arasında yanlı tutum sergileme temayülü olarak tanımlanabilir (Demir, 2013). Başka bir ifade ile aşinalık eğilimi insanların, birini diğerinden daha iyi bildikleri iki seçenekle karşı karşıya kaldıklarında, daha iyi bilinen seçeneği tercih etme eğilimleridir. Bu bağlamda aşinalık genellikle tanıdık olana, olduğundan daha fazla değer vermek ya da aksine tanıdık olmayana olduğundan daha az değer vermek suretiyle iki farklı şekilde rasyonalite anlayışına ters düşmektedir (Demir, 2013).

Aşinalık hissi, tekrarlarla da ilgilidir. Zira tekrarlar bilişsel bir rahatlık ve güven verici bir tanıdıklık duygusu yaratır. Buna salt maruz kalma etkisi denir. Zajonc'a (1968) göre salt maruz kalma etkisinin meydana gelme nedeni, bir uyarana tekrar tekrar maruz kalmanın ardından kötü bir şey olmaması ve böylece ilgili uyarının güvenli olarak işaretlenmesidir.

Aşinalık eğilimi aynı zamanda, daha fazla aşına olunan olayların ya da kişilerin daha kolay hatırlanması nedeniyle (Baddeley, 2012) hafızaya dayalı hazırda bulunanı tercih etme eğilimlerinden birisidir (Wärneryd, 2008).

Yatırımcılar açısından bu eğilim, yatırım alternatifleri arasından bazılarının bir yönüyle ya da bir özelliğiyle yakınlık kurarak, yakınlık kurduğu bu alternatiflerin cazibesine kapılmak suretiyle tercih yapmalarına yol açmaktadır (Sulphey, 2014).

Sonuç olarak aşına olunanı tercih etme eğilimi, bir taraftan yatırımcıların aşına oldukları hisse senetlerine aşırı güvenmeleri (Nofsinger, 2005), beklenen getirilerini ve risklerini değerlendirirken aşırı iyimser olmaları (Li, 2004) ya da aşına olmadıkları hisse senetleriyle

ilgili olarak aşırı karamsar olmaları (Glassman ve Riddick, 2001) gibi nedenlerle salt bir portföy çeşitlendirmesi yapmalarına engel olan *irrasyonelliğin(!)* kaynağı olurken; diğer taraftan hem bilgi edinme maliyetini düşürerek belirsizliği azaltıcı bir rol oynadığı, hem de bilinmeyenin getirebileceği asimetrik bilgi riskini azalttığı için yatırımcının kendi içerisindeki *rasyonelliğinin* bir tezahürü olmaktadır (Demir, 2013).

1.4.7 Belirsizlikten Kaçınma

Finans terminolojisinde *risk*, karar birimlerinin verecekleri kararlar sonucunda elde etmeyi bekledikleri getiriye etkileyebilecek olayların gerçekleşme olasılığı, diğer bir deyişle olayların gerçekleşme olasılığının bilindiği durumu ifade ederken; *belirsizlik* ise olasılıkların bilinmediği bir durumu anlatmaktadır (Knight, 1921/2016). Bu bağlamda belirsizlikten kaçınma, risk ya da belirsizlik içeren kararlarda, muhtemel sonuçları ile ilgili olasılıkların bilindiği, yani riskli seçenekleri, olasılıkların bilinmediği seçeneklere, yani belirsizliğe tercih etme eğilimidir (Colman, 2006).

1.4.8 Pişmanlıktan Kaçınma

Karar alıcılar genellikle seçenekleri ve tercihlerinin sonuçlarını değerlendirirken belirli bir referans noktasını esas alırlar. Bu referans noktası ise ekseriyetle hâlihazırdaki durumdur. Ancak bazı durumlarda insanlar, kararlarının isabetli olup olmadığını değerlendirmek amacıyla, bu kararın sonucunda alacağı muhtemel neticeyle, farklı bir seçim yapılmış olsa alabileceği muhtemel neticeyi ve bunlara bağlı olarak yaşayacağı muhtemel pişmanlığı²³ mukayese etme eğilimi gösterirler (Plous, 1993). Bu bağlamda pişmanlıktan kaçınma, alınan kararların sonucunda yaşanacak muhtemel bir pişmanlığın önsel tahminin kararlar üzerindeki etkisini anlatmakta (Dowling ve Lucey, 2010) ve iki temel varsayıma dayanmaktadır. Bunlardan birincisi birçok insanın pişmanlık ve mutluluk denilen duyguları yaşamış olduğu varsayımı; ikincisi ise insanların belirsizlik koşullarında karar alırken o duyguları göz önünde

²³ *Pişmanlık*, bu mukayese sonucunda yapmış oldukları tercihin yanlış ya da en azından vasatın altında olduğunun anlaşılması sonucunda yaşanan memnuniyetsizlik hissidir.

bulundurarak önceden tahmin edip ona göre davranmaya gayretli oldukları varsayımıdır (Loomes ve Sugden, 1982).

Finansal piyasalar açısından bakıldığında ise pişmanlıktan kaçınma, yatırımcı davranışlarında paradoksal bir şekilde bazen riskten kaçınmaya neden olurken, bazen de risk almayı teşvik edebilmektedir. Birinci durumda pişmanlıktan kaçınma, insanları olumsuz bir netice ihtimalini azaltmak gayesiyle daha az risk almaya sevk etmektedir (Simonson, 1992). İkinci durumda ise pişmanlıktan kaçınma, bir eylemi yerine getirmenin ya da bazı durumlarda da getirmemenin yol açabileceği bir pişmanlık nedeniyle, insanları daha fazla risk almak için cesaretlendirebilmektedir (Zeelenberg, 1999).

1.4.9 Aşırı Tepki Hipotezi

Aşırı tepki hipotezi, yatırımcıların beklenmedik haberlere karşı aşırı tepki vererek hisse senedi fiyatlarında abartılı hareketlere ve ardından düzeltmelere yol açtığı varsayımına dayanmaktadır. Başka bir ifadeyle aşırı tepki hipotezi, daha önce ortalamanın üzerinde getiri sağlayan hisse senetlerinin gelecekte kötü performans sergilediği, daha önce kötü performans sergileyen hisse senetlerinin ise gelecekte iyi performans sergileyerek yatırımcının beklentilerini aştığı fikrine dayanmaktadır (De Bondt ve Thaler, 1985). Etkin piyasalar hipotezinin bu potansiyel ihlali, *aşırı tepki* olgusu olarak adlandırılmaktadır. De Bondt ve Thaler (1985) aşırı tepki hipotezini ortaya koydukları çalışmalarında hisse senedi fiyatlarının sistematik olarak aşırılaşması durumunda herhangi bir muhasebe verisi kullanılmadan, yalnızca geçmiş getiri verilerine dayanarak fiyatların tersine çevrilmesinin tahmin edilebileceğini savunmuşlardır. Bu çerçevede, hisse senedi fiyatlarındaki aşırı hareketleri ters yönde fiyat hareketlerinin izleyeceği ve ilk fiyat hareketi ne kadar aşırı olursa sonraki düzeltmenin de o kadar büyük olacağını ifade eden ve esasında zayıf form piyasa etkinliğinin ihlal edildiğini ima eden iki hipotez öne sürmüşlerdir. Aşırı tepki hipotezinin öngörü gücünü test ettikleri bu çalışmada kazanan ve kaybeden hisse senetlerinin 36 aylık dönemde tersine çevrilme eğilimi gösterdikleri bulgusuna ulaşmışlardır. Üstelik aşırı tepki etkisinin

kaybedenler için kazananlardan çok daha büyük, yani asimetrik olduğunu ortaya koymuşlardır.

Brown ve Harlow (1988) aşırı tepki hipotezini yön, büyüklük ve yoğunluk etkisi bağlamında incelemişlerdir. Araştırmada *yön etkisi*, hisse senedi fiyatlarındaki aşırı hareketleri ters yöndeki hareketlerin izleyeceğini; *büyüklik etkisi*, ilk fiyat değişikliği ne kadar aşırıysa, dengeleme tepkisinin o kadar aşırı olacağını; *yoğunluk etkisi* ise ilk fiyat değişikliğinin süresi ne kadar kısaysa, müteakip tepkinin o kadar aşırı olacağını ifade etmektedir. Brown ve Harlow (1988) araştırmanın sonucunda hem yönde hem de zamanda asimetritler olduğu bulgusuna ulaşmışlardır. Ayrıca piyasanın ekstrem finansal olaylara kısa vadeli olarak tepki verme biçiminde önemli bir asimetri olduğu, özellikle büyüklük ve yoğunluk hipotezlerinin, olumlu olaylardan ziyade, olumsuz olayları izleyen tepkiyi karakterize ettiği, sistematik bir şekilde aşırı tepki verme eğiliminin, olumsuz bir uyarı tarafından tetiklendiğinde çok daha güçlü ve daha öngörülebilir olduğu sonucuna varmışlardır. Brown ve Harlow (1988) bu durumun Kahneman ve Tversky'nin (1979) ortaya koyduğu kayıptan kaçınma, yani kayıpların yatırım kararlarını eşdeğer miktardaki kazançlardan daha fazla etkilediğine dair bulguları ile somutlaştırabileceğini ifade etmişlerdir. Benzer şekilde Bremer ve Sweeney (1991) %10 veya üzerinde günlük fiyat düşüşlerini takip eden hisse senedi getirilerini incelemişler ve düşüşlerin ardından yaklaşık iki gün süren bir fiyat düzeltmesi olduğunu bulmuşlardır. Böylesine yavaş bir toparlanmanın, piyasa fiyatlarının ilgili bilgileri tam ve hızlı bir şekilde yansıttığı fikriyle tutarsız olduğu, yani piyasaların etkin olmadığı sonucuna varmışlardır. Benou ve Richie (2003) ise aşırı tepki hipotezini belirli bir ayda hisse senedi fiyatlarında %20'den fazla önemli düşüşler yaşayan büyük ABD şirketlerinden oluşan bir örneklem için incelemişler ve aşırı tepki hipoteziyle büyük ölçüde tutarlı sonuçlar elde etmişlerdir.

De Bondt ve Thaler (1985) aşırı tepki eğiliminin genellikle yatırımcıların, yeni bilginin kalitesine veya uygunluğuna bakılmaksızın son bilgilere gereğinden fazla ağırlık vermesi sonucu ortaya çıktığını düşünmekte ve Tversky ve Kahneman'ın (1974) ortaya koydukları bilişsel kestirmelerden biri olan örnekleme eğilimi ile ilişkilendirmektedirler. Daniel,

Hirshleifer ve Subrahmanyam (1998) ise aşırı tepki eğiliminin aşırı güven ve kendine atfetme davranışlarının bir sonucu olarak ortaya çıktığını düşünmektedirler.

Bu bulgulardan farklı olarak Zarowin (1990) kaybedenlerin kazananlardan daha iyi performans gösterme eğiliminin, yatırımcıların aşırı tepkisinden değil, kaybedenlerin kazananlardan daha küçük ölçekli firmalar olma eğiliminden kaynaklandığını savunmaktadır. Kaybedenlerin eşit büyüklükteki kazananlarla karşılaştırıldığında, getiri tutarsızlığının ortadan kalktığını ve kazananların kaybedenlerden daha küçük olduğu dönemlerde, kazananların kaybedenlerden daha iyi performans gösterdiğini ortaya koymuştur. Conrad ve Kaul (1993) ise kümülatif anormal getiri yerine dönemsel getiri yöntemini kullanarak yeniden dengeleme yapmadan elde tutma sürelerini üç yıla çıkardıklarında, De Bondt ve Thaler'ın (1985) sonuçlarının Ocak ayı dışında artık geçerli olmadığını, Ocak ayı sonuçlarının da aşırı tepkiye atfedilemeyeceğini savunmuşlardır.

1.4.10 Sosyal Etkileşim

Var olduğundan beri bir topluluğun içerisine yerleşik olan insan, toplumun diğer üyeleriyle birlikte yaşamakta, onları etkilemekte ve onlardan etkilenmektedir. Bu etki bazen çok belirgin bir şekilde ortaya çıkarken, bazen de bireye hissettirmeden arka planda kararlarına şekil verebilmektedir. Bu bağlamda en bağımsız karar alıcılar dahi sosyal faktörlerin tesirinden yalıtılmış değildir (Plous, 1993). Dolayısıyla insanların nasıl karar verdiklerini doğru bir şekilde kavrayabilmek, insanların birbirlerini nasıl etkilediklerini anlamaya da bağlıdır. Literatürde *sosyal etki* kavramına tekabül eden bu durum, kişilerin inançlarının, tutumlarının ya da davranışlarının, genellikle dolaylı ve kasıtsız bir biçimde bir başkasından ya da başkalarından etkilenmesi olarak ifade edilmektedir (Kağıtçıbaşı ve Cemalcılar, 2014). Üstelik sosyal etkinin söz konusu olabilmesi için insanların bir arada bulunması da şart değildir. Diğer bir deyişle sosyal etki bazen insanlar bir aradayken tezahür ederken, bazen (sosyal normlara itaat etmek suretiyle) onların yokluğunda da tesirini göstermeye devam etmektedir.

Sosyal etkinin birey üzerinde gösterdiği tesire göre tepkiler de farklı şekillerde ortaya çıkmaktadır. Bazı durumlarda doğrudan bir iletişim söz konusu olmadığı halde, birlikte bulunmanın sonuçları, yalnız olduğu durumlara göre farklılık arz etmektedir. Bazı durumlarda sosyal etki, bireyi, inanç, tutum ve davranışlarını değiştirmeye sevk ederken, bazen bu etki içselleştirilmemekte, yani sadece davranışlarla sınırlı kalmaktadır (Nolen-Hoeksema, 2009). Kararların birlikte alındığı en demokratik ortamlar bile bazen aşırılığa, bazen taklide yol açabilmektedir. Son olarak ifade etmek gerekirse esasen bu etkileri ve sonuçlarını bu şekilde sınıflandırmış olmak, bunların birbirinden tamamen farklı ve aralarındaki sınırların belirgin olduğu anlamına gelmemekte, sadece bir yönüne dikkat çekmek maksadıyla böyle bir yöntem izlenmektedir. Bu kısımda öncelikle kuramsal izahı yapılmakta, akabinde de bu etkinin yatırımcı davranışlarında nasıl ortaya çıktığı ve finansal piyasalara nasıl yansıdığı üzerinde durulmaktadır.

1.4.10.1 Sürü Davranışı

Sürü davranışı merkezi bir koordinasyon söz konusu olmadan, yerel etkileşimlerin neticesinde gruptaki bireylerin davranışlarının ve düşüncelerinin peş peşe birbirine uyum göstermesiyle ortaya çıkan bir durumdur (Raafat ve diğerleri, 2009). Esasen sürü davranışı, bireyler ardışık olarak karar verdiğinde (Banerjee, 1992), grup kutuplaşmasındaki aşırılığın, patika bağımlılığı ile bir yola sevk olunduktan sonra zirveye çıktığı duruma tekabül etmektedir. Bu durum bazen bir moda, bir heves, bir coşku ya da kitle histerisi şeklinde; bazen de bir korku, bir panik şeklinde tezahür etmektedir.

Finansal piyasalarda sürü davranışı, yatırımcıların varlık fiyatlarının nasıl hareket edeceği konusunda şüpheye düştükleri ve kendi bilgisine güvenmekten ziyade kitlenin ya da liderinin bilgeliğine güvenmek suretiyle karar verdikleri durumlarda (Scharfstein ve Stein, 1990) ve genel olarak iki şekilde ortaya çıkmaktadır. Bunlardan birincisi piyasaların yükseliş trendinde olduğu, coşkunluk ve çılgınlıkların zirveye doğru tırmandığı zamanlarda gerçekleşmektedir (Shiller, 2006). Bunun en klasik örneği “lale çılgınlığı”dır (Mackay, 2012). Bu çılgınlığın akıbeti de esasen sürü psikolojisinin finansal piyasalardaki diğer

yönüne örnek teşkil etmektedir. Başka bir deyişle sürü davranışının ortaya çıkmasıyla ilgili ikinci durum, afaki bir şekilde yükselen fiyatların, bu kez panik ve korku ile oldukça sert bir şekilde düşmesiyle ilgilidir (Kindleberger, 2011). Bu bağlamda sürü davranışının fiyatlara yansımaları da, artışın ya da düşüşün artan oranda artması [price cascades] ve varlık fiyatlarında balonlar oluşması şeklinde olduğu söylenebilir (Altman, 2014).

Finansal piyasalarda sürü davranışı, yatırımcıların sadece birbirlerini doğrudan taklit etmeleri ile ilgili değildir. Zira yatırımcılar, bir finansal varlığın temel değerinin ne olabileceğinin yanı sıra, diğer yatırımcıların gelecekteki fiyatların ne olacağı hakkındaki inançlarını sezerek beklentilerinin gelecekteki fiyatlara etkisini de tahmin etmeye çalışırlar. Keynes'in (1935) *güzellik yarışması* örneğiyle izah ettiği bu durum, diğer yatırımcıların beklentilerini tahmin etmenin gelecekteki fiyatların yönelimlerini tahmin etme konusunda yatırımcılara yardımcı olabileceğini ifade etmektedir. Bu bağlamda, bireysel yatırımcılar beklentilerini genellikle söylentiler ya da uzman görüşleri çerçevesinde inşa ettiklerinden, bu bilgi akışı, diğer insanların nasıl davranacağı hususunda bir fikir verebilmektedir. Daha çok aktif yatırım stratejisini benimseyenlerle ilgili olan bu durum, bu stratejilerde finansal varlığın temel değerinin hemen hemen hiç göz önünde bulundurulmadığını ima etmektedir (Altman, 2014).

Sonuç olarak bireysel yatırımcı açısından sürü davranışı, özünde bireylerin birbirlerinin duygularından, düşüncelerinden ve eylemlerinden, kelimeler, eyleme dönük gözlemler (arz talep miktarları gibi) ve eylemlerin sonuçlarına dönük gözlemlerden (piyasa fiyatları gibi) etkilenmesiyle ortaya çıkan bir davranışsal yakınsama durumudur (Hirshleifer ve Hong Teoh, 2003).

1.4.10.2 Sosyal Medya ve Duygusal Durum

Duygusal durum ya da hâletiruhiye [mood]; geçici olmakla birlikte, duyguya nispetle belirli bir süre devamlılığı olan (Colman, 2006) ve belirginleşmesi halinde kişinin davranışlarını ve dünyayı algılayış biçimini önemli ölçüde etkileyen ruhsal bir durumu ifade etmektedir (Termbank, n.d.). Duygusal durum bu yönüyle kişinin kendi ruh hali ile ilgiliyken diğer

yönüyle kişiden kişiye bulaşarak toplumsal bir boyut kazanmaktadır. Bu anlamda *duygusal bulaşma* [emotional contagion]; bireyin etkileşim içerisinde olduğu kişinin duruşunu, yüz ifadelerini, hareketlerini ve işitsel iletişime dönük seslerini, ekseriyetle gayriiradî bir şekilde, eş zamanlı olarak taklit etmesi sonucunda, taklit ettiği kişinin duygularını kendi iç dünyasında yaşaması eğilimidir (Hatfield ve diğerleri, 1993). Daha kapsamlı bir tanıma göre ise duygusal bulaşma, bireyin ya da bir grup insanın duygusal durumlarının ve davranışsal tutumlarının bilinçli ya da bilinçsiz bir şekilde tetiklemesiyle, bundan diğer bir kişi ya da grubun duygularının veya davranışlarının etkilendiği ve sonunda aynileştiği bir süreç olarak tanımlanmaktadır (Schoenewolf, 1990). Esasında bu durum insanın toplumsallığının da temelini oluşturmaktadır.

Duygusal halin yüz yüze iletişimde kişiden kişiye nasıl geçtiğiyle ilgili ayna nöronlar, yüzle geri bildirim teorisi gibi (Eagleman, 2016; Christakis ve Fowler, 2012) farklı açıklamalar söz konusudur. Ancak kökleri Adam Smith'in (1759/2018) sempati kuramına kadar uzanan duygusal bulaşmanın, yüz yüze iletişimin ötesinde Facebook, Twitter gibi iletişim araçları vasıtasıyla söz konusu olup olmayacağı da aydınlatılması gereken önemli bir konudur. Nitekim bu amaçla yapılan çalışmalar duygusal bulaşmanın söz konusu olabilmesi için, yüz yüze iletişimin şart olmadığını, bireylerin birbirinin durumundan haberdar olmasının yeterli olacağını vurgulamaktadır (Leenders, 1997). Örneğin Facebook kullanıcılarıyla yapılan bir çalışmada bunu destekler nitelikte bulgulara ulaşılırken (Coviello ve diğerleri, 2014), aynı platform kullanıcılarıyla yapılan bir başka çalışma duygusal bulaşmanın olabilmesi için, bireyler arasında doğrudan (şahsen) bir iletişimin de şart olmadığı yönündedir (Kramer, 2014). Benzer bulgular Twitter kullanıcıları için de geçerlidir (Ferrara ve Yang, 2015).

1.4.11 Gürültü Tüccarı Riski

Finansal piyasalar bağlamında ilk kez Black (1986) tarafından kullanılan *gürültü*²⁴ [noise] kavramı, *bilgi* (enformasyon) ile zıt bir anlam taşımaktadır. Black'e (1986) göre gürültü,

²⁴ Gürültü [noise], sözlükte hem istenen bilgilerle birlikte ortaya çıkan ilgisiz veya anlamsız veriler veya çıktılar hem de söylenti anlamına gelmektedir. Söylenti [rumor] söcüğü ise bir anlamda ayırt edilebilir bir kaynak

portföyün beklenen getirisinin bilinmemesine neden olan ve gözlemleri mükemmellikten uzaklaştıran her şeydir. Makroekonomik çerçeveden bakıldığında ise gürültü, beklenen faydanın aksiyomlarını ihlal eden kararların temel nedendir. Aynı zamanda piyasaları etkileyen belirsizlik ile zevk ve teknolojideki beklenmeyen kaymalar da gürültü kavramının kapsamına girmektedir.

Black (1986), gürültünün her ne kadar piyasaları mükemmel olmaktan uzaklaştırsa da piyasaların varlığı için elzem olduğunu ifade etmektedir. Zira bir piyasada gürültüye dayalı işlemler yoksa finansal varlıkların işlem hacmi ve dolayısıyla piyasa likiditesi de çok düşük olacaktır. Buna karşılık bir piyasada ne kadar çok gürültü tüccarı varsa piyasa o kadar likit olacak ve çok sık işlem yapılması hisse senetlerinin fiyat hareketlerini gözlemlenebilir kılacaktır. Bu nedenle bir hisse senedinin fiyatı da, değerinin gürültü içeren bir tahmini olacaktır. Ancak gürültüye dayalı işlemler fiyatların da gürültü içermesine neden olacaktır. Dolayısıyla fiyatlar hem bilgiyi hem de gürültüyü yansıtacaktır. İşlem hacmi arttıkça fiyatların içerdiği gürültü artacağından piyasada bilgiye dayalı işlem yapmak daha kârlı hale gelecektir. Hisse senetlerinin fiyatı gerçek değerinden uzaklaştıkça, bilgiye dayalı yatırımcılar daha iştahlı hale gelecektir. Bu nedenle hisse senetlerinin fiyatı zaman içerisinde gerçek değerine dönme eğilimi gösterecektir. Üstelik hisse senedinin fiyatı gerçek değerinden ne kadar uzaklaşmışsa, gerçek değerine geri dönme eğilimi de o kadar hızlı olacaktır. Ancak bütün değer tahminleri gürültü içerdiği için hisse senedinin gerçek değerinden ne kadar uzaklaştığı hiçbir zaman tam olarak bilinemez.

Gürültü tüccarı, eksik veya yanlış verilere dayanarak, genellikle irrasyonel işlem yapan kişidir. Başka bir ifadeyle gürültü tüccarı, hisse senedi alım-satım kararlarında profesyonel tavsiyelere veya bir işletmenin temel değerine dayalı işlemler yerine rastgele seçimlerden

olmadan geniş çapta yayılan konuşma veya fikir, diğer anlamda ise doğruluğu bilinen bir otorite olmadan geçerli olan bir beyan veya rapordur. Bu tanımlar göz önüne alınarak bu çalışma kapsamında gürültü ve söylenti sözcükleri eş anlamlı olarak kullanılmamıştır. Daha açık bir ifadeyle gürültü, söylenti dışında farklı şekillerde de ortaya çıkabilir. Bu nedenle daha kapsayıcı olduğu düşünülmektedir. Söylenti ise daha çok Shiller'in *anlatı* sözcüğüne yakındır. Nitekim Shiller söylentinin, bulaşıcı anlatılar için kullanılan Latince bir kelime olduğunu ifade etmektedir. Dolayısıyla bu çalışma kapsamında duygu, fikir, beklenti, tecrübe ve bilgelik yüklü daha uzun ömürlü hikayeler *anlatı*, duygu, fikir ve beklenti yüklü doğruğu henüz kesinleşmemiş olan daha kısa ömürlü hikayeler ise *söylenti* sözcüğüne karşılık gelmek üzere kullanılmıştır.

daha iyi getiri sağlamayan faktörlere dayalı olarak işlem yapan yatırımcılara denilmektedir. Gürültü tüccarları duygu odaklı, fevri ve tepkisel kararlar verirler. Bu tür yatırımcılar yatırım kararlarında trendlere göre hareket eden, piyasa ile ilgili haberlere aşırı tepki gösteren ve sürü psikolojisi sergileyen yatırımcılardır (Downey, 2022). Bu tüccarlar, fiyat oynaklığı yaratabilir ve fiyatları etkileyebilecek düzeyde irrasyonel kararlar verebilirler. Bu nedenle finansal piyasalarda gürültü tüccarlarının varlığı, diğer tüm tüccarlar rasyonel olsa bile fiyatların ve risk seviyelerinin beklenen seviyelerden sapmasına neden olabilmektedir. Bu durum gürültü tüccarı riski olarak ifade edilmektedir (De Long ve diğerleri, 1990).

Gürültü tüccarları genellikle yüksek hacimli işlem günleriyle bağlantılıdır. Genellikle yükseliş dönemlerinde menkul kıymetlerin fiyatını aşırı şişirirler ve düşüş dönemlerinde aşırı derecede söndürürler. Ayrıca, medyada giderek daha fazla içerik üretilmesi, hisse senetleri ve piyasalar için *manşet riskini* artırmaktadır (Downey, 2022). Manşet riski [headline risk], bir haber başlığının veya hikâyenin bir hisse senedi, sektör veya daha geniş bir piyasanın fiyatını etkileme riskidir. Buna göre hikâye yanlış veya yanıltıcı olsa bile gazete, televizyon, çevrimiçi medya veya sosyal medya tarafından oluşturulan manşetler hisse senedi fiyatlarını etkileme potansiyeline sahiptir (Chen, 2022).

1.4.12 Anlatı İktisadı

Anlatı, bir durumu ya da olaylar dizisini, belirli bir bakış açısını veya değerler kümesini yansıtacak ve destekleyecek şekilde sunma ya da anlama olarak tanımlanmaktadır (Merriam-Webster, n.d.). Bu çerçevede bir anlatı, olgulardan, duygulardan, insani ilgilerden ve insan zihninde bir izlenim oluşturan diğer dışsal ayrıntılardan oluşan insani temsil, hikâye ve benzeri kurgulardır.

Shiller'ın (2019/2021) ortaya koyduğu anlatı iktisadı [narrative economics] teorisi ise, insan hayatı ve toplumu hakkında anlatılan hikâyelerin kişiden kişiye nasıl yayılabileceğini, olaylara ilişkin ortak algıları nasıl değiştirebileceğini ve nihayetinde ekonomik davranışları nasıl şekillendirebileceğini incelemektedir. Bu bağlamda anlatı iktisadı teorisinin temel

önermesi iktisadi dalgalanmaların, büyük ölçüde iktisadi anlatıların aşırı basitleştirilmiş ve kolayca iletilebilen biçimlerinin bulaşmasıyla tetiklendiği yönündedir. Dolayısıyla bir iktisadi anlatı, insanların iktisadi kararlarını etkileme potansiyeline sahip bulaşıcı bir öyküdür. Üstelik geleneksel finans yaklaşımlarında yok sayılan ya da irrasyonel olarak etiketlenerek analizlerin dışında tutulan duygu, anlatıların yapısı içinde oldukça önemli bir yere sahiptir. Başka bir ifadeyle anlatılar genellikle duygu yüklü bulaşıcı hikâyeler şeklinde kurgulanırlar. Bu suretle duygu yüklü anlatılar insanların duygularına hitap ederek karar süreçlerine müdâhil olurlar.

Shiller'ın (2019/2021) anlatı iktisadi teorisinin merkezinde bulaşıcılık fikri yer almaktadır. Bu bağlamda anlatı iktisadi iki temel unsura odaklanmaktadır. Bunlardan birincisi fikirlerin hikâyeler formunda ağızdan ağıza bulaşması, ikincisi insanların yeni bulaşıcı hikâyeler üretme ya da var olan hikâyeleri daha bulaşıcı kılmaya yönelik çabalarıdır. Anlatı iktisadi, anlatıların bulaşmasının iktisadi olayları nasıl etkilediğini araştırmayı amaçlamaktadır. Zira iktisadi anlatılar öngörülebilir şekillerde işler; bulaşıcıdır, insanların takip etmesi için senaryolar sunarlar, mesajlarını tekrarlar ve insanların ilgi göstermesiyle başarıya ulaşırlar. Bu şekilde toplumu ve iktisadi faaliyetlerin gidişatını son derece güçlü etkileri olacak şekilde yönlendirirler. Dolayısıyla anlatıların ekonomi üzerindeki etkisini anlamak ve böylece iktisadi olayları tahmin etmek mümkündür. Hatta Shiller'a göre salgınların doğasını ve bulaşma faktörleriyle olan ilişkisini anlamak, öngörülerde bulunma açısından salt istatistiksel yöntemler kullanmaktan daha faydalı olabileceği yönündedir. Bu bağlamda Shiller epidemiyoloji alanındaki birikimin önemli katkılarının olacağını düşünmektedir. Ayrıca analizlerde nedensellik ilişkisinin çok iyi bir şekilde ortaya konulması gerektiğine dikkat çekerek anlatılarda iki yönlü bir nedensellik olduğuna işaret etmektedir. Yani yeni bulaşıcı anlatılar iktisadi olaylara neden olurken ve iktisadi olaylar da anlatıların değişmesine neden olmaktadır.

Shiller'ın (2019/2021) dikkat çektiği bir diğer husus ise internetin ve sosyal medyanın yükselişinin, bulaşmanın doğasını temelden değiştirmiş olmasıdır. Zira sosyal medya ve arama motorları, bulaşma sürecini temelden değiştirme potansiyeline sahiptir. Geçmişte

fikirler rastgele, sistematik olmayan bir şekilde yayılıyordu. Sosyal medya platformları benzer şekilde düşünen insanların birbirlerini bulmasını ve inançlarını daha da güçlendirmesini mümkün kılmaktadır. Bu bulaşma bilgi denetçileri tarafından yavaşlatılamamakta, aksine internet ve sosyal medya, fikirlerin varsa bile görünür olmayan bir merkezi kontrolle yayılmasına olanak sağlamaktadır. Ayrıca hikâyelerin ve fikirlerin bulaşma oranlarını artıran bir diğer unsur ise sahip oldukları medya kanalları aracılığıyla belirli bir kitleyi etkileyebilen influencerlardır.

Shiller (2019/2021) psikoloji, sinir bilim ve yapay zekâ sahasındaki gelişmelerin, anlatı iktisadı alanında bir yapı duygusu geliştireceğini savunmaktadır. Özellikle yapay zekâ, örneğin yapılandırılmamış veriler sözkonusu olduğunda araştırmalara önemli katkılar sunma potansiyeline sahiptir. Zira araştırma yöntemleri ilerledikçe ve sosyal medyaya ait daha fazla veri biriktikçe metin analizinin iktisat biliminde daha baskın bir güç haline gelme ihtimali yüksektir. Ekonominin bilinçli, yaşayan insanlardan oluştuğu ve bu insanların kendi eylemlerini duygu ve fikir yüklü hikâyeler ışığında icra ettiği düşünüldüğünde farklı bakış açılarının önemi ortaya çıkmaktadır. Ayrıca bu gelişmeler kasıtlı manipölasyonları ve aldatmacaları daha iyi anlamaya ve anlatıları dikkate alan politikalar oluşturmaya yardımcı olacaktır.

Shiller anlatı iktisadı hipotezinde ampirik kanıtlara çok az atıfta bulunduğu için eleştirilmektedir. Başka bir ifadeyle Shiller'in tezinin test edilebilir hipotezlerle ortaya konulması gerektiğine ilişkin bir eleştiri bulunmaktadır (Portes, 2020).

1.4.13 Geleneksel Medya ve Yatırımcı Eğilimleri

Kitle iletişimi, geniş bir izleyici kitlesi ile bilgi paylaşma sürecini ifade etmektedir. Kitle iletişim yoluyla bilgi, radyo, televizyon, reklam panoları, gazeteler, dergiler, kitaplar, filmler ve internet gibi çeşitli kanallar aracılığıyla birçok kişiye hızlı bir şekilde iletelebilmektedir (Volle, 2023). Kitle iletişiminde mesajlar çoğunlukla profesyonelce kurgulanmakta ve dağınık bir kitleye iletilmektedir. Bu tarz iletişimde karşılıklılık çok sınırlıdır. Başka bir

ifadeyle iletişim kaynaktan alıcıya doğru tek yönlüdür. Mesajların muhatapları çoğu zaman mesajın kaynağına ulaşamazlar (Kıyılar ve Akkaya, 2020).

Medya, toplum ve bireyler üzerinde oldukça etkin bir role, yani insanların bilgi, duygu, tutum ve davranışlarını etkileme gücüne sahiptir. Medyanın bu etkisi yatırımcı davranışları ve hisse senedi fiyatları için de geçerlidir. Örneğin Peress (2014) getiri öngörülebilirliği analizi yaparak, gazetelerin bir önceki günden haberler yaydığını göstermiştir. Ayrıca çeşitli ülkelerdeki ulusal gazete grevlerini inceleyerek işlem ve fiyat oluşumu üzerindeki nedensel etkilerini araştırmış, işlem hacminin grev günlerinde %12 düştüğünü, hisse senedi getirilerinin dağılımının ve gün içi oynaklığının %7 azaldığını bulmuştur. Benzer şekilde Griffin ve diğerleri (2011), kazançların medyada yer almasının, hisse senedi fiyatlarının oynaklığındaki önemli değişikliklerle ilişkili olduğunu bulmuşlardır. Tetlock (2007), haberlerdeki olumsuz duyarlılığın getiriler üzerinde kısa vadeli baskıya ve artan işlem hacmine yol açtığını bulmuştur. Dougal ve diğerleri (2012) bu sonucu doğrulamakla birlikte farklı köşe yazarlarının piyasayı farklı şekillerde etkilediğinin altını çizmişlerdir. Benzer etkiler başlıca kitle iletişim araçlarından biri olan televizyon için de geçerlidir (Aman ve diğerleri, 2018). Televizyon aracılığıyla artan bilgi akışının, yatırımcıların dikkatini çekme etkisiyle tutarlı olarak daha fazla işlem hacmi ve daha büyük fiyat değişimi ile önemli ölçüde ilişkili olduğu görülmüştür.

1.4.14 İnternet ve Yatırımcı Eğilimleri

İnternet, dünyadaki çeşitli bilgisayar ağlarının bilgi taşıma protokolü paketleri (TCP/IP) kullanılarak birbirine bağlanmasına izin veren ve böylece iletişim ve ticaret yöntemlerinde devrim yaratan bir sistem mimarisidir. Ağlar ağı [network of networks] olarak da adlandırılan internet, ABD’de 1970’lerde ortaya çıkmış, ancak 1990’ların başında halk tarafından kullanılabilir hale gelmiştir. İnternet, bilgiye dayalı hemen hemen her amaç için kullanılabilecek kadar güçlü ve genel bir yetenek sağlamakta ve onu oluşturan ağlardan birine bağlanan her bir birey tarafından erişilebilmektedir. Sosyal medya, elektronik posta, sohbet odaları, haber grupları, ses ve görüntü iletimi yoluyla insan iletişimini desteklemekte

ve insanların birçok farklı lokasyonda işbirliği içinde çalışmasına olanak tanımaktadır (Dennis ve Kahn, 2023).

İnternetin ortaya çıkışı, geleneksel kitle iletişiminin “birden çoğa” modelini, “çoktan çoğa” iletişim ağı ile değiştirmiştir. Bir diğer ifadeyle geleneksel iletişim araçlarının yerini alarak tüm dünyada başkalarıyla bağlantı kurma ve bilgi paylaşma biçimini dönüştüren internet günlük hayatın giderek daha önemli bir parçası haline gelmiştir. Dünyada toplam internet kullanıcılarının sayısı 2000 yılında 396 milyon kişi iken, bu sayı 2010 yılında 1,9 milyar kişiye, 2023 yılının başında 5,1 milyar kişiye ulaşmıştır (Digital 2023, 2023). Bu rakam dünya nüfusunun yaklaşık üçte ikisinin internet erişimine sahip olduğunu, yani küresel çapta ağa bağlı olduğunu göstermektedir.

İnternet pek çok işlemi kolay, hızlı ve basit hale getirerek günlük hayatın vazgeçilmezi haline gelmiştir. İnternetin günlük hayatta kullanımı daha çok bireysel gereksinimlerle ilgilidir. Bu gereksinimlerden başlıcası *bilgi bulmaktır*. Buna ilave olarak arkadaşlar ve aile ile iletişimde kalmak, haberler ve güncel olaylardan haberdar olmak, video izlemek de internetin diğer kullanım amaçları arasındandır. Ayrıca internetin finans ve tasarrufların yönetiminde kullanımı da oldukça yaygındır (Digital 2023, 2023).

Yatırımcılar açısından internet, işlem ve bilgi maliyetlerini düşürerek borsaya daha kolay erişim olanağı sağlamakta ve borsalara katılımı artırmaktadır (Bogan, 2008). İnternet aynı zamanda bilginin yatırımcılara iletilme şeklini ve yatırımcıların bu bilgilere göre hareket etme biçimlerini de değiştirmektedir. Bu bağlamda Barber ve Odean’ın (2001) öncü bir çalışmasında internetin yatırımcılar açısından hem faydaları hem de riskleri ortaya konulmuştur. Buna göre internetin daha düşük maliyetler ve daha fazla alternatif sunması yatırımcılar açısından faydalıdır. Ancak çoğu yatırımcının piyasada yeni olması bir takım riskleri de beraberinde getirmektedir. Zira bulgular çevrim içi yatırımcıların düşük işlem maliyetleri, iyileştirilmiş uygulama hızı ve daha fazla erişim kolaylığı nedeniyle daha aktif, daha spekülatif ve öncekinden daha az kârlı işlem yaptıklarını ortaya koymaktadır. Barber ve Odean’a (2002) göre yatırımcıların doğrudan alım satım yapmalarının, işlemlerinin sonucu

üzerinde abartılı bir kontrol duygusu vererek kontrol yanılması ihtimalini güçlendirmektedir. Ayrıca hâlihazırdaki pek çok çevrimiçi yatırım verisi, yatırımcıların önceki inançlarını doğrulamalarını sağlayarak kendilerine aşırı güvenmelerine yol açabilmektedir. Bir diğer ifadeyle hem kendine atıf önyargısı hem de bilgi ve kontrol yanılışlarıyla artan aşırı güven, çevrimiçi yatırımcıların işlemlerinde artışa ve performansında düşüşe yol açmaktadır. Üstelik değerlemelerin belirsiz olduğu, yatırımcıların aktif ve deneyimsiz olduğu ve yatırım için hazır paranın olduğu piyasalar, tüm yatırımcılara zarar verebilecek spekülasyonlara eğilimli olduğu bulgusuna ulaşmışlardır (Barber ve Odean, 2001).

1.5 DAVRANIŞSAL FİNANSTA BİLİŞSEL KESTİRMELER

Davranışsal finans yazınında bilişsel kestirmelerle [heuristics] ilgili iki farklı yaklaşım söz konusudur. Bunlardan bazıları bilişsel kestirmeleri sistematik bilişsel hatanın bir sonucu olarak görürken, bazıları ise bilişsel kestirmeleri yaşamın muhtelif alanlarında isabetli kararlar verebilmek için faydalı araçlar olarak görürler (Ackert, 2014). Bu yaklaşımlardan ilki Kahneman'ın bilişsel kestirmeler ve yanlılıklar, ikincisi Gigerenzer'in hızlı, yalın ve adaptif sezgiler görüşü etrafında toplanmaktadır.

1.5.1 Kahneman'ın Bilişsel Kestirmeleri ve Yanlılıkları

Bilişsel kestirmeleri sezgideki yanlılıklar olarak tanımlayan bu yaklaşım taraftarları, zihnin sistematik yanılığlara açık olduğunu, sistematik hataların, yani rasyonaliteden sapmaların duygulardan ziyade bilişsel mekanizmanın tasarımından kaynaklandığını düşünen Kahneman'ın (2011) görüşleri etrafında öbeklenirler.

Bu çerçevede Kahneman (2011) düşünme biçimlerini biri sezgisel diğeri bilinçli aklı, başka bir ifadeyle birincisi hızlı düşünmeyi, ikincisi yavaş düşünmeyi temsil etmek üzere, Sistem 1 ve Sistem 2'den oluşan ikili bir modele dayandırmaktadır. Bu modele göre Sistem 1, çok fazla çaba sarf edilmeden, hesaplamaların otomatik olarak ve ivedilikle yapıldığı, algoritmik

düşünmeden ziyade duygusal tepkilerin baskın olduğu bir durumu ortaya koyarken Sistem 2, ne hakkında düşünüldüğü ve ne yapılacağı ile ilgili olarak, bilinçli bir şekilde ve belirli inançlar çerçevesinde akıl yürütülen, tercihler yapılan ve karar verilen bir durumu ifade etmektedir (Kahneman ve Frederick, 2005; Slovic ve diğerleri, 2007). Esasen Sistem 1 ve Sistem 2 eşzamanlı olarak aktiftir. Sistem 1 karar sürecine ilişkin sorunlar ortaya çıktığında, süratle sezgisel çözüm önerileri getirirken, Sistem 2 bu önerilerin niteliğini gözlemleyip, ya desteklemekte, ya düzeltmekte, ya da geçersiz kılmaktadır. Başka bir ifadeyle gerek otomatik, gerekse kontrollü bilişsel faaliyetler, probleme ilişkin çözümleri kontrol edebilmek için birbirleri ile rekabet halindedirler. Nihai yargıların belirlenmesinde iki sistemden hangisinin etkin rol oynayacağı, üzerinde düşünmek için yeterince zaman olup olmadığı, bireyin ruh hali, anlama yeteneği ve istatistiksel düşüncenin etkisinde kalmasını da içeren, birey ve problemle ilgili özelliklere bağlıdır. Varılan nihai yargılar, şayet üzerinde çok fazla değişiklik yapılmadan, ortaya konulan ilk haliyle kalmışsa, bunlar *sezgisel* olarak adlandırılırlar.

Tablo 1.2: İki Bilişsel Sistem

Sistem 1 (Sezgiye Dayalı)	Sistem 2 (Düşünceye Dayalı)
Süreçe İlişkin Özellikler	
Otomatik	Kontrollü
Zahmetsiz	Gayret Gerektiren
Çağrışımsal	Çıkarımsal
Hızlı, Eş Anlı	Yavaş, Ardışık
Süreç Şeffaf Değil	Süreç Bilinçli
Maharete Dayalı	Kural Bazlı
Süreçlerin Kapsamı (Content on Which Processes Act)	
Duygusal	Nötr
Nedensel Eğilimler	İstatistikler
Somut, Belirli	Soyut
Prototipler	Kümeler

Kaynak: Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge university press.

Her ne kadar Sistem 1, Sistem 2'den daha ilkel olsa da, bu durum Sistem 1'in kabiliyetlerinin daha az olduğu anlamına gelmemektedir (Kahneman ve Frederick, 2002). Aksine belirli bir uzmanlık ve kabiliyet düzeyine ulaşıldığında, karmaşık bilişsel işlemler Sistem 2'den Sistem 1'e taşınırlar. Nitekim Sistem 1'le temsil edilen aklın işleyişine satranç ustalarının, satranç taşlarının tahtadaki konumunu görür görmez içerisinde bulunduğu durumun güçlü ve zayıf yönlerini algılama hususundaki maharetleri örnek verilebilir²⁵. Bu bağlamda Sistem 1 ve Sistem 2 arasındaki etkileşim durağan olmaktan ziyade, bireyin kendisinin ya da başkalarının tecrübelerinden yararlanarak veyahut belirli bir eğitimden geçmek suretiyle, karar vermede en üstün olan yöntemlerin geliştirildiği dinamik bir süreci ifade eder (Altman, 2014). Diğer bir deyişle bu durum, insanların karar verirken kullandığı *bilişsel kestirmelerin* zaman içerisinde tekâmülüyle birlikte birinden diğerine transfer olabileceğine işaret etmektedir (Hayek, 1945).

Kahneman (2011), Tablo 1.2'de de görüldüğü üzere, Sistem 1 ve Sistem 2'ye atfedilen özellikler ve etkinliklerin bir listesini sunar. Buna göre Sistem 1, Sistem 2 tarafından desteklendiğinde inançlara, tutumlara ve niyetlere dönüşen izlenimler, duygular ve eğilimler üretir. Hemen hemen hiçbir çaba gerektirmeden istemsizce ve hızlı işler. Yeterli bir eğitimden sonra ustaca sezgiler üretebilir. Bilhassa çağrışımsal bellekte etkinleştirilerek fikirlerin tutarlı bir modelini yaratır. Alışılmadık olanı alışlagelenden ustalıklı ayırt eder. İnanmaya ve doğrulamaya eğilimlidir. Hâlihazırdaki kanıtlara odaklanırken noksan delilleri önemsemez. Belirsizliği göz ardı ederken kuşkuyu baskılar. Düşük olasılıklara haddinden fazla ağırlık verir. Sınırlı bir temel değerlendirmeler kümesi üretir. Kümeleri normlar ve prototiplerle temsil eder, kaynaştırmaz. Bazı durumlarda zor bir soruyu daha kolay bir soruyla ikame eder. Kayıplara, kazançlardan daha güçlü tepki verir. Hale etkisi olarak adlandırılan duygusal tutarlılığı abartır. Değişikliklere karşı duyarlılığı durumlara karşı duyarlılığından daha yüksektir. Niceliğe azalan bir duyarlılık gösterir. Amaçlanandan

²⁵ Örneğin insanlar; boş bir yolda araç kullanırken, ses tonundan dostluğu ve düşmanlığı ayırt ederken ya da büyük bir reklam tabelasındaki sözcükleri okurken Sistem 1'i; aracı dar bir alanda park ederken (araba parkedicileri hariç), kalabalık ve gürültülü bir ortamda belirli bir insanın sesine odaklanırken veya bir formu doldururken Sistem 2'yi kullanmaktadırlar.

fazlasını hesaplar. Bilişsel rahatlık duygusuyla doğruluk yanılmasına kapılır. Karar alma sorunlarını birbirinden yalıtarak dar bir çerçeveye oturtur.

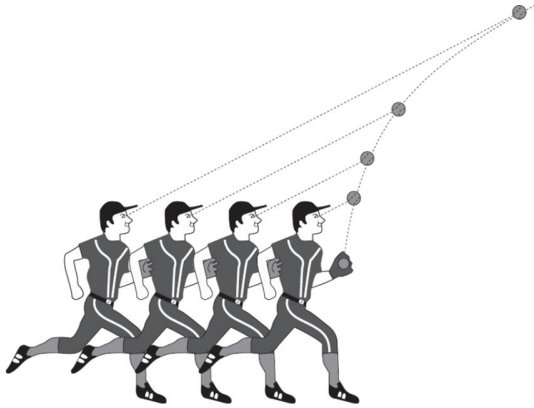
1.5.2 Gigerenzer'in Hızlı, Yalın ve Adaptif Sezgileri

İnsanlar karar verirken *mantık, istatistik* ve *bilişsel kestirmeler* olmak üzere üç farklı yöntemle başvurabilmektedirler (Gigerenzer ve Gaissmaier, 2011). Gigerenzer'e (2018) göre bilişsel kestirmeler ve yanlışlıklar [heuristics-and-biases] programının taraftarları, mantığın ve istatistiğin kurallarını rasyonel akıl yürütmeye, bilişsel kestirmeleri ise hataya yatkın sezgiler ve hatta irrasyoneliteyle ilişkilendirirler. Bu anlamda mantıksal veya istatistiksel ilkelerden sapmalar rutin olarak yargısal yanlışlıklar olarak yorumlanırken kusur bilişsel kestirmelere veya sezgisel özelliklerin merkezi olan “Sistem 1” e atfedilir. Başka bir deyişle bilişsel kestirmeler ve yanlışlıklar [heuristics-and-biases] programında rasyonel seçim teorisinden muhtelif sapmalar, teorisinin değil de insan aklının sistematik kusurları olarak yorumlanır. Sonuç olarak insanların sıklıkla bilişsel kestirmelere güvendikleri, ancak güvenmemelerinin daha iyi olacağı fikri yanlışlıklar programının genel kabulü olur. Zira sezgisel düşünmenin mucizevi yönü olduğu kadar kusurlarının olduğunu ve sistematik hatalar şeklinde sürekli tekrar ettiklerini düşünürler (Kahneman, 2011).

Gigerenzer tam bu noktada Simon'ın “Neoklasik iktisat modelinin öne sürdüğü rasyonellik koşulları karşılanmadığında insanlar nasıl akıl yürütür?” sorusunu gündeme taşır. Zira Simon'a göre klasik rasyonellik modeli ilgili tüm alternatiflerin, sonuçların ve olasılıkların bilindiği, sürprizlere yer olmayan tahmin edilebilir bir dünya hakkında bilgi verebilir. Ancak bu koşullar, bireylerin karşılaştıkları sorunlar için nadiren gerçekleşir. Dolayısıyla rasyonel seçim teorisinin varsayımlarının söz konusu olmadığı gerçek durumlarda insanların hızlı ve yalın [fast and frugal] sezgisel yöntemlere başvurduğunu ifade eder. Bu bağlamda Gigerenzer'e göre bilişsel kestirmeler, kararları daha kompleks yöntemlerden ziyade daha hızlı, yalın ve/veya doğru bir şekilde almak amacıyla bilgilerin bir bölümünü göz ardı eden bir stratejidir. Gigerenzer (2007), zekâyı, sadece mantık yasalarının rehberliğinde bilinçli bir faaliyet olarak görmenin bakış açısını daraltacağını ifade eder. Oysa zihinsel hayatın çoğu,

mantığa yabancı olan süreçlere dayanarak bilinçsiz bir şekilde içgüdüsel duygular [gut feelings] veya sezgilerle işler. Örneğin Resim 1.1, bir beyzbol oyuncusunun yüksekten gelen bir topu yakalamak için bilinçli hesaplamalar yapmak yerine başparmak kurallarına [rule-of-thumb] güvendiğini göstermektedir. Oyuncu topu tutabilmek için bakışlarını topa sabitleyerek topa doğru koşmaya başlamakta ve bakış sabit kalacak şekilde hızını ve elinin yüksekliğini bilinç dışı bir şekilde ayarlamaktadır. Benzer şekilde Resim 1.2’de ise, ayrıntılı matematiksel hesaplamalar yapmak yerine, bilinç dışının kurallarına güvenmek zorunda olan bir başka insan resmedilmektedir (Gigerenzer, 2014). Bu bağlamda Gigerenzer’e (2008) göre hızlı ve yalın bir sezgi, bilinçli veya bilinçsiz, bilişsel kapasitelerin ve çevresel yapının birlikte geliştiği, asgari bilgi ile kararlar verebilen bir stratejidir. Birlikte evrilen beyne ve dış çevreye bağlı olduğundan genelde etkili sonuçlara ulaşabilir. Zihnin ve çevrenin birlikte evrilmesi insanı *ekolojik rasyonel* yapar.

Resim 1.1: Uçan Topu Tutmak



Kaynak: Gigerenzer, G. (2008). Why heuristics work. *Perspectives on psychological science*, 3(1), 20-29.

Resim 1.2: Tehlikeden Kurtulmak



Kaynak: Gigerenzer, G. (2014). *Risk savvy: How to make good decisions*. Penguin Books.

Gigerenzer (2008) rasyonaliteyle ilgili olarak dört farklı yaklaşımdan bahseder. Bunlardan birincisi sınırsız rasyonalitedir. Sınırsız rasyonalite [unbounded rationality] bilişsel süreçleri tasvir etmekten ziyade, insanların her şeyi bilmeleri ve optimize etmek için gerekli tüm

zaman ve hesaplama gücüne sahip olmaları durumunda nasıl davranırlardı sorusuna cevap arar. Psikolojik gerçekliğin noksanlığı nedeniyle sınırsız rasyonaliteye dayanan teoriler *varsayalım ki* [as-if] teorileri olarak adlandırılır. Beklenen değerin ve beklenen faydanın maksimizasyonu şeklinde tecessüm eder. İkinci yaklaşım belirli kısıtlar altında optimizasyon [optimization under constraints] yaklaşımıdır. Sınırsız rasyonalitenin göz ardı ettiği, insana özgü zihinsel ve çevresel kısıtları dikkate alır. Kısıt altında optimizasyonda her şeyi bilmemenin eksikliği, bunun sonucu olarak da bilgi arama ihtiyacı kilit rol oynar. Ancak bilgi arama modellerinin²⁶ yokluğu sınırsız rasyonellik teorilerinin tanımlayıcı bir özelliğidir. Her yeni gerçekçi kısıtlama, optimizasyon hesaplamalarını zorlaştıracığından, gerçekçi olmayan varsayımlar talep ederek teoriyi daha gerçekçi hale getirme girişimini baltalayabilecek paradoksal bir yapıya sahiptir. Kısıtlamalar altında optimizasyon teorileri, sınırsız rasyonelliğin modellerinde olduğu gibi zihinsel süreci değil davranışı öngörme amacı ile *varsayalım ki* teorileri halinde sunulur. Bir diğer yaklaşım bilişsel yanlılıklar, yani mantıksal irrasyonel yaklaşımdır. Bilişsel kestirmeler ve yanlılıklar [heuristics and biases] veya bilişsel yanılsamalar [cognitive illusions] programı çerçevesini oluşturan bu yaklaşım, aslında insanların temel olarak rasyonel olduklarını düşünen teorilere karşı çıkar. Ancak geçerli veya geçersiz kararlar üreten bilişsel süreçleri rasyonaliteden sistematik sapmalar ve yargı hataları olarak takdim eder. Başka bir ifadeyle insanların yargılarının sistematik olarak mantık ya da optimizasyon yasalarından saptığı iddiası, aslında mantık ve olasılık yasalarının gerçek dünyadaki rasyonel davranış için gerekli olduğu yönünde örtük bir kabule dayanır. Sonuç olarak Simon, neoklasik ekonomiye bir alternatif geliştirmeyi hedeflerken, bilişsel kestirmeler ve yanlılıklar [heuristics-and-biases] programı tarafından şekillendirilen davranışsal finans çizgisi, kendisini neoklasik paradigmadan sapmalarla ya da bu paradigmanın ne olduğu ile sınırlar. Bu sınırların içerisinde kalan araştırmacıların amaçları da insan davranışlarında “anomaliler” ve “önyargılar” göstermekle sınırlandırılmış olur (Gigerenzer, 2018). Dördüncü rasyonel yaklaşıma ise ekolojik rasyoneliktir. Ekolojik rasyonelite [ecological rationality] insanların, doğanın sosyal ve fiziksel ortamında evrildiğini, hayatta kalabilmek için bu çevreye uyum sağlamak suretiyle değiştiğini,

²⁶ Arama modelleri, bir arama yönünü (bilgiyi nerede arayacağımızı) ve durma kuralını (aramayı ne zaman durduracağımızı) belirtir.

dolayısıyla insan doğasının yapısının mantıksal değil ekolojik olduğunu savunur. Başka bir deyişle sezgisel araştırmanın başlangıç noktasının zihin ve mantık kuralları değil, *zihin ve çevre arasındaki ilişki* olması gerektiğini düşünür.

Gigerenzer (1991) bilişsel kestirmeler ve yanlılıklar programının yargısal kusurlar olarak ortaya koyduğu pek çok bulgunun yöntemsel düzeltmelerle tekrar edildiğinde ortadan kalktığını ifade eder. Bu çalışma kapsamında davranışsal finans yazınının temel bulgularına ve muhalif görüşlere mümkün olduğunca değinilmeye çalışılmış olmakla birlikte, daha ayrıntılı bilgi için ilgili kaynaklara bakılabilir (Gigerenzer ve Goldstein, 1996; Gigerenzer, 2008; Gigerenzer, 1996; Kahneman ve Tversky, 1996).

Sonuç olarak davranışsal finansın her iki çizgisi de, geleneksel finans kuramının aksine psikolojiyi sürece dâhil ederek insanların iktisadi ve finansal kararlarını verirken nasıl bir mekanizma kullandıklarını soyut modeller yerine insan gerçekliğine dayanan modeller çerçevesinde analiz etmektedir.

İKİNCİ BÖLÜM

BÜYÜK VERİ, SOSYAL MEDYA VE FİNANS

2.1 DİJİTAL ÇAĞDA VERİLEŞTİRME VE SOSYAL BİLİMLER

İnsanlık tarihinde elde edilen en büyük başarılar, dünyayı ölçmeye dönük çabalarla yakından ilgilidir. Bu çabaların altında bir olgunun anlaşılabilmesinin, ölçülebilmesine bağlı olduğu inancı yatmaktadır. Zaman içerisinde bu ölçüm, bilimsel gözlem ve açıklama yöntemine bağlanarak bilim, tekrarlanabilir sonuçları ölçme, kaydetme ve sunma yeteneği olarak şekillenmiştir. Ancak tarihin büyük bir kısmında veri toplama, düzenleme, depolama ve analiz etme araçları yetersiz olduğundan çözümlemeler çok az veriyle yapılmıştır. Mümkün olduğunca az veri kullanarak geneli anlayabilmek için ayrıntılı teknikler geliştirilmiştir. Nitekim bu süreçte doğup gelişen istatistik, bir yönüyle buna hizmet etmektedir. Az veriyle çalışabilmek, bütünü temsil edebilme yeteneğine sahip örneklerle odaklanmayı gerektirmektedir. Veri toplamının ve analiz araçlarının kısıtlı olduğu bir çağın eseri olan örneklemede, örneklemin temsil başarısı, örneklemin boyutunu artırmaktan ziyade rastgelelikle çarpıcı bir biçimde artmaktadır. Dolayısıyla örnekleme, ayrıntıdan ödün vermeyi gerektirmektedir. Diğer bir ifadeyle örnekleme ayrıntıyı, ayrıntı düzeyini ve alt gruplara bakma imkânını ortadan kaldırmaktadır. Bu nedenle çoğu zaman örneklerin yakalayamadığı yerlerdeki önemli bulgulardan mahrum kalılabilmektedir (Mayer-Schönberger ve Cukier, 2013).

Analog bir ortamda bilgileri ölçmenin ve kaydetmenin yüksek maliyeti nedeniyle, büyük çapta verileştirmenin gerçekleşmesi için, daha ekonomik ölçüm araçlarının icat edilmesi gerekmektedir. Bu bağlamda bilgisayarların icadı, veri toplamayı çok daha verimli hale getiren dijital ölçüm ve depolama cihazlarını beraberinde getirmiştir. Örneğin 2000 yılına kadar dünya genelinde toplanan bilgilerin %25'i dijital, %75'i analog iken, 2007 yılına

gelindiğinde dijital verilerin oranı %93'e, 2013'te ise %98'e ulaşmıştır (Mayer-Schönberger ve Cukier, 2013). Bilgisayarlara eşlik eden akıllı telefonlar, çipler ve sensörlerle başlayan dijital dönüşüm, hiçbir zaman veri olarak ele alınmamış ve hatta bilgi niteliğine sahip olduğu dahi düşünülmemiş şeyleri nicel bir forma dönüştürmeye başlamıştır. Üstelik bu dönüşüm, analog çağıdaki pek çok alanda sadece verilerin küçük bir kısmının toplanabilmesinden, olabildiğince çok verinin ve hatta hepsinin toplanabilmesine doğru bir geçiş sağlamıştır.

Dijital devrim, sosyal bilimlerin geçmişte verileştirilemeyen soyut alanlarını gün yüzüne çıkartırken analiz yöntemlerini de temelden etkilemektedir. Sosyal bilimlerdeki dönüşümün temelinde iletişim teknolojilerindeki gelişmeler yer almaktadır. Duygu, düşünce ve bilginin akla gelebilecek her türlü yolla başkalarına aktarılmasının, yani iletişimin (TDK, 2009) teknolojik araçları telgrafla başlayıp, radyo, televizyon ve telefona evrilmiş, ardından pek çok özelliği bünyesinde barındıran taşınabilir ve giyilebilir yeni nesil iletişim araçları ve bunlara hitap eden yeni nesil uygulamalar doğmuştur.

Ortaya çıkış hikâyesi ABD ile o günlerdeki adıyla Sovyetler Birliği arasındaki soğuk savaş dönemine kadar uzanan ve bilgisayarlara iletişim ekleyen internet ise Web 1.0 ile dünyaya gözlerini açmış, Web 2.0 onu takip etmiştir. Bugünlerde ise internet, Web 3.0'ın doğum sancılarını çekmektedir. Web 1.0 dönemi, internet sitelerinin içeriğini oluşturanların sayısının çok az olmasına karşın içeriği kullananların sayısının çok fazla olduğu ve bu kullanıcıların içeriğe müdahale edemedikleri, bir bakıma sadece içeriğin müşterisi konumunda oldukları pasif dönemdir (Cormode ve Krishnamurthy, 2008). Bu dönemde kişisel bir web sayfasına sahip olmak ya da web sayfasına kişisel bir içerik koyabilmek çok kolay olmadığı için genellikle bu tür işler uzmanlar tarafından yerine getirilirdi. Web'in ikinci jenerasyonu olan Web 2.0'da Web 1.0'dan farklı olarak nihai kullanıcılar, aktif bir şekilde içerik oluşturabilmekte ve diğer kullanıcılarla iletişim kurabilmektedirler (Silva ve diğerleri, 2008). Bir başka ifadeyle Web 2.0 dönemi, kullanıcıların içeriğe müdahil olabildikleri, gruplar oluşturup yorum yapabildikleri, metin, ses ve video gibi içerikleri paylaşabildikleri interaktif ve sosyal bir ortamı ifade etmektedir. Nitekim sosyal ağlar ya da sosyal medya platformları bu dönemin başat ürünüdür. Web 3.0 dönemi ise, insan gibi

düşünebilen cihazların, interneti veri tabanı gibi kullanarak internetteki tüm bilgileri ve bunların birbiriyle ilişkilerini organize ettikleri, farklı uygulamalardaki ya da aynı uygulamanın farklı bölümlerindeki veriler arasındaki ilişkiden anlam çıkarabilecek yeteneğe sahip sezgisel yapay zekâ uygulamaların ön planda olduğu dönemdir (Hendler, 2009).

Sosyal bilimler açısından Web 2.0 döneminin en önemli veri kaynaklarından biri sosyal medya platformlarıdır. Sözlükte (Merriam-Webster, n.d.) sosyal medya, bilgi, fikir, kişisel mesajlar ve diğer içerikleri paylaşmak amacıyla kullanıcıların oluşturduğu ya da içerisinde yer aldığı çevrimiçi topluluklar vasıtasıyla iletişim kurma şekli olarak tanımlanmaktadır. Bir diğer tanıma göre sosyal medya, kullanıcıların, kullanıcı tarafından oluşturulan içerikten ve diğerleriyle etkileşim algısından değer elde eden hem geniş hem de dar kitlelerle, gerçek zamanlı veya eş zamansız olarak etkileşime girmesine ve seçici olarak kendini göstermesine izin veren internet tabanlı kanallardır (Carr ve Hayes, 2015). Sosyal medya platformlarının (sosyal ağ siteleri) en önemli özelliği kullanıcıların kendilerini ifade etmelerine ve sosyal ağlarının görünebilir olmasına olanak tanımasıdır (Boyd ve Ellison, 2007). Nitekim çalışmalar sosyal ağ sitelerinin kullanıcılarının bu platformlarda yeni insanlarla tanışmaktan ziyade hali hazırda sosyal ağının bir parçası olan insanlarla (arkadaş, iş arkadaşı, yakın arkadaş vb.) daha uzun süreli ve sık iletişim kurduklarını göstermektedir (Haythornthwaite, 2005). Dolayısıyla insanlar, ayrıntılı ilişkisel kayıtlar sağlayan cihazları kullanarak etkileşimi giderek daha fazla tercih ettikçe, geçmişte sosyal hayatın verileştirilmesinin en zor alanlarını veriye dönüştürme yeteneği üzerindeki uzun süredir devam eden kısıtlar hızla ortadan kalkmaktadır. Diğer bir ifadeyle sosyal etkileşimlerin zaman damgalı dijital izlerini üreten dijital teknolojilerin kullanımındaki hızlı artış, dijital platformlardan ve özellikle sosyal medyadan hem deneysel hem de gözlemsel verilerin hem büyük hem de mikroskobik bir ölçekte toplanması için benzeri görülmemiş fırsatlar sunmaktadır (Golder ve Macy, 2014). Böylece çevrimiçi sosyal ağlardan elde edilen veriler, sosyal yaşamı aktörler arasındaki ilişkiler olarak modellemesine olanak tanırken (Macy ve Willer, 2002), rastgele örneklerle güvenerek dayatılan izole birey varsayımlarının gevşetilmesine de izin vermektedir. Ayrıca gerçek dünyadaki karar verme süreçlerinin daha önce gözlemlenemeyen erken bilgi toplama aşamaları hakkında fikir vermekte ve gelecekteki eylemlerin daha iyi tahmin edilmesine

olanak tanımaktadır (Moat ve diğeri, 2014). Üstelik insanlar zaten normalde yaptıklarını yaparken veriler pasif ve gerçek zamanlı olarak kaydedildiğinden, örnekleme ve anketlerle ilgili önyargılar da ortadan kalkmaktadır (Mayer-Schönberger ve Cukier, 2013).

Özetle bilgisayarlar, taşınabilir ve giyilebilir akıllı cihazlar, internet ve nesnelerin internetinin [internet of things] merkezinde yer aldığı teknolojik devrim, toplum, finans ve ekonomi bilimlerine bakış açısını çarpıcı bir şekilde değiştirme potansiyeline sahip olan benzeri görülmemiş veri çeşitliliğini ve devasa boyutlara ulaşan veri kümelerine erişim imkânını beraberinde getirmektedir. Sosyal bilimler açısından dijital dönüşüm, insanların bilgi arayışını ve sosyal çevreleriyle etkileşimlerini çevrimiçi mecralara taşımasına, dolayısıyla bireysel ya da kolektif davranışların dijital izlerinin büyük veri kümelerine dönüştürülerek yeni bakış açılarıyla çözümlenebilmesine olanak sunmaktadır. Ayrıca dijital dönüşüm, farklı türlerde, hızla üretilen ve geleneksel veri işleme yöntemleriyle ele alınamayacak kadar büyük veri kümelerinin insan gücüyle anlamlandırılması mümkün olmadığından, bilgisayar bilimleri, istatistik, matematik ve alan bilgisini [domain knowledge] kullanan yapay zekâ, makine öğrenmesi, veri analitiği gibi disiplinlerarası bir yaklaşıma ihtiyaç duyulmaktadır (Cao, 2017).

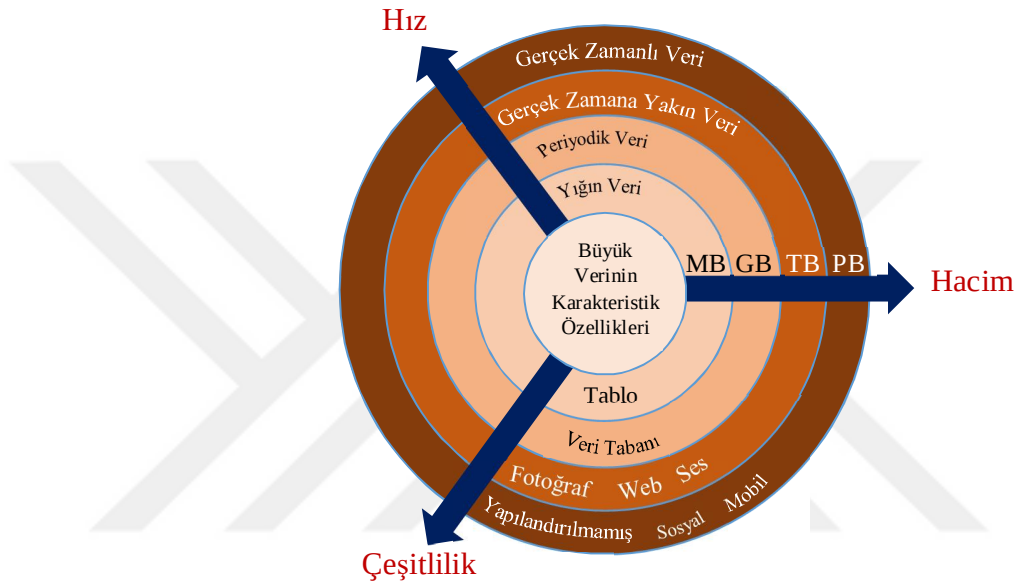
Bu bölümde öncelikle büyük veri ve yapay zekâ konularına kısaca değinilmekte, ardından bu alandaki gelişmelerin finans ile ilişkisi ele alınmaktadır. Daha sonra sosyal medyanın tanımına, kullanım istatistiklerine, kurumsal ve bireysel yatırımcılar açısından önemine, finansal influencerlara ve sosyal medyada sermaye piyasası manipülasyonu riskine değinilmektedir.

2.1.1 Büyük Verinin Tanımı

Literatürde 3V modeli olarak bilinen yaklaşıma göre büyük veri, hızla artan, daha fazla çeşitliğe sahip ve büyük hacimli veri kümeleri olarak tanımlanmaktadır. Gartner'ın (2012) biraz daha kapsamlı bir başka tanımına göre ise büyük veri [big data], gelişmiş içgörü, karar verme ve süreç otomasyonu sağlayan, yenilikçi bilgi işleme biçimleri gerektiren, yüksek

hacimli [volume], yüksek hızlı [velocity] ve yüksek çeşitlilikteki [variety] bilgi varlıklardır. Büyük veriyi farklı kılan bu üç temel karakteristik özellik, İngilizce karşılıklarının ilk harflerine atfen, alanyazınında büyük verinin 3V'si olarak ifade edilmektedir. Şekil 2.1'de, büyük verinin üç temel karakteristiği görülmektedir.

Şekil 2.1: Büyük Verinin 3V'si

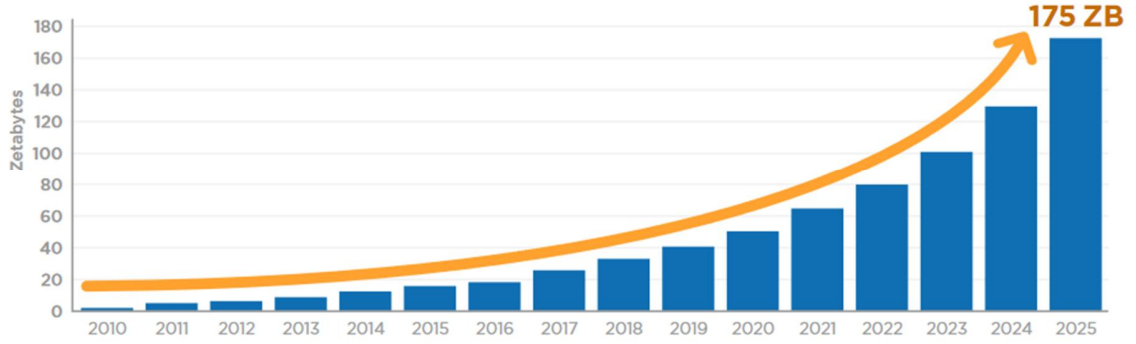


Kaynak: Data science central. 3V of big data. <https://www.datasciencecentral.com> adresinden 08 Mayıs 2020 tarihinde alınmıştır.

Büyük verinin bu ayırt edici özelliklerinden biri verinin miktarıdır. *Hacim* olarak da ifade edilen bu ayırt edici özellik, geleneksel yöntem ve araçlarla baş edilemeyecek miktardaki veriyi ifade etmektedir. Bu veriler taşınabilir ve giyilebilir cihazlar, sosyal medya platformları ve benzeri dijital araçların milyonlarca ya da milyarlarca kullanıcısı ile IoT (Internet of Things) cihazlar gibi çeşitli makineler tarafından oluşturulmakta ve her geçen gün hızla artmaktadır. Grafik 2.1 küresel veri hacminin yıllara göre genişlemesini göstermektedir. Buna göre veri hacminin 2025 yılına gelindiğinde dünya genelinde 175 zetabaytlık²⁷ bir boyuta ulaşacağı tahmin edilmektedir.

²⁷ 1 Zetabayt (ZB) 10²¹ bayttır.

Grafik 2.1: Küresel Verinin Boyutu



Kaynak: MacFeely, S., Me, A., Fu, H., & Schweinfest, S. (2020, Nov 13). *We urgently need a Global Data Convention. Here's why.* <https://www.weforum.org/agenda/2020/11/global-data-convention> adresinden 13 Mart 2022 tarihinde alınmıştır.

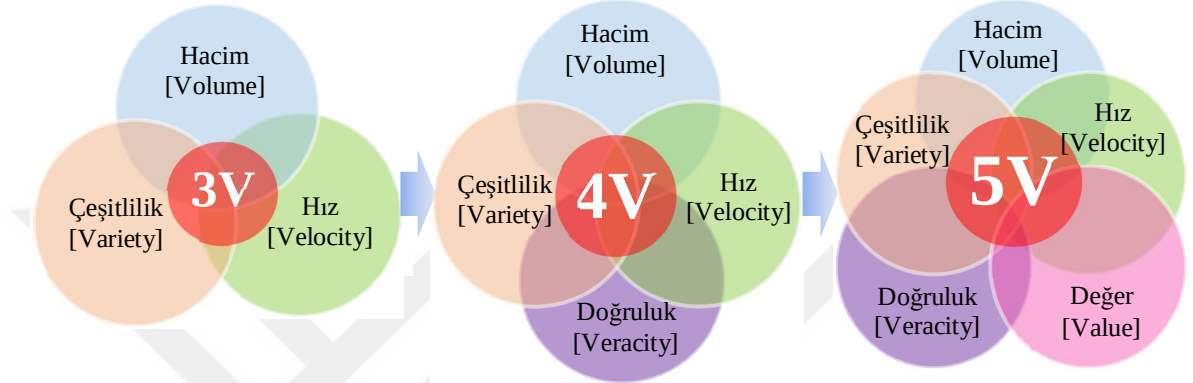
Büyük verinin ayırt edici bir diğer özelliği hızdır. Büyük veriyi geleneksel veriden farklı kılan iki tür *hız* söz konusudur. Bunlardan biri verinin *oluşma hızı*yla, diğeri *işlenme hızı*yla ilgilidir. Dijital araçların ve platformların yaygın bir şekilde kullanılmasıyla birlikte verilerin oluşma hızı giderek artmaktadır. Geleneksel veride örnekleme yoluyla ve zamansal aralıklarla toplanan verilerden farklı olarak büyük veri, sürekli olarak ve artan bir hızla üretilmektedir. Örneğin 2019 yılında her bir dakikada Twitter’da 511 bin tweet atılırken, Instagram’da 55 bin fotoğraf paylaşılmış, Google’da 4,5 milyon arama yapılmıştır (World Economic Forum, 2019, April 17; World Economic Forum, 2019, July 24). Bu rakamlar 2021 yılına gelindiğinde Twitter’da 575 bin tweete, Instagram’da 65 bin fotoğrafa, Google’da 5,7 milyon aramaya ulaşmıştır. Yine benzer şekilde 2021 yılında 694 bin saat Youtube içeriği izlenmiş, 6 milyon insan online alışveriş yapmıştır (Domo, 2021). Şekil 2.1’de de görüldüğü üzere verinin oluşumu, birikmiş olanlardan, belirli aralıklarla üretilenlere, oradan da gerçek zamanlı üretilenlere doğru hızla evrilmektedir. Verinin oluşumundaki bu hız aynı zamanda verinin hızla büyüdüğü, bazılarının hızla eskidiği ve dolayısıyla hızla, yani gerçek zamanlı olarak işlenmesi gerektiği anlamına da gelmektedir (Kitchin ve McArdle, 2016). Diğer bir ifadeyle büyük veri, hızla akan verinin hızla işlenmesi gerekliliğini de ortaya çıkartmaktadır.

Büyük verinin üçüncü ayırt edici niteliği *çeşitliliğidir*. Geleneksel durumda veriler genellikle yapılandırılmış biçimdeyken, büyük veri yapılandırılmış, yarı yapılandırılmış ve yapılandırılmamış olarak çeşitlenmektedir. *Yapılandırılmış veri* [structured data] satırlar ve sütunlar biçiminde kaydedilip geleneksel veri tabanlarında (SQL) ilişkisel olarak tutulabilen veri biçimidir. Günümüzde tüm verilerin yaklaşık %10-%20'ini yapılandırılmış verilerin oluşturduğu tahmin edilmektedir (IBM Cloud Education, n.d.). *Yapılandırılmamış veriler* ise [unstructured data], önceden belirlenmiş bir veri modeline veya şemasına göre düzenlenmemiş, bu nedenle de geleneksel bir ilişkisel veri tabanında saklanamayan veri türleridir. Diğer bir ifadeyle yapılandırılmamış veriler, satırlar ve sütunlar şeklinde tutulamayan metin, video, fotoğraf, ses dosyası, sosyal medya ve web sitesi verileri gibi verilerdir. Büyük verinin yaklaşık %80-%90'ının yapılandırılmamış verilerden oluştuğu tahmin edilmektedir. Bu tür veriler ilişkisel olmayan veri tabanlarında (NoSQL) saklanmaktadır (MongoDB, n.d.). Üçüncü bir veri kategorisi ise yapılandırılmış ve yapılandırılmamış veriler arasında karma bir nitelik taşıyan (json, csv, xml gibi) yarı yapılandırılmış [semi-structured] verilerdir. *Yarı yapılandırılmış veriler* ilişkisel veri tabanlarında saklanamasa da ayırtırmayı ve analiz etmeyi kolaylaştıran kısmen düzenli özelliklere sahiptir. Diğer bir ifadeyle yarı yapılandırılmış veriler bazı tanımlayıcı veya tutarlı özelliklere haiz, ancak ilişkisel veri tabanlarının beklediği katılıkta bir yapıya sahip değildirler. Yarı yapılandırılmış veriler genellikle gruplandırmaya ve hiyerarşilere izin veren dâhili etiketler ve işaretler içermekte, belirli veri özelliklerini tanımlamak ve verileri önceden belirlenmiş alanlara ölçeklemek için meta verileri kullanmaktadır. Meta veriler, yarı yapılandırılmış verilerin yapılandırılmamış verilere göre daha iyi kataloglanmasını, aranmasını ve analiz edilmesini sağlamaktadır (IBM Cloud Education, n.d.).

Büyük veriyi tanımlamak için başvurulmuş 3V yaklaşımı literatürde oldukça kabul görmüş ve zaman içerisinde Şekil 2.2'de görülen iki karakteristik özellik daha eklenerek tanım 5V modeline dönüşmüştür (Buyya ve Ramamohanarao, 2016). Sonradan eklenen bu ayırt edici özelliklerden biri değerdir. *Değer* [value], verinin içgörüye ve bilgiye dönüştürülmesi sonucunda yaratabileceği sosyal ve ekonomik faydayı ifade etmektedir (Mayer-Schönberger

ve Cukier, 2013). Dolayısıyla büyük veriden elde edilen bilginin karar süreçlerine etki ederek sosyo-ekonomik bir değere dönüşmesi beklenmektedir.

Şekil 2.2: Büyük Verinin Ayırt Edici Özellikleri



Kaynak: Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big data analytics= machine learning+ cloud computing. In R. Buyya, R. N. Calheiros, & A. V. Dastjerdi (Eds.), *Big Data: Principles and Paradigms* (pp. 3-38). Morgan Kaufmann.

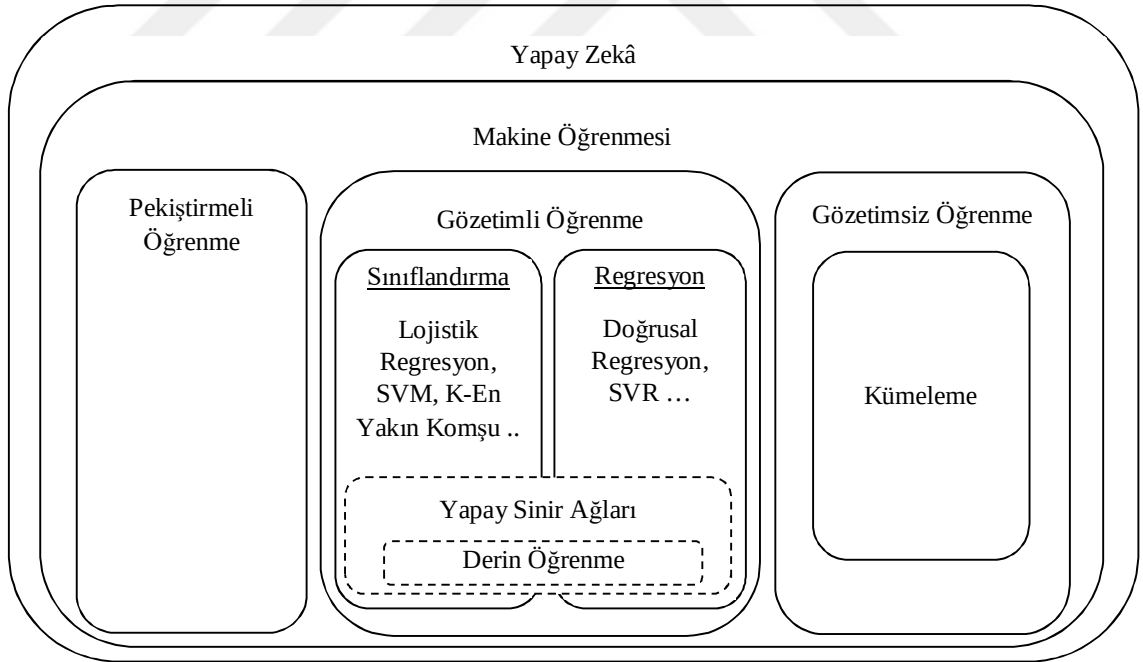
Büyük verinin beşinci ayırt edici özelliği ise verilerin doğruluğudur. *Doğruluk* [veracity], toplanan verilerin kalitesini veya güvenilirliğini ifade etmektedir (Frizzo-Barker ve diğerleri, 2016; Onay ve Öztürk, 2018). Büyük veri, birden çok veri türü ve kaynağına bağlı olarak çok sayıda veri boyutunu içerdiğinden, toplanan verilerin bazı tutarsızlıklar ve belirsizlikler barındırma olasılığı söz konusu olabilmektedir. Bu nedenle doğruluk, büyük veri analizinde değer yaratabilmek için verinin doğru ve kaliteli, veri kaynağının, türünün ve işlenmesinin güvenilir olması gerektiğini ifade etmektedir (Hariri ve diğerleri, 2019) ²⁸.

²⁸ Büyük veri literatürü gözden geçirildiğinde, büyük veriyi tanımlama çabalarının 5V ile de sınırlı kalmadığı, farklı çalışmalarda farklı ayırt edici özellikler (6V, 11V gibi) eklendiğinden bu sayının artarak devam ettiği görülmektedir. Ancak bu çalışma kapsamında büyük verinin beş ayırt edici özelliğine değinmenin yeterli olduğu düşünülmektedir.

2.1.2 Yapay Zekâ, Makine Öğrenmesi ve Finansal Sektör

Yapay zekâ terimi, Türkiye Ulusal Yapay Zekâ Stratejisinde, dinamik ve belirsiz ortamlarda akıl yürütme, anlam keşfetme, genelleme veya geçmiş deneyimlerden öğrenme gibi insanlara özgü bilişsel kabiliyetlerle donatılmış sistemleri ifade etmektedir. Bu bağlamda yapay zekâ da bir bilgisayarın veya bilgisayar kontrolündeki bir robotun çeşitli faaliyetleri zeki canlılara benzer şekilde yerine getirme yeteneği olarak tanımlanmaktadır (Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı ve Sanayi ve Teknoloji Bakanlığı, 2021). OECD'nin (2019) tanımına göre ise yapay zekâ, değişen seviyelerde özerklikle çalışabilecek şekilde tasarlanmış, makine ve/veya insan temelli girdileri/verileri kullanarak belirli hedefler kümesi için gerçek veya sanal ortamları etkileyen tahminler, öneriler veya kararlar verebilen makine tabanlı sistemlerdir. Yapay zekânın kapsamı Şekil 2.3'te gösterilmektedir.

Şekil 2.3: Yapay Zekânın Kapsamı



Yapay zekânın bir alt kümesi olan makine öğrenmesi ise, istatistik, yapay zekâ, felsefe, bilgi teorisi, biyoloji, bilişsel bilim, hesaplama karmaşıklığı [computational complexity] ve

kontrol teorisi dâhil olmak üzere birçok alanın kavram ve sonuçlarından yararlanarak, deneyimlemek suretiyle otomatik olarak gelişen bilgisayar programlarıyla ilgilidir (Mitchell, 1997). Başka bir ifadeyle makine öğrenmesi makinelere, insanlar tarafından programlanmaksızın öğrenerek kendini geliştirebilme [self-improve] yeteneği kazandırmak için kullanılan bilgisayar bilimi teknikleridir.

Makine öğrenmesinde temel amaç veriden öğrenme ve veriye dayalı içgörü, tahmin ve karar yeteneği kazanmaktır. Verilerin doğası göz önünde bulundurulduğunda, gözetimli öğrenme, gözetimsiz öğrenme ve pekiştirmeli öğrenme gibi farklı öğrenme yaklaşımları söz konusudur. *Gözetimli öğrenme* [supervised learning], hem girdilerin hem de çıktıların bilindiği durumlarda kullanılan ve girdilerin çıktılarıyla eşleştirilmesinin öğrenildiği yaklaşımdır (Apaydın, 2012). Sınıflandırma ve regresyon, gözetimli öğrenme yöntemleridir. Sınıflandırmada çıktılar ayrık değerler (sınıf etiketleri) alırken, regresyonda çıktılar sürekli. Sınıflandırma algoritmalarının örnekleri arasında k-en yakın komşu, lojistik regresyon ve destek vektör makinesi (SVM) bulunurken, regresyon örnekleri arasında destek vektörü regresyonu, doğrusal regresyon ve polinom regresyonu bulunmaktadır. Yapay sinir ağları ya da derin öğrenme gibi bazı algoritmalar hem sınıflandırma için hem de regresyon için kullanılabilir (L'Heureux ve diğerleri, 2017). *Gözetimsiz öğrenme* [unsupervised learning] sadece girdi verisi kullanılarak ve herhangi bir geribildirim olmadan girdilerdeki düzenliliklerin bulunması amaçlanmaktadır. En yaygın gözetimsiz öğrenme yaklaşımı, girdileri yerleşik benzerlik ölçütlerine göre gruplayan kümeleme yaklaşımıdır (L'Heureux ve diğerleri, 2017). Kümele algoritmalarına örnek olarak k-means ve hiyerarşik kümeleme örnek verilebilir. Gözetimli öğrenme ile gözetimsiz öğrenme arasında bir yerde duran bir diğer yaklaşım ise pekiştirmeli öğrenmedir. *Pekiştirmeli öğrenme* [reinforcement learning] ise dizgenin çıktısı hedefe ulaşmak için gerçekleştirilen doğru eylem dizisidir. Yapay öğrenme uygulamasına sadece eylemin sonucu belirtilmekte, bir dizi deneyimin her aşamasında olası eylemlerden hangilerinin seçilmesi gerektiğinin öğrenilmesi beklenmektedir (Apaydın, 2012). Diğer bir ifadeyle pekiştirmeli öğrenmede durumdan duruma geçilerek ödülün elde edileceği duruma ulaşılması öğrenilmektedir (Nilsson, 2011).

Yapay zekâ, makine öğrenmesi ve büyük verinin güçlü bir şekilde etkilediği alanlardan biri finanstır. Özellikle finansla ilgili hızla artmaya devam eden büyük veri, uygun maliyetli bilgi işlem kapasitesiyle bir araya geldiğinde, verilerdeki temel ilişkileri insanların yeteneklerinin ötesinde bir başarıyla yakalamak için yapay zekâ ve/veya makine öğrenmesi modellerini ön plana çıkarmaktadır. Dolayısıyla yapay zekâ, makine öğrenmesi ve büyük verinin finans sektöründe aktif olarak kullanılmasının hem maliyetleri düşürme ve üretkenliği yükseltme yoluyla verimliliklerini artırmak, hem de müşterilere sunulan hizmet ve ürünlerin kalitesini artırmak suretiyle finans sektöründe faaliyette bulunan firmalara rekabet avantajı sağlaması beklenmektedir.

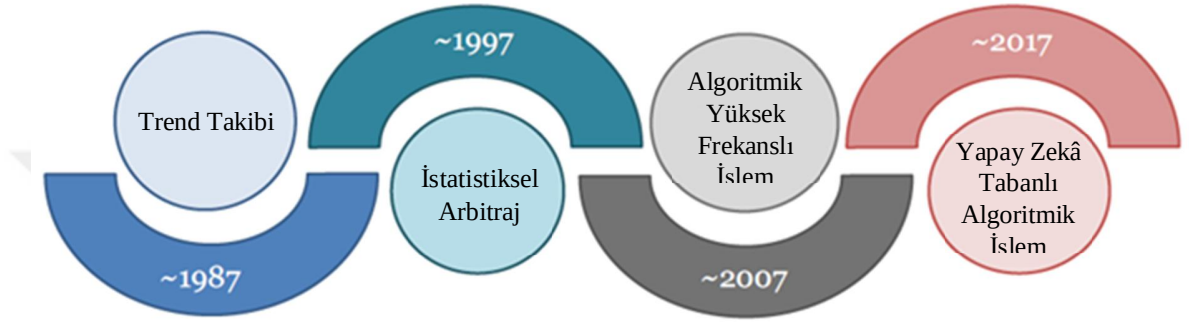
Finansta yapay zekâ tekniklerinin özellikle, varlık yönetimi, algoritmik işlemler, kredi aracılığı ve blokzincir tabanlı finans alanlarında oldukça etkin bir rol oynadığı görülmektedir. Varlık yönetiminde yapay zekâ ve makine öğrenmesi, müşteri deneyimini iyileştirmek, operasyonel iş akışlarının verimliliğini ve doğruluğunu artırmak ve yatırım süreçlerinin birçok yönünü destekleyerek performansını artırmak için kullanılmaktadır (Novick ve diğerleri, 2019). Makine öğrenmesi modelleri binlerce risk faktörünü izleyebildiğinden ve portföy performansını binlerce ekonomik koşul senaryosu altında test edebildiğinden, varlık yöneticileri ve diğer büyük kurumsal yatırımcılar için iyi bir risk yönetimi vadetmektedir. Benzer şekilde büyük veri ile beslenen makine öğrenmesi modelleri portföy tahsisi ve/veya hisse senedi seçimi konusunda da kararları etkileyen öneriler sağlayabilmektedir. Geleneksel stratejilerin merkezindeki yapılandırılmış verilere ilave olarak büyük miktarda ham veya yapılandırılmamış ya da yarı yapılandırılmış verilerle, stratejilerinin uygulanmasında yapay zekâyı kullanan yatırımcılara yeni bilgi avantajı sağlamayı da vaat etmektedir. Diğer bir ifadeyle yapay zekâ, varlık yöneticilerinin çok sayıda kaynaktan gelen büyük miktardaki verileri analiz ederek ve çok kısa zaman dilimlerinde stratejileri için verilerden yeni iç görüler elde etmesine olanak tanımaktadır. Bununla birlikte yapay zekâ modellerinin çok sayıda varlık yöneticisi tarafından kullanılması, özellikle stres zamanlarında likidite ve sistemin istikrarı için potansiyel riskleri artırabilecek sürü davranışına ve tek yönlü piyasalara yol açabilme potansiyeline de sahiptir (Organisation for Economic Co-operation and Development, 2021a).

Yapay zekânın finasta etkin rol oynadığı alanlardan biri de algoritmik işlemlerdir. Yapay zekâ tabanlı finansal işlem sistemleri [algorithmic trading], derin öğrenme, olasılıklı mantık gibi yapay zekâ tekniklerini kullanarak herhangi bir insan müdahalesi olmadan al-sat işlemlerini yürütmektedirler (Metz, 2016). Başka bir deyişle finansal işlemlerde [trading] yapay zekâ, hem sağladığı al-sat stratejisi önerileri hem de yaptığı hızlı tahminlerle hareket tarzını seçerek al-sat yapan otomatik finansal işlem sistemlerini güçlendirmek için kullanılmaktadır. Ayrıca düşünce süreçlerinin “eğer/durumunda [if/then]” prosedürünü kullanarak, finansal piyasalardaki gelişmelere göre işlemlerin sistematik bir şekilde stratejilendirilmesine de yardımcı olmaktadır (Basar, 2021). Örneğin yapay zekâ tabanlı uygulamalar, kendi kendilerini eğiterek değişen piyasa koşullarına uyum sağladıkları için, tamamen otomatik bir şekilde ve yeniden programlamaya gerek kalmadan yatırımcının ihtiyaçlarına, varlıkların risk durumlarına ve piyasalardaki gelişmelere göre pozisyonları güncelleyebilmekte veya pozisyondan çıkabilmektedirler. Ayrıca hisse senedi fiyatı üzerinde etkili olabilecek haberler, raporlar ya da sosyal medya gönderileri gibi yapılandırılmamış verilerdeki eğilim ve sinyalleri de doğal dil işleme ya da metin madenciliği algoritmalarıyla analiz ederek karar süreçlerine dâhil edebilmektedir. Nitekim günümüzde yapay zekânın daha gelişmiş biçimleri çoğunlukla, yüksek frekanslı algoritmik işlemlerden ziyade, tanımlanması ve değer çıkarılması daha zor olan “düşük bilgi değeri” olaylarından gelen sinyalleri belirlemek ve bu bilgileri işlemlerle ilgili kararlara dönüştürmek için kullanılmaktadır. Daha az gelişmiş algoritmalar ise çoğunlukla, yürütme hızının önemli olduğu durumlar için kullanılmaktadır (Organisation for Economic Co-operation and Development, 2021a).

Şekil 2.4, yapay zekâ ve finansal işlemlerin tarihsel gelişimini göstermektedir. Finansal işlemlerde yapay zekâ kullanımı 2000’li yıllarda eylem hızından kâr sağlamayı hedefleyen yüksek frekanslı işlem stratejilerine dayalı, basit parametrelerle alış-satış emri verebilen birinci nesil algoritmalarından oluşmaktaydı. İkinci nesil algoritmalar ise 2017 yılından itibaren, verilerdeki örüntüleri yakalayarak algoritmaların karar mantığını geliştirmek için kullanılmakta ve böylece öğrenme becerisine sahip ve daha az insan müdahalesine dayalı algoritmalara dönüşmektedirler. Diğer bir ifadeyle yeni nesil algoritmalar, yapay zekâ

tekniklerinin kendi karar mantıklarını özerk ve dinamik olarak geliştirmesine izin vermek suretiyle alım satımların otomatik olarak yürütülmesi için artırılmış yetenekler sunarak stratejiler tasarlamaya ve bunları yürütmeye kadar tüm eylem aşamalarına hizmet etmesi beklenmektedir.

Şekil 2.4: Yapay Zekâ ve Finansal İşlemlerin Tarihsel Gelişimi



Kaynak: Organisation for Economic Co-operation and Development. (2021a). *Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers*. OECD Publishing. <https://www.oecd.org/finance/artificial-intelligence-machine-learning-big-data-in-finance.htm>

Yapay zekâ modellerinin finasta kullanıldığı alanlardan bir diğeri kredi aracılığı ve kredi riskinin değerlendirilmesidir. Günümüzde yapay zekâ, bankalar tarafından olası borçluların kredibilitesinin değerlendirilmesinde, sigortalamayla ilgili kararlarda ve dolandırıcılık tespitinde giderek daha aktif bir rol üstlenmektedir (Organisation for Economic Co-operation and Development, 2021a). Bu çerçevede makine öğrenmesi modelleri, özellikle sınırlı bilgiye sahip olduğu durumlarda, istatistiksel modellere kıyasla temerrüt riskinin tahmininde daha başarılı sonuçlar sağlamaktadır (Moscatelli ve diğerleri, 2020; Liu ve diğerleri, 2022). Büyük veri kümelerinin ve yapay zekânın finasta kullanılması geleneksel verilerle temerrüt riskinin tahmin edilmesine ek olarak sosyal medya verileri (Kulkarni ve Dhage, 2019), metin tabanlı veriler (Netzer ve diğerleri, 2019), sosyal ağ verileri (Niu ve diğerleri, 2019) ve benzeri dijital izleri de analize dâhil ederek tahmin başarısını artırmaktadır. Kredibilitenin belirlenmesinde yapay zekâ modellerinin kullanılması, kredi piyasalarında hüküm süren bilgi asimetrisini azaltarak teminat ihtiyacını asgari düzeye çekme ve teminatla ödeme gücünü kanıtlayamayan sürdürülebilir şirketlere kredi verilmesini

sağlayarak krediye erişimi artırma potansiyeline de sahiptir. Böylece özellikle KOBİ finansmanında karşılaşılan kısıtları hafifleterek reel ekonominin büyümesini desteklemesi beklenmektedir (Organisation for Economic Co-operation and Development, 2021a). Yapay zekâ tabanlı kredi puanlama modellerinin büyük potansiyellerine rağmen taşıdıkları bazı riskler de gözden uzak tutulmamalıdır. Özellikle temel insan önyargılarını yansıtan, doğru etiketlenmemiş ya da veri kalitesinin düşük veya yetersiz olduğu durumlarda belirli insan sınıflarına (örneğin ırk, cinsiyet, etnik köken, dine dayalı) karşı ayrımcılığa neden olan önyargılı sonuçlar üretebilme veya mevcut ön yargıları şiddetlendirebilme potansiyeline de sahiptir (Fuster ve diğerleri, 2022; Petrasic ve diğerleri, 2022). Makine öğrenmesinin, finanstaki diğer uygulamalarına benzer şekilde, açıklanabilirliklerinin olmaması, yani karar verme sürecini anlama, takip etme veya tekrar etme zorluğu nedeniyle şeffaflık sorunlarını da gündeme getirmektedir (Organisation for Economic Co-operation and Development, 2021a).

2.2 SOSYAL MEDYA VE YATIRIMCILAR

Geleneksel medya döneminde finansal piyasalarda yatırım kararları, aracı kurumlar ve finans uzmanlarının yanı sıra radyo, televizyon, gazete ve dergi gibi kitle iletişim araçları vasıtasıyla ulaşılan bilgiye dayalı olarak veriliyordu. Teknolojideki gelişmeler bu trendi değiştirmiştir. Dolayısıyla günümüzde finansal piyasalar ile ilgili araştırmalar, sosyal medya platformları da dâhil olmak üzere dijital mecralardaki mevcut bilgilere dayalı olarak yapılmaktadır. Özellikle sosyal medya, ürettiği ilgi çekici içerikler aracılığıyla farklı arayışlara hitap edebilme, en son gelişmeler hakkında çok daha güncel bilgileri çok daha hızlı sunabilme yeteneğine sahiptir. Ayrıca yatırımcıların reklamlardan ziyade yatırımlar hakkında bilgi sahibi olan arkadaşlardan ve aileden gelen tavsiyelere güvendiği, akran değerlendirmesine göre iyi olana yatırım yapmayı tercih ettiği ve bu etkileşimin özellikle sosyal medya platformları aracılığıyla gerçekleştiği göz önünde bulundurulduğunda, sosyal medya çoğu yatırımcı için bilgi arayışlarında başvuru olan önemli bir kaynak olma potansiyeline sahiptir.

Bununla birlikte sosyal medya sahte ve yanıltıcı haberlerin yayılmasına da aracılık edebilmektedir. Bu durum yatırımcılar açısından manipülasyon riskini beraberinde getirmektedir.

Bu başlıkta sosyal medyanın tanımına, kurumsal ve bireysel yatırımcılar açısından önemine yer verilmektedir. Ayrıca sosyal medyada yatırım fikirleriyle ilgili, eğitim, bilgi ve tavsiyelerde bulunan finansal etkili kişiler konusuna değinilmekte ve son olarak sosyal medyada sermaye piyasası manipülasyonu riski ele alınmaktadır.

2.2.1 Sosyal Medyanın Tanımı ve Kullanım İstatistikleri

Sosyal medya, sanal ağlar²⁹ ve topluluklar aracılığıyla fikirlerin, düşüncelerin ve bilgilerin paylaşımını kolaylaştıran bilgisayar tabanlı teknolojileri ifade etmektedir (Dollahide, 2021). Bir başka tanıma göre sosyal medya, paylaşılan içeriğin yayılmasını, diyalog oluşturmayı ve daha geniş bir kitle ile iletişimi kolaylaştıran, kullanıcı odaklı çeşitli platformlardan oluşmaktadır. Esasen insanlar tarafından ve insanlar için yaratılan dijital bir alandır ve etkileşimlerin ve ağların farklı seviyelerde (kişisel, profesyonel, iş vb.) gerçekleşmesine elverişli bir ortam sağlamaktadır (Kapoor ve diğerleri, 2018). Bu çerçevede sosyal medya, insanların dijital platformlarda sosyal çevreleri ile etkileşim kurmasına imkân sunan, duygu ve düşüncelerin ifade edilmesini ve yayılmasını sağlayan ve dijital olarak izlenebilen platformlar olarak tanımlanabilir. Daha kapsayıcı bir tanımda ise sosyal medya, bireyler arasındaki etkileşimi kolaylaştırmak için tasarlanmış her türlü çevrimiçi kaynak olarak ifade edilmektedir (Bishop, 2019).

Sosyal medya, kullanıcıların içerik oluşturmasını ve paylaşmasını veya sosyal ağlara katılmasını sağlayan web sitelerini ve uygulamaları ifade etmektedir. Bu platformlar düşünceleri, görüşleri, fotoğrafları, videoları vb. içerikleri gerçek zamanlı olarak paylaşma

²⁹ Sosyal ağ oluşturma ve sosyal medya örtüşen kavramlardır. Ancak sosyal ağ genellikle kullanıcıların kendi aralarında topluluklar oluşturması olarak anlaşılırken, sosyal medya daha çok bir kitle oluşturmak için sosyal ağ sitelerini ve ilgili platformları kullanmakla ilgilidir.

imkânı vermektedir. Sosyal medyayı geleneksel medyadan ayıran temel farklılıklar Tablo 2.1’de yer almaktadır. Bu bağlamda sosyal medyayı geleneksel medyadan ayıran en önemli özellik iletişimin tek yönlü değil karşılıklı olmasıdır. Bu nedenle geleneksel medyanın tek yönlü etkisinin yerini etkileşim almaktadır. Başka bir ifadeyle sosyal medyada kullanıcı iletişim ortamının içeriğini ve dolayısıyla iletişimde bulunduğu kişilerin davranışlarını değiştirebilme özelliğine sahiptir. Bu nedenle geleneksel medyadaki *izleyicilerin* yerini sosyal medyada *kullanıcılar* almıştır. Bu farklılıklardan bir diğeri iletişim aracının internet tabanlı olup olmamasıyla ilgilidir. Bir diğer deyişle sosyal medya internete bağlı iken geleneksel medya internete bağlı değildir. Ayrıca geleneksel medyada bir mesaj yayımlandıktan sonra değişiklik yapmak zor iken sosyal medyada mesaj yayımlandıktan sonra değişiklik yapmak kolaydır. Son olarak geleneksel medyada kontrol belirli bir grubun elinde iken sosyal medyada kontrol genele daha yaygındır.

Tablo 2.1: Geleneksel Medya Sosyal Medya Karşılaştırması

GELENEKSEL MEDYA	SOSYAL MEDYA
Geleneksel medya, dijital medyanın ortaya çıkmasından önce mevcut olan kitle iletişim araçlarını ifade etmektedir.	Sosyal medya, kullanıcıların içerik oluşturmasını ve paylaşmasını veya sosyal ağlara katılmasını sağlayan web sitelerini ve uygulamaları ifade etmektedir.
Televizyon, radyo, gazete ve dergiler geleneksel medya örnekleridir.	Facebook, YouTube, Instagram, Twitter, LinkedIn, Wikipedia, Pinterest sosyal medya örnekleridir.
Tek yönlü iletişim sunar.	İki yönlü iletişim sunar.
İzleyicilere hitap eder.	Kullanıcılara hitap eder.
Etki	Etkileşim
İnternete bağlı değildir.	İnternete bağlıdır.
Bir mesaj yayımlandıktan sonra değişiklik yapmak zordur.	Bir mesaj yayımlandıktan sonra değişiklik yapmak kolaydır.
Paylaşılacak istenilen mesaj üzerinde daha az kontrole sahip olunur.	Paylaşılacak istenilen mesaj üzerinde daha fazla kontrole sahip olunur.

Kaynak: <https://pediaa.com/difference-between-social-media-and-traditional-media/>

İnternetin yaygınlaşmasıyla birlikte iletişimde geleneksel medya kullanımı giderek azalmaya devam ederken dijital medya kullanımı hızla artmaktadır. Örneğin ABD’de 2011 yılında dijital medyada günlük 214 dakika ve geleneksel medyada 453 dakika harcanırken, 2023 yılında, dijital medyada harcanan süre 500 dakikaya çıkmış, geleneksel medyada harcanan süre ise 285 dakikaya düşmüştür (Guttmann, 2023). Ayrıca 2019 yılının sonunda ortaya

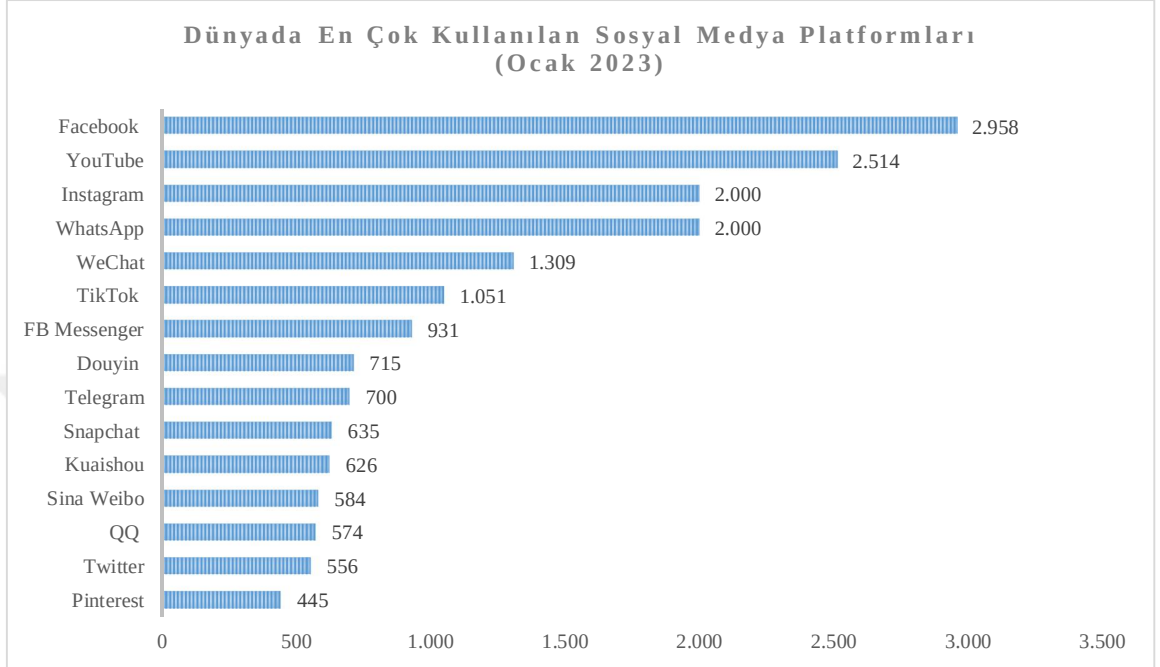
ıkarak tm dnyaya yayılan Covid-19 pandemisi de insanların dijital medya kullanımını artırmıştır (Guttmann, 2023).

İnternet ve teknoloji­deki hızlı deęişimler dnya genelinde sosyal medya kullanımını oldukça yaygınlaştırmıştır. Örneğin 2013 yılında 1,7 milyar olan sosyal medya kullanıcı sayısı, 2018 yılında 3,2 milyara, 2023 yılında ise 4,7 milyara ulaşmıştır (Digital 2023, 2023).

Sosyal medya kullanımının yaygınlaşması bu platformların kullanım amaçlarıyla ilgili istatistikleri de önemli hale getirmektedir. Bu bağlamda yapılan araştırmalar (Digital 2023, 2023) insanların genellikle arkadaşlar ve aile ile iletişimde kalmak, boş zamanlarını doldurmak, haber okumak, içerik bulmak, konuşulnlardan haberdar olmak ve başkalarıyla görüşlerini paylaşmak amacıyla sosyal medyayı kullandıklarını göstermektedir. Dikkat edilecek olursa bütün bu amaçlar temelde insanların bilgi arayışı/paylaşımı ile ilgili olduğu görölmektedir.

Sosyal medya kullanımıyla ilgili bir dięer konu ise insanların bir birleriyle etkileşim kurarken en çok hangi platformları tercih ettikleriyle ilgilidir. Grafik 2.2, küresel düzeyde en çok kullanılan sosyal medya platformlarını göstermektedir. Buna göre Facebook 2,9 milyar YouTube 2,5 milyar, WhatsApp 2 milyar, Instagram 2 milyar, WeChat 1,3 milyar ve Tiktok 1 milyar kullanıcı tarafından tercih edilmektedir. Bu istatistiklere göre Twitter ise 556 milyon kullanıcıya hitap etmektedir.

Grafik 2.2: Dünyada Genelinde En Çok Kullanılan Sosyal Medya Platformları

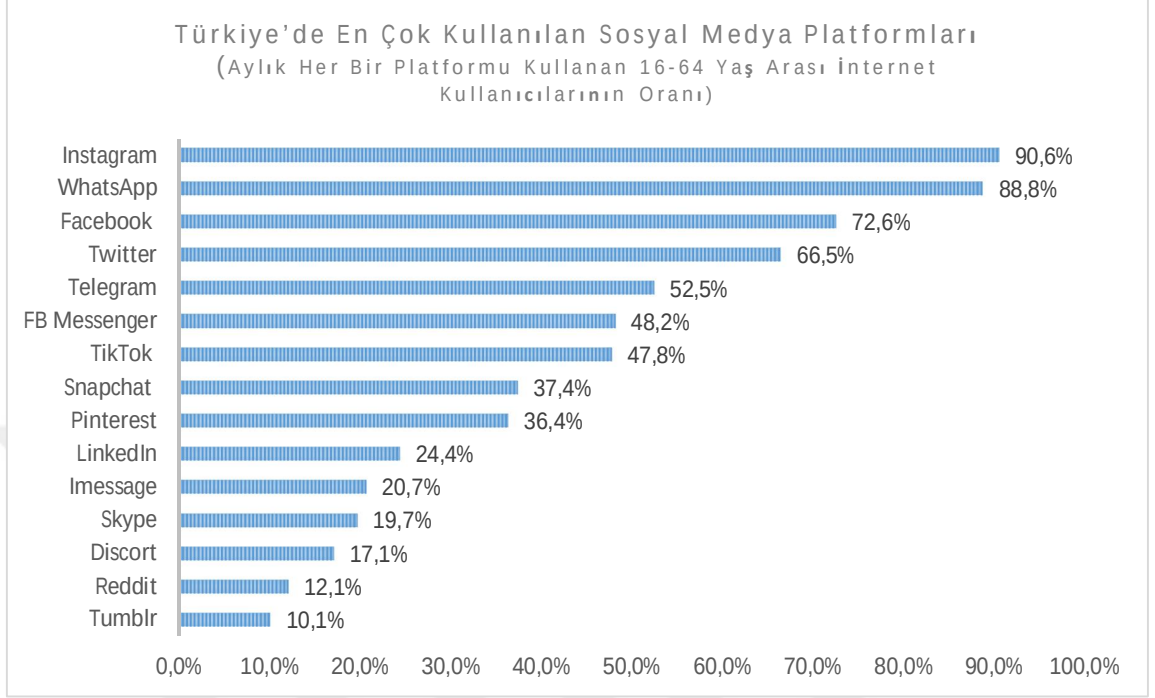


Kaynak: Digital 2023: Global Overview Report

Sosyal medya platformlarının kullanımı yaş gruplarına göre farklılık gösterebilmektedir. Bu bağlamda 16-24 ile 25-34 yaş gruplarında Instagram ve TikTok'un daha fazla kullanıldığı söylenilebilir. Ayrıca WhatsApp hemen hemen bütün yaş gruplarında kullanılırken, Facebook özellikle 16-24 yaş grubunda diğer yaş gruplarına kıyasla çok daha az kullanılmaktadır. Bu durum platformların gelecekteki potansiyellerine ışık tutması açısından önemli ipuçları barındırmaktadır (Digital 2023, 2023).

Sosyal medya platformlarının Türkiye'de kullanımı, küresel ölçekteki kullanımdan kısmen farklılaşmaktadır. Türkiye nüfusunun %73'ü sosyal medyayı kullanmaktadır. Grafik 2.3, Türkiye'de en çok kullanılan sosyal medya platformlarını göstermektedir. Bu grafiğe göre Türkiye'de sosyal medya kullananların %90,6'sı Instagram'ı, %88,8'i WhatsApp'ı, %72,6'ı Facebook'u, %66,5'i Twitter'ı ve 52,5'i Telegram'ı tercih ettiği görülmektedir.

Grafik 2.3: Türkiye’de En Çok Kullanılan Sosyal Medya Platformları



Kaynak: Digital 2023: Turkey

Özetle sosyal medya, insanları fikir alışverişinde bulunmaya, bağlantı kurmaya, ilişki kurmaya, bir amaç için harekete geçmeye, tavsiye aramaya ve rehberlik sunmaya olanak tanıyan yeni bir iletişim aracıdır. Sosyal medyanın dünya genelinde oldukça yaygın bir şekilde kullanılıyor olması, bu platformların topluma etkisini de ön plana çıkarmaktadır. Zira günümüz dünyasında sosyal medyanın kültür, ekonomi, finans ve genel dünya görüşünü etkilemede önemli bir rol oynadığı yadsınamaz bir gerçekliktir. Daha açık bir ifadeyle toplumsal olayların örgütlenmesinden (Stepanova, 2011; Wolfsfeld ve diğerleri, 2013; Rane ve Salem, 2012; Eltantawy ve Wiest, 2011), seçimlerin öngörülmesine, ekonomiden finansal piyasalara (Schoen ve diğerleri, 2013) kadar pek çok alanda sosyal medyanın etkisi bulgularla ortaya konulmaktadır.

2.2.2 Sosyal Medya ve Kurumsal Yatırımcılar

Kurumsal yatırımcılar; sigorta şirketleri, mevduat kuruluşları, sosyal güvenlik kuruluşları, emeklilik fonları, yatırım şirketleri, yatırım fonları, vakıflar, sendikalar gibi kurumlar ya da bireyler adına fon aktarım yetkisi verilmiş tüzel kişilerdir. Sosyal medya sadece iletişim kurma şeklini değil, aynı zamanda iş yapma şeklini de değiştirdiğinden, sosyal medyanın etkilediği alanlardan biri kurumsal yatırımcılarla ilgilidir.

Sosyal medyanın kurumsal yatırımcılar üzerindeki etkisini araştırmak amacıyla Brunswick (2023) tarafından dijital yatırımcı anketi yapılmıştır. Bu anket kurumsal yatırımcıların çevrimiçi olarak bilgilerini nereden aldıklarını, bir şirketin dijital varlığını nasıl değerlendirdiklerini ve bu analizin yatırım kararları üzerindeki etkisini araştıran kapsamlı bir çalışmadır. Dijital Yatırımcı Anketinde ABD, İngiltere, Kanada ve AB'de faaliyet gösteren büyük ve küçük yatırım firmalarından 257 kurumsal yatırımcının küresel bir örneği incelenmiştir.

Bu kapsamda yapılan dijital yatırımcı anketine göre yatırımcıların %88'i dijital platformlardan veya sosyal medyadan gelen bilgilere dayanarak tavsiye veya karar verdiklerini belirtmişlerdir. Başka bir ifadeyle bu araştırma sonuçlarına göre yatırımcıların araştırmalarını bir dizi dijital kaynağa dayandırdıkları görülmektedir. Bu nedenle ankette katılımcılara bir hisse senedinin araştırılması ve değerlendirilmesi için hangi dijital ve sosyal medya platformlarının ne kadar önemli olduğu da sorulmuştur. Bu sonuçlara göre birinci sırada şirketlerin yatırımcı ilişkileri web sayfaları yer almaktadır. Onu LinkedIn, Google aramaları ve e-mail haber bültenleri takip etmektedir. Benzer şekilde YouTube, Instagram, Facebook ve Twitter da yatırımcıların başvurdukları önemli bilgi kaynakları olarak yerine almaktadır (Brunswick, 2023). Bu sonuçlar yatırımcıların, bilgi toplamak için daha geniş bir kaynak yelpazesine yöneldiğini göstermektedir. Başka bir ifadeyle yerleşik platformların çoğu, hisse senetlerinin araştırılması ve değerlendirilmesi için önemli kaynak niteliği taşımaktadır.

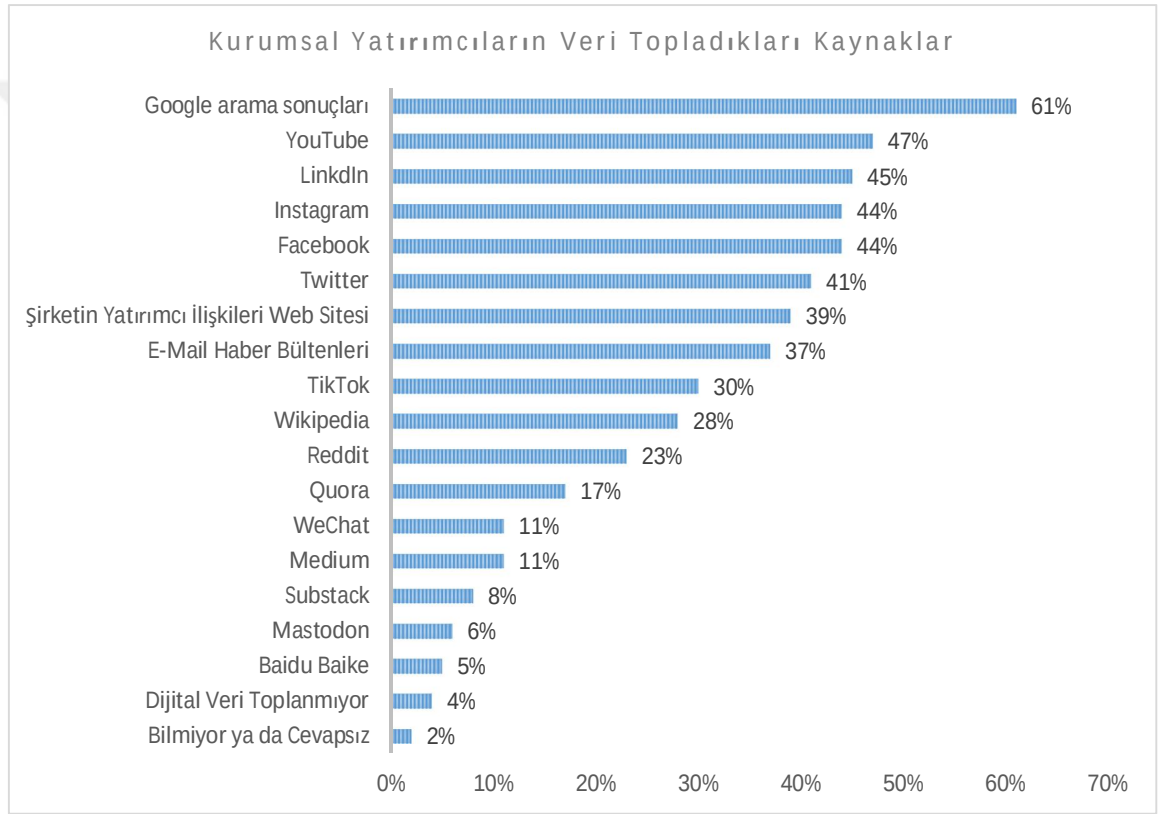
Kurumsal yatırımcıların sosyal medyayı kullanma amaçları ise araştırmanın bir başka boyutunu oluşturmaktadır. Bu çerçevede elde edilen sonuçlara göre kurumsal yatırımcılar, rekabet avantajını artırmak için çok çeşitli dijital ve sosyal medya platformlarından bilgi topladıkları görülmektedir. Bunun temel nedenleri arasında, haber ve bilgileri diğer medya kaynaklarından daha hızlı alıyor olmaları, başka yerde bulunmayan bilgileri öğreniyor olmaları ve takip ettikleri şirketlerle ilgili güncel bilgilere erişim imkânı elde ediyor olmalarıdır (Brunswick, 2023). Ayrıca yatırımcılar, tek bir platformda bile, çok çeşitli hesaplardan farklı bilgilere ulaşabilmektedirler. Buna ek olarak yatırımcıların yarısından fazlası, sadece takip ettiği hisse senetlerini değil, aynı zamanda hisse senetlerini takip ettikleri şirketin CEO'sunu ve diğer kilit yöneticilerini de takip edebilmektedirler. Örneğin Twitter'da, yatırımcıların yaklaşık yarısı hisse senetlerinin hashtaglerini takip etmekte ve hisse senetleriyle ilgili Twitter paylaşımlarını izlemektedir. Bu sonuçlar, bir yatırımcı web sitesini takip eden yatırımcılar için gerçeğin tek bir versiyonuna sahip olma günlerinin bittiğini göstermektedir.

Dijital yatırımcı anketinde araştırılan bir diğer konu, dijital ve sosyal medya kaynaklarının kurumsal yatırımcılar bağlamında geleceğine ışık tutmaktadır. Ankette katılımcılara son 12 ayda dijital ve sosyal medya kaynaklarını kullanımlarının ne ölçüde değiştiği ve sonraki 12 ayda, dijital ve sosyal medya kaynaklarını kullanımlarının ne ölçüde değişeceğine dair beklentileri sorulmuştur. Bu anketten elde edilen sonuçlara göre yatırımcıların kullandıkları ve gelecekte de kullanmaya devam edecekleri platformlar söz konusu olduğunda tanıdık olana bağlı kaldıkları görülmektedir. Bulgular, Google arama sonuçları, yatırımcı ilişkileri web siteleri, LinkedIn, Instagram, Facebook ve Twitter gibi platformların, yatırımcılar için en önemli dijital bilgi kaynakları olmaya devam edeceğine işaret etmektedir. Bununla birlikte, yatırımcıların TikTok, e-posta bültenleri ve podcast'ler gibi daha yeni platformlara çok aşina oldukları, gelecek yıl bunların kullanımını artırmayı planladıkları ve görece daha yeni olan platformlar büyümeye devam edeceği sonucuna ulaşılmıştır (Brunswick, 2023).

Araştırmada son olarak kurumsal yatırımcılara, şirketlerle ilgili araştırmalarının bir parçası olarak dijital ve sosyal medya verilerini toplamalarına ve analiz etmelerine yardımcı olarak

yapay zekâ ve makine öğrenimi gibi araçları kullanıp kullanmadıkları da sorulmuş, katılımcıların büyük çoğunluğu (%94) bu araçları veri toplamak için kullandıklarını ifade etmişlerdir. Grafik 2.4, yatırımcıların veri topladıkları kaynakları göstermektedir. Bu sonuçlara göre Google, YouTube, Instagram, Facebook, Twitter ve şirketlerin yatırımcı ilişkileri web sitelerinin kurumsal yatırımcıların başlıca veri kaynakları olduğu söylenebilir.

Grafik 2.4: Kurumsal Yatırımcıların Veri Topladıkları Kaynaklar



Kaynak: Brunswick. (2023)

Bu sonuçlar, kurumsal yatırımcıların hisse senetleri hakkında veri toplamak ve analiz etmek için çeşitli araçları kullanma şekillerinin nasıl değiştiğine ilişkin önemli ipuçları vermektedir. Bir diğer ifadeyle sonuçlar kurumsal yatırımcıların çok sayıda platformdan çok miktarda dijital ve sosyal medya bilgisini alarak yapay zekâ ve makine öğrenmesi yardımıyla verileri

anlamlandırmak için çeşitli platformlara ve araçlara yöneldiğini göstermektedir. Bu durum sosyal medyanın, kurumsal yatırımcıların yatırım süreçlerinde kullandıkları önemli bir bilgi kaynağı olarak geleneksel finans haber medyasına katıldığı anlamına gelmektedir. Ayrıca araştırmada kurumsal ve bireysel yatırımcı davranışında örtüşmenin giderek arttığı da gözlemlenmiştir. Zira geleneksel olarak bireysel yatırımcılarla ilişkilendirilen bu platformlar (Reddit ve bir dereceye kadar Twitter dahil), kurumsal yatırımcılar arasında popülerlik ve önem kazanmaya başlamıştır. Bu durumun merkezi olmayan bilgi paylaşım kanalları üzerinden bireysel yatırımcı davranışını takip etmenin rasyonelitesinden kaynaklandığı düşünülmektedir. Bir diğer ifadeyle kurumsal yatırımcılar, yeni haberlere anında tepki vermek için sosyal medya platformlarını izlemenin yanı sıra, belirli bir alanda artan ilgi veya duyarlılığa dayalı olarak gelecekteki piyasa hareketlerini tahmin amacıyla da izleyebilmektedirler.

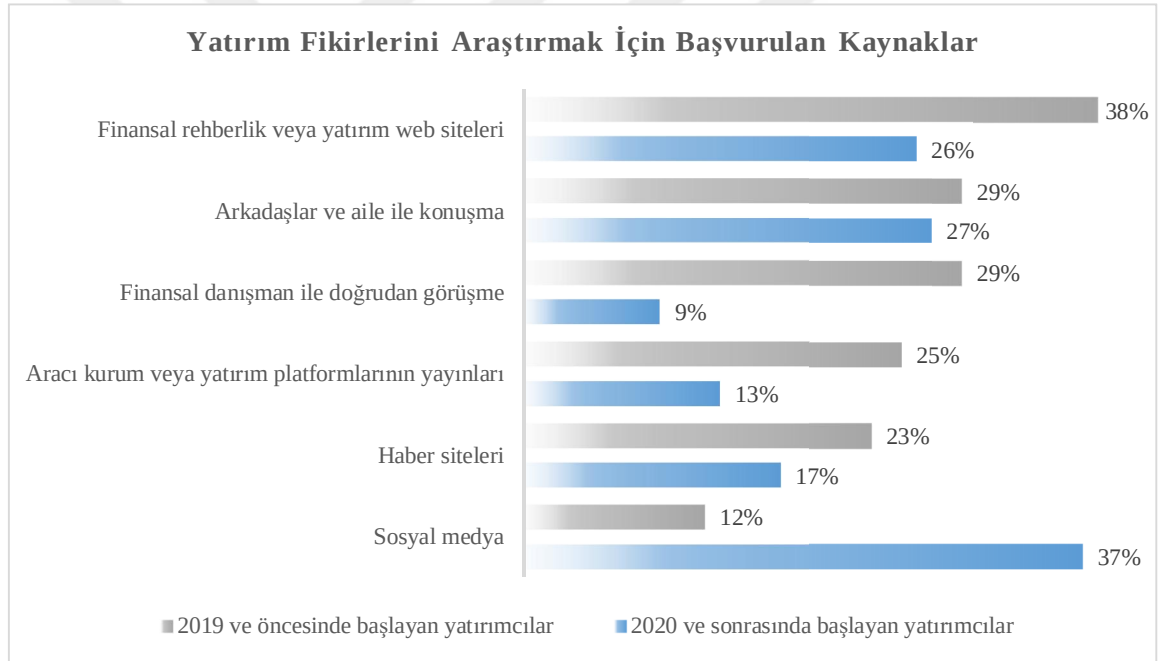
2.2.3 Sosyal Medya ve Bireysel Yatırımcılar

Bireysel yatırımcılar, kendi nam ve hesaplarına menkul kıymet veya fon alıp satan yatırımcılardır. Bir başka tanıma göre ise bireysel yatırımcılar, profesyonel olmayan, menkul kıymet alıp satmak için kendi kaynaklarını kullanan yatırımcılardır. Sosyal medyanın etkisi ve dijital alım satım uygulamaları aracılığıyla borsaya erişilebilirliğin artması, bireysel yatırımcıların borsaya girişini hızlandırmıştır (Arora, 2022). Bir araştırmaya göre bireysel yatırımcılar, 2021'de yönetilen küresel varlıkların %52'sine sahip iken, bunun 2030'a kadar %61'in üzerine çıkması beklenilmektedir (Guild ve Johnson, 2022).

Küresel finansal sistemin içinden geçtiği dijital dönüşüm ve sosyal medya piyasaların işleyişini yeniden şekillendirirken bireysel yatırımcılar da bu dönüşümden derinden etkilenmektedirler. Özellikle 2019 yılında başlayan ve dünya geneline yayılan Covid 19 pandemisi nedeniyle dijital medyanın kullanımında önemli bir artış yaşanmıştır. Bu durum bireysel yatırımcıların yatırım fikirlerini araştırmak için başvurdukları kaynakları temelden etkilemiştir. Grafik 2.5, 2019 yılı ve öncesi ile 2020 yılı ve sonrasında bireysel yatırımcıların yatırım fikirlerini araştırmak için başvurdukları kaynakları göstermektedir. 2019 ve

öncesinde bireysel yatırımcılar için en önemli veri kaynaklarının, yatırım web siteleri, arkadaşlar ve aile, finansal danışman, aracı kurum ve haber siteleri olduğu görülmektedir. 2020 yılından itibaren yatırım web sitelerinin, aracı kurum yayınlarının ve haber sitelerinin görece önemi azalırken sosyal medyanın önemini çarpıcı bir şekilde artırmıştır. Daha açık bir ifadeyle 2019 yılı ve öncesinde yatırım fikirleri ile ilgili bilgiyi sosyal medya platformlarında arayanların oranı %12 iken bu oran 2020 yılı ve sonrasında %37'ye çıkmıştır. Bu durum bireysel yatırımcıların yatırımlarla ilgili bilgi ve fikir aramak için etkileşimli platformları tercih ettiğine işaret etmektedir.

Grafik 2.5: Yatırım Fikirlerini Araştırmak İçin Başvurulan Kaynaklar

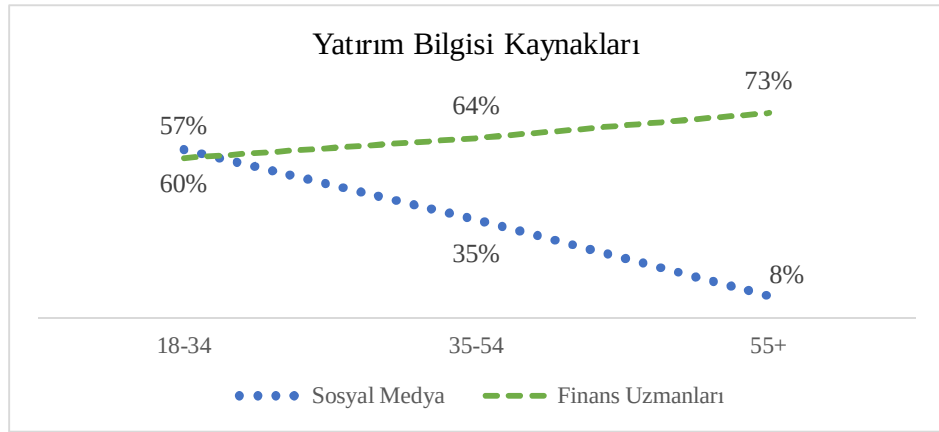


Kaynak: <https://www.cnbc.com/2021/08/26/social-media-top-pick-of-young-investors-for-ideas-cnbc-survey-finds.html>

ABD'de 2.824 yetişkinle Temmuz ve Aralık 2021 tarihleri arasında yapılan, *Amerika Birleşik Devletleri'ndeki Yatırımcılar: Değişen Manzara* (Lin ve diğerleri, 2022) araştırması da, yatırım davranışları ve tutumları açısından eski kuşaklardan çok farklı olan, daha genç ve

daha az deneyimli yeni bir yatırımcı kuşağının bazı ayırt edici özelliklerini ortaya koymaktadır. Grafik 2.6, bireysel yatırımcıların yatırımla ilgili başvurdukları bilgi kaynaklarını yaş gruplarına göre göstermektedir. Buna göre yatırım bilgisi kaynağı olarak daha ileri yaştaki yatırımcıların finans uzmanlarına, daha genç yaştaki yatırımcıların sosyal medyaya başvurma olasılıklarının daha yüksek olduğu görülmektedir. 18-34 yaş aralığındaki yatırımcılar için finans uzmanlarına başvurma olasılığı %57 iken sosyal medyadan bilgi arama olasılığı %60'tır. Bu durum genç yatırımcılar için sosyal medyanın finans uzmanları kadar etkili olabileceğine işaret etmektedir. 35-54 yaş aralığındaki yatırımcıların finans uzmanlarına başvurma olasılığı (%64), sosyal medyaya başvurma olasılığından (%35) yaklaşık iki kat daha fazladır. 55 yaş ve üzeri yatırımcıların yatırım fikirleri ile ilgili bilgi almak için finans uzmanlarına başvurma olasılıkları (%73) sosyal medyaya başvurma olasılıklarından (%8) yaklaşık 10 kat daha fazladır. Bu sonuçlar çoğu Y ve Z Kuşağının, aracılık veya yatırım danışmanlığı firmaları gibi daha geleneksel yatırım bilgileri kaynaklarına güvenmek yerine yatırım hakkında bilgi almak için sosyal medya kaynaklarını kullandıklarını göstermektedir.

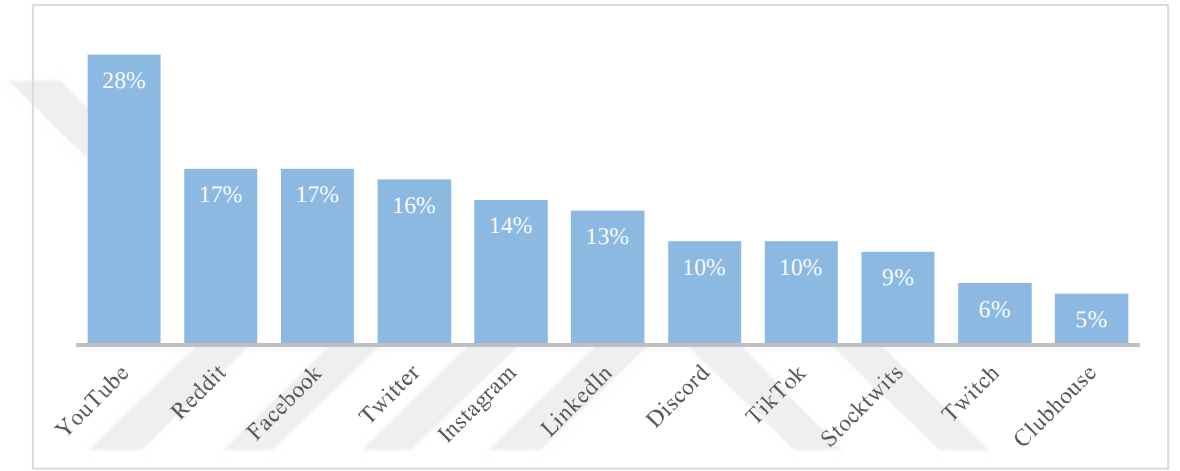
Grafik 2.6: Yatırım Bilgisi Kaynağı



Kaynak: Lin ve diğerleri, 2022.

Bireysel yatırımcıların bilgi kaynağı olarak sosyal medyayı tercih etmeleri, bilgiyi hangi platformlarda aradıklarını önemli hale getirmektedir. Grafik 2.7, yatırımla ilgili bilgi almak için kullanılan sosyal medya platformlarını göstermektedir. Yatırımcıların yatırım bilgileri için kullandıkları sosyal medya platformları arasında YouTube (%28) en popüler olurken, onu Reddit (%17), Facebook (%17) ve Twitter (%16) takip etmektedir.

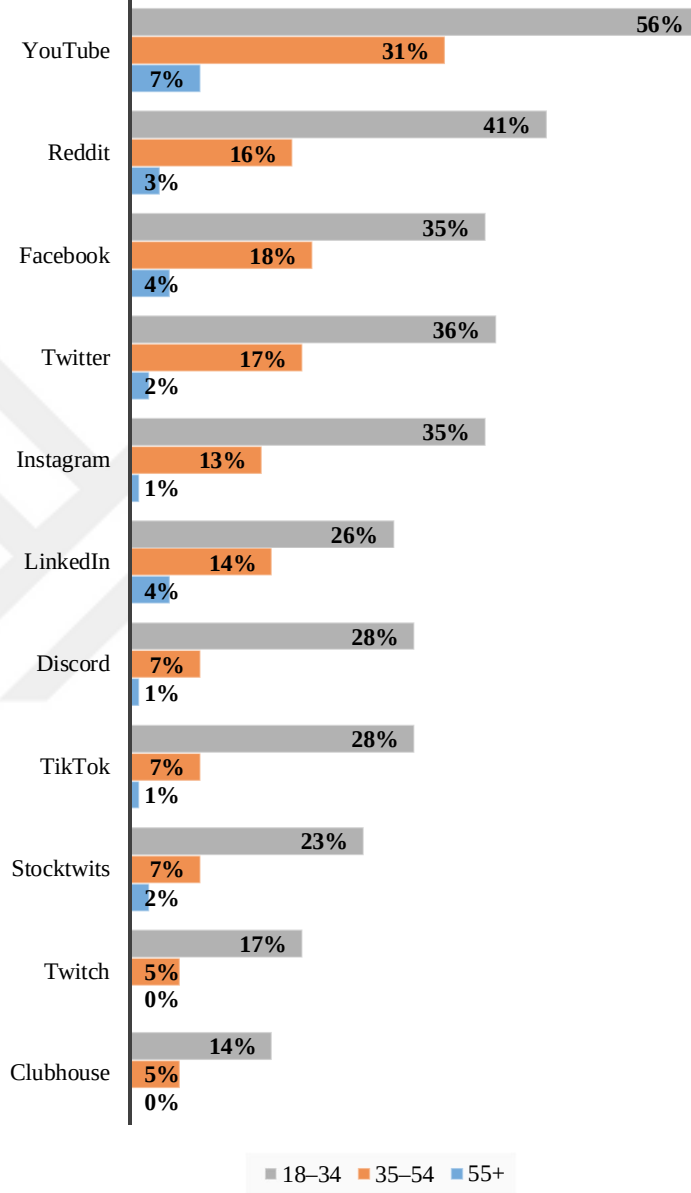
Grafik 2.7: Yatırımla İlgili Bilgi Almak İçin Kullanılan Sosyal Medya Kanalları



Kaynak: Lin ve diğerleri, 2022.

Yatırım bilgisi kaynağı olarak sosyal medya platformlarının tercihi, yaş gruplarına göre önemli ölçüde değişim göstermektedir. Grafik 2.8, yatırımlara ilişkin bilgi ararken başvuru yapılan sosyal medya platformlarını yaş gruplarına göre sınıflandırmaktadır. Özellikle 18-34 ve 35-54 yaş grupları için YouTube, Reddit, Facebook, Twitter ve Instagram en popüler bilgi kaynaklarıdır. 55 yaş ve üzeri yatırımcıların sosyal medyada bilgi arayışları oldukça düşüktür.

Grafik 2.8: Yaş Gruplarına Göre Yatırım Bilgisi Almak İçin Başvurulan Sosyal Medya Platformları



Kaynak: Lin ve diğerleri, 2022.

Bu bağlamda sosyal medya ve yatırımcılar ile ilgili bir diğer konu ise yatırımcıların hangi sosyal medya platformunu hangi amaçla kullandıklarıyla ilgilidir. Zira yatırımcıların ilgi alanlarına göre kullandıkları sosyal medya mecraları farklılaşabilmektedir. Örneğin YouTube ve Facebook daha çok çevrim içi işlemler ve yatırımlar hakkında bilgi edinilmek amacıyla kullanıldığı söylenebilir. Instagram ve TikTok'un daha çok gençler arasında yatırım bilgileri, uygun maliyetli yatırım platformları ve yorumlanması kolay bilgilere erişmek için, Twitter'ın ise özellikle piyasaya özel yorumlardan ve piyasa olaylarından haberdar olmak, fikir ve tavsiye almak için kullanıldığı ifade edilebilir. LinkedIn ise daha çok kurumsal yatırımcılar tarafından kullanılmaktadır (Rawell, 2023).

Bu sonuçlar, çevrimiçi platformlarda yer alan çok miktarda bilgi ve bu bilgiye erişimin çok düşük maliyetle sağlanabilmesinin bireysel yatırımcılar açısından çığır açıcı olduğuna işaret etmektedir. Başka bir ifadeyle sosyal medya, bireysel yatırımcıların yatırım alternatifleri hakkında etkileşimli bir şekilde bilgi toplaması ve yatırımın temellerini öğrenmesi için çeşitli araçlar sunmaktadır. Bu platformlar üzerinden bilgi genellikle finansal influencerlar olarak ifade edilen kullanıcılar tarafından sağlanmaktadır. Bu nedenle sosyal medya platformları kadar bu tür platformlarda bilginin paylaşılmasına aracılık eden etkili kişiler de oldukça önemli bir rol üstlenmektedirler.

2.2.4 Sosyal Medyada Finansal Etkili Kişiler

Günümüz borsalarının bilgi ekosistemi köklü bir değişim geçirmektedir. Hem bireysel işlemlerin hisse senedi fiyat hareketlerindeki etkisi, hem de sosyal medyanın yatırım eğilimlerini şekillendirme ya da yatırım kararlarını yönlendirmedeki rolü giderek artmaktadır. Başka bir ifadeyle yatırımcıların ihtiyaç duyduğu bilgiye geleneksel kanalların yanı sıra geleneksel olmayan platformlar ya da kuruluşlar da etkin bir şekilde aracılık etmektedir. Bütün bu gelişmeler finfluencerların hisse senedi piyasalarında bilgi ve fiyat keşfi üzerindeki etkisinin yadsınamaz boyutlara ulaştığına işaret etmektedir.

Finansal etkili kiři, finansal influencer ya da finfluencer [financial + influencer] terimi, sosyal medya aracılıęıyla yatırımcı kararları üzerinde etkisi olan kiři veya kuruluş anlamında kullanılmaktadır. Bir dięer ifadeyle finfluencerlar, bireysel yatırımcılara borsada işlem yapma, kişisel finans ve yatırım fonları gibi bir dizi finansal konuda bilgi ve tavsiye veren kişilerdir. Popüler veya kültürel statüsü sayesinde, sosyal medyada tanıtımlar veya tavsiyeler yoluyla başkalarının finansal karar verme sürecini etkileme yeteneğine sahiptirler. Genel olarak finfluencerlar sosyal medya hesaplarında, gönderi veya videonun dięer potansiyel alıcılarla paylaşılması için eğlenceli ve/veya bilgilendirici olacak şekilde hazırlanan içerikler yayımlayarak potansiyel alıcıları etkileyebilmektedirler (Nasaa, 2022). Tanınmış kişilerden, YouTube, Twitter, TikTok ve dięer sosyal medya platformlarında takipçileri olan sıradan kullanıcılar ya da yatırımcılara kadar çeşitli finfluencer örnekleri vardır. Bu mecralarda yer alan finfluencerlar ekseriyetle geleneksel finansal analistler değildirler. Hedef kitleleri çoğunlukla bireysel yatırımcılardan oluşmakta ve mesajları genellikle bilgiye erişimin artırılması üzerine odaklanmaktadır.

Genel olarak finfluencerlar sosyal medya aracılıęıyla takipçilerinin bilgiye ulaşmasına aracılık ederler, yatırım tercihlerini ve finansal işlem örüntülerini şekillendirirler ve büyük işlemci grupları arasında onların ve takipçilerinin alım satım motivasyonlarının hisse senedi fiyat hareketlerini etkilemesine izin veren güçlü koordinasyon mekanizmaları sağlarlar (Guan, 2022).

Esasında sosyal medyanın yaygın bir şekilde kullanılmaya başlamasından önce de, popüler bir analist geleneksel kitle iletişim araçları vasıtasıyla bir şirketten bahsettiğinde hisse senedi fiyatının geçici olarak yükselmesine veya düşmesine neden olabildięi çeşitli çalışmalarla ortaya konulmuştur (SEC, 2010; Liu ve dięerleri, 1990). Sosyal medya ise bu anlamda isteyen herkesin güçlü bir finfluencer olabilmesine imkân tanımaktadır. Başka bir deyişle sosyal medya, sıradan kişilerin finfluencer olabilmesinin kapılarını aralamaktadır.

Finfluencerlar genellikle yatırımlarla ilgili *algılanan uzmanlıklarını* sergileyerek etki yaratmaya çalışmakta ve çoğunlukla bilgi yayma iddiasıyla takipçileri ile aralarında bir

güven inşa etmektedirler. Özellikle genç bireysel yatırımcıların eğlenceli eğitim arayışlarına hitap ederek zaman içerisinde geniş bir nüfuza sahip olurlar (Narula, 2022). Bir finansal influencerın takipçilerinin giderek artması, sosyal medya platformunun aynı zamanda bireysel yatırımcıların koordine edildiği bir mekanizmaya dönüşmesine yol açar. Zira bireysel yatırımcıların bilgiye doğrudan erişimi ve düşük maliyetle birbirleriyle iletişim kurabilmeleri yatırım güçlerinin birleştirilebilmesine olanak tanır. Bu çerçevede özellikle sosyal işlem [social trading] teknolojisi *ayna finansal işlem*³⁰ [mirror trading] ve *kopya finansal işlem* [copy trading] gibi tekniklerle bireysel yatırımcıları koordine bir şekilde işlem yapmasını ve fiyatları bir grup olarak etkilemesini giderek daha kolay hale getirmektedir(Guan, 2022).

Bazı durumlarda bu koordinasyon gayri iradi tezahür etmektedir. Herhangi bir kasıt olmadan ortaya çıkan bu koordinasyon bireysel yatırımcıların aynı kaynaklardan bilgi almasından veya bireysel yatırımın sosyal yönlerinin bilişsel eğilimleri ya da sürü psikoloji gibi olgularla artırmaktan kaynaklanabilmektedir (Bikhchandani ve Sharma, 2000). Ayrıca araştırmalar, yatırımcıların sosyal ağlarındaki diğer insanları dinleyerek inançlarını güncellediklerini ortaya koymaktadır. Bu nedenle sosyal ağ yayılmaları, influencerların ve kanaat önderlerinin etkisiyle balonlar, yüksek hacimler ve tersine dönüşler yaratma potansiyeline sahiptir (Pedersen, 2022).

Buna ek olarak bireysel yatırımcılar sosyal platformlarda *yankı odaları*³¹ [echo chambers] oluşturma eğilimindedirler. Zira bulgular yatırımcı sosyal ağlarındaki kullanıcıların doğrulayıcı bilgilere seçici olarak maruz kaldıklarını ortaya koymaktadır. Başka bir ifadeyle kendilerini alıcı [bulls] olarak tanımlayanların, aynı hisse senedinin yükseliş görünümüne sahip bir kullanıcıyı takip etme olasılığı, kendilerini satıcı [bears] olarak tanımlayanlardan

³⁰ Ayna finansal işlem, bireysel yatırımcıların deneyimli ve başarılı yatırımcıların işlemlerini kopyalayarak aynı işlemleri neredeyse gerçek zamanlı olarak kendi hesaplarında gerçekleştirmelerine olanak sağlayan bir stratejidir. Kopya finansal işlemde, yatırımcı başarılı bir yatırımcının hareketlerini doğrudan kopyalar; ayna finansal işlemde ise yatırım kararları, başarılı yatırımcıların işlem modellerinden geliştirilen algoritmalara dayanır.

³¹ Yankı odaları, kişinin sadece kendisinininkine benzer görüş ve inançlarla karşılaştığı ve alternatifleri düşünmek zorunda olmadığı ortamları ifade etmektedir.

beş kat daha fazladır. Yankı odalarında oluşan inançlar, daha düşük getiri, daha fazla bilgi yığını ve daha fazla işlem hacmi ile ilişkilidir (Cookson ve diğerleri, 2023). Bu bağlamda yankı odaları bireysel yatırımcıların doğrulama eğilimini besleyerek gayri iradi koordinasyona yol açabilmektedir.

2.2.5 Sosyal Medya ve Sermaye Piyasası Manipülasyonu

Yatırımcılar, yatırım fikirleri hakkında bilgi almak için etkileşimli iletişime izin veren Twitter, Facebook, YouTube, LinkedIn ve diğer çevrimiçi ağlar dâhil olmak üzere sosyal medyaya giderek daha fazla yönelmektedir. Başka bir ifadeyle sosyal medya, belirli hisse senetleri hakkında araştırma yapılmasına, yatırımın temelleri ve genel bir yatırım stratejisi hakkında rehberliğe, yatırım alternatifleri hakkında etkileşimli bir şekilde bilgi toplamasına, güncel haberleri anlık olarak takip etmeye ve başkalarıyla piyasaları tartışmaya olanak tanıdığı için yatırımcılar açısından oldukça önemli bir araç haline gelmiştir. Bu bağlamda çevrimiçi platformlarda yer alan bol miktarda bilgi ve bu bilgiye erişimin çok düşük maliyetle sağlanabilmesi, teorik olarak bireylerin daha bilgili olmalarına ve yeni fikir ve bilgilere erişerek yatırım fikirlerinin çeşitlenmesine ve dolayısıyla finansal okuryazarlığın artırılmasına katkı sağlamaktadır. Bu durum aynı zamanda sosyal medyadaki söylemlerin kişisel finansal kararları giderek daha fazla etkilediği anlamına da gelmektedir.

Sosyal medyanın yatırımcılar için sağladığı faydalar ve yatırım kararlarına etki potansiyeli, aynı zamanda bu platformları dolandırıcılar için de cazip hale getirmektedir. Zira sosyal medya ve genel olarak internet, suçluların çekici bulabileceği bir dizi özelliklere sahiptir. Örneğin sosyal medya, dolandırıcıların meşru görünen bir site, hesap, e-posta, doğrudan mesaj veya web sayfası oluşturarak nispeten düşük bir maliyetle birçok farklı kişiyle iletişim kurmasına olanak tanımaktadır. Güvenilir piyasa bilgisi kaynağı kimliğine bürünerek oluşturduğu meşruiyet algısı, dolandırıcıların hedeflerindeki ikna potansiyellerini artırmaktadır. Ayrıca, sosyal medyada sahte hesap kullanan dolandırıcılar için anonimlik potansiyeli, hukuki olarak sorumlu tutulma ihtimallerini de azaltmaktadır. Bütün bunlar

sosyal medya aracılığıyla sermaye piyasalarında manipülasyon riskine işaret etmektedir (U.S. Securities and Exchange Commission, 2015).

Manipülasyon ya da sermaye piyasası dolandırıcılığı, belirli bir sonucu elde etmek amacıyla piyasadaki iktisadi karar birimlerinin kararlarını etkilemek olarak tanımlanmaktadır (TDK, 2011). Bir başka tanıma göre ise manipülasyon, piyasada yanlış bilgi yaymak, yanıltıcı davranışlarda bulunmak ve benzeri etik dışı yollarla fiyatları kendi çıkarları doğrultusunda yönlendirmektir (Basık, 2011).

Sosyal medya aracılığıyla sermaye piyasası dolandırıcılığı farklı şekillerde ortaya çıkabilmektedir. Bu bağlamda sermaye piyasası dolandırıcılarının sosyal medyayı istismar etme yollarından biri, hisse senedi fiyatını etkilemek için bir şirket hakkında yanlış ve yanıltıcı bilgiler yayarak piyasa manipülasyonu yapmaktır. Bu yöntemde yanlış veya yanıltıcı söylentiler olumlu veya olumsuz olabilmektedir. Literatürde pompala ve boşalt [pump and dump] olarak bilinen bu yöntemde manipülatörler hisse senedini düşük fiyattan topladıktan sonra satın alma çılgınlığı yaratmak için yanlış ve yanıltıcı beyanlarda bulunarak ya da olumlu söylentiler yayarak hisse senedi fiyatını yükseltmekte ve aldatmaca sona ermeden önce elindeki hisse senetlerini yüksek fiyattan hızla satmaktadır. Bu satış baskısıyla birlikte hisse senedi fiyatları düştüğünden yatırımcılar zarar etmektedirler. Diğer durumda ise piyasa dolandırıcıları hisse senedi fiyatının düşmesi için yatırımcıları hisseleri satmaya teşvik eden olumsuz söylentiler başlatmakta ve ardından yapay olarak düşük fiyattan hisse satın almanın avantajını kullanmaktadırlar (U.S. Securities and Exchange Commission, 2021).

Sosyal medya aracılığıyla yapılan bir diğer sermaye piyasası manipülasyonu güvenilir piyasa bilgisi kaynağı kimliğine bürünerek yapılmaktadır. Bu manipülasyon şeklinde sermaye piyasası dolandırıcıları, sosyal medyada meşru brokerleri, yatırım danışmanlarını veya diğer piyasa bilgisi kaynaklarını taklit etmektedirler. Bir diğer ifadeyle bu yöntemde dolandırıcılar, belirli bir bireyi veya firmayı taklit etmek için tasarlanmış bir hesap adı, profil veya web sayfası oluşturabilmektedirler. Dolandırıcılar ayrıca brokerlerin, yatırım danışmanlarının

veya diğerk piyasa bilgisi kaynaklarının sosyal medya hesaplarına yorum yazarak yatırımcıları sahte bir web sitesine yönlendirebilmektedirler (U.S. Securities and Exchange Commission, 2022).

Sosyal medya aracılığıyla yapılan bir başka dolandırıcılık ise topluluk temelli sermaye piyasası dolandırıcılığıdır. Birçok topluluk, bağlantıda kalmanın ve bilgi paylaşmanın bir aracı olarak sosyal medyayı kullandığından, topluluk temelli sermaye piyasası dolandırıcılığı da sosyal medya üzerinden etnik köken, milliyet, din ve benzeri ortak bağlara sahip grupların üyelerini hedef almaktadır. Bu dolandırıcılar, dolandırmaya çalıştıkları grubun bir üyesiymiş gibi davranarak grup içinde var olan güven ve dostluğu istismar etmektedirler (U.S. Securities and Exchange Commission, 2022).

Sosyal medyada fikir birliği olduğuna dair yanlış izlenimler uyandırmaya dayalı piyasa manipölasyonu da yapılabilmektedir. Bu piyasa dolandırıcılığı şekli, esasında fikir birliği olmadığı halde çok sayıda insanın bir yatırım satın alıyormuş gibi görünmesine neden olarak yatırımcıları yanıltılmaya dayalıdır. Zira daha küçük piyasa kapitalizasyonuna sahip, sosyal medyada, haber platformlarında, yatırım araştırma web sitelerinde, çevrimiçi yatırım haber bültenlerinde, sohbet odalarında ve tartışma forumlarında yoğun bir şekilde tanıtılan veya tartışılan viral popüler [meme stocks] hisse senetleri yatırımcıların dikkatini çekebilmektedir. Bu bağlamda kalabalığın yaptığını düşündüğü şeyi takip etmek yatırımcıya cazip gelebilmektedir (U.S. Securities and Exchange Commission, 2021). Bu yöntemde manipölâtörler ayrıca fırsatı kaçırmadan “hemen al” ya da “hemen sat” baskısı da oluşturabilmektedir (U.S. Securities and Exchange Commission, 2015).

Yapay zekâdaki gelişmelerle ortaya çıkan bir manipölasyon şekli de sosyal medya yatırım fikirleriyle ilgili bilgi arayan kullanıcılarının *bot* olarak bilinen otomatik hesaplar aracılığıyla manipölle edilmesidir. Yapay zekâ tarafından yönetilen, anonimleştirilmiş yapıları nedeniyle kimlikleri tespit edilmeden hareket edebilecekleri binlerce sahte kullan-at hesap ağları olan botlar, sermaye piyasalarıyla ilgili kötü niyetli davranışlar için etkili bir araç olarak hizmet edebilmektedirler (Blackbird.AI, 2022).

Sosyal medya aracılığıyla yapılan diğer bir sermaye piyasası manipölasyonu ise finfluencerlar ile ilgilidir. Esasında sosyal medya ve finfluencerların yatırımcıların karar alma süreçlerinde giderek artan bir etkiye sahip olması, bireysel yatırımcıların finansal okuryazarlığını belirli bir dereceye kadar geliştirmede olumlu bir rol oynarken, aynı zamanda, finansal düzenleyicilerin iki temel odağı olan *yatırımcıların korunması* ve *finansal istikrar* için bir takım riskler oluşturabilmektedirler (Deloitte, 2022). Zira finfluencerlar sosyal medyadaki itibarları ve etkileri nedeniyle fiyat hareketleri üretebilme potansiyeline sahiptirler. Dolayısıyla bir finfluencerın hisse senedi fiyatını etkileyebileceği ve bu değişiklikten kazanç sağlayabileceği gerçeği bir çıkar çatışması yaratmaktadır. Başka bir deyişle bir finfluencer, dolandırıcılık veya manipölasyona girişmesi durumunda pompala ve boşalt şemasını rahatlıkla kullanarak önemli kazanımlar elde etme olanağına sahiptir. Ayrıca, belirli bir hisse senedi hakkında bilgi yaymak için şirketler tarafından finfluencerlara ödeme yapıldığı durumlar da olabilmektedir. Dolayısıyla finansal etkili kişiler bağlamında düşünüldüğünde esasında potansiyel faydaları ve potansiyel zararları ile ilgili gri bir alan olduğu görülmektedir. Bu kapsamda bir düzenleme yapılırken, onların ifade özgürlüklerini ve finansal okuryazarlığa katkılarını olumsuz yönde etkilemeyecek fakat aynı zamanda hem bireysel yatırımcıları piyasa dolandırıcılığına karşı koruyacak hem de piyasa istikrarını sağlayacak şekilde dengeleyici bir duruş benimsemesi uygun olacaktır.

2.2.6 Sosyal Medya ve Sermaye Piyasası Regölasyonları

Sosyal medyada yer alan bilgiler maliyetsiz, kolay erişilebilir, daha eğlenceli ve daha anlaşılabilir olduğundan bireysel yatırımcıların karar alma süreçlerinde giderek artan bir etkiye sahiptirler. Bu durum bir taraftan finansal okuryazarlığın artırılması anlamında önemli bir avantaj sağlarken, diğer taraftan sermaye piyasası dolandırıcılarının iştahını kabartabilmektedir. Bu anlamda sosyal medya platformları üzerinden finans içerikli bilgi paylaşımları düzenleyicilerin de ilgisini çekmektedir. Zira sosyal medya platformları, yayılma hızı ve sahip olduğu etkinin büyüklüğü göz önünde bulundurulduğunda, iletişim ve haberleşme yoluyla işlenen piyasa bozucu eylemlerin aracı haline gelebilme potansiyeline sahiptir.

Bu bağlamda dünya genelinde düzenleyici kurumlar yatırımcıları ve piyasayı korumaya dönük yatırımların yanı sıra, sosyal medya içerik oluşturucularına *yatırım tavsiyesi* olmayan içeriğin nasıl üretileceği konusunda da rehberlik sunmaktır. Örneğin Yeni Zelanda Mali Piyasa Otoritesi (FMA, 2021) Para Hakkında Çevrimiçi Konuşma Rehberi; Avrupa Menkul Kıymetler ve Piyasalar Otoritesi (ESMA, 2021), Sosyal Medyada Yapılan Yatırım Tavsiyelerine İlişkin, yatırım tavsiyelerinin tanımını, sosyal medyada nasıl yayınlanacağını açıklayan ve ihlalin sonuçlarını ortaya koyan bildiri; Avustralya Menkul Kıymetler ve Yatırım Komisyonu (ASIC, 2022), içeriklerinde finansal ürün ve hizmetleri tartışan sosyal medya fenomenleri için ilgili yasalar çerçevesinde vaka incelemeleri ile olası ihlal örneklerinin yer aldığı bir bilgi sayfası (INFO 269) yayınlamıştır.

Bu çerçevede, örneğin Sermaye Piyasası Kurulu (SPK), sermaye piyasalarında piyasa dolandırıcılığını *işlem bazlı piyasa dolandırıcılığı* ve *bilgi bazlı piyasa dolandırıcılığı* olarak iki farklı şekilde ele almaktadır. İşlem bazlı piyasa dolandırıcılığı, sermaye piyasası araçlarının fiyatlarının doğru biçimde oluşmasına engel olan işlemlerle ilgilidir. Bilgi bazlı piyasa dolandırıcılığı ise “sermaye piyasası araçlarının fiyatlarını, değerlerini veya yatırımcıların kararlarını etkilemek amacıyla yalan, yanlış veya yanıltıcı bilgi verilmesi, söylenti çıkarılması, haber verilmesi, yorum yapılması veya rapor hazırlanması ya da bunların yayılmasına” dönük fiiller olarak tanımlanmaktadır³². Bununla birlikte SPK, yönlendirici olmayan ve yorumsuz bilgi paylaşımı ile ilgili bir kısıtlama getirmemiştir.

³² 6362 Sayılı Sermaye Piyasası Kanunu’nun 39’uncu maddesi uyarınca yatırım hizmetlerinin ve faaliyetlerinin düzenli uğraşı, ticari veya mesleki faaliyet olarak icra edilebilmesi için SPK’dan izin alınması zorunlu olup, yatırım hizmetleri ve faaliyetleri (-ki yatırım danışmanlığı, aynı kanunun 38’inci maddesinde yatırım hizmetleri ve faaliyetlerinden bir olarak sayılmaktadır) ancak yatırım kuruluşları tarafından yerine getirilebilmektedir. Sermaye Piyasası Kurulu’ndan izin alınmadan icra edilen sermaye piyasası faaliyetinin yaptırımı, Sermaye Piyasası Kanunu’nun 109’uncu maddesinin ikinci fıkrasında düzenlenmiş olup; anılan maddede “...Sermaye piyasasında izinsiz olarak faaliyette bulunanlar **iki yıldan beş yıla kadar hapis ve beş bin günden on bin güne kadar adli para cezası** ile cezalandırılırlar...” hükmü yer almaktadır (<https://spk.gov.tr/yatirimcilar/izinsiz-sermaye-piyasasi-faaliyetleri>, Erişim Tarihi:18.12.2022). Buna ek olarak 6362 sayılı Sermaye Piyasası Kanunu’nun “Piyasa Dolandırıcılığı” başlıklı 107’nci maddesinde, işlem bazlı piyasa dolandırıcılığı ve bilgi bazlı piyasa dolandırıcılığı suçları düzenlenmiştir. Anılan düzenlemeler uyarınca, “sermaye piyasası araçlarının fiyatlarına, fiyat değişimlerine, arz ve taleplerine ilişkin olarak yanlış veya yanıltıcı izlenim uyandırmak amacıyla alım veya satım yapanlar, emir verenler, emir iptal edenler, emir değiştirenler veya hesap hareketleri gerçekleştirenler” ile “sermaye piyasası araçlarının fiyatlarını, değerlerini veya yatırımcıların kararlarını etkilemek amacıyla yalan, yanlış veya yanıltıcı bilgi veren, söylenti çıkaran,

Diğer taraftan manipölasyondan korunmak için öncelikli olarak yatırımcıların çevrimiçi platformlarda piyasa manipölasyonu potansiyelinin farkında olmaları önemlidir. Ayrıca manipölasyon riski özellikle kısa vadeli yatırımlarda söz konusu olduğundan, yatırımcıların uzun vadeli yatırımlara teşvik edilmesi bu riski azaltacaktır. Buna ek olarak manipölasyonla mücadelede gelişmiş yapay zekâ teknolojilerinden yararlanmak da oldukça önemli bir avantaj sağlayacaktır. Başka bir deyişle yapay zekâ ve makine öğrenmesi sayesinde düzenleyiciler ve finansal kurumlar sosyal medya platformlardaki etkinlikleri izleyerek önemli bir zarar gerçekleşmeden önce hileli ve manipölatif faaliyetleri daha hızlı ve daha kolay bir şekilde tespit edebilme imkânına sahip olabileceklerdir.

haber veren, yorum yapan veya rapor hazırlayan ya da bunları yayan ve bu suretle menfaat sağlayan” kişilerin 3 yıldan 5 yıla kadar hapis ve beş bin güne kadar adli para cezası ile cezalandırılacağı hükme bağlanmıştır (<https://www.spk.gov.tr/duyurular/basin-duyurulari/basin-aciklamasi-04-06-2020>, Erişim Tarihi: 18.12.2022). Bu yasal düzenlemeler nedeniyle, çeşitli sosyal medya mecralarında ve iletişim platformlarında sermaye piyasası araçlarına ilişkin paylaşım yapan hesaplardan (kişiler ya da bot hesaplar) bazıları, paylaşımlarının yatırım danışmanlığı kapsamında olmadığını, bazıları da paylaşımlarının bir eğitim çalışması olduğunu ifade etmektedirler. Bunun için de bu platformlardan elde edilen verilerde, YTD, yani “yatırım tavsiyesi değildir” ibaresiyle sıkça karşılaşmaktadır.

ÜÇÜNCÜ BÖLÜM

LİTERATÜR TARAMASI

3.1 FİNANSAL METİNLERİN YATIRIMCI DUYARLILIĞI VE FİNANSAL VARLIK FİYATLARI ÜZERİNDEKİ ETKİSİ

Finansal metinler bilgi, fikir ya da duygu içeren, genellikle yapılandırılmamış, nitel veri kaynaklarıdır. Bu tür içerikler bir yatırımcının piyasa ya da finansal ürünler hakkındaki anlayışını ve bilgisini güncelleyerek yatırımcı duyarlılığını [investor sentiment] etkilemektedir. Yatırımcı duyarlılığı, genel olarak, yatırımın getirileri ve riskleri ile ilgili inançlar olarak ifade edilmektedir. Finansal metinlerin sayısal olmaması, içerdiği bilginin karar verme süreçlerine dâhil edilebilecek şekilde sayısallaştırılmasını gerektirmektedir. İnsan yeteneklerini aşan bir hızla artan kaynak çeşitliği ve veri hacmi, otomatik veri toplama (web scraping, API vb.) ve sınıflandırma yapmayı mümkün kılan teknolojilerin ve makine öğrenmesi yöntemlerinin kullanılmasını zorunlu kılmaktadır. Ayrıca finansal metin analitiğiyle elde edilen bilgiler, algoritmik finansal işlemlerin stratejilerinin belirlenmesinde ve al-sat kararlarının verilmesinde kilit bileşenlerden biridir. Bu teknolojiler rutin izleme yükünü azaltırken, metinlerin analiz edilerek yatırım kararlarının alınmasına ve hemen eyleme geçilmesine imkân sunmaktadır. Bu yönüyle metin analitiği, finansal ekonomi, finans mühendisliği, davranışsal finans ve yapay zekâdan (özellikle doğal dil işleme) yararlanan çok disiplinli bir alandır. Bu bağlamda hem geleneksel medyanın hem de sosyal medyanın yatırımcı duyarlılığını, dolayısıyla varlık fiyatlarını, varlık fiyatlarındaki oynaklığı ve riski etkilediğini savunan ve hem nicel modellere, hem de yatırım karar sürecine eşzamanlı [synchronous] veya eşzamansız [asynchronous] olarak nasıl etkin bir şekilde dâhil edileceğini sorgulayan ve giderek büyüyen bir araştırma literatürü vardır (Mitra ve Mitra, 2011). Bu bölümde söz konusu literatür finansal haberlerin hisse senetleri fiyatlarıyla ilişkisi

ve sosyal medya paylaşımlarının finansal piyasalarla ilişkisi olmak üzere ayrı ayrı ele alınmaktadır.

3.2 FİNANSAL HABERLER VE HİSSE SENEDİ FİYATLARI İLİŞKİSİNE İLİŞKİN LİTERATÜR

Haberlerin finansal piyasalarda önemli bir rol oynadığı yaygın olarak kabul edilmektedir. Literatürde; finansal haberlerin analiz edilmesinde kullanılan teknikler (Das, 2011), gerçek zamanlı risklerin ve getirilerin yönetiminde haber akışlarının kullanılması (Healy ve Lo, 2011), haberlerin hisse senedi getirileriyle ilişkisi (Leinweber ve Sisk, 2011), firmaya özgü haber akışları ile gün içi hisse senedi endeksi ve vadeli işlem getirilerinin oynaklığı (Kalev ve Duong), piyasa bilgilerini ve duyarlılığının hisse senedi portföyü risk tahmininde kullanılması (Mitra ve diğerleri, 2011) ve haberlerin algoritmik finansal işlem stratejilerine dâhil edilmesi (Brown, 2011) gibi konuların araştırıldığı görülmektedir.

Finansal haberler ile hisse senedi fiyatları arasındaki ilişkiyi inceleyen çalışmalardan biri Gidofalvi ve Elkan'a (2001) aittir. Bu çalışmada araştırmaya konu şirketler hakkında çıkan haberler ile bu şirketlerin hisse senedi fiyatları arasındaki ilişki incelenmiştir. Çalışmada kullanılan sınıflandırıcının öngörü gücü her ne kadar düşük olsa da, haber makaleleri ile hisse senedi fiyat hareketleri arasında, haberin yayımlanmasının 20 dakika öncesinden 20 dakika sonrasına kadarki süreçte güçlü bir korelasyon olduğu bulgusuna ulaşılmıştır. Wuthrich ve diğerlerinin (1998) webde yayımlanan içeriğe dayanarak hisse senedi piyasalarını öngörebilmeyi amaçladıkları bir çalışmada günlük finansal analiz raporları ile hisse senedi, döviz ve tahvil piyasalarına ilişkin değerlendirmeleri içeren metinsel verileri, anahtar sözcükler yardımıyla, kural tabanlı olarak, en yakın komşu ve yapay sinir ağı teknikleri ile analiz etmişlerdir. Sonuçta kural tabanlı tekniklerin uygulaması daha karmaşık olsa da diğerlerine kıyasla daha başarılı olduğu ve bu sistemden elde edilen öngörü sonuçlarının karar destek aracı olarak hizmet edebileceği bulgusuna ulaşmışlardır.

Peramunetilleke ve Wong (2002), TF x IDF ve TF x CDF [Term Frequency x Category Discrimination] yöntemlerine dayanan bir model geliştirerek, para piyasasıyla ilgili haber

başlıklarının gün içi döviz kuru hareketlerini tahmin etmek için nasıl kullanılabileceğini araştırmışlardır. Neticede araştırmacılar metinlerin analizine dayalı modelin, zaman serisi analizleri ve teknik analiz gibi klasik tahmin metotlarına kıyasla daha başarılı sonuçlar elde ettiği bulgusuna ulaşmışlardır. Mittermayer (2004) da, benzer şekilde basın bültenlerinin yayınlanmasından hemen sonra hisse senedi fiyat eğilimlerini tahmin etmek için uygulanan üç aşamalı bir öngörü sistemi oluşturmuştur. Sistemin ilk bileşeni metin ön işleme tekniklerinin uygulanması yoluyla basın bültenlerinden ilgili bilgileri toplarken, ikinci bileşen basın bültenlerini önceden tanımlanmış kategorilere ayırmaktadır. Son aşamada ise bu sınıflandırma esas alınarak yatırım stratejileri üretilmiştir. Metinlerin temsilde TF-IDF [Term Frequency - Inverse Document Frequency] yönteminin, sınıflandırma algoritması olarak da destek vektör makinelerinin [Support Vector Machines] kullanıldığı bu çalışmada bulgular, basın bültenlerinin sınıflandırılmasının, hisse senedi fiyat eğilimlerini tahmin etmek için kullanılabilecek ek bilgiler sağlayabildiğini, ancak sınıflandırma sonuçlarından tam olarak yararlanılması için bir al-sat stratejisinin gerekli olduğunu göstermektedir.

Fung ve diğerlerinin (2005) gerçek zamanlı metinsel haber verileri ile gün içi hisse senedi fiyat hareketleri arasındaki ilişkiyi inceledikleri bir çalışmada, haberlerin etkileri pozitif, negatif ve nötr olarak sınıflandırıldıktan sonra destek vektör makineleri yöntemi ile analiz edilmiş, analiz sonucunda metinsel haberler ile hisse senetleri fiyat hareketleri arasında güçlü bir ilişki olduğu bulgusuna ulaşılmıştır.

Schumaker ve Chen (2006), metinlerin temsilde kullanılan sözcük torbası [bag of words], adlandırılmış varlıklar [named entities] ve ad öbeği [noun phrase] formlarının hisse senedi fiyat öngörü gücünü kıyaslamak amacıyla yaptıkları bir çalışmada, metnin ad öbeği şeklinde temsiline standart sözcük torbası yaklaşımına kıyasla daha iyi bir performans ortaya koyduğunu bulmuşlar, ayrıca doğrusal regresyona kıyasla destek vektör makineleri yönteminin tahmin başarısının daha iyi olduğu sonucuna ulaşılmıştır. Schumaker ve Chen (2009), metinlerin temsilde özel isimlerin [proper name] kullanıldığı bir başka çalışmalarında da benzer sonuçlar ortaya koymuştur.

Groth ve Muntermann (2011) finansal risk yönetiminde metinsel verilerin katkısını araştırdıkları bir çalışmada en yüksek olağandışı [anormal] volatilité seviyeleri ile ilgili kurumsal açıklamaları belirlemeye odaklanan bir metin madenciliđi tekniđi geliřtirmişlerdir. Artan risk baskısını metinsel veriler bağlamında açıklamaya çalışırken, NB [naive bayes], kNN [k-nearest neighbors], NNet [neural network] ve SVM [support vector machines] yöntemlerini kullanmışlar, gerek sınıflandırma sonuçları gerekse hesaplama etkinliđi bağlamında SVM yönteminin daha başarılı olduđu bulgusuna ulaşmışlardır. Sonuç olarak çalışmada metinsel verilerin risk yönetiminde önemli bir bilgi kaynađı olarak kullanılabileceđi kanaatine varmışlardır.

Schumaker ve diđerleri (2012), hisse senedi fiyat hareketlerinin, finansal haber yazarlarının sözcük tercihleri ve kullandıkları dilin tonu ile iliřkili olup olmadıđını, iliřkili ise bu deđişkenler kullanılarak fiyat hareketlerinin tahmin edilip edilemeyeceđini arařtırmak amacıyla yaptıkları bir çalışmada, finansal haberleri pozitif ve negatif olarak sınıflandırıldıktan sonra, destek vektör makineleri yöntemi ile analiz etmişlerdir. Arařtırmanın sonucunda olumsuz haberlerle fiyat hareketlerinin iliřkisinin daha kuvvetli olduđu bulgusuna ulaşmışlardır.

Ming ve diđerleri (2014), Wall Street Journal'da yayımlanan günlük makaleleri esas alarak aynı günün hisse senedi kapanıř fiyatlarının tahmin edilmesine odaklanmışlar ve hisse senedi fiyatları arasındaki, haber makaleleri arasındaki ve hisse senedi fiyatları ile haber makaleleri arasındaki korelasyonları eşanlı olarak göz önünde bulunduran bir model geliřtirmişlerdir. Model, S&P 500, DJIA ve NASDAQ'daki her bir hisse senedi fiyatının yönünü, yukarı veya ařađı olacak řekilde %55 dođrulukla tahmin edebilmektedir. Dolayısıyla Wong vd.'lerinin geliřtirdiđi bu modelin tahmin başarıısı, yaygın olarak kullanılan zaman serileri yöntemlerinin yaklaşık %51.5'lik başarıısına kıyasla daha yüksektir. Ayrıca Wong vd.'leri, bu modele dayalı bir yatırım stratejisinin baz endekslere kıyasla daha yüksek getiri ve Sharpe oranına sahip olduđu bulgusuna ulaşmışlardır.

Shynkevich ve diğerleri (2016), belirli hisse senetleri ile farklı derecelerde alaka düzeyine sahip finansal haber makalelerinin, aynı anda ve uygun şekilde kullanıldığında finansal haberlere dayalı tahminde avantaj sağlayıp sağlayamayacağını araştırdıkları bir çalışmada, hisse senetlerini Küresel Endüstri Sınıflandırma Standardına [Global Industry Classification Standard] göre alt sektörlerle ayırmışlardır. Daha sonra bu hisse senetleri ile ilgili haberler, *hisse senedine özel* [stock-specific], *alt endüstriye özel* [sub-industry-specific], *endüstriye özel* [industry-specific], *grup endüstriye özel* [group-industry-specific] ve *sektöre özel* [sector-specific] haberler olarak kategorilere ayrılmıştır³³. Çalışmanın sonuçları, SVM ve kNN tekniğinin kullanıldığı tek bir haber kategorisinin fiyatları öngörü gücüyle mukayese edildiğinde, haber makalelerini farklı kategorilere ayırma, ayrı ayrı ön işleme koyma, onlardan öğrenme ve tahminlerini tek bir tahmin kararına entegre etme girişiminin, tek bir haber alt kümesine dayalı yaklaşımlara kıyasla tahmin performansını iyileştirdiğini göstermektedir.

Picasso ve diğerleri (2019), veri bilimi ve makine öğrenimi tekniklerinin uygulanması yoluyla hem teknik hem de temel analizi birleştirmeyi amaçladıkları bir çalışmada, teknik analiz göstergeleri ile haber makalelerinin duygusal yönelimini girdi olarak kullanmak suretiyle gelecekteki piyasa eğilimlerini tahmin edebilen bir modelin geliştirilip geliştirilemeyeceğini araştırmışlardır. NASDAQ 100'de listelenen ve aktif büyüklüğü en yüksek [most capitalized] yirmi şirketin hisse senetlerinden oluşan bir portföyün eğilimini tahmin edebilmek için rassal orman [Random Forest, RF], destek vektör makineleri ve ileri beslemeli sinir ağı olmak üzere üç farklı model uygulamışlar, neticede hem istatistiksel hem de finansal açıdan modelin etkili olduğunu kanıtlayarak incelenen portföyde yıllık %85,2'lik bir getiri elde etmişlerdir.

Huang ve Zhao (2021), ABD ve Çin'in önemli haber sitelerinde yayınlanan haber başlıkları ile borsa endekslerinin tahmin edilip edilemeyeceğini araştırdıkları bir çalışmada haberleri TF-IDF yöntemiyle kodlanmışlardır. ABD için haber verileri Reddit World News Channel'ın

³³ Endüstri, çok daha spesifik bir şirket veya işletme grubunu ifade ederken, sektör terimi ekonominin geniş bir bölümünü tanımlamaktadır (Langager, 2022).

08 Haziran 2008 ile 01 Temmuz 2016 tarihleri arasındaki haber başlıklarını kapsamaktadır. ABD hisse senedi verileri için Dow Jones Endüstriyel Ortalaması (DJIA) seçilmiştir. Çin için haber verileri 26 Haziran 2009 ile 08 Kasım 2020 tarihleri arasında yayımlanan günlük haberleri kapsamaktadır. Hisse senetleri için Shanghai Shenzhen CSI 300 Endeksinin günlük kapanış fiyatları kullanılmıştır. Huang ve Zhao bu çalışmada TF-IDF yöntemiyle kodlanmış haber metinlerine dayalı olarak bir sonraki günün hisse senedi fiyat hareketlerini tahmin etmek için çeşitli makine öğrenmesi algoritmalarını test etmişler, iki sınıflı yapay sinir ağı [two-class neural network] modelinin Çin piyasası için, iki sınıflı karar ormanı [two-class decision jungle] modelinin ABD piyasası için en iyi tahmin modeli olduğu bulgusuna ulaşmışlardır.

3.3 SOSYAL MEDYADA FİKİR VE DUYGU ANALİZİ: TWITTER VE HİSSE SENEDİ FİYATLARI İLİŞKİSİNE İLİŞKİN LİTERATÜR

İnsanların duygu ve düşüncelerini sosyal medya platformları aracılığıyla ve bir kısmını metin mesajları şeklinde paylaşmaları, gerek paylaşılan içeriklerin, gerekse sosyal etkileşimlerin hızla artan devasa veri yığınları şeklinde dijital olarak izlenebilmesine imkân tanımaktadır. İnsanların karar verme süreçlerinin bazen diğer insanların bazen de kanâat önderlerinin görüşlerinden etkilendiği göz önünde bulundurulduğunda, bu geniş bilgi yığınlarının sistematik bir şekilde toplanıp analiz edilerek sosyal fenomenlerin açıklanması ve tahmin edilmesine önemli katkılar sunması beklenmektedir. Bu bağlamda fikir ve duygu analizi³⁴

³⁴ *Fikir* ve *duygu* sözcükleri arasındaki anlamsal farklılık, araştırmacılar arasında bu yöntemin kavramlaştırılmasında da farklılığa yol açmış; bazıları fikir madenciliğini [opinion mining] tercih ederken, bazıları duygu analizi [sentiment analysis] kavramını tercih etmiştir. Merriam-webster sözlüğüne bakıldığında “sentiment” [duygu] sözcüğü bir anlamda **duygu kaynaklı görüş, düşünce ve tutum** olarak, diğer bir anlamda ise **duyguların etkilediği fikir** şeklinde tanımlanmıştır. “Opinion” [fikir] sözcüğü **belirli bir konu hakkında zihinde oluşan görüş, yargı veya değerlendirme** olarak tanımlanmıştır. Duygu [emotion] sözcüğü ise **genellikle belirli bir nesne, olay ya da bireye yönelik güçlü ve öznel olarak deneyimlenen ve vücutta fizyolojik ve davranışsal değişikliklerin eşlik ettiği bilinçli zihinsel tepki (öfke veya korku gibi)** olarak tanımlanmaktadır. Dolayısıyla duygu analizi [emotion analysis] denildiğinde, öfke [anger], üzüntü [sadness], korku [fear], şaşkınlık [surprise], umut [anticipation] serin kanlılık [calm], tetikte olma [alert], emin [sure], mutlu [happy], iğrenç [disgust], neşe [joy] gibi duygular kastedilmektedir. Bütün bunlar gözönünde bulundurularak, burada bahsi geçen sözcüklerin anlamlarını kapsayabilmesi açısından, yani hem duyguyu hem de düşünceyi ifade edecek şekilde bu tezde “fikir-duygu” sözcükleri birlikte kullanılmış, “duygu analizi” veya “fikir madenciliği” yerine “fikir ve duygu analizi” kavramı tercih edilerek İngilizce’deki “sentiment” sözcüğünün yerine kullanılmıştır.

[sentiment analysis], yapılandırılmamış metinlerden karmaşık içgörüye geçebilen temsiller, modeller ve algoritmalar sağlayarak insanların duygu ve düşüncelerini sayısallaştırarak açıklığa kavuşturmaya çalışmaktadır. Fikir madenciliği [opinion mining] olarak da adlandırılan duygu analizi, bir karar verici (insan ya da makine) tarafından kullanılmak üzere doğal dildeki metinlerden öznel enformasyonu (fikri ve duygusal yönelimi / anlamı) çıkartarak yapılandırılmış ve eyleme geçirilebilir bilgiler üreten otomatik araçlardır. Başka bir ifadeyle metin temelli fikir ve duygu analizi, insanların belirli bir şey hakkında sahip oldukları fikir, duygu ve düşüncelerini ifade ettikleri yazılı metinlere bir takım teknikler uygulayarak metindeki öznelliği ortaya çıkarma çabaları olarak tanımlanmaktadır (Feldman, 2013).

Sosyal medya platformları, kullanıcıların finansal ürünler hakkında da duygu, düşünce ve beklentilerini paylaşabildikleri bir mecradır. Bu nedenle finansal ürünlerle ilgili sosyal medyada paylaşılan duygu ve düşüncelerin, bireysel yatırımcıların yatırım kararlarını ve dolayısıyla finansal ürünlerin fiyatlarını etkileme olasılığının yüksek olduğu düşünülmektedir. Twitter dünyanın her yerinden milyonlarca kullanıcının olayları neredeyse gerçek zamanlı raporlamasına imkân sunarak bilginin hızlı bir şekilde paylaşılmasını sağlamaktadır. Bu düşünceden hareketle Twitter paylaşımlarından ölçülen duygular ile hisse senedi piyasaları arasındaki ilişkiyi araştıran pek çok çalışma yapılmıştır. Bu çalışmalardan bazıları özellikle Twitter'a odaklanarak, Twitter kullanıcılarından toplanılan verilere kullanıcı düzeyinde, bazıları, hisse senedi düzeyinde, bazıları ise toplulaştırılmış bir şekilde duygu analizi yapmış buradan elde ettiği duygu skorunu da ya borsanın genel durumunu temsil eden endeks verileriyle ya da belirli hisse senetlerinin fiyatlarıyla karşılaştırmıştır. Burada bu çalışmaların bir kısmının özeti sunulmaktadır.

3.3.1 Twitter'da Paylaşılan Mesajların (Tweetlerin) Kullanıldığı Çalışmalar

Twitter kullanıcılarının paylaşımlarının finansal piyasalar ile ilişkisini araştıran pek çok çalışma yapılmıştır. Bu çalışmaların büyük bir kısmı hisse senedi piyasaları ile ilgili iken,

Twitter’da paylaşılan mesajların döviz kurları, değerli metaller, emtia fiyatları ve kripto paralar ile ilişkisini araştıran çalışmalar da mevcuttur.

3.3.1.1 ABD Hisse Senedi Piyasaları İçin Yapılan Çalışmalar

Bollen ve diğerleri (2011), duyguların bireysel davranışlar ve kararlardaki etkisinin toplumlar için de geçerli olup olmadığını araştırdıkları öncü bir çalışmada, ABD'nin önemli hisse senedi endekslerinden biri olan ve en büyük otuz şirketi içeren Dow Jones Borsası Endüstri Endeksi’nin [DJIA-Dow Jones Industrial Average] yönü ile tweetlerden elde edilen kolektif ruh halleri [mood] arasındaki ilişkiyi analiz etmişlerdir. Analizde 28 Şubat 2008’den 19 Aralık 2008’e kadar, yaklaşık 2,7 milyon kullanıcı tarafından paylaşılan 9.853.498 tweet kullanılmıştır. Tweetlerin duygusal yöneliminin sınıflandırılmasında iki farklı duygu endeksi kullanılmışlardır. Bu duygu endekslerinden ilkinde duygular *pozitif* ve *negatif* olarak ikili bir ayrıma tabi tutulurken diğerinde ruh hali durumları *endişesiz/sakin* [calm], *tetikte* [alert], *emin* [sure], *hayati* [vital], *mutlu* [happy], *hoşgörülü* [kind] şeklinde altı gruba ayrılmıştır. Çalışmada ilk olarak çoklu doğrusal regresyon analizi yapmışlar, *emin*, *hayati*, *mutlu* ruh halleri ile borsa verileri arasındaki ilişkinin anlamlı olduğu bulgusuna ulaşmışlardır. Granger nedensellik analizi sonucunda ise, özellikle *sakin* ruh hali değişkeninin Dow Jones Borsası Endüstri Endeksinde 3-4 gün sonra meydana gelen kaymalarla eşleştiğini ortaya koymuşlardır. Ayrıca nedensellik analizi bağlamında *mutlu* ruh halinin de öngörü gücüne sahip olduğu sonucuna ulaşmışlardır. Ancak şaşırtıcı bir biçimde *olumlu* ve *olumsuz* duygularla temsil edilen genel ruh hali değişkeni için benzer bir sonuç gözlemlenememiştir. Çalışmada doğrusal analizlere ilave olarak, doğrusal olmayan etkiyi görebilmek amacıyla özdüzenleyici bulanık yapay sinir ağları [self-organizing fuzzy neural network, (SOFNN)] yöntemine başvurmuşlardır. 28 Şubat 2008’den 28 Kasım 2008’e kadar olan veriler *eğitim* için, 1 Aralık 2008’den 19 Aralık 2008 tarihine kadar olan veriler ise *test* için kullanılmıştır. Bu tarihler belirlenirken herhangi bir olağandışı veya önemli sosyo-kültürel olayın olmadığı zaman dilimlerinin tercih edilmesine dikkat edilmiştir. SOFNN analizi sonucunda, *pozitif* ve *negatif* duygularla temsil edilen genel ruh hali değişkeninin, DJIA endeksinin sadece tarihsel değerleri kullanılarak yapılan tahmin başarısına ilave bir katkı sağlamadığı bulgusuna

ulaşmışlardır. Analiz, diğer altı ruh hali değişkenleriyle tekrar edildiğinde, DJIA'nın yukarı ve aşağı hareketini tahmin etmede *mutlu* ruh halinin %80 doğruluk skoruna, *sakin* ruh halinin ise %87,6 ile en yüksek doğruluk skoruna sahip olduğunu ortaya koymuşlardır. Bollen vd.'leri özetle tweetler aracılığıyla ölçülen *sakin* ve *mutlu* ruh hallerinin DJIA'nın kapanış değerlerinin yönünü tahmin etmedeki başarıyı önemli ölçüde artırdığı sonucuna ulaşmışlardır.

Zhang ve diğerleri (2011), Twitter paylaşımları aracılığıyla ölçtükleri kolektif umut ve korku düzeyleri ile ABD'nin Dow Jones, NASDAQ ve S&P 500 hisse senedi endekslerini tahmin etmeyi amaçladıkları çalışmalarında öncelikle korku ve umudu temsil eden anahtar sözcükleri belirlemişlerdir. Daha sonra 30 Mart 2009'dan 7 Eylül 2009'a kadar altı ay boyunca günlük paylaşılan tweetleri, tweetlerin toplam hacminin (ki bu rakam 2009 yılında günlük 2.5 milyon tweete tekabül etmektedir) yaklaşık %1'ini (günlük 8.100 ila 43.040 tweet arasında değişmektedir) rassal olarak örneklemek suretiyle toplamışlardır. Analizi üç temel düzeyde gerçekleştirmişlerdir. Bu düzeylerden birinde belirli bir gündeki (t günü) olumlu kelimelerin (umut [hope] ve mutluluk [happy]) ve olumsuz kelimelerin (korku [fear], endişe [worry], gergin [nervous], kaygılı [anxious], üzgün [upset]) sayısını o günkü toplam tweet sayısına oranlamak suretiyle her bir duygu durumunu ölçmüşlerdir. Her bir gün için (t günü) elde edilen duygu durumu sonuçlarının Dow Jones, NASDAQ ve S&P 500 endekslerinin ertesi günkü (t+1 günü) değerleri ile ilişkisini ortaya koyabilmek amacıyla korelasyon analizi yapmışlardır. Ayrıca yatırımcı korkusunun dışsal ölçütü olarak, Dow Jones, S&P 500 ve NASDAQ ile güçlü bir şekilde negatif korelasyon gösteren Chicago Board Opsiyon Borsası Volatilite Endeksini (VIX) kullanmışlardır. Analizlerin sonucunda, iyimser ruh hali ile borsa göstergeleri arasındaki ilişkinin pozitif, kötümser ruh hali ile negatif ilişkili olduğu beklentisinin aksine, bütün duygu durumlarının VIX ile pozitif, Dow Jones, NASDAQ ve S&P500 ile negatif korelasyona sahip olduğunu ortaya koymuşlardır. Bu sonucu, olumlu veya olumsuz bir bağlama sahip olup olmalarından bağımsız olarak, insanların ekonomik belirsizlik zamanlarında daha fazla duygusal kelime kullanmaya başladıkları şeklinde yorumlamışlardır. Zhang vd.'leri analizin ikinci düzeyinde, iyimser ya da kötümser bir kullanıcının izleyici kitlesi ne kadar büyükse o kadar çok insanın etkilenebileceği

varsayımına dayanarak potansiyel duygusal izleyici yüzdesi ile borsa göstergeleri arasındaki ilişkiyi analiz etmek amacıyla popülerlik ölçülerinden biri olan takipçi sayısını kullanmışlardır. Analizde, belirli bir gündeki (t günü) duygu tweetlerinin tüm takipçi sayılarını o günkü toplam takipçi sayısına oranlamışlar, buradan elde ettikleri skorların Dow Jones, NASDAQ ve S&P 500 endekslerinin takip eden günkü (t+1 günü) değerleri ile ilişkisini korelasyon analizi yöntemiyle çözümlemişlerdir. Analizin sonucunda korelasyon katsayıları düşük çıktığından takipçi sayısına dayalı duygusal ölçümün borsa endeksleri için iyi bir gösterge olmadığı sonucuna varmışlardır. Zhang vd.'leri analizin üçüncü düzeyinde belirli bir konunun başkaları tarafından ne kadar çok retweetlenirse o kadar belirgin olacağı hipotezine dayanarak duygusal retweet yüzdesi ile borsa göstergeleri arasındaki ilişkiyi analiz etmişler ve retweet sayısının, takipçi sayısına kıyasla daha iyi bir gösterge olduğu sonucuna ulaşmışlardır. Ayrıca daha büyük bir örneklem kümesiyle çalışılması durumunda korelasyonların daha yüksek beklentilerini ifade etmişlerdir. Zhang vd.'leri çalışmalarında son olarak tahminlerin zamansal gecikmesini ortaya koymak amacıyla tweetlerin ardışık üçer günlük ortalamasını alarak basit bir Twitter volatilité endeksi [Twitter volatility index] oluşturmuşlar ve borsa endeksleriyle karşılaştırmışlardır. Analizin sonucunda Twitter volatilité endeksinin Dow Jones, NASDAQ ve S&P500 ile önemli negatif korelasyon ve VIX ile önemli pozitif korelasyon gösterdiği bulgusuna ulaşmışlardır.

Rao ve Srivastava (2012), portföy düzeltmelerinde [portfolio adjustments] yönsele doğruluk tahminlerini en üst düzeye çıkarabilecek bir sistem oluşturmak ve riskten korunma [hedging] kararlarının alınmasında etkili bir uygulama geliştirmek amacıyla toplumsal duygu [sentiment] ile piyasa hareketleri arasındaki ilişkiyi incelemişlerdir. Araştırma problemini, tweet duygu analizi [tweet sentiment analysis], hisse senedi hareket tahmini [stock movement prediction] ve korunma modeli [working hedging model] olmak üzere üç bileşene ayırmışlardır. Araştırmanın ilk aşamasında toplanan tweetleri %82,7 doğruluk oranı ile eğitilmiş olan Naive Bayes yöntemi kullanarak *pozitif* ve *negatif* olarak sınıflandırmışlardır. Sınıflandırma sonucunda veri setindeki tweetlerin %67,14'ü pozitif, %32,86'sı negatiftir. Rao ve Srivastava çalışmalarında piyasa verileri olarak ABD'nin DJIA endeksinde yer alan hisse senetlerinin *açılış*, *kapanış*, *düşük*, *yüksek*, *getiri*, *işlem hacmi* ve *volatilité* niteliklerini

esas almışlardır. Ayrıca günlere göre açılış fiyatlarındaki değişimin yönünü, başka bir deyişle yükseliş ve düşüşü temsilen +1 ve -1 olacak şekilde ikili bir sınıflandırmaya [binary classification] tabi tutulmuş, ardından modelin eğitimi aşamasına geçilmiştir. Tweet tabanlı riskten korunma stratejisine dayalı modelin amacı, belirli bir haftadaki (w haftası) tweetlerin özelliklerine dayalı olarak bir sonraki haftanın (w+1 haftası) piyasa yönünü tahmin etmektir. Model, 14 Kasım 2010'dan 30 Haziran 2011'e kadarki zaman diliminin %76'sına karşılık gelen verilerle eğitilmiştir. Destek vektör makineleri yöntemi ile eğitilen model 8 Mayıs 2011'den 30 Haziran 2011'e kadarki verilerle test edilmiştir. Modelin test verileri için sınıflandırıcı doğruluğu %91 olmuştur. Sonuç olarak araştırmacılar, Twitter'dan gelen işlem sinyalleri kullanılarak öncelikle piyasanın yönünün tahmin edilebileceği, buna bağlı olarak da korunma stratejileri gözden geçirilerek portföyde gerekli değişikliklerin yapılabileceği, böylece yatırımcıların portföylerini olağanüstü piyasa koşullarından minimum risk ile koruyabileceği bulgusuna ulaşmışlardır.

Mao ve diğerleri (2012), ABD'nin S&P 500 endeksinde işlem gören hisse senetlerinden bahseden tweetlerin duygusal içeriğinden ziyade sayısına odaklandıkları bir çalışmada, günlük tweet sayısının endeks, sektör ve hisse senedi düzeylerinde işlem hacmi ve hisse senedi fiyatları ile ilişkisini doğrusal regresyon yöntemi ile araştırmışlardır. Twitter'da S&P 500 endeksinde yer alan hisse senetlerinin kodları (örneğin \$AAPL gibi) kullanılarak 16 Şubat-10 Mayıs 2012 tarih aralığı için arama yapılmış, hisse senetlerinden bahseden tweetler (günlük ortalama 9434) toplanarak veri seti oluşturulmuştur. Araştırmanın sonucunda günlük tweet sayısının S&P 500 endeksinin kapanış fiyatları ile %1 anlamlılık düzeyinde, günlük fiyat değişimi ile %10 anlamlılık düzeyinde ilişkili olduğu bulgusuna ulaşmışlardır. Ayrıca sektör düzeyinde yapılan analizler on sektörden sekizi için günlük tweet sayısı ile günlük işlem hacmi arasında anlamlı bir korelasyon olduğunu göstermektedir. Bu sekiz sektör arasında Finans, Enerji, Sağlık ve Malzeme [materials] sektörleri için anlamlılık düzeyi %1'dir. Hisse düzeyinde yaptıkları analizlerle ise günlük tweet sayısı ile günlük işlem hacminin ve günlük mutlak fiyat değişiminin %1 anlamlılık düzeyinde ilişkili olduğunu ortaya koymuşlardır. Mao vd.'leri çalışmalarında son olarak Twitter verileri ile borsa göstergelerinin tahmin edilip edilemeyeceğini doğrusal regresyon modeli kullanarak analiz

etmişler, hem endeks hem sektör için tahmin doğruluğu %68 olurken, hisse senedi düzeyinde bu oran %52 olmuştur. Sonuç olarak günlük tweet sayısının her üç düzeyde de borsa göstergeleri ile ilişkili olduğunu ve Twitter'ın borsayı tahmin etmede yardımcı olabileceği bulgusuna ulaşılmıştır.

Smailović ve diğerleri (2014), ABD piyasalarında işlem gören sekiz şirket ve bu şirketlerin ürünlerinden bahseden tweetlerdeki duygunun hisse senedi fiyatlarındaki etkisini araştırmak amacıyla 11 Mart 2011 - 9 Aralık 2011 tarihleri arasında 152.572 tweet toplanmışlardır. Bu tweetleri, Stanford Üniversitesi'nin 1.600.000 tweetten oluşan duygu etiketli veri setini kullanarak destek vektör makineleri yöntemiyle eğittikleri duygu sınıflandırıcısı yardımıyla pozitif, negatif ve nötr olarak sınıflandırılmıştır. Çalışmalarında statik durum için uyguladıkları Granger nedensellik testi, hisse senediyle ilgili tweetlerdeki duyguların birkaç gün önceden hisse senedi fiyat hareketlerinin göstergesi olarak kullanılabileceğini ortaya koymuştur. Bunun üzerine Smailović vd.'leri duygu analizine yönelik, sürekli değişen finansal tweet akışlarından artımlı öğrenebilen akış tabanlı aktif öğrenme yaklaşımını geliştirmişlerdir. Analizlerin sonucunda aktif öğrenme yaklaşımının duygu sınıflandırıcının tahmin gücünü daha da iyileştirdiği, tweetlerin pozitif duygu olasılığındaki değişimin hisse senedi kapanış fiyatlarındaki değişimin göstergesi olarak kullanılabileceği bulgusuna ulaşmışlardır. Ayrıca tweetlerdeki duyguların gerçek zamanlı olarak hisse senedi fiyatlarını tahmin edip edemediğini test etmişler, pozitif duygu olasılığı değerlerindeki değişiklikleri dikkate alan bir işlem stratejisinin getirileri iyileştirilebileceğini ortaya koymuşlardır.

Si ve diğerleri (2014), hisse senedi korelasyonlarını ağ tabanlı bir görünümde incelemek amacıyla yaptıkları çalışmada, ABD'de bulunan S&P 100 endeksinde yer alan şirketlerle ilgili 2 Kasım 2012 - 3 Nisan 2013 tarih aralığında firma kodları yardımıyla 629.977 tweet toplamışlardır. Hisse senetlerinin tarihsel fiyat serilerinin ikili korelasyonunu hesaplayıp hisse senetlerini korelasyon güçlerine göre bağlayan bir ağ oluşturmak yerine, hisse senedi kodlarının tweetlerde birlikte bulunma sıklığına dayalı bir korelasyon ağı oluşturmuşlardır. Ayrıca her bir düğüm [node] ve her bir kenar [edge] için Labeled LDA yöntemiyle duygu skoru hesaplamışlar, buradan elde edilen zaman serisi ile hisse senedi fiyat serilerini VAR

yöntemiyle analiz etmişlerdir. Araştırmanın sonucunda bir hisse senedinin yakın komşularından gelen sosyal duygunun [social sentiment], hisse senedi piyasasının tahminini önemli ölçüde iyileştirmeye yardımcı olabileceği bulgusuna ulaşmışlardır.

Sprenger ve diğerleri (2014), tweetlerin hacmi, içerdiği mesajın duygusal yönü ve tweetler arasındaki uzlaşma ya da anlaşmazlık düzeyinin getiri, işlem hacmi ve volatilité üzerindeki etkisini araştırmak amacıyla yaptıkları çalışmada bilginin etkin bir şekilde toplanmasının arkasında yatan mekanizma ile yatırım tavsiyelerinin kalitesi ve bahsedilme düzeyi, retweet oranı ve tweet paylaşımcısının takipçileri arasındaki ilişkilerini de incelemişlerdir. 1 Ocak 2010 ile 30 Haziran 2010 tarihleri arasındaki 6 aylık döneme odaklandıkları çalışmada S&P 100 Endeksine yer alan şirketlerin hisse senedi kodlarını içeren yaklaşık 250.000 tweet toplamışlardır. Tweetleri “al”, “sat” ve “tut” sinyalleri olarak kategorize edebilmek için sınıflandırıcıyı mesajların %49,6’sının “tut”, %35,2’sinin “al” ve %15,2’sinin “sat” sinyali olan 2.500 tweetten oluşan bir eğitim seti ile Naive Bayes yöntemini kullanarak eğitmişler, %81,2 sınıflandırma doğruluğu elde etmişlerdir. Analizin sonucunda tweet duygusu ile hisse senedi getirileri, tweet hacmi ve işlem hacminin yanı sıra anlaşmazlık ve volatilité arasında bir ilişki olduğu sonucuna ulaşmışlardır. Ayrıca, ortalamanın üzerinde yatırım tavsiyesi veren kullanıcıların daha sık retweetlendiğini ve daha fazla takipçiye sahip olduğunu, bu da onların seslerini daha fazla duyurduğunu göstermektedir.

Ruan ve diğerleri (2015), Twitter kullanıcıları arasındaki etkileşimlere göre tanımladıkları güven ağı ile borsayı tahmin etmeyi amaçladıkları çalışmalarında, borsa ile ilgili oldukları varsayımıyla FinancialTimes adlı halka açık Twitter hesabının takipçilerinin 11 Ekim 2013 tarihinden önce atılan tüm tweetlerini toplamışlardır. 11 Ekim 2013 itibarıyla bu gruptaki kullanıcı sayısı 1.576.889, bu gruptan toplanan tweet sayısı yaklaşık 17.000.000’dur. Kullanıcılar arasındaki etkileşime dayalı güven ağını oluşturmak için etkileşimli tweetlere, yani mentionlara (Twitter’da “@” sembolü ile gösterilen) başvurmuşlardır. Kullanıcıların hedefe yönelik attığı tweetleri analiz ederek hedef hakkındaki tutumunu veya düşüncesini öğrenmek için bu tweetleri, bir duygusal sınıflandırma aracı olan Viralheat yardımıyla olumlu ve olumsuz olarak sınıflandırmışlardır. Ruan vd.’lerine göre bir kullanıcının diğer

birçok kullanıcı tarafından takip ediliyor ve güveniliyor olması, bu kullanıcının büyük bir etki gücüne sahip olduğuna işaret etmektedir. Dolayısıyla çalışmada kullanıcıların sahip oldukları etki gücünü de bir önceki adımda hesaplanan güven değerini esas alarak hesaplamışlardır. Çalışmanın son aşamasında ABD borsalarından DJIA'nın 1 Ocak 2013'ten 11 Ekim 2013'e kadarki fiyat değişimleri ve işlem hacmiyle tweetler arasındaki ilişkiyi incelemişlerdir. Pearson korelasyon katsayısı yöntemini kullanarak yaptıkları analizlerin sonucunda kullanıcıların güven bilgisine dayalı etki gücünü hesaba katmayan tweetler ile kapanış fiyatları ve işlem hacimleri arasında ilişki olmadığı bulgusuna ulaşmışlardır. Analizi güven bilgisine göre ağırlıklandırılmış tweetler ile tekrar ettiklerinde işlem hacmi ile ağırlıklandırılmış tweetler arasında 0,25 korelasyon olduğunu, ancak fiyat değişimleri ile anlamlı bir korelasyona sahip olmadığını ortaya koymuşlardır. Analizi, ağırlıklandırılmış tweetleri “#” ve “\$” sembolleri yardımıyla filtreleyerek sadece borsa ile ilgili olanlarla tekrarladıklarında, işlem hacmi ile ağırlıklandırılmış tweetler arasındaki korelasyonun 0,31'e yükseldiği sonucuna ulaşmışlardır.

Ranco ve diğerleri (2015) finansal tweetlerdeki duygusal yönelimler ile hisse senedi getirileri arasındaki ilişkiyi incelemek amacıyla yaptıkları çalışmada ABD'de faaliyet gösteren DJIA endeksinde yer alan firmalarla ilgili 1 Haziran 2013 - 18 Eylül 2014 tarihleri arasında 15 aylık bir zaman dilimi için yaklaşık 1,5 milyon tweet toplamışlardır. Duygu sınıflandırıcısını destek vektör makineleri yöntemiyle 103.262 etiketli tweet kullanılarak eğitmişler, tweetleri bu sınıflandırıcı ile pozitif, negatif ve nötr olarak sınıflandırmışlardır. Pozitif tweetlerin negatif tweetlerden farkının toplam pozitif ve negatif tweetlere oranı şeklinde hesaplanan duygu kutupluluğu [sentiment polarity] skoru ile getiri arasındaki ilişki Pearson korelasyon testi ve Granger nedensellik testi ile analiz etmişler, kutupluluk değişkeninin getiriye tahmin etmek için kullanışlı olmadığı bulgusuna ulaşmışlardır. Kutupluluk değişkeni yerine tweet sayısı ile getiri arasındaki ilişkiyi incelediklerinde şirketlerin üçte biri için tweet sayısının getiri için Granger nedeni olduğunu ortaya koymuşlardır. Bu sonuçları tatmin edici bulmadıkları için Ranco vd.'leri olay incelemesi [event study] tekniğine başvurmuşlardır. Bu teknikte Twitter kullanıcılarının artan aktivitesi ile işaretlenen olayları (kazanç duyuruları gibi) tespit etmişler ve olayları takip eden günlerde piyasa davranışını gözlemlemişlerdir. Analizin

sonucunda pozitif duygu olaylarından sonraki 10 gün boyunca birikimli normalüstü getiri [cumulative abnormal return] değerlerinin %1 düzeyinde anlamlı olduğu, bu durumun olumsuz duygu olayları için de geçerli olduğu bulgusuna ulaşmışlardır.

Nasseria ve diğerleri (2015), StockTwits'ten ölçülen duygulardaki değişikliğin borsada al, sat, tut kararlarını nasıl etkilediğini araştırmak amacıyla yaptıkları çalışmada, 3 Nisan 2012–5 Nisan 2013 dönemi için bir yıllık tweet toplamışlardır. Dow Jones [DJIA] Endeksindeki 30 hisse senedinin kodlarını içeren 289.443 tweete ulaşmışlar, bu tweetlerden 2.892'sini rassal örnekleme yöntemiyle seçerek al, tut veya sat şeklinde elle etiketlemişlerdir. Duygu sınıflandırıcısını elle etiketlenmiş bu veri setini kullanarak karar ağacı yöntemi ile eğitmişler ve %60 doğruluk oranına ulaşmışlardır. Ayrıca karar ağacı modeli, özellik seçimi ile belirlenen en alakalı terimlerin nasıl etkileşime girdiğine bağlı olarak belirli bir terimin veya terim kombinasyonlarının bir sermaye piyasasında bir yatırımcının kararını tahmin etme gücüne sahip olup olmadığını belirlemek için kullanmışlar ve önceden belirlenmiş bir yatırım hipotezine dayalı işlem stratejisi oluşturmuşlardır. Analizin sonucunda terimlerin hacimlerindeki değişiklikler ile piyasa hareketleri arasında bir ilişki olduğu bulgusuna ulaşmışlardır. Pozitif duygusal terimlerin, doğası gereği yatırımcıların finansal piyasalarda işlem gören belirli hisse senetlerine yönelik iyimserliğini yansıtan “satın alma” kararıyla; negatif duygusal terimlerin yatırımcıların söz konusu hisse senedi hakkında kötümserliğini gösteren “sat” kararıyla daha fazla ilişkili olduğunu bulmuşlardır. Ayrıca yatırımcılar tarafından bir “tut” kararı verilecekse, belirli bir hisse senedini tartışan bir tweet mesajında “nötr” terimlerin bulunma olasılığının daha yüksek olduğu sonucuna ulaşmışlardır. Buna göre negatif kelimelerdeki görünürlüğün azalmasının piyasa yükselmeden önce iyi bir sinyal anlamına geldiğini; pozitif kelimelerdeki görünürlüğün artmasının uzun pozisyon alma eğilimleri ile sonuçlandığını OLS [Ordinary Least Square] yöntemiyle ortaya koymuşlardır.

Yang ve diğerleri (2015) sosyal medyada tüm kullanıcıların eşit derecede etkili olmadığını varsayımına dayanarak daha etkili Twitter kullanıcılarının ruh hali ile piyasa fiyat hareketleri arasında bağlantı olup olmadığını araştırmak amacıyla yaptıkları çalışmada aynı zamanda finansal toplulukların Twitter'daki varlıklarını da incelemişlerdir. İlk olarak finans haber

sağlayıcılarından ve Twitter'da finans piyasasıyla ilgili en etkili kullanıcılardan seçilmiş elli Twitter hesabı belirlenmiş, sonra çekirdek hesapları takip edenler ve çekirdek hesapları takip edenleri takip edenler şeklinde elli çekirdek hesabın takipçilerinden oluşan iki adımda 154.327 Twitter hesabı içeren finans topluluğunu oluşturmuşlardır. İkinci adımda finans topluluğunun sosyal ağ analizi ile takipçi sayısı, arasındalık merkeziliği ve yakınlık merkeziliği hesaplanarak sosyal ağın karakteristiğini belirlemişlerdir. Üçüncü adımda ABD borsalarında (NASDAQ, NYSE, AMEX) listelenen şirketlerin borsa kodları ile tweetleri filtreledikten sonra SentiWordNet yardımıyla kalan 1.606.104 tweetin duygu skorlarını -1 ile +1 arasında değişen bir aralıkta hesaplamışlardır. Dördüncü adımda sosyal ağ teorisine göre belirlenen kritik kullanıcıların paylaştıkları tweetlerin duygusal durumu ağ merkeziliği ölçütleri ile ağırlıklandırılarak 15 Şubat 2014 - 15 Haziran 2014 tarihleri arasındaki hisse senedi getirileri ile ilişkisini regresyon yöntemiyle analiz etmişlerdir. Analizin sonucunda Twitter evrenindeki finans toplulukları içerisinde yer alan etkili Twitter kullanıcılarının toplumsal duygu [social sentiment] ve finansal piyasa hareketi arasındaki ilişkiyi temsil edebildiği bulgusuna ulaşmışlardır. Başka bir deyişle finansal topluluk içindeki bu kritik düşümlerden oluşturulan ağırlıklı duygunun, genel duygudan daha güçlü bir finansal piyasa öngörüsü sağladığı sonucuna varmışlardır.

Cazzoli ve diğerleri (2016), şirketlerle ilgili tweetleri kullanarak ABD'deki piyasalarda işlem gören hisse senedi fiyatlarının tahmin edilemeyeceğini araştırdıkları çalışmalarında, duygusal bir sınıflandırmadan ziyade tweet sayıları ile işlem hacimleri arasındaki ilişkiyi analiz etmişlerdir. Araştırmada ABD piyasalarında işlem gören şirketlerin hisse senedi kodlarını içeren, 11 Mart 2011 tarihinden 11 Haziran 2013 tarihine kadarki tweetlere odaklanmışlardır. Analizde kullandıkları veri seti yaklaşık 6 milyon tweetten ve 1.723 şirketin hisse senedi fiyat bilgisinden oluşmaktadır. Analiz sonucunda tweetler ile hisse senedi işlem hacimleri arasındaki korelasyonun 0,29 olduğu, holding ve teknoloji sektörünün diğer sektörlerle kıyasla tweetler ile daha güçlü bir korelasyona sahip olduğu ve son olarak takipçi sayısı arttığında kullanıcıların piyasalara etkisinin çok daha fazla olduğu bulgusuna ulaşmışlardır.

He ve diğerleri (2016), ABD'nin en büyük yedi finansal hizmet şirketinin hisse senedi fiyatları ile tweet sayısı ve tweetlerdeki duygular arasındaki ilişkiyi firma düzeyinde analiz etmek amacıyla yaptıkları çalışmada, 2013 yılı Ekim ayında paylaşılan 37.657 tweet toplamışlardır. Her bir tweetin pozitif ya da negatif olarak duygu yönelimini tespit etmek için Lexalytics adlı bir duygu analiz aracı kullanmışlardır. Analizde ilk olarak, tek değişkenli doğrusal regresyon [univariate linear regressions] yöntemiyle tweet sayısının hisse senedi fiyatları üzerinde istatistiksel olarak anlamlı bir etkisinin olduğunu, duygular ve hisse senedi fiyatları arasında eşzamanlı bir ilişki olmadığını ortaya koymuşlardır. Daha sonra çok değişkenli regresyon yöntemiyle, tweetlerde yer alan yedi finansal hizmet şirketinin eşzamanlı duygularının veya gecikmeli duygularının hisse senedi fiyatlarını etkileyip etkilemediğini test etmişlerdir. Twitter'dan gelen bir günlük negatif duygunun, hisse senedi fiyatı üzerinde hem istatistiksel hem de ekonomik olarak önemli bir olumsuz etkiye sahip olduğu, hem eşzamanlı hem de gecikmeli pozitif duyguların ise hisse senedi fiyatları üzerinde önemli bir etkisinin olmadığı bulgusuna ulaşmışlardır.

Li ve diğerleri (2016), çeşitli duygu türleri ile ilgili sözcüklerin borsadaki genel eğilim üzerine etkisini araştırmak amacıyla yaptıkları çalışmada, 9 Kasım 2015'ten 20 Kasım 2015'e kadar iki haftalık zaman dilimi için tweetleri toplamışlar ve toplanan tweetlerdeki altı duygu modunu temsil eden sözcüklerin frekanslarını belirlemişlerdir. Daha sonra duygu sözcüklerinin sıklığı ile NASDAQ Composite arasındaki ilişkiyi analiz etmişler, sonuçların hisse senedi piyasasını tahmin edecek kadar güçlü olmadığı bulgusuna ulaşmışlardır. Bununla birlikte mutlu, üzgün ve öfkeli gibi ruh hallerini temsil eden sözcüklerin nispeten güçlü bir korelasyona sahip olduğunu göstermişlerdir. Özellikle üzgün ruh halini temsil eden sözcüklerin, borsa üzerinde diğer sınıflara göre önemli ölçüde daha fazla etkiye sahip olduğu sonucunu ortaya koymuşlardır.

Sun ve diğerleri (2016), hisse senedi fiyatlarını tahmin etmek için metin analizine dayalı bir model ve bu modeli kullanan bir işlem stratejisi ortaya koyabilmek amacıyla, StockTwits platformundan 1 Ocak 2011'den 31 Ağustos 2015'e kadarki dönemi kapsayan yaklaşık 45 milyon paylaşımın S&P 500 şirketlerinin hisse senedi fiyatları ile ilişkisini araştırmışlardır.

Analizin ilk aşamasında her yıl en çok konuşulan sözcükler incelenip hisse senetlerinin kodları ile birleştirilerek terimler sözlüğü oluşturmuşlar, terim ağırlıklandırma şemaları yardımıyla ağırlıklandırılmıştır. Ağırlıklandırılmış sözcük sıklığı, günlük hisse senedi kapanış fiyatları ve getiri değişkenleri ile seyrek matris çarpanlarına ayırma [sparse matrix factorization, SMF] tekniğini kullanarak yapılan analizin sonucunda modelinin yön doğruluğunu tüm yıllarda temel modellerden yüzde bir puan kadar daha iyi tahmin ettiğini bulmuşlardır. Modelin temel regresyondan daha iyi performans gösterdiğini, cazip bir yıllık getiri ve Sharpe oranı sağlayan bir işlem stratejisi sunduğunu ortaya koymuşlardır.

Zhao ve diğerleri (2016), tweetler ile getirinin tersine çevrilmesi³⁵ [return reversal] arasındaki ilişkiyi analiz amacıyla yaptıkları çalışmada 11 Haziran 2009 - 31 Aralık 2009 tarihleri arasında toplanan tweetleri OpinionFinder'ın 4.912 olumsuz sözcük ve 2.717 olumlu sözcük içeren sözlüğüne dayanarak, sözlük tabanlı yöntemle sınıflandırmışlardır. ABD borsasının günlük getirinin tersine çevrilmesi arasındaki korelasyonları DJIA üzerinde inceleyen çalışmada duygu faktörü olarak olumlu ve olumsuz duygular, dikkat faktörü olarak ise olarak DJIA için günlük tweet oranları analize dâhil etmişlerdir. İlk aşamada tweet faktörleri seti arasındaki dinamik karşılıklı ilişkileri ve getirinin tersine çevrilmesiyle ilgili ekonomik faktörleri tasvir etmek için dinamik Bayes faktörü grafiğini [Bayesian factor graph] kullanmışlardır. Ardından kayan bir zaman penceresinde getirinin tersine çevrilmesini tahmin etmek için destek vektör makinesine dayalı iki model oluşturmuşlar, çalışmanın sonucunda tweet duygu faktörlerinin getirinin tersine çevrilmesiyle yakından ilişkili olduğunu ortaya koymuşlardır. Spesifik olarak, tweetlerdeki düşük seviyeli pozitif duygunun veya yüksek seviyeli negatif duygunun, daha büyük olasılıkla getirinin tersine çevrilmesine

³⁵ Getirinin tersine çevrilmesi [return reversal], yatırımcıların aşırı alınan hisse senetlerini satmaları veya aşırı satılan hisse senetlerini satın alarak hisse senedi fiyatlarını tersine çevirmeleri durumunda ortaya çıkmaktadır. Bu durum genellikle aşırı tepki hipotezi, likidite etkisi ve ocak ayı etkisiyle açıklanır. Aşırı tepki hipotezi yatırımcıların genel olarak son ekonomik durumlara aşırı tepki verdiğini, hisse senedi fiyatlarının kısa vadede aşırı derecede hareket ettiğini ve uzun vadede normal seviyelere geri döndüğünü belirtmektedir. Likidite etkisi, getirinin tersine çevrilmesini, aciliyet gerektiren ve hisse senedi fiyatlarını temellerden sapmaya iten bilgi dışında likidite işlemleriyle ilişkilendirmektedir. Ocak ayı etkisi ise yatırımcıların aralık ayında kötü performans gösteren hisse senetlerini yoğun bir şekilde satma ve Ocak ayında iyi performans gösteren hisse senetlerini satma eğiliminde olduğu görüşünü savunmaktadır. Getirinin tersine çevrilmesinden yararlanmak için geçmiş getirilerin sıralamasına göre seçici olarak düşük getirili hisse senetlerini satın almak ve yüksek getirili hisse senetlerini satmak tavsiye edilmektedir.

yol açtığı ve bu duygu faktörlerinin, getirinin tersine çevrilmesi tahminini önemli ölçüde iyileştirme yeteneğine sahip olduğu bulgusuna ulaşmışlardır.

Pagolu ve diğerleri (2016), kolektif duygular ile bir şirketin borsa hareketleri arasındaki ilişkiyi analiz etmek amacıyla yaptıkları 31 Ağustos 2015 - 25 Ağustos 2016 aralığını kapsayan çalışmada, 3.216 tweeti elle etiketledikten sonra n-gram ve word2vec yöntemleriyle eğitmişler %70 başarı oranıyla duyguları pozitif, negatif ve nötr olarak sınıflandırmıştır. Daha sonra otomatik olarak sınıflandırılan tweet duygularının hisse senetleri fiyatları ile ilişkisini araştırmak için lojistik regresyon ve LibSVM yöntemleriyle analiz etmişler, lojistik regresyon için yaklaşık %69, LibSVM için yaklaşık %72 başarı oranına ulaşmıştır. Çalışmanın sonunda, hisse senedi fiyatlarındaki artış ve düşüşler ile tweetlerdeki kolektif duygular arasında güçlü bir ilişki olduğunu göstermişlerdir.

Wei ve diğerleri (2016) Twitter hacim artışları ile hisse senedi opsiyonlarının fiyatlandırması arasındaki ilişkiyi araştırmak amacıyla yaptıkları çalışmada 1 Ağustos 2013 - 6 Ağustos 2014 tarihleri arasında bir yıl boyunca S&P 500 hisseleriyle ilgili tweetleri cashtagler yardımıyla toplamışlardır. Twitter hacim artışını, hisse senediyle ilgili tweet sayısının normalden önemli ölçüde fazla olduğu durum olarak tanımlanmışlar, Twitter hacim artışlarını belirlerken biri tweet sayısı, biri tweet gönderen benzersiz kullanıcı sayısı ve üçüncüsü tweet gönderen kullanıcı çeşitliliği olmak üzere üç koşul kullanmışlardır. Analiz kapsamındaki 500 hisse senedinin tamamında en az bir Twitter hacim artışı ve toplamda ise 3.288 Twitter hacim artışı tespit etmişlerdir. Daha sonra hisse senedi opsiyonları fiyatlandırması için yaygın olarak kullanılan Black – Scholes modelinin temel varsayımlarının, yani hisse senedi fiyatının geometrik bir Brownian hareketi izlediği ve bu nedenle hisse senedi getirisinin lognormal bir dağılım izlediği varsayımının Twitter hacim artışlarına sahip hisse senetleri için ne zaman geçerli olduğunu araştırmışlar, varsayımın bir Twitter hacmi artışından önceki zaman diliminde geçerli olma olasılığının düşük olduğu ve sonrasında geçerli olma olasılığının daha yüksek olduğu bulgusuna ulaşmışlardır. Ayrıca, bir hisse senedinin volatilitésinin, bir Twitter hacmi artışından sonra, artıştan öncekine göre önemli ölçüde daha düşük olduğu, zımnî volatilitenin bir Twitter hacmi artışından önce keskin bir şekilde arttığı ve sonrasında hızla

azaldığı sonucuna ulaşmışlardır. Buna ek olarak, satım opsiyonlarının, alım opsiyonlarından daha yüksek fiyatlı olma eğilimine sahip oldukları ve Twitter hacmindeki artışın hemen ardından opsiyonların aşırı pahalı olabileceği bulgusuna varmışlardır. Araştırmanın sonunda bu bulgulara dayanarak, hisse senedi opsiyon işlemleri için bir satış yayılımı stratejisi önermişler, bir yıllık borsa verilerini kullanan bir portföyün gerçekçi simülasyonunu yaparak, ihtiyatlı bir ortamda bile, bu stratejinin komisyonlar ve satış-teklif farkı [ask-bid spread] dikkate alındığında %34,3'lük bir kazanç elde ettiğini, S&P 500'ün ise aynı dönemde yalnızca %12,8 arttığını göstermişlerdir.

Domeniconi ve diğerleri (2017), DJIA'nın artmasına ve azalmasına yol açan ilgili tweetleri tespit etmek ve özelliklerini belirlemek amacıyla yaptıkları çalışmada basit madencilik teknikleri ve metin benzerliği ölçümlerine dayalı iki aşamalı bir model geliştirmişlerdir. Birinci aşamada vektör uzay modeli [vector space model] temsiline dayalı gözetimli bir sınıflandırıcı [supervised classifier], ikinci aşamada ise ilgisiz tweetleri filtrelemek için kullanılan bir gürültü algılama tekniği kullanmışlardır. Kosinüs benzerliği iyi prototiplerin (TP ve TN) belirli bir eşiğin altında kalanları ve kötü prototiplerin (FP ve FN) belirli bir eşiğin üstünde olanların arındırılmasına dayalı bir gürültü algılama tekniği uyguladıktan sonra modeli yeniden eğitmişler, 2008 yılı için bir yıla yayılan 10 milyon tweet ile test ederek DJIA'nın günlük tahmininde %88,9 doğruluk oranına ulaşmışlardır.

Sul ve diğerleri (2017), S&P 500 endeksinde listelenen firmalarla ilgili tweetler ve bunların kümülatif duygusal değeri, yani gönderilerin genel olarak pozitif ya da negatif bir duygu içerip içermediği ve bunun firmaların ortalama günlük borsa getirileriyle ilişkisini incelemek amacıyla yaptıkları çalışmada, Mart 2011-Ekim 2011 tarihleri arasında 2.503.385 tweet toplamışlardır. Firmalar hakkındaki tweetleri duygusal bir değer değişkeni olarak ölçmek için kelime analizi yöntemini kullanmışlar, Harvard-IV sözlüğünde yer alan “NEG” etiketli kelimeleri olumsuz değer taşıyan kelimeler, “POS” etiketli kelimeleri olumlu değer taşıyan kelimeler olarak kabul etmişler ve tüm tweetlerdeki bu tür kelimelerin sayısını belirlemişlerdir. Araştırmanın sonucunda belirli bir firma hakkındaki tweetlerin kümülatif duygusal değerinin (pozitif veya negatif) o firmanın hisse senedi getirileriyle önemli ölçüde

ilişkili olduğu bulgusuna ulaşmışlardır. Ayrıca çok sayıda takipçisi (ortalamadan daha fazla – ki bu çalışmadaki örneklem için ortalama 177 takipçi) olan kullanıcılardan gelen tweetler ağda daha hızlı yayıldığından, bu tweetlerin duygusal değerinin hisse senedi fiyatlarına daha hızlı yansıdığı ve aynı gün getirileri üzerinde daha güçlü bir etkiye sahip olduğunu ortaya koymuşlardır. Buna karşılık, az takipçisi olan kullanıcılardan gelen tweetlerin duygusal değerinin, gelecekteki hisse senedi getirileri (10 günlük getiriler) üzerinde daha güçlü bir etkiye sahip sonucuna ulaşmışlardır.

Alostad ve Davulcu (2017), metin analizine dayalı olarak saatlik hisse senedi fiyatlarının yönünü tahmin etmek amacıyla Ekim 2009 ile Eylül 2014 arasındaki dönem için Dow Jones Endeksi'ndeki 30 hisse senedi sembolü ile ilgili NASDAQ web sitesinden 53.641 haber makalesi, ayrıca Mart 2014 ile Eylül 2014 arasındaki dönem için hisse senedi kodlarıyla eşleşen 780.139 tweet toplamışlardır. Twitter hacim artışlarının da kullanıldığı çalışmada son dakika haberine dayalı sistemin, tüm haberleri kullanan sisteme kıyasla tahmin doğruluğunda istatistiksel olarak önemli bir artış sağladığı bulgusuna ulaşmışlardır.

Bujari ve diğerleri (2017), tweetlerin pozitif, negatif ya da nötr olması ile Dow Jones Borsasında işlem gören belirli şirketlerin hisse senetlerinin gelecekteki eğilimini tahmin etmenin mümkün olup olmadığını araştırmak amacıyla yaptıkları çalışmada, genel endeks yerine teknoloji, hizmet ve sağlık olmak üzere üç sektöre odaklanmışlardır. Her bir hisse senedi için 5 farklı metriği (işlem hacmi, en yüksek değer, en düşük değer, açılış fiyatı ve kapanış fiyatı) ve trendi dikkate almışlar, şirketler hakkındaki tweetler için şirketlerin borsa kodlarını içeren cashtag (\$) ve hastagleri (#) kullanarak 70 günlük analiz dönemi için 825.987'den fazla tweet toplanmıştır. Tweetlerin duygusal olarak sınıflandırılmasında kelime listesi tabanlı bir yaklaşım olan AFINN (Nielsen, 2011) yöntemini kullanmışlar, sınıflandırmadan sonra tweetler için 13 farklı değişken (örneğin, hacim, pozitif tweet hacmi, negatif tweet hacmi, nötr tweet hacmi, mention içeren pozitif ve negatif tweetler gibi) elde etmişlerdir. Analizin sonucunda, önerilen geçici tahmin yöntemlerinden bazılarının belirli şirketlerin hisse senedi değerlerinin bir sonraki gün trendini %82 başarı oranıyla tahmin ettiği

sonucuna ulaşmışlardır. Ancak tweetlerle ilgili değişkenleri kullanarak yapılan tahminlerin bir sonraki günden öteye taşınmadığına da dikkat çekilmiştir.

Li ve diğerleri (2014; 2017), Twitter verileri kullanılarak ölçülen toplumsal duygu durumunun [public sentiment] hem belirli bir şirketin hisse senedi fiyatlarını öngörmek için kullanılıp kullanılmayacağını hem de öngörü gücünün şirketlere ya da sektörlerle göre farklılık arz edip etmeyeceğini araştırdıkları bir çalışmada ABD’de faaliyet gösteren NASDAQ ve New York Menkul Kıymetler Borsası’nda listelenen 30 şirketi analiz etmişlerdir. Tweetlerin duygusal yönelimlerini pozitif⁺, pozitif, nötr, negatif, negatif⁻ olmak üzere beş kategoride sınıflandırabilmek için NLP [natural language processing] tekniklerinden, her bir sözcüğün duygu değerinin hesaplanmasına dayalı ve metin bilgisini tanımlayıcıların vektörleri olarak temsil edebilen cebirsel bir model olan vektör uzay modeli [Vector Space Model, (VSM)] ve TF-IDF modelini kullanmışlar, ayrıca belirli kelimelerin duygu değerini hesaplamak için sözlük tabanlı bir uygulama olan SentiWordNet 3.0’a da başvurmuşlardır. Farklı sektörleri temsilen seçilen 30 şirket ve bu şirketlerin ürünleri ile ilgili anahtar kelimeler kullanarak Ekim 2011’den Mart 2012 tarihine kadar 200 milyon tweet toplamışlardır. Şirketlerin hisse senedi fiyatlarıyla ilgili verileri ise Yahoo’dan toplamışlar, analizde sadece kapanış fiyatlarını kullanmışlardır. Bing vd.’leri araştırmanın sonucunda, önerdikleri algoritma (Sosyal Medya Veri Analizcisi – Duyarlılık Analizi (SMeDA-SA)) ile bazı şirketlerin hisse senedi kapanış fiyatlarının %76,12 doğrulukla tahmin edilebildiğini, bilişim teknolojileri ve medya gibi sektörlerde bu algoritmanın tahmin başarısının diğer sektörlerle kıyasla daha başarılı olduğunu ve mevcut tweetlerin duygusal yönelimini 3 gün sonraki hisse senedi fiyatlarını tahmin etmek için kullanmanın daha iyi bir performans sergilediğini bulmuşlardır.

Urolagin (2017), ABD Borsalarına kote şirketlerden bazıları hakkındaki tweetler ve hisse senedi fiyatları arasındaki ilişkiyi analiz etmek amacıyla yaptığı çalışmada, 31 Temmuz 2017 - 7 Ağustos 2017 tarihleri arasında şirketlerle ilgili paylaşılan (314 ila 719 arasında değişen sayılarda) tweetleri toplamıştır. Her tweetteki önemli sözcükleri belirlemek ve bu sözcüklerin duygu puanını hesaplamak için AFINN yöntemini kullanmıştır. Her bir tweetin

duygu puanını, tweetlerdeki her bir sözcüğün duygu puanının toplamı olarak hesaplamış ve her tweeti özellik vektörü cinsinden temsil etmek için N-gram tabanlı bir özellik çıkarma tekniği kullanmıştır. Tweetlerin duygu sınıfının tahmini için Naïve Bayes ve SVM sınıflandırıcıları kullanmış, SVM için %81,51 ve Naïve Bayes için %80,042 tahmin doğruluğu elde etmiş ve tweetleri puanlarına göre pozitif, negatif veya nötr olarak kategorize etmiştir. Daha sonra bir şirketin hisse senedi fiyatları ile tweetlerin duygu durumu arasındaki ilişkiyi SVM yöntemiyle incelemiştir. Araştırmanın sonucunda bir hisse senedinin açılış ve kapanış fiyatlarında, tweetlerin duygusal yönelimi ile birlikte kademeli bir değişim gözlemlemiştir. SVM yönteminin, tweetlerin özelliklerini kullanarak hisse senedi fiyatlarını tahmin etmedeki başarısını %84,8 olarak ortaya koymuştur.

Bernardo ve diğerleri (2017) belirli şirketlerle ilgili tweetleri kullanarak bireysel hisse senedi fiyatlarını tahmin etmek amacıyla yaptıkları çalışmada kısa bir zaman dilimi içinde hisse senetleriyle ilişkilendirilmesi muhtemel olan en yeni haberleri yaydığı için bir günlük, bir saatlik ve üç dakikalık periyotlarda tweetlerdeki ortalama duyguları hesaplamışlardır. ABD Borsalarında işlem gören 16 şirketle ilgili 11 Şubat 2014 – 28 Şubat 2014 tarihleri arasında 1.535.167 tweet toplamışlar ve bu şirketlerin tarihsel hisse senedi fiyat verilerini temin etmişlerdir. Tweetleri anahtar kelime tabanlı sınıflandırıcı ve Bayes sınıflandırıcı ile pozitif, negatif ve nötr olarak sınıflandırmışlar, gözetimli öğrenme yaklaşımına dayalı Bayes sınıflandırıcının anahtar kelime tabanlı sınıflandırıcından daha başarılı bir sonuç verdiği bulgusuna ulaşmışlardır. Çalışmada ayrıca tweetlerdeki duygular ile bazı şirketlerin hisse senedi fiyatları arasında anlamlı bir ilişki olduğu sonucuna varılmışlar, bu ilişkinin kısa zaman dilimleri için test edildiğinde kullanılan zaman aralığına çok bağlı olduğu, bazı şirketler için tepki süresinin çok kısa, diğerleri için daha uzun olduğu sonuçlarını elde etmişlerdir. Son olarak hisse senedi fiyatlarının nadiren de olsa, bazı şirketlerle ilgili (British Petroleum) tweetlerin duygusal içeriğini etkilediği bulgusuna ulaşmışlardır.

Xun ve Guo (2017), müşterilerin doğrudan hizmet deneyimleri konusunda elektronik ağızdan ağıza iletişimi (eWOM) ile firma değeri arasındaki ilişkiyi araştırmak amacıyla yaptıkları çalışmada yedi Amerikan havayolu şirketinin Ocak 2012 - Aralık 2012 tarihlerini kapsayan

52 haftalık dönemi analiz etmişlerdir. Analizde müşterilerin şirketlerle tweetlerini toplamışlar ve Hu ve Liu ile Loughran ve McDonald duygu kelimelerine dayanan bir duygu analizi algoritması ile sınıflandırmışlardır. Daha sonra müşterinin elektronik kulaktan kulağa iletişimi ile bu firmaların hisse senedi fiyat volatilitesi ve getirileri arasındaki ilişkiyi vektör otoregresif (VAR) yaklaşımı ile analiz etmişler, müşterinin bir firmaya ilişkin elektronik ağızdan ağıza iletişiminin, firmanın hisse senedi getirisiyle pozitif, hisse senedi volatilitesi ile negatif bir şekilde ilişkili olduğu bulgusuna ulaşmışlardır.

Daniel ve diğerleri (2017), hisse senedi fiyatını etkileyebilecek finansal anlamda özel olayların duygu analizi yoluyla popüleritesini pozitif ya da negatif olarak sınıflandırılan Twitter'dan tespit etmek amacıyla yaptıkları çalışmada Dow Jones Endeksini oluşturan şirketlerle ilgili ürün lansmanları ve kampanyalar gibi geçmiş olaylara odaklanmışlardır. Çalışma finansal toplulukların oluşturulmasından tweetlerin toplanmasına, tweetlerin içeriğinin analiz edilmesinden olayların tespit edilip analizine kadar birkaç aşamadan oluşmaktadır. İlk olarak Twitter içerisinde bir finans topluluğu tanımlayan ve her kullanıcı için tüm tweetleri çıkaran bir algoritma geliştirmişlerdir. Çalışmada finans topluluğu, yatırım uzmanlarını, finansal haber sağlayıcılarını, yöneticileri ve şirket kurucularını temsil eden on Twitter kullanıcı hesabı ve bu on hesabın takipçileri içerisinde, takip ettiği kişi sayısından daha fazla takipçisi olan ve takipçi sayısı 100'den fazla olan hesaplar şeklinde tanımlanmıştır. Takip ettiğinden daha fazla takipçisi olan kullanıcılar finansal topluluğa dâhil edilmemiştir. İkinci olarak finans topluluğunun tweetlerini, investopedia (<https://www.investopedia.com/>) adlı finansal sözlüğe dayanarak hazırlanan finans piyasası ve borsa ile ilgili 865 sözcük ve şirketlerin borsa kodlarını veya isimlerini içeren cashtagler ve hashtagler aracılığıyla filtrelemişlerdir. Üçüncü olarak tweetlerdeki duygu analizi için Sentistrength, TextBlob, Affin gibi araçların yanı sıra kendi geliştirdikleri ve MySentimentApi adını verdikleri aracı kullanmışlar ve bu dört algoritma aracılığıyla, tweetleri pozitif, negatif ve nötr puanlarla değerlendirmişlerdir. Dördüncü olarak finansal anlamda özel olayları, zaman serilerindeki çok negatif veya çok pozitif zirveler olarak tanımlamışlardır. Bu zirveler, şirketle ilgili paylaşılan tweetlerden çıkarılan günlük duygu zaman serisinin analizine dayanmaktadır. Bir önceki aşamada dört farklı araçla gerçekleştirilen duygu analizi sonuçlarını bu aşamada

zirvelerle ilgili olayların belirlenmesinde tutarlılık olup olmadığını test etme amacıyla kullanmışlardır. Analiz sürecinde (Eylül 2013'ten Eylül 2015'e kadar) toplamda 9.011 kullanıcı hesabından 12 milyondan fazla tweet toplanmış ve filtrelendikten sonra analize 192.935 tweet ile devam etmişlerdir. Araştırmanın sonucunda geliştirilen modelin, tanımlanmış finans topluluğunun şirketler hakkında paylaştığı tweetlere dayalı olarak finansal olayları tespit etmede iyi bir performans sergilediğini ortaya koymuşlardır.

Liew ve Budavari'nin (2017) hisse senedi fiyatlarının hem geleneksel kaynaklardan hem de tweetler gibi geleneksel olmayan kaynaklardan halka açık tüm bilgileri içerip içermediğini araştırmak amacıyla yaptıkları çalışmada Fama-French'in (2015) beş faktör modelini esas almışlardır. Fama-French'in bir portföyün getirisinin risksiz faiz oranını aşan kısmının beş faktörden etkilendiği tezine altıncı faktör olarak şirkete özgü sosyal medya değişkenini ilave etmişlerdir. S&P 500 endeksinde yer alan, Ocak 2013 - Kasım 2015 döneminde 15 hisse senedine ait günlük veriler ile bu şirketler hakkında içerdiği duygunun yukarı yönlü [bullish] ya da aşağı yönlü [bearish] olarak etiketlendiği tweetler arasındaki ilişkiyi analiz etmişlerdir. Analizin sonucunda yukarı yönlü duygu değişkeni ile AAPL'in fazla getirileri [excess return] arasında 0,35, piyasanın fazla getirileri arasında 0,16 pozitif korelasyon olduğu bulgusuna ulaşmışlardır. Ayrıca hem CAPM hem Fama-French'in üç faktörlü modeli hem de Fama-French'in beş faktörlü modelinde sosyal medya faktörünün düzeltilmiş R^2 'yi %1'den %10'a kadar pozitif ve tutarlı bir şekilde artırdığı sonucuna ulaşmışlardır.

Jong ve diğerleri (2017) hisse senedi getirileri ve volatilité ile tweetler arasında çift yönlü bir gün içi ilişkisi olup olmadığını, diğer bir ifadeyle hisse senetlerinin Twitter'da daha fazla anılmasının hisse senetlerinin getirilerini ve / veya volatilitésini etkileyip etkilemediğini ve hisse senetlerinin volatilitésinin hisse senetlerinin Twitter'da bahsedilme sayısını etkileyip etkilemediğini araştırmak amacıyla yaptıkları çalışmada DJIA Endeksindeki 30 hisse senedine odaklanmışlardır. 2 Haziran - 18 Haziran 2014 tarihleri arasında 13 günlük bir dönem için tweet verilerini ve DJIA'daki 30 hisse senedi için dakika bazında 150.180 hisse senedi fiyat verisini kullanarak BEKK-MVGARCH yöntemiyle analiz etmek için öncelikle Augmented Dickey ve Phillips-Perron birim kök testleri yapmışlar ve hisse senedi getirileri

ile tweet serilerinin durağan olduđu sonucuna varmışlardır. Analizin sonucunda bulgular hisse senedi getirilerinin %87'sinin tweet verilerindeki gecikmeli yeniliklerden etkilendiğini, ancak yönün karşılıklı olduğunu destekleyen çok az kanıt olduğunu, tweetlerin yalnızca %7'sinin hisse senedi getirilerindeki gecikmeli yeniliklerden etkilendiğini göstermektedir. Sonuçlar ayrıca hisse senedi getirilerinin %40'ından kaynaklanan gecikmeli yeniliklerin tweetlerin mevcut koşullu volatilitelerini etkilerken, tweetlerin %73'ü hisse senedi getirilerinin mevcut koşullu volatilitelerini etkilediğine işaret etmektedir. Dahası, tweetlerin getirilerinden kaynaklanan volatilitenin, hisse senetlerinin %33'ü için devam ettiğini gösteren güçlü kanıtlar bulmuşlardır. Tweetlerden getirilere doğru dalgalanma hisse senetlerinin %73'ü için devam etmektedir. Son olarak, hisse senetlerinin %53'ü getirilerden tweetlere hem anlık hem de kalıcı etkiler gösterirken, hisse senetlerinin %90'ı tweetlerden getirilere hem anlık hem de kalıcı etkiler sergilemektedir.

Tabari ve diğerleri (2018) borsa ile ilgili tweetlerin duygusal olarak sınıflandırılması için özel olarak eğitilmiş bir model geliştirmek amacıyla yaptıkları çalışmada 01.01.2017 - 03.31.2017 tarihleri arasında hisse senedi sembollerini içeren 20.000 tweet toplamışlar ve toplanan bu tweetleri Amazon Mechanical Turk kullanarak manuel bir şekilde negatif, pozitif ve nötr olarak etiketlemişlerdir. Etiketlenen tweetleri makine öğrenimi yöntemine dayalı sınıflandırıcının eğitiminde kullanmışlar, rassal orman yönteminde %78,9, destek vektör makineleri yönteminde %79,9 doğruluk skoruna ulaşmışlardır. Sınıflandırıcıyı yapay sinir ağı temelli CNN ve LSTM yöntemleriyle de eğitmişler, en yüksek doğruluk skorunu %92,7 ile LSTM yöntemiyle yapılan eğitiminden elde etmişlerdir. Son aşamada ise tweetler ile borsa getirileri arasındaki nedensel bağlantıyı analiz etmişlerdir. Getiriler ve duygular 15 dakikalık, 30 dakikalık, bir saatlik, üç saatlik ve 1 günlük olmak üzere beş aralığa bölündükten sonra her aralık ve üç şirket (Apple, Facebook ve Amazon) için getiri değerleri ve duygu değerleri arasındaki nedensel ilişkileri belirlemek için Granger nedensellik analizi yapmışlardır. Analizin sonucunda Amazon ve Facebook için üç saat ve 1 günlük bir gecikmeyle önemli nedensellik bağlantıları tespit etmişlerdir. Bu iki hisse senedi için en güçlü nedensel ağırlık, üç saatlik bir gecikmeyle gerçekleştiği sonucuna ulaşılmış ve daha da önemlisi, nedensel bağlantının her iki yönde de ortaya çıktığı, yani tweetlerin gelecekteki

borsa getirilerini, borsa getirilerinin de gelecekteki tweetleri etkilediği bulgusuna ulaşmışlardır.

Rakowski ve diğerleri (2018) Twitter’da şirketlere olan ilgilinin [attention] ABD hisse senedi piyasalarından Russell 3000 Endeksindeki hisse senedi fiyatları üzerindeki etkisini araştırmak amacıyla yaptıkları çalışmada tweetler cashtagler (\$) yardımıyla 2011'den 2015'e kadar günlük frekanslarda toplanmıştır. Beş yıllık süre boyunca şirketlere özel 21 milyondan fazla tweeti incelemişler, Twitter etkinliğinin, işlem hacmi ve varlık fiyatları ile istatistiksel açıdan anlamlı bir ilişkisi olduğu, dolayısıyla Twitter’ın yatırımcıların dikkatini artırdığı ve yatırımcılar arasında bilginin yayılmasını kolaylaştırdığı bulgusuna ulaşmışlardır. Ayrıca Twitter etkinliğinin, yatırımcıların düşük düzeyde ilgi gösterdiği menkul kıymetler üzerinde daha büyük bir etki gösterdiği, pozitif normalüstü getirilerle ilişkili olduğu ve tweetlerin geleneksel haber etkinlikleriyle birlikte ortaya çıktığı durumlarda yatırımcılara daha fazla bilgi yayıldığı sonuçlarına ulaşmışlardır. Rakowski vd.’leri çalışmada son olarak Twitter ile hisse alım satım işlemleri arasındaki bağlantıyı kurumsal yatırımcıların mı yoksa bireysel yatırımcıların mı yönlendirdiğini araştırmışlar, etkilenen hisse senetlerinin küçük, volatil ve arbitraj yapılması zor olan hisse senetleri olmasından dolayı kurumsal yatırımcıların bu hisse senetlerinin alım satım işlemlerinde yer alma ihtimalinin düşük olduğunu ortaya koymuşlardır. Dolayısıyla kurumsal yatırımcıların yokluğunda, bu hisse senetlerinin bireysel yatırımcıların alım satım modellerine karşı daha savunmasız olduğu ve fiyatları temel değerlerden daha kolay saptığı bulgusuna ulaşmışlardır.

Bartov ve diğerleri (2018) bir firmanın kazanç duyurusundan hemen önce tweet atan kişilerin görüşlerinin firmanın kazançlarını ve getiri duyurularını tahmin edip etmediğini araştırmak amacıyla yaptıkları çalışmada, geçmiş Twitter verilerini, yetkili sosyal medya veri satıcısı olan GNIP'den almışlardır. Russell 3000 Endeksinde yer alan firmaların borsa kodlarını kullanarak 1 Ocak 2009 ile 31 Aralık 2012 tarihleri arasında borsa ile ilgili 8.462.761 tweet toplamışlar ve 4.132 firma için 54.906 çeyrek dönemi almışlardır. Ayrıca kazançların açıklanmasından hemen önceki tweetlerin tahmin yeteneğini ölçmek için kazanç duyurusundan önceki dokuz işlem gününe odaklanmışlar ve 3.604 farklı Russell 3000

firmasından 33.186 çeyrek dönemi kapsayan 869.733 tweet toplamışlardır. Tweetlerin duygusal yönelimini iki farklı yaklaşımla sınıflandırmışlardır. İlk yaklaşım (OPI_BAYES) Naive Bayes yöntemini kullanarak bireysel tweetleri pozitif, negatif veya nötr olarak sınıflandırırken ikinci yaklaşım (OPI_VOCAB), negatif kelimelerin sayısını tanımlayan ve yaygın olarak kullanılan sözlüklerden türetilmiş bir yöntem kullanmaktadır. Ayrıca iki sınıflandırma şeması kullanarak farklı tweet türleri arasında da ayırım yapmışlardır. İlk sınıflandırma şemasını orijinal bilgileri aktaran tweetleri mevcut bilgileri yayan tweetlerden ayırmak için, ikinci sınıflandırma şemasını kazançlar, şirket temelleri [fundamentals] ve/veya hisse senedi işlemleri ile ilgili bilgileri ileten tweetleri diğer bilgileri ileten tweetlerden ayırmak için kullanmışlardır. Araştırmanın sonucunda ilk olarak, bireysel tweetlerden alınan toplu görüşlerin bir firmanın gelecek üç aylık kazançlarını ve getiri duyurularını başarıyla tahmin ettiği sonucuna ulaşılmıştır. İkinci olarak, üç aylık kazanç duyurusuna anında hisse senedi fiyat tepkisi ile hem OPI_BAYES hem de OPI_VOCAB arasında pozitif bir ilişki olduğunu bulmuşlar ve bu durumu Twitter'ın hisse senedi fiyat beklentileri ve değerleri hakkında orijinal bilgileri iletenlerin yanı sıra mevcut bilgileri de yaydığı şeklinde yorumlamışlardır. Araştırmanın üçüncü bulgusu Twitter'dan ölçülen duygu ile kazanç duyurularına hisse senedinin fiyatı tepkisinin, zayıf bilgi ortamlarındaki firmalar için daha güçlü olduğu yönündedir.

Li ve diğerleri (2018a) tweetlerdeki duygu ile 5 firmanın (Apple, Google, Alibaba, Amazon, Microsoft) günlük getirileri arasındaki korelasyonu incelemek amacıyla yaptıkları çalışmada her bir gün ve her bir firma için 100'er tweet toplanmış ve eğitim seti hazırlanırken anahtar kelimelerin duygu puanlarını hesaplayan AFINN kütüphanesi kullanılmıştır. Çalışmada tweetler SVM ve Naive Bayes Bernoulli yöntemleriyle pozitif, negatif ve nötr olarak sınıflandırılmıştır. SVM yöntemi %85,34 doğruluk oranı ile sınıflandırma yaparken Naive Bayes Bernoulli yöntemi %81,03 doğruluk oranı ile sınıflandırma yaptığı, SVM yönteminin doğruluğunun, özellikle “pozitif” ve “negatif” sınıflar açısından Bernoulli NB'den daha iyi olduğu sonucuna ulaşmışlardır. Ayrıca geçmiş fiyatlara dayalı tahminde Random Forest ve SVM yöntemini karşılaştırmışlar, analize dahil olan firma sayısının ilave bir katkı sağlamadığı sonucuna vardıkları için seçilen bu 5 firmayla analize devam etmişlerdir. Tweet

duyguları ile borsa fiyatları arasındaki korelasyonu incelemek için SVM yöntemini kullanmışlar ve bu model ile %56,5’lik bir tahmin doğruluğuna ulaşmışlardır.

Li ve diğerleri (2018b) hisse senetleriyle ilgili tweetlerin bilgi içeriğinin borsa performansı üzerindeki etkilerini ve sosyal etkinin rolünü araştırmak amacıyla yaptıkları çalışmada hisse senedi bazında tweet özellikleri (mesaj yükseliş sinyali, mesaj hacmi ve mesaj uzlaşısı skoru) ile günlerarası [interday] ve gün içi [intraday] piyasa performansı (hisse senedi getirileri, işlem hacmi ve volatilité) arasındaki ilişkiye odaklanmışlardır. Ayrıca iyi yatırım tavsiyelerinin ne ölçüde dikkat çektiğini inceleyerek hisse senetleriyle ilgili paylaşımların yapıldığı sosyal medya platformlarından verimli bilgi toplamanın olası mekanizmalarını incelemiştirler. Çalışmada ABD hisse senedi piyasalarını yansıtmaya açısından S&P 100 Endeksine odaklanmışlar, 20 Ağustos 2012 ile 25 Mart 2013 tarihleri arasında hisse senedi sembolünü içeren cashtagler vasıtasıyla hisse senediyle ilgili toplam 1.278.604 tweet toplamışlardır. Tweetlerdeki duyguyu “al”, “sat” ve “tut” şeklinde kategorize edebilmek için Naive Bayes yöntemiyle eğitilmiş ve %86,3 sınıflandırma doğruluğuna sahip olan sınıflandırıcıyı kullanmışlardır. Tweet özelliklerinden *mesaj yükseliş sinyalini* [bullishness] belirli bir zaman aralığında satışla ilgili mesajların alışla ilgili mesajlara oranı, *mesaj hacmini* [message volume] belirli bir zaman aralığındaki tweet sayısının doğal logaritması, *uzlaşısı skorunu* [agreement] ise fikir birliği düzeyini yansıtacak şekilde hesaplamışlardır. Bir kullanıcının sosyal etkisini ölçmek için retweetler, mentionlar, takipçi sayısı ve Kred etkisi³⁶ ölçütlerini kullanarak tweetleri ağırlıklandırmışlardır. Analizin sonucunda mesaj yükseliş sinyalinin eşzamanlı günlük normalüstü getirileri etkilediği, mesajlarına yansıyan yeni bilgilerin gün içi sonuçların da gösterdiği gibi hızla piyasa fiyatlarına dahil olduğu, mesaj hacminin 15 dakikalık periyotta getirileri, işlem hacmini ve volatilitéyi tahmin edebildiği ve uzlaşısı skorunun, hem günlük hem de gün içi analizlerde hisse senedi özelliklerini olumlu yönde etkilediği bulgularına ulaşmışlardır. Ayrıca retweetler, takipçiler, mentionlar ve Kred etki puanı şeklindeki bilgi yayılımına ilişkin analizlerde, ortalamanın üzerinde yatırım

³⁶ Çevrimiçi sosyal etkiyi ölçmek için kurulan PeopleBrowsr tarafından oluşturulan bir web uygulamaları ekosistemidir.

tavsiyesi saęlayan kullanıcıların, daha yüksek düzeyde retweet ve Kred etkisi yoluyla hisse senedi mikrobloęları üzerinde daha fazla sosyal etkiye sahip olduęu sonucuna varmışlardır.

Strauß ve dięerleri (2018) Reuters ve Bloomberg hesapları tarafından yayınlanan tweetler kapsamında ekonomik haberler ile borsa arasındaki ilişkiyi incelemek amacıyla yaptıkları çalışmada ABD piyasalarından Dow Jones Endüstriyel Ortalama Endeksine odaklanmışlardır. Analiz kapsamında Reuters ve Bloomberg'in Twitter hesaplarından 2015 yılının Eylül ayındaki tüm işlem günlerinde paylaşılan 2.440 tweet toplamışlar, ayrıca aynı dönemde yayımlanan 50 Bloomberg piyasa haberini analize dâhil etmişlerdir. Analizi 5 dakikalık, 20 dakikalık ve 1 saatlik zaman dilimleri için VAR yöntemiyle yapmışlar, ekonomik tweetlerin DJIA dalgalanmaları üzerindeki etkisinin daha uzun zaman aralıklarında, yani 1 saatte en güçlü olduęu bulgusuna ulaşmışlardır. Bu durumu bireysel yatırımcıların bilgileri deęerlendirmek ve bir alım satım kararı vermek için profesyonel yatırımcılardan daha fazla zamana ihtiyaç duyduklarının bir göstergesi olarak yorumlamışlardır. Ayrıca Reuters ve Bloomberg tweetlerinin tüm ekonomik haberleri, piyasa genelini veya firmaya özgü bilgileri ele alması arasında farklılıklar olup olmadıęı araştırıldığında, firmaya özgü bilgiler için etkinin olmadıęı, yalnızca bir bütün olarak piyasa hakkında bilgi içeren tweetler dikkate alındığında güçlü bir etkinin olduęu bulgusuna ulaşmışlardır. Son olarak haber hacminin, haber ilgi düzeyinin ve uzman görüşünün kısa vadeli aralıklarla DJIA dalgalanmalarını olumlu yönde etkilediğini ortaya koyarak, tweetlerin bireysel yatırımcılar [public investors] için bilgi sunabileceęi sonucuna varmışlardır.

Garcia-Lopez ve dięerleri (2018), sosyal medya mesajları ile borsa arasındaki ilişkiyi ölçmek amacıyla ABD Borsalarında işlem göre dokuz bilişim teknolojisi firmasıyla ilgili 23 Mart 2017'den 3 Temmuz 2017'ye kadarki 3 aylık bir dönemi kapsayan tweetlere odaklanarak şirketlerin borsa kodlarını içeren 141.007 tweet toplamışlardır. Bu tweetler kelime çantası [Bag of Words, BOW] ve kelime gömülmeleri [word embeddings] yaklaşımlarıyla pozitif, negatif ve nötr olarak sınıflandırılmıştır. Tweetlerin sınıflandırılmasında kelime gömülmeleri yaklaşımı kullanılarak elde ettikleri sonuçların, kelime çantası yöntemiyle elde edilen

sonuçlardan önemli ölçüde daha iyi performans gösterdiği bulgusuna ulaşmışlardır. Daha sonra dokuz bilişim teknolojisi firmasının günlük açılış fiyatları, kapanış fiyatları ve işlem hacimlerinin günlük tweet hacmi ile ilişkisi analiz edilmiş, günlük tweet hacmi ile finansal işlemlerin hacmi arasında korelasyon olduğunu ortaya koymuşlar, iki şirket için tweet hacminin işlem hacmini tahmin etmek amacıyla bir veya üç günlük gecikmeyle kullanılabileceğini göstermişlerdir. Ayrıca MAP [Moving Approximation] ve LTAM [Local Trend Association Measure] yöntemleriyle, korelasyon analizi ile keşfedilemeyen ters ilişkiler, yani daha spesifik olarak yerel trend ilişkisi keşfedilmiştir.

Xu ve Cohen (2018), bir sonraki gün fiyat hareketini tahmin etmek için geçmiş fiyatları tweetlerle birleştiren yapay sinir ağlarına dayalı bir tahmin modeli geliştirmek amacıyla yaptıkları çalışmada, 8 sektörde faaliyet gösteren, sermaye büyüklüğüne göre ilk 10 sırada yer alan ve NASDAQ listelenen 88 şirketin hisse senetlerinin 01 Ocak 2014 - 01 Ocak 2016 tarihleri arasındaki iki yıllık fiyat hareketlerine ve şirketlerin borsa kodlarına göre toplanan tweetlere odaklanmışlardır. Tweetler NLTK paketi kullanılarak ön işlemden geçirilerek analize hazır hale getirilmiştir. Önerdikleri model (StockNet) ile yaptıkları analizin sonucunda, sosyal medya verileri ile hisse senedi fiyatlarının tahmininde %58,23 doğruluk oranına ulaşmışlardır. Bu çalışmada karşılaştırma yapabilmek amacıyla ortaya konan diğer doğruluk oranları örneğin ARIMA için %51,39, rassal orman [Random Forest] için %53,08 ve HAN [Hierarchical Attention Network] için %57,64'tür.

Batra ve Daudpota (2018), tweetlerdeki duyguyu analiz ederek yatırımcıların karar süreçlerini destekleyici bir model ortaya koyabilmek amacıyla yaptıkları çalışmada, StockTwits adlı sosyal medya platformundan Apple ürünleriyle ilgili 2010 ile 2017 yılları arasında alınan 300.000 tweet SVM yöntemiyle %63,5 doğruluk skoruyla yükseliş [bullish] veya düşüş [bearish] olarak kategorize etmişler ve tweetlerdeki duygusal yönelimini gün bazında toplulaştırmışlardır. Ardından tweetlerin duygu puanı ve piyasa verilerini temel alarak bir sonraki günün hisse senedi hareketini tahmin etmek için SVM yöntemini kullanmışlardır. Analizin sonucunda insanların Twitter'dan ölçülen görüşleri ile piyasa

verileri arasında pozitif bir ilişki olduğu ve tahmin modelinin %76,68 test doğruluğu elde ettiği bulgusuna ulaşmışlardır.

Roslan ve Daudpota (2018), tweetlerin borsaya gün içerisindeki etkisini araştırmak amacıyla yaptıkları çalışmada, NASDAQ'ta listelenen Apple şirketiyle ilgili tweetlere odaklanmışlar, 6 günlük bir zaman dilimi için hisse senedi ve tweet verileri toplamışlardır. 6 gün için saatlik düzeyde 42 gözleme ulaşmışlardır. Apple ile ilgili her bir tweetin SentiStrenth adlı bir uygulama kullanarak hesaplanan duygusal puanı ile Apple'ın hisse senedi fiyatı arasında korelasyon analizi yapmışlar ve 0,33 pozitif korelasyon olduğu bulgusuna ulaşmışlardır. Analizde ayrıca tweetlerdeki duygulara dayalı olarak hisse senedi fiyatlarını tahmin etmek için yapay sinir ağı yöntemi kullanılarak tweetlerdeki duygusal durum ile hisse senedi fiyatının %99,95 doğrulukla tahmin edilebileceği bulgusuna ulaşmışlardır.

Bhatia ve diğerleri (2018), ABD borsalarında işlem gören hisse senetlerinin (Microsoft, Apple, Exxon, Walmart, IBM) hareketinin analiz edildiği çalışmada hem hisse senedi hem de daha geniş açıdan ekonomi ile ilgili duygusal içeriğe odaklanmışlardır. Ayrıca bir hisse senedinin işlem gördüğü piyasa ile hisse senedinin fiyat hareketleri arasındaki göreceli ilişkinin bir sonraki gün için yükseliş ya da düşüş olarak tahmin edilip edilemeyeceğini araştırmışlardır. Bu amaçla Haziran 2009'dan Aralık 2009'a kadar 5 büyük hacimli hisse senedinin (Microsoft, Apple, Exxon, Walmart, IBM) fiyat hareketinin analiz edildiği çalışmada örneğin, bir "ABC" hisse senedi bir günde %0,5 artarken, piyasa endeksi "SPY" aynı gün %1 artarsa - o gün için "ABC" hissesinin göreceli hareketi o gün için piyasa endeksinin altında performans gösterdiğinden "düşüş" olarak işaretlenmiştir. Analizin sonucunda sadece fiyata dayalı sınıflandırıcıların ortalama doğruluğunun %55,1 iken, fiyat ve duygu temelli sınıflandırıcıların ortalama doğruluğunun %59 olduğu bulgusuna ulaşmışlardır. Diğer bir ifadeyle sonuçlar, tweetlerden alınan duygularla ilgili bilgileri kullanan sınıflandırıcıların, dayanak varlık fiyatları için daha yüksek tahmin doğruluğu elde etmeye yardımcı olduğunu göstermektedir.

Wu (2019), şirketlerin üç aylık kazanç duyurularına yanıt olarak sosyal medyanın hisse senedi fiyatlarını ve kazanç duyurusu sonrası kaymayı ³⁷ nasıl etkilediğini araştırmak amacıyla yaptığı çalışmada üç aylık kazanç verilerinin yanı sıra Twitter ve StockTwits verilerini kullanmıştır. Çalışmada S&P 500 şirketlerini sektörlerine göre sıralayarak ve 2010-2014 dönemi boyunca ortalama tweet hacimlerini incelemiştir. Popüler endüstrileri (en yüksek sıralamaya sahip endüstriler) popüler olmayan endüstrilerle (en düşük sıralamaya sahip endüstriler) karşılaştırarak, yeni dikkat etkilerinin hem büyüklük hem de istatistiksel anlamlılık açısından farklı olduğu bulgusuna ulaşmıştır. Buna ilave olarak Twitter temelli dikkat etkilerinin çeşitli sektörlerde farklılık gösterdiği, yeni dikkat etkilerinin popüler endüstrilerden ziyade popüler olmayan endüstriler arasında önemli olduğu sonucuna varmıştır. Son olarak, kötü haber bildiren şirketlerin bile, bir kazanç duyurusundan sonra yatırımcılardan yeterince dikkat çekmeleri halinde, pozitif ani normalüstü getirilere sahip olabilecekleri dair kanıt sunmuştur.

Guijarro ve diğerleri (2019) Twitter'ın likidite ve işlem maliyetleri üzerindeki etkisini araştırmak amacıyla yaptıkları çalışmada Twitter'dan ölçülen duygu analizi ve farklı likidite risk göstergeleri arasındaki ilişkiyi analiz etmişlerdir. Asimetrik bilginin likidite ve işlem maliyetlerinin belirleyici unsuru olduğu kabulüyle sosyal medya katılımcılarının finansal piyasalar ve bir bütün olarak belirli menkul kıymetler hakkında bilgi alışverişine açık oldukları, bu nedenle, bu tür bilgi ve görüşlerin finansal piyasa likiditesi üzerinde etkili olabileceği görüşüne dayanan çalışmada, S&P500 Endeksi ile ilgili 3 Temmuz 2019 – 1 Ekim 2019 döneminde günlük olarak toplanan 23.008 gözlemden oluşan tweetleri dikkate almışlardır. Önişlemden geçirilmiş her bir tweetin pozitif veya negatif yönelimini -5 ile 5 arasında değişen sayısal bir değer olarak hesaplamışlardır. Ayrıca varlık fiyatında önemli bir değişikliğe neden olmadan varlığın hızlı bir şekilde alınıp satılabilmesi özelliği olarak tanımlanan piyasa likiditesini, en yüksek ve en düşük fiyat arasındaki fark olan spread, kote edilmiş spread, efektif spread gibi çeşitli göstergelerle ölçerek analize dahil etmişlerdir.

³⁷ Kazanç duyurusu sonrası kayma veya PEAD [post-earnings announcement drift] olarak da adlandırılan durum, bir hisse senedinin kümülatif normal üstü getirilerinin, takip eden birkaç hafta boyunca (hatta birkaç ay) bir kazanç sürprizine doğru kayma eğilimidir.

Analizin sonucunda günlük olarak regresyon tahminlerinin pozitif olmasına rağmen istatistiksel olarak anlamlı olmadıkları, bununla birlikte, ilgili tüm değişkenler için iki günlük bir hareketli ortalama hesaplandığında sonuçların biraz iyileşeceği bulgusuna ulaşmışlardır. Buna ek olarak yatırımcıların ruh halinin likidite ölçülerinden efektif spread ile anlamlı ve pozitif ilişkili olduğu sonucuna ulaşmışlardır.

Shi ve diğerleri (2019) hisse senedi fiyat hareketlerini tahmin etmede metin tabanlı derin öğrenme modellerini görsel olarak yorumlayan bir model geliştirmek amacıyla yaptıkları çalışmada, finansal haberleri, grafikleri ve tweetleri birlikte kullanmışlardır. ABD Borsalarından S&P 500 hisselerinin 2006 yılından 2015 yılına kadarki tarihsel fiyat verilerini, Reuters ve Bloomberg'den 341.310 ana akım finans haberini ve Twitter'dan Nisan 2015 - Kasım 2015 tarih aralığı için 6.869.771 hisse senedi kodlarıyla etiketlenmiş şirketlerle ilgili tweetleri toplamışlardır. Önerilen algoritma, semantiği ara işleme katmanlarına yerleştiren hiyerarşik bir derin sinir ağı modeli kullanmakta ve tahmin kararını tek tek belgelere, bigramlara ve kelimelere etkili bir şekilde dağıtan geri yayılım benzeri bir yaklaşıma dayanmaktadır. Analizin sonucunda önerilen algoritmanın hisse senedi hareketlerini tahmin performansını derin öğrenme yöntemlerinden LSTM ve CNN ile karşılaştırmışlardır. Karşılaştırmanın sonucunda ortalama test tahmin doğruluğu LSTM için %48,3, CNN için %53,3 ve önerilen algoritma için %55,9 olduğu bulgusuna ulaşmışlardır.

Liu (2020), sosyal medya kullanıcılarının oluşturduğu içeriğin firmaların hisse senedi fiyat performansı üzerindeki etkisini araştırdığı çalışmada, aynı zamanda bu etkinin işletmeden işletmeye (B2B) firmalar ile işletmeden müşteriye (B2C) firmaların hisse senedi fiyat performansları açısından farklı olup olmadığını da incelemiştir. Bu amaçla S&P 500 endeksinde yer alan ve 11 sektördeki 407 şirketle ilgili Temmuz 2009'dan Haziran 2017'ye kadar uzanan günlük hisse senedi bilgilerinden ve 20,3 milyon Twitter hesabından toplanan 84 milyon tweetten oluşan veri kümesi oluşturmuştur. Öncelikle, spam mesajları ve botlar tarafından oluşturulan tweetleri, sosyal botları tespit etmek için ayrıntılı prosedürlere işaret eden literatürdeki yöntemleri kullanarak veri setinden çıkarmıştır. Metinlerin konularını belirlemek için LDA yöntemine, duyguları pozitif, negatif ve nötr olarak sınıflandırmak için

ise maksimum entropi yöntemine başvurmuştur. Metinsel içerikler, retweet sayısı, beğeni sayısı, olumsuz duygu, olumlu duygu ve mentionlar gibi değişkenleri çok yıllık aylık panel veriler haline getirerek sabit etki modeli [fixed effects model] ile analiz etmiştir. Araştırmanın sonucunda kullanıcı tarafından üretilen içeriğin firmaların hisse senedi performansı üzerinde önemli bir etkiye sahip olduğu ve hisse senedi performansı üzerindeki etkisinin B2C firmalarında B2B firmalara göre çok daha güçlü olduğu bulgusuna ulaşmıştır. Tüketicilerin olumlu duygularının hisse senedi performansında önemli bir rol oynamazken, tüketicilerin olumsuz duygularının ve ağızdan ağıza söylentilerin [Word-of-mouth, (WOM-)] hisse senedi fiyatlarını önemli ölçüde etkilediğini ortaya koymuştur.

Bouktif ve diğerleri (2020) artırılmış metin ve duygu analizlerinin borsanın yönünü daha doğru tahmin edip etmeyeceğini analiz etmek amacıyla yaptıkları çalışmada NASDAQ-100'de yer alan Google, Yahoo, Amazon, Apple, Alibaba, Tesla, Microsoft, IBM, Facebook ve Bitcoin verilerine ve bunlarla ilgili tweetlere odaklanmışlardır. Verileri 2008'den 2018'e kadarki dönemi kapsayacak şekilde 2800 işlem günü için toplamışlardır. Metinlerdeki duygusal özelliklerin çıkarımı için denetimli öğrenmeye dayalı TextBlob ve kural tabanlı yaklaşıma dayanan VADER [Valence Aware Dictionary and Sentiment Reasoner] yöntemleri kullanmışlardır. Daha ayrıntılı bir analiz için hisse senedi fiyatları geçmişi, tweetlerdeki duygusal yönelim, öznellik, N-gramlar gibi metin tabanlı pek çok özelliği analize dahil etmişlerdir. Nedensellik analizi, algoritmik özellik seçimi ve düzenli model istifleme yöntemleri dâhil olmak üzere çeşitli makine öğrenimi teknikleri ile yapılan analizin sonucunda bir şirketin hisse senedi fiyatlarındaki yükseliş ve düşüşün, Twitter'da dile getirilen kamuoyu görüşlerinden veya duygularından etkilendiği sonucuna ulaşmışlardır. Ayrıca çalışmada kullanılan yaklaşımın yeterli performans gösterdiğini ve hisse senedi hareketlerini %60 gibi bir doğrulukla tahmin ettiğini ortaya koymuşlardır.

Schnaubelt ve diğerleri (2020) S&P 500 şirketleri için Twitter verilerinin gün içi gün içi tahmin değerini olay tabanlı geriye dönük test yaklaşımı [event-based backtesting approach] kullanarak kapsamlı bir şekilde incelemeyi amaçladıkları çalışmada üç aşamalı bir analiz gerçekleştirmişlerdir. Analizde 1 Ocak 2014'ten 31 Aralık 2015'e kadar tüm S&P 500 şirket

adları arama terimleri olarak kullanılmış ve 1.042.670 tweet toplanmıştır. Metinleri Natural Language Toolkit ile işlemişler ve duygu analizini ise R programı için geliştirilmiş olan duygu analiz kütüphanesi aracılığıyla yapmışlardır. Analizin ilk aşamasında metinsel veri ile hisse senedi getirileri arasındaki işlevsel ilişkiyi yön göstergeleri [directional indicators], alaka göstergeleri [relevance indicators] ve meta özellikleri [meta features] kullanarak sinyal-gürültü ayırımına dayanan bir makine öğrenimi sistemi geliştirmişlerdir. Gürültülü tweetler bu şekilde işlenebilir hale getirildikten sonra rassal orman (RF), sinir ağı (NN), K-en yakın komşu (KNN) ve naif [naive] duygu tabanlı model (N) yöntemleri kullanılarak mesajın potansiyel yönü ve alaka düzeyi belirlenmiştir. İkinci aşamada işlem maliyetleri ve sermaye gereksinimleri gibi kısıtları da göz önünde bulundurarak geriye dönük bir test yapmışlardır. Tweet tabanlı alım satım stratejisinin, işlem maliyetlerini hesaba kattıktan ve hisse senedi risklerini makul bir seviyeyle sınırladıktan sonra, 2,2'ye yakın bir Sharpe oranıyla yüzde 6,4'lük istatistiksel olarak anlamlı yıllık getiri sağlayabildiği bulgusuna ulaşmışlardır. Üçüncü aşamada getirilerin tweetlerin zamansal dağılımı tarafından yönlendirildiğini ortaya koymuşlar ve kârların çoğunun zaman içinde birbirine yakın kümelenmiş tweetlerden kaynaklandığı bulgusuna ulaşmışlardır. Ayrıca olumlu çağrışım içeren tweetlerin genellikle hisse senedi fiyatlarında bir artış, olumsuz çağrışım içeren tweetlerin genellikle hisse senedi fiyatlarında bir düşüş izlediği sonucunu elde etmişlerdir. Son olarak farklı elde tutma süreleri göz önünde bulundurularak haberlerin piyasa fiyatlarına yansımaya bakmışlar, bilginin kademeli olarak yayıldığı ve açığa satışları tetikleyen tweetler için daha güçlü ve daha kalıcı bir fiyat reaksiyonu olduğu sonucuna ulaşılmıştır.

3.3.1.2 Türkiye Hisse Senedi Piyasası İçin Yapılan Çalışmalar

Eliçık ve Erdoğan (2018), etkili kullanıcıların [influential user] konuya dayalı bir mikroblog topluluğunun duygu kutupluluğu üzerindeki etkisinin finansal piyasalarla ilişkisini araştırmak amacıyla yaptıkları çalışmada, BloombergHT and CNBCE'nin Twitter hesaplarına ve bu hesapları takip edenlere ait 23 Şubat 2015 ve 22 Şubat 2016 tarihleri arasında 2.044.967 Türkçe tweet ve sosyal ağ verisini haftalık olarak toplamışlardır. Her topluluk üyesinin etki düzeyini hesaplamak için PageRank algoritması kullanmışlar,

topluluğun duygusal yönelimini buradan elde edilen skor ile ağırlıklandırmışlardır. Çalışmada finansal sosyal çevrenin duygusal durumunu, hem PageRank tabanlı farklı algoritmalarla genişletilmiş duygu analizi yöntemi ile hem de yalnızca metin içeriğini dikkate alan klasik duygu analizi yöntemi ile belirlemişlerdir. Daha sonra elde edilen duygu skorları ile BIST 100 endeksinin haftalık değişimi arasındaki ilişkiyi Pearson korelasyon katsayısı yöntemi ile analiz ederek sonuçları karşılaştırmışlardır. Araştırmanın sonucunda PageRank benzeri algoritmaları kullanan etkili kullanıcı ağırlıklı duygu analizinin, yalnızca metinsel özelliklere dayalı klasik duygu analizine kıyasla daha başarılı olduğu bulgusuna ulaşmışlardır.

Bozma ve Kul (2020), tweetlerin duygu analizi ile hisse senetleri arasındaki ilişkiyi analiz etmek amacıyla yaptıkları çalışmada Borsa İstanbul'da listelenen üç telefon markasıyla (Alcatel, Turkcell ve Vestel) ilgili tweetleri Şubat 2017 – Temmuz 2017 tarihleri arasında 6 aylık bir dönemi kapsayacak şekilde toplamışlardır. Duygu sınıflandırmasını Naive Bayes yöntemiyle yapmışlardır. Daha sonra duygu puanlarının hisse senedi fiyat volatilitesine etkisini BEKK-GARCH(1,1)-X yöntemi ile incelemişlerdir. Analizin sonucunda Turkcell ve Vestel için elde edilen duygu puanlarının Alcatel'in koşullu varyansını istatistiki olarak anlamlı bir şekilde artırdığını tespit etmişlerdir.

Othan ve Kilimci (2021), yüksek hacimli hisse senetleri ile ilgili tweetleri inceleyerek borsanın yönünün tahmin edilip edilemeyeceğini araştırmak amacıyla yaptıkları çalışmada Borsa İstanbul'un BIST100 endeksine odaklanmışlardır. İşlem hacmi en yüksek 9 banka ile ilgili 1 Eylül 2018 - 1 Eylül 2019 tarihleri arasında 117.136 Türkçe tweet toplamışlardır. Twitter duygusunu Word2Vec, GloVe, FastText, CNN, RNN, LSTM, BERT, ELMo ve ULMFiT kelime gömülme modelleri ile sınıflandırmanın BIST100'ün yönünü tahmin etme başarısını kıyaslamışlardır. Analizin sonucunda ELMo'nun Twitter veri setinde yeni nesil derin bağlamsal kelime gömülme modeli olarak kullanımının, borsa yönünü tahmin etmede dikkate değer sonuçlar sergilediği bulgusuna ulaşmışlardır.

Yeşiltaş ve diğerleri (2021), Türkiye’deki ekonomik politika belirsizliğini ölçmek amacıyla yaptıkları çalışmada Twitter tabanlı bir toplumsal duygu [public sentiment] ölçümüne odaklanmışlardır. Tweetlerinin halk tarafından konuyla ilgili uzman görüşünü yansıttığı düşünülen, seçilmiş bir dizi Twitter kullanıcı hesabına dayalı olarak Twitter tabanlı bir Ekonomik Politika Belirsizliği (TEPU) Endeksi oluşturmuşlardır. Analizin sonucunda TEPU Endeksinin CDS ile %82, VIX ile %69, BIST 100 ile %33 ve TRY-USD kuru ile %27 korelasyona sahip olduğu bulgusuna ulaşmışlardır.

Ateş ve Güran (2021), BIST30 endeksiyle ilgili tweetlerden pozitif, negatif ve nötr olarak elde edilen üç farklı duygu değeri ile endeksin günlük değer değişimleri arasındaki korelasyon ve Granger nedensellik ilişkilerini araştırmak amacıyla yaptıkları çalışmada 07.05.2018 - 30.04.2019 tarih aralığında paylaşılan tweetlere odaklanmışlardır. Analizde pozitif, negatif ve nötr olarak etiketlenmiş 57.933 kısa iletiyi içeren bir eğitim veri kümesi, aynı döneme ait 200.000 etiketsiz veri kümesi ve BIST30 endeksi günlük değer değişimlerini kullanmışlardır. Hem uzun dönem hem de kısa dönem için gerçekleştirilen korelasyon analizlerinde, tweetlerden elde edilen duygu değerleri ve BIST30 endeksi hisselerinin günlük değer değişimleri arasında orta düzeyde ve pozitif yönde bir korelasyon ilişkisi olduğu bulgusuna ulaşmışlardır. Daha sonra Granger nedensellik analizi yapmışlar, uzun dönem için gerçekleştirilen analizlerde endeksin değer değişiminden duygu değerlerine doğru tek yönlü bir Granger nedenselliği olduğu sonucuna varmışlardır. Öte yandan duygu değerlerinden endeksin hareket yönüne doğru herhangi bir Granger nedensellik ilişkisine rastlamadıklarını raporlamışlardır. Ayrıca, özellikle olağan dışı olayların gerçekleştiği dönemlerdeki duygu değerlerinden BIST30 değer değişimlerine doğru bir Granger nedensellik ilişkisinin ortaya çıktığı sonucuna ulaşmışlardır.

Sarıyer ve diğerleri (2022) şirketlerin kamuya yaptıkları açıklamalardan ve tweetlerden gelen duyguları kullanarak bireysel hisse senedi fiyatını ve hacim değişikliklerini makine öğrenmesi yöntemiyle tahmin etmeyi amaçladıkları çalışmada, 1 Ocak 2016 ile 1 Ocak 2021 tarihleri arasında BIST30 şirketleri ile ilgili toplam 1.494.108 tweet, Kamuyu Aydınlatma Platformundan (KAP) 20.667 özel durum açıklaması toplamışlardır. Hem KAP hem de

Twitter verilerinin duygu puanını BERT modeli ile hesaplamışlardır. Tahmin doğruluğunu artırmak ve piyasaların genel durumunu ve risk iştahını modele entegre etmek için BIST30, DJIA, USD/TRY kuru ve ons altının günlük fiyatlarından oluşan piyasa koşullarına ilişkin finansal verileri de analize dahil etmişlerdir. Dokuz farklı makine öğrenmesi yöntemiyle yaptıkları analizin sonucunda sosyal medyada varlığı [presence] ve kamuyu aydınlatma sayısı yüksek olan şirketler için maksimum %67 bireysel hisse senedi fiyatı tahmin doğruluğu, BIST30 şirketleri için %74,7 ortalama hacim tahmin doğruluğu elde etmişlerdir.

3.3.1.3 Diğer Ülke Hisse Senedi Piyasaları İçin Yapılan Çalışmalar

Gomez-Carrasco ve Michelon (2017), toplumsal hareket teorisini³⁸ temel alarak sosyal medya aktivizminin hedeflenen firmaların borsa performansı üzerindeki etkisini değerlendirmek amacıyla yaptıkları çalışmada Twitter'da paylaşım yapan tüketici dernekleri ve sendikalara odaklanmışlardır. İspanyol Menkul Kıymetler Borsasında listelenen sekiz bankayı kapsayan çalışmada, 14 Kasım 2013 - 19 Mayıs 2014 tarihleri arasında 127 işlem günü için bu bankalara atıfta bulunan 1.534.435 tweet toplamışlardır. Protesto ifadesi olarak olumsuz bir ton sunan tweetleri manuel olarak sınıflandırmışlardır. Daha sonra her bir firma ve gün için toplam tweet sayısı, hashtag içeren toplam tweet sayısı, bir web sitesine bağlantı içeren toplam tweet sayısı, toplam retweet sayısı, tweet gönderen Twitter hesaplarının takipçi sayısı gibi verileri de analize dâhil etmişlerdir. Ayrıca paydaşların tweetleri sendikalar ve sivil-tüketici dernekleri olarak da iki sınıfa ayırmışlardır. Bulgular, tüketici dernekleri ve sendikalar gibi kritik paydaşların Twitter aktivizminin yatırımcıların kararları üzerinde önemli bir etkisi olduğunu göstermektedir. Spesifik olarak, sendikaların tweetlerinin hem hisse senedi fiyatı hem de işlem hacmi üzerinde önemli bir etkisi olduğu, ancak sendika takipçilerinin sayısının hisse senedi fiyatlarını veya işlem hacmini etkilemediği sonucuna ulaşmışlardır. Bu sonuçlar yatırımcıların, görünürlüğünden çok sendika tweetlerinin içeriğiyle ilgilendiğine, takipçileri az da olsa tweetlerine tepki verdiklerine işaret etmektedir.

³⁸ Toplumsal hareketler [social movement], toplumun yapısında veya değerlerinde değişim gerçekleştirmek isteyen veya meydana gelebilecek değişimi engellemeye çalışan gevşek bir şekilde organize edilmiş oluşumlardır.

Sivil demeklerin ve tüketici derneklerinin etki mekanizmaları ise bunun tersidir. Bu kuruluşlar yatırımcıların kararlarını, paylaşımlarının yoğunluğundan ziyade görünürlükleri ve sosyal medya üzerindeki nüfuzlarıyla etkilemektedir.

Bharathi ve diğerleri (2017) Amman Menkul Kıymetler Borsasında (ASE) işlem gören bir şirketin hareketli ortalama, MACD, stokastik RSI gibi göstergeler ile ilgili tweetlerdeki toplumsal duygu durumu [public sentiment] ve RSS ³⁹ yayınları arasındaki ilişkiyi araştırdıkları bir çalışmada RSS yayınları ve tweetlerden ölçülen duygularla alış satış sinyallerini tahmin etmenin mümkün olduğu bulgusuna ulaşmışlardır.

Tiwari ve diğerleri (2017) zaman serisi modeli, yapay sinir ağları, regresyon gibi çeşitli yöntemleri kullanarak en büyük Hint şirketlerinin yer aldığı NIFTY 50 endeksindeki firmaların 9 yıllık verisinin tweetler de dâhil edildiği çalışmada düğümleri artırılmış yapay sinir ağlarının hisse senedi fiyat tahmininde ve trend tahmininde daha iyi sonuçlar ürettiği bulgusuna ulaşmış, özellikle ileri beslemeli yapay sinir ağlarının [feed forward neural network] hisse senedi açılış fiyatlarının tahmininde en yüksek doğruluğu sağladığına dikkat çekmişlerdir. Ayrıca farklı hisse senetleri için farklı yöntemlerin daha iyi sonuç verebileceğini ve bu durumun farklı fiyat türleri (açılış, düşük, yüksek ve düşük gibi) için de geçerli olduğunu ortaya koymuşlardır.

Hegde ve diğerleri (2018) hisse senedi fiyatlarını etkileyen tarihsel hisse senedi verilerini, tweetleri ve haberleri dikkate alan ve belirli bir süre için hangi hisse senetlerine yatırım yapılacağına dair öneriler sunan bir model geliştirmek amacıyla yaptıkları çalışmada, Hindistan Ulusal Menkul Kıymetler Borsası'nda (NSE) işlem gören firmalara odaklanmışlardır. Yinelemeli sinir ağları (RNN) ve uzun ömürlü kısa dönem bellek (LSTM) ve evrimsel sinir ağları (CNN) yöntemlerini kullanan modellenmiş, 50 şirketin hisse senedinin gelecekteki eğilimini tahmin ederek, belirli bir şirket hissesine ne zaman yatırım yapılacağına dair öneriler sağlamak için hissenin trendinin zirve ve dip noktaları üzerine eğitmişlerdir.

³⁹ Genellikle haber sağlayıcıları tarafından kullanılan ve yeni eklenen içeriğin kolaylıkla takip edilmesini sağlayan bir web sayfası bildiricisidir.

Modelde tweetlerin ve haber başlıklarının duygusal olarak sınıflandırılması için eğitimi LSTM yöntemiyle gerçekleştirmişlerdir. Model bir haftalık gerçek hisse senedi fiyatları ile test edildiğinde, hisse senedi eğilimlerini tahmin etmede yüksek doğruluk skoru elde ettiği sonucuna ulaşmışlardır.

Zhang ve diğerleri (2018) kurumsal yatırımcılardan ziyade bireysel yatırımcıların daha fazla olduğu Çin hisse senedi piyasasında [Shanghai Stock Exchange (SHCOMP)] Twitter benzeri bir sosyal medya platformu olan Xueqiu'nun etkisini araştırmak amacıyla yaptıkları çalışmada kolektif duygu [collective sentiment] ve algı ile ilgili özelliklere odaklanmışlardır. Kasım 2014 – Mayıs 2015 tarihleri arasında 6.8 milyon tweet toplamışlar, Naive Bayes yöntemiyle pozitif, negatif ve nötr olarak sınıflandırmışlardır. Daha sonra sosyal ağ verileri kullanılarak her bir kullanıcının PageRank skorunu hesaplamışlar ve bu PageRank skorunu tweetlerin duygusal yönelimini ağırlıklandırmak için kullanmışlardır. Bu değişkenlerle hisse senedi fiyat hareketlerini tahmin etmek için destek vektör makineleri (SVM) ve çok katmanlı algılama (MLP) yöntemlerini kullanmışlar ve araştırmanın sonucunda Xueqiu'dan elde edilen kolektif duygunun Çin borsasındaki hisse senedi hareketleriyle ilişkilendirilebileceği bulgusuna ulaşmışlardır. Ayrıca sonuçlar, tahmin performansının, duygu ve algılanan hisse senedi ile ilgili özelliklerin dahil edilmesiyle geliştirilebileceğini göstermektedir.

Nisar ve Yeung (2018), sosyal medyadan ölçülen siyasetle ilgili duygunun borsa ile ilişkini olay çalışması yöntemiyle analiz ettikleri çalışmada İngiltere'de 2016 yerel seçimleri öncesi, sırası ve sonrasını içeren 6 gün boyunca 3 anahtar hashtag kullanarak 60.000'den fazla tweet örneği toplanmışlardır. Tweetlerdeki duyguyu sözlük tabanlı duygu sınıflandırıcısı ile kategorize ettikten sonra günlük ruh hallerinin Londra Borsası'nın (FTSE 100) fiyatlarındaki günlük değişimlerle ilişkisini incelemek için korelasyon ve regresyon analizleri yapmışlardır. Araştırmanın sonucunda kısa vadede toplumsal ruh hali ile yatırım davranışı arasında ilişki olduğu ancak istatistiksel olarak anlamlı olmadığı bulgusuna ulaşmışlardır. Ayrıca günlük kapanış fiyatı ve gecikme süresi bulguları açısından toplumsal ruh hali ile borsa hareketleri arasında nedensellik kanıtı olduğunu da ortaya koymuşlardır.

Groß-Klußmann ve diğerleri (2019), sosyal medya paylaşımlarından elde edilen sinyaller ile beş büyük ekonomik bölgenin finansal piyasa gelişmeleri arasındaki uzun vadeli ilişkiyi incelemek amacıyla yaptıkları çalışmada beş büyük ekonomik bölgeyi temsilen Avustralya için ASX200, Çin için HangSeng, Avrupa için EURO STOXX 50, Japonya için Nikkei225 ve ABD için S&P 100 hisse senedi endekslerini belirlemişlerdir. Çalışmada 01 Ocak 2010 - 30 Eylül 2018 dönemine ait dokuz yıllık Twitter verisi içerisinde finansla ilgili 102 milyon tweet içeren bir alt veri seti oluşturmuşlardır. Tweetlerdeki duyguları sınıflandırmak için sözlük yaklaşımını benimsenmişler, yani duygu çıkarımını pozitif ve negatif sözcüklerle yapmışlardır. Ayrıca Twitter mesajlarındaki gürültüyü gidermek ve sinyal kesinliğini artırmak için gözetimli ve gözetimsiz istatistiksel öğrenme yöntemlerinden k-means, LDA ve destek vektör makinelerini kullanarak tüm kullanıcıları uzmanlar ve uzman olmayanlar şeklinde gruplandırmışlardır. Analizin sonucunda uzman kullanıcıların Twitter'daki duygusal yönelim ve finansal piyasalar arasındaki karşılıklı bağımlılığın arkasındaki ana itici güç olmakla birlikte, hisse senedi endeksi getirileri için uzman duygu ölçümlerinin doğrudan tahmin değerinin anlaşılması zor ve kısa ömürlü olduğu bulgusuna ulaşmışlardır. Fakat yine de negatif piyasa getirilerinin olduğu zamanlarda kıyaslama modellerine göre önemli tahmine dayalı kazançlar elde edilebildiğini tespit etmişlerdir. Uzman duygu ölçütlerinin büyük ekonomik bölgelerdeki geniş finansal gelişmeleri iyi bir şekilde yakaladığını ve ilgili hisse senedi piyasası endeksi vadeli işlemleri için kârlı yatırım kararlarını destekleyebileceğini ortaya koymuşlardır. Başka bir deyişle uzman duygu sinyallerinin klasik fiyat bazlı sinyallerden daha yüksek riske uyarlanmış getiri [risk-adjusted returns] sağlayabildiğini göstermişlerdir.

Pasupulety ve diğerleri (2019), Hindistan Ulusal Menkul Kıymetler Borsası'nda (NSE) yer alan firmalar üzerine yapılan bir çalışmanın ilk aşamasında destek vektör makineleri ve rassal orman yöntemleriyle teknik göstergelere dayalı bir özellik seçimi [feature selection] yapmışlar, daha sonra Word2Vec modelini kullanılarak Twitter'dan gelen şirkete özgü şirket sembolü etiketli gönderileri, pozitif veya negatif olarak sınıflandırmışlardır. Çalışmada tahmin modelini, teknik gösterge verilerini bir şirketin toplam pozitif ve negatif tweet sayısı ile birleştiren yeni bir veri seti ile eğitmişlerdir. Analiz sonucunda teknik gösterge verilerini

pozitif ve negatif tweet sayılarıyla birleştirmenin, topluluk modelinin [ensemble model] performansı üzerinde ihmal edilebilir bir etkisi olduğu bulgusuna ulaşmışlardır. Ancak bu sonucun elde edilmesinde, alana özgü etiketli bir eğitim veri seti kullanmak yerine, Stanford Üniversitesi'nin sinema yorumları veri setinin kullanılmış olmasının etkili olabileceğini ifade etmişlerdir.

Duz Tan ve Taş (2020), sosyal medyanın ABD, Avrupa ve gelişmekte olan piyasalar için 2015-2017 örneklem dönemi arasında S&P 500, S&P 350 Avrupa ve S&P Gelişen Piyasalar endeks bileşenleri üzerindeki etkisini araştırmak amacıyla yaptıkları çalışmada firmaya özgü Twitter duygusu ve aktivitesi ile işlem hacmi ve getiriler üzerindeki etkisini uluslararası yatırımcı bakış açısıyla analiz etmişlerdir. Bloomberg'den⁴⁰ aldıkları Twitter ve StockTwits duygu verileri ile uluslararası borsalardan aldıkları hisse senedi verilerini kullanarak yaptıkları analizin sonucunda, günlük Twitter etkinliğinin ve duygusal yönelimin işlem hacmiyle ilişkili olduğu ve sonraki gün işlem hacmini tahmin ettiği, ayrıca günlük tweet sayısının ve duygusal yönelimin daha yüksek hisse senedi getirileri ile ilişkili olduğu bulgusuna ulaşmışlardır. Buna ek olarak günlük firmaya özgü tweetlerdeki duygusal yönelimin, gelecekteki hisse senedi getirilerini tahmin etmek için bilgi içerdiğini, ancak tweet sayısında ile böyle bir ilişkinin olmadığını ortaya koymuşlardır. Çalışmada ayrıca Twitter'daki duygusal yönelimin olumlu tonunun, küçük ve gelişmekte olan piyasa firmalarında daha fazla tahmin gücüne sahip olduğunu tespit etmişlerdir.

Chun ve diğerleri (2020) mikroblog⁴¹ paylaşımlarındaki duygu analizine dayalı olarak Kore'deki hisse senedi vadeli işlemleri piyasasının (KOSPI 200) yönünü tahmin etmek amacıyla yaptıkları çalışmada 9 Mart 2012'den 8 Aralık 2016 tarihlerini arasında hisse senetleriyle ilgili olarak paylaşılan 1.602.880 mesaj toplanmışlardır. Toplanan bu mesajları Kore duygusal sözlüğünü esas alarak yedi duygu sınıfına ayırmışlardır. Duygu göstergeleri

⁴⁰ Bloomberg Nisan 2013'te Twitter paylaşımlarını kendi platformuna entegre etmiştir. Bloomberg, şirketlerle ilgili paylaşımları Cashtag'ler yardımıyla izlemek için hem StockTwits hem de Twitter'dan gelen ham mesaj akışlarını girdi olarak kullanmakta ve ayrıca tescilli bir doğal dil işleme algoritması ile her tweeti sınıflandırmaktadır.

⁴¹ Twitter benzeri bir sosyal medya platformu olan paxnet.co.kr kastedilmektedir.

daha sonra pozitif duygu göstergeleri (neşe, ümit, hayret) ve negatif duygu göstergeleri (iğrenme, öfke, üzüntü, korku) olarak ikiye grupta toplanmıştır. Pozitif duygu göstergeleri ve negatif duygu göstergelerinin model bağımsız değişkenler olarak kullanıldığı ve öğrenme süresine göre 3 ay, 6 ay, 9 ay ve 12 aylık dönemleri kapsayan çalışmalar yapmışlardır. Analizde sonucunda, yükselme ve düşme tahminleri sunan sistemin doğruluk skoru tüm öğrenme dönemleri için %90,74'ün üzerinde iken, en iyi tahmin doğruluk skoru ise %96,67 olarak gerçekleştiğini raporlamışlardır.

Hassanein ve diğerleri (2021) piyasa değerine göre dünyanın en büyük 100 firmasının tweetlerine ilişkin yatırımcıların düşüncelerine odaklandıkları çalışmada 3 Eylül - 8 Eylül 2019 tarihleri arasındaki dönemde 423.495 tweet toplamışlar, tweetlerin duygusal yöneliminin sınıflandırılması için Naive Bayes yöntemini kullanmışlardır. Duygu analizine dayalı olarak piyasa yönünü tahmin etmek için yaptıkları analizin sonucunda Twitter duygularının piyasa kümülatif normalüstü getirisini ve al-ve-tut getirisini etkilediği ve bu etkinin kısa bir süre devam daha uzun bir süre içinde azalacağı bulgusuna ulaşmışlardır. Ayrıca duyguların etkisinin örfe dayalı hukuk [common law] sistemine sahip ülkelerde medeni hukuk [civil law] sistemine sahip ülkelere kıyasla daha gözlemlenebilir olduğunu ortaya koymuşlardır.

3.3.1.4 Diğer Piyasalar İçin Yapılan Çalışmalar

Zhang ve diğerleri (2012) Twitter gönderilerini analiz ederek altın fiyatı, ham petrol fiyatı, döviz kurları ve borsa göstergeleri gibi finansal piyasa hareketlerini tahmin etmek amacıyla yaptıkları çalışmada 15 Kasım 2010 - 20 Nisan 2011 tarihleri arasında “dolar”, “\$”, “altın”, “petrol”, “iş”, “ekonomi” anahtar kelimelerini kullanarak ilgili retweetleri toplamışlardır. Analizin sonucunda ekonomi ile ilgili retweetlerin günlük hacmindeki değişikliklerin ertesi güne karşılık gelen piyasada değer kaymasıyla eşleştiği bulgusuna ulaşmışlardır. Ayrıca “\$” dışında, diğer beş anahtar kelimeye dayalı zaman serisi ile seçilen varlık değerlendirme hareketleri arasında Granger nedenselliği olduğu sonucuna varmışlardır.

Rao ve Srivastava (2013), sosyal medya ve webde arama davranışlarından finansal içgörüler elde amacıyla yaptıkları çalışmada sosyal medya paylaşımlarına ve finansal piyasa araçlarıyla ilgili yapılan aramaların hacimlerine odaklanmışlardır. 2 Haziran 2011'den 13 Eylül 2011 tarihine kadarki zaman diliminde yaklaşık 2.000.000 tweet toplamışlar ve sınıflandırma doğruluğu %82,7 olan Naive Bayesian sınıflandırıcı ile pozitif ve negatif olarak sınıflandırmışlardır. Arama davranışlarıyla ruh halini bütünleştirerek petrol, altın, Euro ve hisse senedi piyasaları ile arasındaki ilişkiyi araştırmışlar, analizin sonucunda hem emtia, hem de hisse senedi için hem değişen oranların hem de yükselme ve düşüş olasılıklarının tahmin edilebileceği, dolayısıyla portföy kararlarında kullanılabileceğini bulgusuna ulaşmışlardır.

Kaminski (2014), Twitter paylaşımlarındaki duygusal sinyaller ile Bitcoin piyasası arasındaki ilişkiyi ve nedenselliği incelemek amacıyla yaptığı çalışmada, duygu göstergelerini temsilen olumlu, olumsuz ve belirsizlik durumları ile ilgili anahtar sözcükler belirlenmiş, 23 Kasım 2013 - 7 Mart 2014 tarihleri belirli terimleri içeren yaklaşık 160.000 Twitter gönderisi toplamışlardır. Bu terimlerden “mutlu”, “aşk”, “eğlenceli”, “iyi”, “kötü”, “üzgün” ve “mutsuz” terimleri olumlu ve olumsuz duygusal sinyalleri temsil ederken, “umut”, “korku” ve “endişe” belirsizliğin göstergesi olarak değerlendirilmiş, ayrıca “bitcoin” anahtar kelimesiyle de arama yapmışlardır. Daha sonra durağan bu verileri günlük bazda Pearson korelasyon testi ile analiz etmişler, duygu sinyalleri barındıran tweetler ile Bitcoin'in gün içi fiyat aralıkları [price spread], işlem hacmi ve kapanış fiyatları arasında pozitif bir ilişki olduğu bulgusuna ulaşmışlardır. Ayrıca dinamik Granger nedensellik analizi yapmışlar, duygusal Tweetlerin Bitcoin piyasa değerleri üzerinde istatistiksel olarak anlamlı bir etkisi olduğuna dair herhangi bir kanıtı ulaşamamışlardır. Aksine analiz, daha yüksek Bitcoin işlem hacminin Granger'ın 24 ila 72 saatlik bir zaman diliminde daha fazla belirsizlik sinyaline neden olduğuna işaret etmektedir. Dolayısıyla duyguların piyasayı öngörülebilir kılmaktan çok onu yansıttığı sonucuna ulaşmışlardır.

Matta ve diğerleri (2015) Bitcoin fiyatının tweet hacimleriyle veya web arama davranışlarıyla ilişkili olup olmadığını araştırmak amacıyla yaptıkları çalışmada Ocak 2015

ile Mart 2015 arasındaki 60 günlük bir dönemi kapsayacak şekilde 1.924.891 tweet toplanarak duygusal yönelimlerini sınıflandırmışlardır. Daha sonra fiyat verilerini Google trend verileriyle, tweet hacmiyle ve özellikle olumlu bir duygu ifade edenlerle karşılaştırmışlar, özellikle Bitcoin fiyatı ile Google trend verileri arasında önemli çapraz korelasyon olduğu bulgusuna ulaşmışlardır.

Oussalah ve Zaidi (2018), haftalık ham petrol fiyatlarının yönünü tahmin etmede Twitter'ın rolünü araştırdıkları çalışmada ABD dış politikası ve petrol şirketleri ile ilgili duyguyu kullanan bir model geliştirilmiştir. Analiz için öncelikli olarak ABD dış politikası ve petrol şirketleriyle ilgili tweetleri toplanmışlar, daha sonra tweetlerin duygusal kategorisini SentiStrength ve Stanford NLP Sentiment araçları ile belirlemişlerdir. İstatistiksel olarak anlamlı korelasyona sahip özellikleri tahmin modelinde girdi olarak kullanmışlardır. Tahmin modelini Naive Bayes, SVM ve ANN olmak üzere üç farklı sınıflandırıcı ile test etmişler, SVM'nin %74 ile en yüksek sınıflandırma doğruluğunu verdiği bulgusuna ulaşmışlardır.

3.3.2 Twitter Dışındaki Diğer Sosyal Medya Mesajlarının Kullanıldığı Çalışmalar

Das ve Chen (2007), hisse senedi mesaj panolarından küçük yatırımcı duygu puanını çıkarmak için bir metodoloji geliştirmek amacıyla yaptıkları çalışmada Morgan Stanley Yüksek Teknoloji Endeksi'nde (MSH) bulunan 24 teknoloji sektörü hissesine odaklanmışlar, Temmuz ve Ağustos 2001 olmak üzere iki aylık bir süre boyunca mesaj panolara gönderilen 145.110 mesaj toplamışlardır. Analizin toplu duygu göstergeleri ile endeks getirileri arasındaki ilişkinin, bireysel hisse senetleri ile korelasyondan çok daha güçlü olduğu bulgusuna ulaşmışlardır. Ayrıca mesaj sayısı ile piyasa getirileri arasında eşzamanlı bir ilişki olduğu, ancak bir öngörü ilişkisinin olmadığı sonucunu da ortaya koymuşlardır.

Siganos ve diğerleri (2014), Facebook'taki durum güncellemelerinin metin analizine dayalı alternatif bir duygu ölçütü olan Facebook'un Gayri Safi Milli Mutluluk Endeksinden [FGNHI] yararlanarak duygu ve borsa getirileri arasındaki ilişkiyi analiz etmek amacıyla yaptıkları çalışmada uluslararası piyasalara odaklanmışlardır. Analizde Facebook'un günlük

duygu endeksinin yirmi uluslararası piyasada⁴² hisse senedi getirileri, işlem hacmi ve hisse senedi fiyat volatilitesi ile ilişkisini incelemişlerdir. Günlük düzeyde yapılan analizin sonucunda, duygu durumunun borsa getirileri ile pozitif ilişkili, işlem hacmi ve volatilitesi ile negatif ilişkili olduğu bulgusuna ulaşmışlardır. Ayrıca, olumsuz duyguların işlem hacmindeki ve getiri volatilitesindeki artışlarla ilişkili olduğunu ortaya koymuşlardır. Son olarak duygu ile getiri arasındaki ilişkinin ilerleyen haftalarda tersine döndüğünü ve temel değerlerde bir düzeltmeye işaret ettiğini gözlemlemişlerdir. Siganos ve diğerleri (2017) bir başka çalışmalarında, pozitif ve negatif duygu kutuplarının ıraksamasının borsa hareketleri ile ilişkisini analiz etmek amacıyla Kasım 2007 ve Mart 2012 tarihleri arasında 20 uluslararası piyasa için Facebook durum güncellemelerinden elde edilen pozitif ve negatif duygulara ilişkin günlük verileri toplamışlardır. Analizin sonucunda pozitif ve negatif duygu kutuplarının ıraksamasının, hem işlem hacmi ile hem de hisse senedi fiyat volatilitesi ile pozitif ilişkili olduğu bulgusuna ulaşmışlardır. Çalışmada ayrıca duygu ayrışmasının etkilerini hem ülke düzeyinde ve hem de küresel düzeyde karşılaştırmışlar, duygu ıraksamasının fiyat oynaklığı ile ilişkisinin her iki düzeyde de olduğu sonucuna varmışlardır. Ancak piyasaların daha entegre olması durumunda yerel etkinin önemli düzeyde zayıfladığı bulgusuna ulaşmışlardır.

Rautiainen ve Jokinen (2017), şirketlerin birden çok sosyal medya aracı kullanması ile hisse senedi fiyat değişimleri arasındaki bağlantıları analiz etmeyi amaçladıkları çalışmalarında Finlandiya borsasına kayıtlı şirketlerin 31 Aralık 2017 - 31 Aralık 2018 tarihleri arasında Facebook (FB), Instagram, LinkedIn, Twitter ve YouTube etkinliklerine odaklanmışlardır. Nasdaq Helsinki Borsasında işlem gören 105 Finlandiya şirketinin seçildiği çalışmada, şirketlerin sosyal medya faaliyetlerinin hisse senedi fiyatlarındaki değişim ile bağlantılı olduğu bulgusuna ulaşmışlardır. Özellikle Instagram takipçileri, Twitter takipçileri ve şirket tarafından yayınlanan YouTube videolarının sayısı ile hisse senedi fiyat değişimini açıklayan doğrusal regresyon modelinin R^2 değerini yaklaşık %39 olarak bulmuşlardır. Bununla

⁴² Bu ülkeler Arjantin, Avustralya, Avusturya, Belçika, Kanada, Şili, Kolombiya, Almanya, Hindistan, İrlanda, İtalya, Meksika, Hollanda, Yeni Zelanda, Singapur, Güney Afrika, İspanya, Birleşik Krallık, Amerika Birleşik Devletleri ve Venezuela'dır.

birlikte pozitif ya da negatif duygunun hisse senedi fiyat değışikliklerini açıklamada anlamlı olmadığı sonucuna varmışlardır.

Teplova ve diğeri (2022), bireysel yatırımcıların sosyal medyadaki duygularının Rusya hisse senedi piyasasına etkisini analiz etmek amacıyla yaptıkları çalışmada Telegram ve mfd.ru platformlarındaki mesajlara odaklanmışlardır. Mesajların duygusal tonunu geliştirdikleri algoritma ile sınıflandırmışlar, 2013'ten 2020'ye kadarki dönem için aylık getiri oranları ve işlem hacimlerini analize dâhil etmişlerdir. Analizin sonucunda sosyal medyadan ölçülen duygunun fiyat ve işlem etkinliklerinde açıklayıcı bir faktör olduğu bulgusuna ulaşmışlardır. Ayrıca hem olumlu hem de olumsuz duygu mesajlarının sayısındaki artışın işlem hacminin artmasına katkıda bulunduğu sonucuna varmışlardır.

3.4 LİTERATÜR DEĞERLENDİRMESİ

Tweetler ile finansal piyasalar arasındaki ilişkileri araştıran çalışmalar analiz kapsamına göre değerlendirildiğinde genel olarak üç grupta toplanabilirler. Bu gruplardan biri tweetler ile finansal piyasalar arasındaki ilişkileri makroskobik ölçekte ele alan çalışmalardan oluşmaktadır. Bu çalışmaların ilk versiyonlarında tweetlerin borsa ile ilgili olup olmadığına bakılmaksızın, toplumsal ruh hali ile borsa endekslerinin yönü tahmin edilmeye çalışılmıştır. Sonraki çalışmalarda ise anahtar kelimeler yardımıyla sadece borsa ile ilgili tweetler filtrelenerek bunlara odaklanıldığı görülmektedir.

İkinci grupta yer alan çalışmaların tweetlerdeki duygu düzeyi ile borsa göstergeleri arasındaki ilişkileri mikroskobik ölçekte ele aldığı görülmektedir. Bu tür çalışmalarda belirli hisse senetleri seçilmiş ve bu hisse senetleri ile ilgili tweetlere odaklanılmıştır. Analize konu hisse senetleri ile ilgili tweetlerin fikri ve duygusal yönelimleri ile hisse senedi göstergeleri arasındaki ilişkiler analiz edilmiştir. Bu gruptaki çalışmalardan bazılarında ayrıca borsa ile ilgili tweetler paylaşan ve bu Twitter hesaplarını takip eden kullanıcılar da mercek altına alınmıştır.

Üçüncü grupta yer alan çalışmalar ise hem makroskobik hem de mikroskobik bakış açısıyla endeks düzeyinde, sektör düzeyinde ve hisse senedi düzeyinde analizler yapmışlardır. Ayrıca bu çalışmalardan bazıları hisse senetleri ile ilgili tweetler paylaşan Twitter hesaplarını, kullanıcı düzeyinde (mikro), topluluk düzeyinde (mezo) ve kolektif düzeyde (makro) analiz etmişlerdir.

Literatürdeki çalışmaların büyük çoğunluğunun ABD borsalarından DJIA, NASDAQ ve S&P 500 endeksleri ile bu endekslerde yer alan hisse senetleri için yapıldığı görülmektedir. Bunun temel nedeninin tweetlerin dili ile ilgili olduğu düşünülmektedir. Zira İngilizce yazılan tweetlerin fikri ve duygusal açıdan sınıflandırılabilmesi için gerekli olan araçlar, yazılım kütüphaneleri, leksikonlar ve etiketli veriler diğer dillere kıyasla çok daha fazladır.

Literatürdeki çalışmalara odaklandıkları zaman periyodu açısından bakıldığında ise yoğunlukla günlük periyotlarda analiz yapıldığı görülmektedir. Genellikle çalışmalar bir iki yıl ve bir kaç milyon tweet ile sınırlıdır. Ulaşılabildiği kadarıyla, hisse senedi düzeyinde analiz yapan sadece iki çalışmada 10 yıla yakın gözlem söz konusudur. Bu çalışmaların ise veri seti küçüktür. Örneğin birinde analize konu 10 hisse senedinden her bir hisse senedi için yaklaşık 100 bin tweete ulaşılmış, diğerinde 8 yıllık bir dönem için yaklaşık 20 milyon tweet analiz edilmiştir. Endeks düzeyinde yapılan ve 8 yıllık bir dönemi kapsayan bir başka çalışmada ise 102 milyon tweete ulaşılmıştır. Günlük periyottan farklı olarak gün içi veriler ile analiz yapan çalışmalarda mevcuttur. Ayrıca daha seyrek de olsa haftalık veri ile analizler yapan çalışmalar olduğu da görülmektedir.

Literatürdeki çalışmalarda karşılaşılan bir diğer durum ise fikir ve duyguların analizine ilişkin yaklaşım farklılıklarıdır. Çalışmaların büyük çoğunluğunun pozitif, negatif ve nötr şeklinde sınıflandırılan fikir ve duyguya [sentiment] odaklanırken bazılarının mutluluk, hüzn, endişe gibi duygu boyutlarına [emotion] odaklandığı görülmektedir. Bazı çalışmalarda ise sadece tweet sayıları kullanılmıştır.

Türkiye hisse senedi piyasası için yapılan çalışmalara bakıldığında ilk göze çarpan unsur tweetlerden elde edilen veriler ile Borsa İstanbul arasındaki ilişkiyi araştıran çalışmaların oldukça sınırlı sayıda olmasıdır. Bu durumun Türkçe için geliştirilmiş duygu sınıflandırıcıların İngilizce için geliştirilmiş olanlara kıyasla daha az olmasından kaynaklandığı düşünülmektedir. Bu çalışmalarda dikkat çeken diğer unsur ise yabancı piyasalar için yapılan çalışmalara kıyasla daha küçük veri setlerinin kullanılmış olmasıdır. Ayrıca bu çalışmalarda ya çok sınırlı sayıda (üç ila dokuz) hisse senedinin ya da iki endeksin (BIST 30 ve BIST 100) tweetlerden elde edilen duygu skorları ile ilişkisinin araştırıldığı görülmektedir.

Genellikle bulgular tweet sayısının, retweet sayının, takipçi sayısının ve tweetlerdeki duygusal yönelimin hisse senedi piyasaları ile ilişkili olduğunu ortaya koymaktadır. Bu bağlamda çalışmalardan bazıları endeks düzeyinde ilişkiler bulurken bazıları hisse senedi düzeyinde ilişkiler bulmuştur. Örneğin getiri ilişkisi daha çok hisse senedi düzeyinde yapılan çalışmalarda ya da olay çalışması türü analizlerde ortaya konulmuştur.

Bu çalışmayı literatürdeki çalışmalardan farklı kılan hususlar, literatürdeki çalışmaların kapsadığı dönem ortalamasının üstünde, **7 yıllık bir gözlem dönemini kapsaması**, bu dönem için **anakütleyle yakın bir veri setine**, yani borsa ile ilgili yaklaşık **1 milyar tweetten oluşan gözleme** sahip olması ve borsa ile ilgili tweetleri hem anahtar kelime yaklaşımı ile hem de makine öğrenmesi yaklaşımı ile sınıflandırmasıdır. Ayrıca Türkçe tweetlerin bağlamsal ve duygusal yöneliminin belirlenmesinde yeterli etiketli veri olmadığı için sınıflandırıcılar çeşitli Türkçe dijital platformlardan veri kazıma tekniği ile oluşturulmuş büyük bir derlem ile eğitilmiştir. Buna ek olarak özellikle fikir ve duygu sınıflandırıcının eğitimi için ön eğitilmiş modeller kullanılmış ve öğrenme aktarımı [transfer learning] yöntemi ile sınıflandırıcının başarısı artırılmıştır. Her iki sınıflandırıcının eğitiminde de diğer yöntemlere göre başarısı daha yüksek olan derin öğrenme yöntemleri tercih edilmiştir. Tweetler sınıflandırıldıktan sonra diğer pek çok çalışmayı kapsayacak şekilde tweetlerden elde edilen özniteliklerin genel endeksler ve sektörel endeksler düzeyinde açılış, işlem hacmi, volatilité ve getiri göstergelerini tahmin başarısı teknik nitelikleri yönünden farklı avantaj ve

dezavantajlara sahip farklı makine öğrenmesi yöntemleri ile analiz edilerek ulaşılan sonuçlar ve yöntemlerin performansları karşılaştırılmıştır.



DÖRDÜNCÜ BÖLÜM

SOSYAL MEDYANIN HİSSE SENEDİ FİYAT

ÖNGÖRÜLERİNDE KULLANILABİLİRLİĞİNE İLİŞKİN

BORSA İSTANBUL ÜZERİNE BİR UYGULAMA

4.1 ARAŞTIRMANIN AMACI

Geleneksel finans kuramı üç temel varsayım üzerine inşa edilmektedir. Bu varsayımlardan ilki *tam bilgi* varsayımdır. Bilindiği gibi etkin piyasalar hipotezine göre bir piyasanın tam anlamıyla etkin olabilmesi için, hisse senedi fiyatlarının piyasada dolaşan mevcut tüm bilgileri tam olarak yansıtması gerekmektedir. Başka bir deyişle güçlü formda etkin bir piyasanın oluşabilmesi için yatırımcıların tüm bilgilere sahip ve bu bilgileri, rasyonel bir akıl yürütme neticesinde aldığı rasyonel kararlarla fiyatlara yansıtması beklenmektedir. Ancak literatürde ortaya konan pek çok anomali, piyasaların güçlü formda etkinliğini ifade eden bu durumun genellikle teorik bir idealden öteye geçemediğini göstermektedir. Dolayısıyla piyasaların etkinliği ekseriyetle zayıf veya yarı güçlü formda tezahür etmektedir. Piyasaların fiili durumunu yansıtan ve eksik bilgi ile karakterize edilen bu formlar, aslında analizlerde *tam bilgi* varsayımıyla hareket etmenin analizin içeriğindeki önemli bir unsurun gözden kaçırılmasına yol açabileceğine işaret etmektedir.

Geleneksel finans kuramının ikinci temel varsayımı *rasyonel insan* varsayımdır. Esasen tam bilgi varsayımı ile iç içe geçmiş olan rasyonel insan varsayımı, analize konu olan bireyin, alternatifler hakkında eksiksiz bilgiye sahip, kişisel çıkarları için kusursuz maksimizasyon hesapları ve belirsizlik koşullarında mükemmel olasılık tahminleri yapabilen ve nihayetinde duygularından ve duyu deneyimlerinden arınmış karar vericiler olduğunu kabul etmektedir. Ancak fiili durumda insan, alternatifler hakkında *tam bilgi* sahibi olmayan, isabetli bir karar

verebilmek için gerekli olan tüm hesaplamaların yapılmasında *bilişsel kısıtlara* sahip (sınırlı rasyonel) bir model ortaya koymaktadır.

Geleneksel finans kuramının üçüncü varsayımı ise bireylerin sosyal yönünün göz ardı edilerek sosyal etkiden bağımsız karar vericiler olarak kabul edilmesine yol açan *izole birey* varsayımdır. Bu varsayım ise bireylerin kararlarında ve davranışlarında, içerisinde yaşadığı sosyal çevrenin etkisinin olmadığını farz etmektedir. Yani bu varsayımın insan modelinde bireyler toplumdan ve sosyal çevresinden soyutlanmış bağımsız karar vericilerdir. Ancak insanın sosyal yönü göz önünde bulundurulduğunda, davranışın şekillenmesinde sosyal ağların, sosyal kimliklerin, sosyal statünün ve insanlar arasındaki etkileşim modellerinin pay sahibi olduğu açıktır. Dolayısıyla insanların karar verme biçimlerinin ve davranışlarının analizinde insanın sosyalliğinin de analize dâhil edilmesi gerektiği düşünülmektedir.

Geleneksel finansın yukarıda ifade edilen üç varsayımına alternatif olarak, bu araştırmanın dört hipotezi bulunmaktadır. Araştırmanın hipotezlerinden birincisi rasyonalite varsayımıyla ilgilidir. Bu hipoteze göre insanlar/bireysel yatırımcılar sınırlı ve ekolojik rasyonaliteye sahiptir. İnsanların rasyonalitesini ulaşabildikleri bilgi miktarı ve bilişsel kapasiteleri sınırlar. Dolayısıyla insanlar eksik bilgiye sahiptir ve eksik bilgisini olabildiğince tamamlamak isterler. Sosyal medya insanların eksik bilgilerini tamamlamak için bilgi arayışlarında başvurdukları başlıca kaynaklardan biridir. Araştırmanın ikinci hipotezi davranışsal eğilimlerle ilgilidir. Buna göre insanlar/bireysel yatırımcılar davranışsal eğilimlerinin etkisi altında karar verirler. Davranışsal eğilimler ve yanlılıklar kararlarında etkin rol oynar. İnsanların sosyal medya paylaşımları davranışsal etkiyi kısmen gözlemleme imkânı sunar. Araştırmanın üçüncü hipotezi insanın bilişsel yönüyle ilgilidir. Buna göre insanlar/bireysel yatırımcılar bilişsel eğilimlere sahiptir. Edindikleri bilgilere göre duygu, düşünce, inanç ve beklenti inşa ederler. İnşa ettikleri duygu, düşünce, inanç ve beklentinin etkisi altında karar verirler. İnsanların duygu, düşünce ve beklentilerine ilişkin sosyal medya paylaşımları bilişsel etkiyi kısmen gözlemleme imkânı sunar. Araştırmanın dördüncü hipotezi insanın toplumsal yönüyle ilgilidir. Bu hipoteze göre insanlar/bireysel yatırımcılar toplumsal varlıklardır. Kararlarını toplumdan izole bir şekilde vermezler. Sosyal çevresinden bilgisel

ve duygusal bulaşa maruz kalırlar. İnsanların sosyal medya paylaşımları sosyal etkiyi kısmen gözlemlene imkânı sunar.

Sonuç olarak, dört hipotezin bir bileşimi olarak bu çalışmanın ana hipotezi; davranışsal, bilişsel ve sosyal etkilerin yatırımcıların kararlarını şekillendirdiği ve sosyal medya aracılığıyla paylaşılan mesajlardan elde edilecek bilişsel, davranışsal ve sosyal özniteliklerin gözlemlenerek insanların/bireysel yatırımcıların yatırım kararlarının tahmin edilebileceğidir.

Bu hipotezler çerçevesinde, sosyal medyadan toplanan büyük veri kümesinin, yapay zekâ algoritmaları yardımıyla çözümlenmesine dayalı bu araştırma, davranışsal finans varsayımları üzerine inşa edilerek, finans içerikli bilginin sosyal medyada yayılma mekanizmasını analiz etmeyi, bunun bireysel yatırımcıların finansal kararları üzerindeki etkisini ortaya koymayı ve çözümlenen mekanizmanın hisse senedi fiyat öngörülerinde kullanılıp kullanılamayacağını araştırmayı amaçlamaktadır. Daha açık bir ifadeyle, davranışsal finans varsayımları bağlamında araştırmanın temel amacı, geleneksel finansın genel kabul görmüş “tam bilgi”, “rasyonel insan” ve “izole birey” kabulleri yerine, eksik bilgisini tamamlamaya çalışan, sınırlı rasyonaliteye sahip, sosyal bir çevreye yerleşik, sosyal ağının ve duygularının karar süreçlerinde pay sahibi olabildiği bir insan modelini esas alarak, bireysel yatırımcıların karar verme eylem ve davranışlarını çözümlemeye çalışmaktır. Bu amaçla çalışmada bireysel yatırımcıların karar süreçlerinin bilişsel, davranışsal ve sosyal boyutlardan bütünleşik bir etkileşim ve sosyal psikolojik bir karar modeli elde edilmekte ve bu modelin Borsa İstanbul pay piyasası endeksleri ile ilişkisi incelenmektedir.

Bu araştırmanın teorik katkısı bireysel yatırımcıların eksik bilgilerini tamamlamak amacıyla sosyal medya platformlarından yatırım fikirleri arayışlarının kararlarına etkisini sosyolojik çerçevede ele alarak davranışsal finans literatürüne kazandırmaktır. Başka bir ifadeyle teorik katkı, yatırımcı beklentilerinin inşasında bilginin, bilgi iletim ve edinme yollarının ve dolayısıyla sosyal çevrenin karar süreçlerine etkisini ve davranışsal eğilimleri tetikleme potansiyelini sosyolojik bir perspektifle çözümleyen bir bakış açısı kazandırmaktır. Bu çalışmanın pratik katkısı ise sosyal medyanın hisse senedi fiyat öngörülerinde

kullanılabilirliğini ortaya koymak ve böylece bireysel ve kurumsal yatırımcılar için hem yatırımcıların kolektif düzeyde fikri ve duygusal yönelimlerinin izlenmesi hem de buradan elde edilen izlenimlerin yatırım kararlarında bir girdi olarak kullanılmasını mümkün hale getiren bir araç ortaya koymaktır.

4.2 ARAŞTIRMANIN KAPSAMI

Araştırma, karar vermede psikolojik ve sosyolojik süreçlerin kısmen gözlemlenebildiği bir mecra olan sosyal medya platformlarından Twitter'daki insan davranış ve eylemlerine, finans içerikli bilgi paylaşımı ve arayışı çerçevesinde odaklanmaktadır. Twitter, insanları gelişmelerden haberdar etmek veya bilgilerini güncel tutmak için her türlü bilgi, düşünce, görüş ve fikrin kısa ve öz bir şekilde paylaşılmasına izin veren bir platformdur. Diğer pek çok konunun yanı sıra, borsa konusunda uzmanların ya da o alana ilgi duyan kişilerin çeşitli analizler yaparak fikir, duygu ve yorumlarını paylaşabilmesine de imkân sunan bu platform, aynı zamanda finansal bilgi arayışı içerisinde olan bireysel yatırımcıların bu paylaşımlara erişmesine ve platformdaki kullanıcılarla etkileşim kurabilmesine de olanak sağlamaktadır. Sosyal etkileşimlerin kısmen bu tür dijital mecralarda gerçekleşmesi, araştırmacılar açısından dijital izlerin çözümlenerek fikri ve duygusal etkileşimlerin piyasa ile ilişkisinin gözlemlenebilmesine olanak vermektedir. Araştırma kapsamında Twitter'ın tercih edilmesinin temel nedeni, Twitter paylaşımlarının (varsayılan olarak) platformun kayıtlı bir kullanıcısı olmaya gerek kalmadan herkes tarafından görülebilir olmasıdır.

Bu kapsamda araştırmada 01.01.2012 - 29.02.2020 tarihleri arasında⁴³ Borsa İstanbul'da işlem gören şirketlerin hisse senetleriyle ilgili Türkçe dilinde paylaşımlar yapan veya bunları takip eden Twitter kullanıcılarının etkileşimlerinin *bilişsel, davranışsal ve sosyal*

⁴³ Analiz dönemi geriye doğru tweetlerin toplanabildiği en stabil yıl olduğu için 1 Ocak 2012'den başlatılmış, dünya geneline yayılan Covid 19 salgının Türkiye'de tespit edilen ilk vakasının 11 Mart 2020 tarihinde açıklanmasının, yatırımcı davranışlarını önemli derecede etkileyebileceği düşüncesiyle, 1 Mart 2020 tarihi itibarı ile sonlandırılmıştır.

boyutlarının finansal piyasalar ile ilişkilerine ve finansal kararlar üzerindeki etkilerine odaklanılmaktadır.

4.3 ARAŞTIRMADA KULLANILAN VERİ SETLERİ

Araştırma kapsamında farklı analiz türlerine özgü, çeşitli kaynaklardan, muhtelif türlerde veriler derlenmiştir. Bu veriler; Türkçe dilinde paylaşımlar yapan veya bunları takip eden Twitter kullanıcılarının tweetleri, etkileşim verileri, finansal metinler içeren raporlar ve web sitelerinden toplanan metinsel veriler ile BIST30, BIST100, BISTTUM, BIST HİZMETLER, BIST MALİ, BIST SİNAİ, BIST TEKNOLOJİ endekslerinin günlük değerlerini gösteren verilerdir. Çeşitlilik, hacim ve hız açısından büyük veri niteliği taşıyan bu verilerden bir kısmının geleneksel yöntemlerle analiz edilmesi mümkün değildir. Zira bu tür veriler (örneğin metin, video, fotoğraf gibi) yapılandırılmamış verilerdir. Öncelikle bu tür verilere, değişken niteliği kazandırılmalıdır. Başka bir ifadeyle, örneğin bir metnin taşıdığı duygunun, fikrin ya da anlamın nicelleştirilerek analize konu bir öznelilik haline getirilmesi gerekmektedir. Dolayısıyla bu verilerin analizinde makine öğrenmesi yöntemleriyle geliştirilen yapay zekâ modellerine ihtiyaç duyulmaktadır. Ancak bu modellerin eğitilebilmesi için de yine araştırmanın amacı ve yöntemine uygun niteliklere sahip veriye gereksinim duyulmaktadır.

Bu kapsamda eğitilen modellerden, metinsel nitelikteki tweetlerin ilk olarak borsa ya da borsada işlem gören bir şirketle ilgili olup olmadığını, ikinci olarak fikri ve duygusal [sentiment] yönelimini sınıflandırması beklenmektedir. Diğer bir deyişle metinsel tweetleri sınıflandırabilmek için biri *bağlam sınıflandırıcı*, diğeri *fikir ve duygu sınıflandırıcı* olmak üzere iki ayrı yapay zekâ modeli eğitilmesi gerekmektedir. Modellerin eğitilmesinde gözetimli öğrenme yöntemi tercih edildiğinden araştırmanın amacına uygun şekilde etiketlenmiş eğitim ve test veri setlerine ihtiyaç duyulmaktadır. Dolayısıyla bağlam sınıflandırıcının eğitimi için *borsa ile ilgili* ya da *değil* şeklinde sınıflandırılmış bir veri seti, fikir ve duygu sınıflandırıcının eğitimi için *pozitif*, *negatif* ve *nötr* şeklinde sınıflandırılmış bir başka veri seti gerekmektedir.

Twitter paylaşımlarının genel olarak bir değerlendirmesinin yapılması, analizde ihtiyaç duyulacak verilerin belirlenmesinde önemli bir yol gösterici olacaktır. Nitekim bu amaçla Twitter'ın bir sosyal ağ mı yoksa bir haber yayın aracı mı olduğunu araştıran, çokça atıf almış bir eserde, esasen Twitter paylaşımlarının oldukça büyük bir çoğunluğunun (yaklaşık %85'inin) haber başlığı niteliği taşıdığı ortaya konulmuştur (Kwak, 2010). Bu durum göz önünde bulundurulduğunda, Twitter verilerini analiz etmek için geliştirilen modelin öncelikle analize konu olan alanın terminolojisine ve bu alanla ilgili gelişmeleri haberleştirirken kullanılan dile hâkim olması gerekmektedir. Biçimsel dilin kullanıldığı bu tür paylaşımlar genellikle kurumsal kullanıcılar tarafından hazırlanmaktadır. Yatırım kuruluşlarının ya da aracı kurumların günlük bültenlerini paylaşması, ekonomi ve finans üzerine odaklanmış web sitelerinin ilgili herhangi bir gelişmeyi okuyucularıyla paylaşması buna örnek olarak verilebilir. Üstelik burada paylaşımın doğrudan ilgili kurumsal yapı tarafından gerçekleştirilmesi de gerekmemektedir. Zira herhangi bir kullanıcı da bülteni ya da haberi paylaşabilmektedir. Twitter'da biçimsel dilin yanı sıra, bazı kullanıcıların gündelik bir dili tercih ettiği, bazılarının ise argo ifadeler kullandığı görülmektedir. Ayrıca bu tweetlerden bazıları doğrudan duygu ve düşünce ifade eden metinler ihtiva ederken bazıları sadece haber niteliğinde olabilmektedir. Dolayısıyla başarılı bir sınıflandırma algoritmasının eğitilebilmesi için biçimsel dilin, gündelik dilin, argo ifadelerin ve finansal jargonun kullanıldığı kaynaklardan bir derlem [corpus] oluşturulması gerekmektedir. Bu bağlamda aracı kurumların raporları, finansal haber ve yorum siteleri ile Twitter platformu, araştırma kapsamında beklentileri karşılayan kaynaklar sunmaktadır.

Araştırmanın amaçlarından birinin insanın sosyal yönünün bireysel yatırımcıların karar süreçlerindeki etkisinin ortaya konulması olduğundan, Twitter kullanıcılarının etkileşim verilerinin de elde edilmesi gerekmektedir. Twitter platformundan sosyal etkileşim için retweet ve beğeni gibi veriler elde edilmiştir.

Son olarak, elde edilen bu verilerin finansal piyasalarla etkileşimini ölçebilmek amacıyla Borsa İstanbul pay piyasası endekslerinin değerlerine ilişkin veriler kullanılmaktadır.

4.3.1 Biçimsel Dilin Kullanıldığı Kaynakların Derlenmesi

Yatırım kuruluşlarının raporları, gazete, dergi ve web sayfalarında yer alan finansal haberler, analizler, raporlar ve köşe yazıları biçimsel dilin kullanıldığı kaynaklardır. Bu kaynakların bir avantajı, sınıflandırma algoritmasının eğitiminde kullanılan metinlerin, uzmanlarca etiketlenmiş olmasıdır. Diğer bir ifadeyle bu tür kaynaklarda metinler olabildiğince nesnel ve alanın uzmanları tarafından etiketlenmiş veri seti sunmaktadır.

Tablo 4.1: Biçimsel Dil Modeli İçin Finansal Derlem Kaynakları

Yatırım Kuruluşlarının Raporları [21]	Diğer Günlük Ekonomi ve Finans Haber Kaynakları [37]
akbank	aa.com.tr
deniz	demokrathaber.org
garanti	gazetevatan.com
gedik_gb	haber3.com
gedik_uzman	habervaktim.com
halk	iha.com.tr
investaz_analiz	milatgazetesi.com
investaz_uzman	posta.com.tr
metro	takvim.com.tr
oyakyatirim_analiz	turkiyegazetesi.com.tr
oyakyatirim_yorum	yenimesaj.com.tr
qnbfi_bist	ntv.com.tr
qnbfi_blog	ahaber.com.tr
qnbfi_makroekon	cnnturk.com
qnbfi_piyasalar	ensonhaber.com
sirket_haber	gercekgundem.com
odeabank	ekonomi.haber7.com
tacirler_gb	habervitrini.com
vakif_yatirim	internethaber.com
ziraat_sabah	karar.com
ziraat_yatirim	milligazete.com.tr
Ekonomi ve Finans Web Siteleri [Efws] [18]	sabah.com.tr
bigpara.hurriyet.com.tr	tgrthaber.com.tr
bloomberght.com	yeniasya.com.tr
borsagundem.com	yenisafak.com
capital.com.tr	724aktuel.com
dunya.com	cumhuriyet.com.tr
ekonomi7.com	evrensel.net
ekonomihaber.com	gazetebirlik.com
finans.haberler.com	gunes.com
finans.mynet.com	habererk.com
hisse.net	istiklal.com.tr
kanalfinans.com	medyahaber.com

paraborsa.net	milliyet.com.tr
paragaranti.com	sozcu.com.tr
paralimani.com	yenicaggazetesi.com.tr
paraanaliz.com	yurtgazetesi.com.tr
fortuneturkey.com	
ekonomist.com.tr	
tr.tradingview.com	

Bu aşamada derlenen veriler Tablo 4.1’de görüldüğü üzere [i] yatırım kuruluşlarının raporları, [ii] ekonomi ve finans web siteleri ve [iii] diğer günlük ekonomi ve finans haber kaynakları olmak üzere üç grupta toplanmıştır. Araştırma amacına uygun şekilde etiketli veriler içeren bütün kaynaklar finansal derleme dâhil edilmeye çalışılmıştır.

4.3.1.1 Yatırım Kuruluşlarının Raporları

Yatırım kuruluşlarının raporları hem makroekonomi hem de finansal piyasalar ile ilgili alanın uzmanlarınca hazırlanan, içerik açısından oldukça zengin, formel dille yazılmış bilgi kaynaklarıdır. Bu raporların analize dâhil edilmesinin iki temel sebebi vardır. Bunlardan birincisi, bu raporlarda duru bir finans terminolojisi kullanıldığı için bağlam sınıflandırıcısının terminolojiye aşina olması; ikincisi ise, bu raporlardan bazılarında şirketlerle ilgili haberlerin uzmanlarca etiketlenerek yatırımcıların istifadesine sunulmuş olmasıdır. Dolayısıyla bu raporlar fikir ve duygu etiketçisinin eğitiminde kullanılmak üzere, doğrudan borsada işlem gören bir şirketle ilgili herhangi bir haberin muhtemel etkisinin alanın uzmanları tarafından etiketlendiği veriler de sunmaktadır.

Araştırma kapsamında, fikri ve duygusal etiket içeren tüm yatırım kuruluşlarının raporları analize dâhil edilmiştir. Tablo 4.1’de görüleceği üzere, hiçbir ayırım gözetmeksizin, bu nitelikleri taşıyan 21 farklı yatırım kuruluşunun 2020 yılından geriye doğru⁴⁴ web sayfaları üzerinden erişilebilen tüm raporlarına ulaşılmıştır. Bu bağlamda, Tablo 4.2’den de görüleceği

⁴⁴ Bu raporların geçmişe dönük erişim tarihleri aracı kurumdan aracı kuruma farklılık göstermektedir. Bazı aracı kurumlar için 2010 yılına kadar geriye dönük raporlara ulaşılabilirken bazıları için 2015, bazıları için 2016 yılına kadar geri gidilebilmiştir.

üzere, 21 farklı kaynaktan 19.323 etiketli veriye ulaşılmıştır. Yatırım kuruluşlarının raporları genelde pdf formatında hazırlanmaktadırlar. Dolayısıyla analizin bu aşamasında öncelikle yatırım kuruluşlarının erişime açık olan raporlarının pdf uzantılı bağlantıları indekslenmiştir. İndekslenen bu bağlantılar daha sonra R programı için geliştirilmiş pdftools (Ooms, J., 2022) kütüphanesi aracılığıyla otomatik olarak indirilerek arşivlenmiş ve PDF2text (Marwick, n.d.) kodları aracılığıyla pdf formatından R ortamına aktarılarak analizde kullanılabilecek biçimde yapılandırılmıştır.

4.3.1.2 Ekonomi ve Finans Web Siteleri

Yatırım kuruluşlarının analiz raporları ve bültenleri gibi ekonomi ve finans web siteleri de hem ekonomi hem de finansal piyasalar ile ilgili genellikle alanın uzmanlarınca hazırlanan oldukça zengin bilgi kaynaklarıdır. Analiz raporlarından ve bültenlerinden farklı olarak ekonomi ve finans web siteleri, bu bağlamdaki günlük gelişmeleri haber diliyle okuyucularına ulaştırdığından *sınıflandırıcıların* bu dili de öğrenebilmesi açısından oldukça önemlidir. Zira daha önce de ifade edildiği gibi ya bu web sayfalarının yöneticileri ya da doğrudan okuyucular ilgili haberi Twitter hesaplarında paylaşmakta ve ardından Twitter’da haberle ilgili bir etkileşim söz konusu olabilmektedir. Dolayısıyla sınıflandırıcının bu tür tweetlere de aşina olması gerekmektedir. Üstelik bu kaynaklarda yayımlanan içeriklerin sonunda yer alan haber etiketleri de, gözetimli öğrenmede ihtiyaç duyulan etiketli veri setini sağlamaktadır. Bu tür metinsel veri kaynaklarında kullanılan biri örtük diğer açık olmak üzere iki farklı etiketleme şekinden söz edilebilir. *Örtük etiketleme*, haber sitelerinde yayınlanan haberlerin belirli kategoriler altında yayınlanmasıyla ilgilidir. Bu etiketleme türünde yayımcısı tarafından haberin önsel ve örtük olarak hangi kategoriye ait olduğu belirlenmekte ve içerik bu kategori altında yayımlanmaktadır. Bazı kaynaklar da ise, haberleri bu şekilde ilgili kategorilerde yayımlamak yerine, haber metninin sonunda haberin ilgili olduğu kategorileri ya da konuları göstermek için “etiketler” ya da “anahtar kelimeler” listesi verilmekte, böylece açık bir bağlam etiketlemesi yapılmaktadır. Burada göz önünde bulundurulması gereken en önemli husus, örtük etiketlemede genellikle kategori tek bir

sözcükle temsil edilirken, açık etiketlemede birden fazla anahtar sözcük kullanılabilmesidir. Bunların dışında her iki yöntemi birlikte kullanan kaynaklar da söz konusudur.

Web sayfalarının taranıp indekslenmesi ve akabinde indekslenen bu bağlantılardan konularına göre metinlerin derlenmesi R programı için yazılan Rcrawler ile gerçekleştirilmiştir (Khalil ve Fakir, 2017; Munzert ve diğerleri, 2015). Bu aşamada öncelikle ekonomi ve finansla ilgili Türkçe yayın yapan, Tablo 4.1’de ayrıntılarının yer aldığı 18 farklı dergi, web sayfası ve haber sitesinin ekonomi ve finansla ilgili haberleri, köşe yazıları, uzman yorumları ve analizleri ile bunların etiketleri, html koddan arındırılarak yapılandırılmış ve böylece 561.404 adet belge sınıflandırıcının eğitiminde kullanılmak üzere hazır hale getirilmiştir⁴⁵.

4.3.1.3 Diğer Günlük Ekonomi ve Finans Haber Kaynakları

Diğer günlük ekonomi ve finans haber kaynakları olarak sınıflandırılan kaynaklar ise, odaklandıkları alanların doğrudan ekonomi ve finans olmadığı, ancak ekonomi ve finansla ilgili haberleri de okuyucularına sunan web sayfaları ve haber siteleridir. Bu kaynakların en önemli özelliği ise ekonomi ve finans terminolojisinin yoğunluğunun nispeten daha düşük kaldığı, politik bakış açılarından yansımalarının söz konusu olabildiği bir dil sunma potansiyelidir. Nitekim Twitter’da bu web sayfalarının haberleri de paylaşılmakta ve bunlara ilişkin etkileşimler söz konusu olabilmektedir. Dolayısıyla bu tür kaynaklarda, modelin yanlış olmaması için toplumun tüm renklerini temsil edecek çeşitlilikte 38 farklı gazete ve haber sitesinde yayımlanan içerik, 01-30 Kasım 2017 tarihleri arasında derlenmiş, bu kapsamda 158.887 adet belge elde edilmiştir.

Biçimsel dilin temsili için ulaşılan kaynak ve belge sayısı Tablo 4.2’de gösterilmektedir. Buna göre 21 farklı yatırım kuruluşundan toplam 19.323 belge, 17 farklı ekonomi ve finans

⁴⁵ Ekonomi ve finansla ilgili haber içeriklerine erişim tarihleri, her bir web sitesi için farklılık arz etmektedir. Bazı web siteleri 2002 yılındaki haber içeriklerine erişme imkanı sunarken bazıları 2015, 2016, 2017 yılları gibi daha yakın tarihli arşiv haberlerine erişme imkanı sunmaktadır.

web sitesinden 561.404 belge ve son olarak 38 farklı gazete ve haber sitesinden günlük ekonomi ve finans haberi içeren 158.887 belge derlenmiştir.

Tablo 4.2: Finansal Derlem Kaynağına Göre Belge Sayısı

Derlem Kaynağı / İçerik Bilgileri	Kaynak Sayısı	Belge Sayısı
Yatırım Kuruluşlarının Raporları	21	19.323
Ekonomi ve Finans Web Siteleri	17	561.404
Diğer Günlük Ekonomi ve Finans Haber Kaynakları	38	158.887
Toplam	68	739.614

Sonuç olarak modelin formel dili tanıyabilmesi amacıyla, eğitiminde kullanılmak üzere, toplamda 739.614 belge toplanmıştır.

4.3.2 Twitter Verilerinin Derlenmesi

Twitter, 21 Mart 2006 tarihinde ABD’de kurulan, kullanıcılarının herhangi bir konudaki düşüncelerini en fazla 140, 2017 yılından itibaren de 280 karakterden oluşan tweetlerle paylaşımlarına ve diğerleriyle *yanıtlar*, *retweetler* ve *beğeni tweetler* yoluyla iletişim kurmasına imkân sunan popüler sosyal ağ ve microblog hizmetlerinden birisidir. *Tweet* adı verilen paylaşımlar kullanıcının Twitter’daki profilinde otomatik olarak akış halinde (ve herkesin erişebileceği şekilde) yayınlanmakta ve anında kullanıcının takipçi ağına gönderilmektedir. Böylece Twitter kullanıcıları, takip ettikleri diğer kullanıcılardan gelen tweetleri kronolojik bir akış sunan zaman çizelgesinde anlık olarak okuyabilmektedirler. Özellikle mobil cihaz kullanımının yaygınlaşmasıyla birlikte anlık bilgi yayılımı ve erişimi için verimli bir ortam sağlamaktadır (Atefeh ve Khreich, 2015).

Twitter, aynı zamanda insanların platformu kullanan diğer kişilerle iletişim kurmasına olanak tanıyan bir çevrimiçi sosyal ağ hizmetidir. Twitter’da bir kullanıcı, (korunmalı bir hesap değilse) takip edilen kullanıcıların onayına gerek kalmadan başka bir kullanıcıyı takip edebilmektedir. Başka bir ifadeyle Twitter’da sosyal ilişki asimetriktir. Kullanıcılar gizlilik

ayarlarını onları kimin takip edebileceğine karar vermek için değiştirebilse de, genellikle mesajın daha çok kişiye ulaşması istendiği için bu kısıtlama çoğu kullanıcı tarafından tercih edilmemektedir. Bu durum platformdaki paylaşımların daha genele yayılmasını sağlamaktadır. Twitter ayrıca, kullanıcıların başka bir kullanıcının tweetini takipçilerine iletmelerine, yani retweetlemelerine de olanak tanımaktadır. Böylece bir kullanıcının bir paylaşımı *retweetleyerek* ya da *beğenerek* mesajın, mesajı paylaşan kullanıcının ağının çok ötesine yayılmasına olanak tanımaktadır. Dolayısıyla bir mesajın yüksek oranda retweetlenmesi veya beğenilmesi paylaşımların Twitter topluluğunun görüşlerini genel olarak yansıttığı anlamına gelmektedir. Twitter ayrıca kullanıcıların mesajlarında birbirlerine atıfta bulunarak sohbet etmeleri ve etkileşimde bulunmalarına da olanak sağlamaktadır. *Mention* olarak ifade edilen bu durum, bir kullanıcı adının önüne @ simgesinin yerleştirilmesi suretiyle yapılmaktadır. Böylece bahsi geçen kullanıcı hesabına bir yanıt bağlantısı oluşturulmaktadır. Yanıt ise, bir kullanıcının başka bir kullanıcının mesajına doğrudan cevap olarak yazdığı özel bir mesajdır. Ayrıca Twitter belirli olay ve konuların takip edilmesi için *hashtag* olarak adlandırılan bir başlıkta (etikette) toplanmasına izin vermekte ve kullanımını teşvik etmektedir. # simgesi ile temsil edilen hashtagler, mesajların konusunu belirtme, paylaşılan bir konuda farklı kullanıcıların tweetlerini bir araya getirme ve belirli olayları gerçek zamanlı olarak düzenli olarak takip etme imkânı sunmaktadır (Atefeh ve Khreich, 2015).

Her bir tweet, oldukça geniş bir üstveri içeriğine sahiptir. Bu üstveriler; tweeti paylaşan kullanıcı bilgisi, konum bilgisi, retweet sayısı, beğenilme sayısı, etiket listesi ya da bahsi geçen kişiler (mention) gibi pek çok ilave bilgi sağlayan 150'nin üzerinde öznitelik içerebilmektedir.

Twitter, uygulama programlama arayüzü [API: Application Programming Interface] ile geliştiriciler ve araştırmacılar için bu verilerin bir kısmına erişim imkânı sunmaktadır (<https://developer.twitter.com>). Çalışma kapsamında Twitter verilerinin bir kısmı, açık kaynak kodlu R istatistiksel programlama dili için geliştirilmiş olan “rtweet” kütüphanesi kullanılarak bu yöntemle toplanmıştır (Kearney, 2020). Ancak Twitter API'daki kısıtlar,

örneklemin boyutu büyüdüğünde kapsamlı bir veri elde etmenin zaman maliyetini oldukça yükseltmektedir⁴⁶. Başka bir ifadeyle Twitter'ın hem bir Twitter hesabına bağlı olarak kullanılabilecek uygulama sayısına getirdiği kısıtlar, hem de uygulamaların *tweet*, *retweet*, *beğeni*, *arkadaş*, *takipçi* gibi farklı türlerdeki verilerin toplanmasıyla ilgili zamana bağlı sınırlandırmaları, Türkçe dilinde yazan tüm Twitter hesaplarının verilerini toplamayı zorlaştırmaktadır. Dolayısıyla bu kısıtları kısmen aşarak daha kapsamlı veri toplamaya imkân sunan ve Python programlama dili için geliştirilmiş olan “Twint” (n.d.) ve “Scweet” (n.d.) adlı kütüphanelere de başvurulmuştur. Açık kaynak kodlu bu kütüphaneler, Twitter API'ı kullanmadan herkese açık Twitter profillerinden web kazıma mantığıyla tweetlerin ve diğer verilerin derlenmesine imkân sunmaktadır. Bu çalışma kapsamında kullanılan *tweet*, *retweet*, *beğeni tweet*, *arkadaş sayısı*, *takipçi sayısı* gibi verilerin bir kısmı da bu araçlar kullanılarak temin edilmiştir.

Veri toplama araçlarının belirlenmesinden sonraki adımda, Türkçe dilinde yazan tüm Twitter kullanıcılarına ulaşabilmek amacıyla öncelikle bir çekirdek liste oluşturulmuştur. Bu çekirdek listeye Türkiye’de (2017 yılında) en çok takipçiye sahip ilk 3 kullanıcının (Recep Tayyip Erdoğan, Cem Yılmaz, Abdullah Gül) Twitter hesapları; Türkiye’de yayın yapan gazete, dergi, haber sitesi, haber ajansı gibi kuruluşların Twitter hesapları; siyasi partiler ile 24, 25 ve 26. dönem milletvekilleri de dâhil olmak üzere politikacıların Twitter hesapları; yatırım uzmanları, portföy yönetim şirketleri, aracı kurumlar, bankalar, bağımsız denetim kuruluşları, derecelendirme kuruluşları, ekonomi ve finans alanında yayın yapan dergi, gazete ve internet medyasının kurumsal hesapları ve bu medya kuruluşlarındaki köşe yazarlarının Twitter hesapları dahil edilmiştir. Ayrıca ekonomi ve finansla ilgili anahtar sözcükler ile Twitter’da arama yapılarak bu alanla bir şekilde ilişkili olduğu düşünülen hesaplar da listeye ilave edilmiştir. Böylece, Şekil 4.1’de görüldüğü üzere, toplamda 1.976

⁴⁶ Bu konunun ayrıntısı için araştırmanın kısıtları başlığına bakınız.

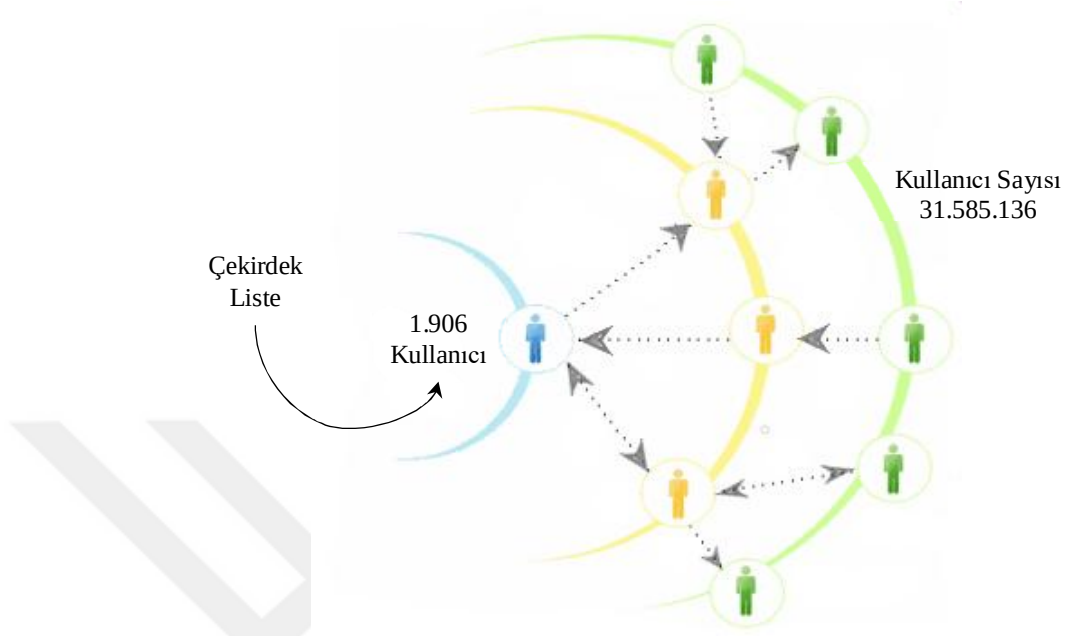
kullanıcı belirlenmiş, ancak korumalı hesaplar⁴⁷ analiz dışı bırakılarak 1.906 Twitter hesabının yer aldığı bir çekirdek liste 23.02.2017 tarihinde oluşturulmuştur.

Çekirdek liste oluşturulduktan sonra, sosyal ağ teorisine göre, çekirdek listedeki Twitter kullanıcılarının bir derece uzağında yer alan diğer hesaplar, yani çekirdek listedeki kullanıcı hesaplarının takipçileri ve arkadaşlarının hesap isimleri toplanmıştır. Ardından bir kullanıcının hem arkadaş hem takipçi olabilmesi, ya da birden fazla hesabı takip edebilmesi nedeniyle ağın birinci düzeyinde toplanan kullanıcı hesaplarından mükerrer olanlar arındırılarak benzersiz hale getirilmiştir. Son aşamada ise bu listede yer alan kullanıcı hesaplarının takipçileri ve arkadaşları da toplanarak hem Türkçe dilinde yazan hesaplar, hem de bu hesapların sosyal ağı toplanmıştır. Bu şekilde, aktif ve pasif hesaplar dahil olmak üzere, 35 milyon kullanıcı hesabına ait profil bilgileri Twitter API aracılığıyla toplanmıştır. Bunlar içerisindeki korumalı hesaplar dışarıda bırakıldıktan sonra, kalan **31.585.136 profilin tweetleri, retweetleri, beğeni tweetleri, profil bilgileri** yine Python ve R programlama dillerindeki istemciler aracılığıyla toplanmış ve bir veri tabanında depolanmıştır⁴⁸. Twitter'dan toplanan verilerin yapılandırılmamış niteliğe sahip olması nedeniyle ilişkisel veri tabanı yerine, yapılandırılmamış verileri saklamaya imkân veren ve bir NoSQL veri tabanı olan MongoDB tercih edilmiştir.

⁴⁷ Bir kullanıcının tweetleri varsayılan olarak herkese açıktır, yani Twitter hesabı olsun ya da olmasın herkes kullanıcının paylaştığı tweetleri görüntüleyebilmektedir. Korumalı tweetler ise yalnızca kullanıcının takipçileri tarafından görüntülenebilmektedir.

⁴⁸ Twitter'dan veri toplamaya 23.02.2017 tarihinde başlanmıştır. Tüm kullanıcıların verilerini (tweetler, retweetler, arkadaşlar, takipçiler, beğeniler) toplamak yaklaşık 1 yıl sürmüştür. Twitter platformunun hızla büyüyen yapısı nedeniyle tüm veriler 13.02.2018 tarihinde ve 01.02.2020 tarihinde tekrar tekrar toplanarak veri seti Twitter kullanıcılarının durumlarındaki değişimi yansıtacak şekilde güncellenmiştir. Verilerin toplanması ve güncellenmesi 3 yılda tamamlanmıştır.

Şekil 4.1: Araştırma Kapsamında Toplanan Twitter Kullanıcı Veri Ağı



Bu çerçevede toplanan Twitter veri türlerinden birisi kullanıcı hesaplarına ilişkin profil bilgilerinin yer aldığı verilerdir. Profil bilgilerinin yer aldığı veri setinde 36 farklı öznitelik yer almaktadır. Bu özniteliklerden bir kısmı Tablo 4.3'te görülmektedir. Bu öznitelikler kullanıcı hesap numarası, kullanıcı ismi, ekran ismi, konum, hesap kullanıcısının biyografisi, hesabın korumalı olup olmadığı, takipçi sayısı, arkadaş sayısı, hesabın hangi tarihte oluşturulduğu, attığı tweet sayısı, beğendiği tweet sayısı gibi pek çok bilgi içermektedir. Bu bilgilerden hesap numarası, ekran ismi gibi veriler statik, biyografisi yarı dinamik, takipçi sayısı, arkadaş sayısı, tweet sayısı gibi bilgiler ise dinamik bir yapı arz etmektedir.

Tablo 4.3: Twitter Kullanıcısı Profil Bilgileri Örneği

name	username	bio	location	join_date	join_time	tweets	following	followers
borsagundem	borsagundem	Türkiye'nin en çok takip edilen borsa haber sitesi. Anlık bors...	Türkiye	14 Jun 2009	03:25:00	273934	11	124115
bigpara	bigparacom	Türkiye'nin Bir Numaralı Finans Portalı	NA	21 Oct 2010	04:53:00	36824	17	213128
NTV Para	NTVPara	Ekonomiden son gelişmeler http://ntv.com.tr 'de	NA	15 Dec 2010	11:22:00	84058	25	441447
Bloomberg HT	BloombergHT	Türkiye'nin Ekonomi Haberleri Platformu http://www.facebo...	Turkey	18 Jan 2011	03:25:00	181008	7	593197
Habertürk Ekonomi	ekonomiHTcom	NA	NA	1 Dec 2015	07:44:00	33819	33	61938
Forbes Türkiye	ForbesTurkey	Forbes Türkiye Resmi Twitter Hesabıdır. Official Twitter acc...	İstanbul	4 Oct 2010	01:42:00	1321	111	7316
hbrturkiye	HBRTurkiye	Harvard Business Review Türkiye	İstanbul	16 Apr 2012	05:39:00	4258	2	36359
Vakıf Yatırım	VakifYatirimAS	Yasal uyarı notu için tıklayınız. https://www.vakifyatirim.com...	NA	27 May 2015	04:04:00	7104	1	9760
Capital	CapitalDergisi	Türkiye'nin en yenilikçi iş ve ekonomi dergisi Capital, 1993 yı...	İstanbul, Turkey	5 Dec 2010	10:35:00	9174	101	54639
Ekonomist Dergisi	EkonomistDergi	Ekonomi'nin dünü, bugünü, yarını EKONOMİST'te	Türkiye	8 Dec 2010	05:21:00	15563	165	201252
Para Dergisi	Para_Dergi	Türkiye'nin haftalık ekonomi dergisi	NA	2 Nov 2011	08:35:00	2	1	401

Twitter'dan toplanan son veri türü ise sosyal ağ verileridir. Sosyal ağ verileri, bir Twitter hesabının takipçileri ve arkadaşları, tweetlerinde bahsettiği kişiler [mentions] gibi sosyal ilişkileri gösteren verilerdir. Sosyal ağ verileri, dinamik bir nitelik taşıdığı ve yaklaşık 31 milyon hesabın sosyal ağındaki değişimleri günlük olarak güncellemenin zaman maliyeti çok yüksek olduğu için, araştırma boyunca üç kez (23.02.2017, 13.02.2018, 01.03.2020 tarihlerinde) toplamıştır. Bu üç güncelleme sayesinde, ağa yeni dâhil olanlar, ağdan ayrılanlar, yeni takip etmeye başlayanlar ve takibi bırakanlar gibi bilgiler yarı dinamik bir şekilde araştırmaya dâhil edilmiştir. Toplamda yaklaşık 31 milyon kullanıcı hesabı elde edilmiştir.

Yaklaşık 31 milyon hesabın benzer tarih aralıklarındaki güncellemelerle tweetleri de toplandıktan sonra, hem tweetlerden, hem de çeşitli haber kaynaklarından toplanan metinleri içeren bir finansal derlem⁴⁹ oluşturulmuştur. Eğitilen modelin anakütleyi temsil edebilmesi için finansal derlemde yer alan ve farklı kaynaklardan ve farklı Twitter hesaplarından toplanan verilerle eğitim seti oluşturulmuştur. Diğer bir ifadeyle modelin eğitim ve test aşamasında kullanılan veri seti, geliştirilen sınıflandırıcının anakütleye karşılaşılabileceği çeşitli cümle ve ifade örneklerini temsil edebilecek şekilde, yani haber kaynakları ve muhtelif

⁴⁹ Derlem Türkçe sözlükte (2009), bir dilin muhtelif kullanım alanlarından derlenmiş örneklerinin dil bilimi araştırmalarında kullanılmak üzere bilgisayar tarafından okunabilecek şekilde bir araya getirilmiş kümesi olarak tanımlanmaktadır.

Twitter hesaplarındaki metinleri temsil edecek şekilde rassal olarak seçilmiştir. Dolayısıyla modelin eğitiminden elde edilen başarının, eğitim setine dâhil edilmemiş diğer metinlerin sınıflandırılmasında da benzer bir performans sergileyeceği varsayılmıştır. Bu varsayım niteliksel verilerle geliştirilen testin geçerliliğinin istatistiksel olarak belirlenmesinde karar geçerliği sürecine dayanmaktadır (Alpar, 2022). Eğitilmiş model bu kabulle birlikte sonraki aşamada Twitter’den 01.01.2012-29.02-2020 tarih aralığında derlenen 4.257.094.438 tweetin sınıflandırılması için kullanılmıştır.

4.3.3 Finansal Verilerin Derlenmesi

Bu çalışma kapsamında toplanan diğer veri ise finansal verilerdir. Çalışmanın temel amacı Twitter verileri ile hisse senedi piyasası ilişkisini makroskobik bir bakış açısı ile çözümlemek olduğundan bireysel hisse senetleri yerine Borsa İstanbul pay piyasası endekslerine odaklanılmıştır. Bu çalışmada hisse senetleriyle ilgili Twitter paylaşımlarının Borsa İstanbul’a etkisinin olup olmadığını ve bu etkinin genel endeksler ve sektörel endeksler düzeyinde farklılaşıp farklılaşmadığını araştırmak amacıyla BIST30 (XU030), BIST100 (XU100), BISTTUM (XUTUM), BIST HİZMETLER (XUHIZ), BIST MALİ (XUMAL), BIST SINAİ (XUSIN) ve BIST TEKNOLOJİ (XUTEK) endekslerine ilişkin günlük veriler 01.01.2012 – 28.02.2020 tarihlerini kapsayacak şekilde Matriks veri platformundan temin edilmiştir.

4.4 ARAŞTIRMANIN KISITLARI

Araştırma sırasında karşılaşılan en önemli kısıt Twitter API ile ilgilidir. Zira Twitter, elinde bulundurduğu verileri, geliştiriciler ve araştırmacıların kullanımına bir takım zamansal ve miktarsal kısıtlarla sunmaktadır. Örneğin Twitter API limitlerine göre 1 token ile 15 dakikada 15 sorgu yapılabilir. Diğer 15 sorgu için 15 dakika beklemek, yani sistemin 15 dakika uyutulması gerekmektedir. Bunu somut bir örnek üzerinden ifade etmek gerekirse, bir kullanıcının takipçileri ve arkadaşlarına ilişkin sosyal ağ verilerini almak için, her sorguda takipçi ya da arkadaşlarının kullanıcı id’leri 5.000’lik paketler halinde elde edilebilmektedir.

Başka bir deyişle 1 token ile 15 sorguda (ya da 15 dakikada) 1 kullanıcının 75.000 takipçisi ya da arkadaşına ilişkin veri alınabilmektedir. Dolayısıyla 1 token ile 150.000 takipçisi ve 150.000 arkadaşı olan bir kullanıcının takipçi ve arkadaş verilerini temin etmenin zaman maliyeti 60 dakika olmaktadır. Normal bir dağılım sergilediği varsayımı altında 30.000.000 kullanıcıdan oluşan bir veri seti için, toplam zaman maliyeti 30.000.000 saat olmaktadır. Bu zaman maliyeti birden fazla bilgisayar kullanarak eş anlı sorgu ile kısmen azaltılabilmektedir⁵⁰. Diğer taraftan Twitter verileri, web kazıma [web scraping] tekniğiyle toplandığında da, sayfa içeriğinin yüklenip verilerin kazınması oldukça yavaş gerçekleşmektedir. Bu nedenlerle Twitter verileri, gerçek zamanlı olarak güncellenememiş, çalışmanın başında (2017), ortasında (2018) ve sonunda (2020) olmak üzere, üç kez toplanabilmiştir. Ayrıca, bazı kullanıcıların zaman içerisinde hesaplarını kapattığı, bazı kullanıcıların tweetlerini sildiği, bazı kullanıcılarında erişime açık hesaplarını korumalı duruma getirebildiği de göz önünde bulundurulmalıdır. Doğal olarak bu kapsamdakiler, analizin dışında kalmıştır. Twitter verileriyle ilgili bir diğer kısıt ise tweetlerin toplanmasıyla ilgilidir. Twitter API, bir kullanıcının geçmişe dönük son 3.200 tweetini alabilmeye imkân tanımaktadır. Bu kısıt ise özellikle çok sık tweet atan kullanıcılar için tarihsel olarak tweet arşivinde çok geriye gidilememesine neden olmuştur. Bu sorunu aşabilmek amacıyla da analizin zaman aralığı oldukça geniş tutulmuş, uzun süre (2017-2020) tweetlerin toplanmasına devam edilmiştir.

Araştırmanın kısıtlarından bir tanesi de korumalı tweetlerle ilgilidir. Twitter kullanıcılarının tweetleri varsayılan olarak herkese açıktır. Yani herkes kullanıcının paylaştığı tweetleri görüntüleyebilmekte ve onlarla etkileşim kurabilmektedir. Ancak bazı kullanıcılar tweetlerini korumaya alabilmektedir. Bu durumda insanlar kullanıcıyı takip etmek istediğinde onaylayabileceği veya reddedebileceği bir istek gönderirler. İstek onaylandığında, istek gönderen kullanıcı takipçi olur ve sadece takipçiler hesabın paylaştığı tweetleri görebilir ve kullanıcıyla etkileşim kurabilir. Araştırma kapsamında korumalı

⁵⁰ Bu araştırmada sırasında bilgisayar laboratuvarlarını kullanma imkânı veren Yozgat Bozok Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dekanlığı'na ve İşletme Bölüm Başkanlığına teşekkür ederim.

hesaplar, paylaşımlarının herkese açık olmamasından dolayı analizin dışında tutulduğu için, etkileşimlerinin ve tweetlerinin etkisi bulgulara yansımamıştır.

Araştırma sırasında karşılaşılan bir diğer kısıt ise web sayfalarından veri kazanmasıyla ilgilidir. Web sayfalarından bazıları, belirli bir zaman dilimindeki istek sayısı belirli bir eşiği aştığında, belirli IP'ler için siteye erişimi durdurabilmektedir. Dolayısıyla finansal derlemin oluşturulması için gerekli olan içerikler, bu tür web sayfaları için zamana yayılarak toplanmıştır.

Analizde karşılaşılan bir diğer kısıt ise, Türkçe dili için geliştirilen duygu analiz araçlarının İngilizce dili için geliştirilmiş olanlara kıyasla daha sınırlı kalmasıdır. Bu anlamda Türkçe için etiketli veri setleri de oldukça sınırlıdır. Bu nedenle veri setlerinin bir kısmı araştırma özelinde oluşturulmuş, bu durum da araştırmaya ilave bir zaman maliyeti olarak yansımıştır.

Araştırmanın bir başka kısıtı ise metinlerin duygusal ve fikri yönelimlerinin *pozitif*, *negatif* ve *nötr* olmak üzere üç sınıfla temsil edilmesidir. Büyük verinin anlamlandırılmasında maruz kalınan, özellikle zaman maliyeti nedeniyle literatürdeki çalışmaların büyük bir çoğunluğunun ya *pozitif* ve *negatif* gibi ikili bir sınıflandırmayı, ya da bunlara *nötr* sınıfını da ekleyerek üçlü bir sınıflandırmayı tercih ettiği görülmektedir. Esasında bir renk skalasındaki çeşitliliğe ve geçişkenliğe benzeyen duygu ve fikirlerin sadece üç boyutla temsil edilmesi önemli bir kısıttır. Ancak çeşitlilik nedeniyle analizin ulaşacağı boyut, analiz araçlarının kapasitesi ve performansı düşünüldüğünde bu yaklaşım makul karşılanabilir. Ayrıca temsil çeşitliliğinin artması, bir önceki paragrafta ifade edildiği gibi, gözetimli öğrenme yaklaşımlarında ihtiyaç duyulan, her bir sınıfı temsil edebilecek niteliklere sahip, etiketli verinin elde edilebilmesiyle de yakından ilgilidir. Bununla birlikte kategorik bir veri yerine -1 ile +1 arasında değişen duygu olasılığını gösteren bir skalada da kullanılabilmektedir. Ancak duygu olasılığı her ne kadar -1 ile +1 arasında olsa da eğitim verisi yine kategorik olacağından duygunun çeşitliliği ve geçişkenliğindeki kayıp tümüyle ortadan kaldırılamamaktadır.

Araştırmanın bir diğer kısıtı da bireysel yatırımcı karar modeli ile ilgilidir. Bu model bireysel yatırımcıların bu üç boyutun etkisi altında karar verdiklerini varsaymakta, ancak bilişsel, davranışsal ve sosyal bileşenlerin bazı durumlarda sınırlarını net bir şekilde çizerek ayırtmak pek mümkün olamamaktadır. Başka bir ifadeyle araştırmada kullanılan bazı öznitelikler hem davranışsal hem bilişsel hem de sosyal nitelikte olabilmektedir. Bu kısıt nedeniyle nihai çözümleme bütünlük karar modelleri ile gerçekleştirilmekte, özniteliklerin modeldeki etki düzeyleri açıklanabilir yapay zekâ [explainable AI (XAI)] yardımıyla analiz edilerek yorumlanmaktadır.

4.5 ARAŞTIRMADA KULLANILAN YÖNTEMLER VE ANALİZ AŞAMALARI

Bu araştırma kapsamında geliştirilen çözümleme modeli üç katmanlı bir mimariye sahiptir. Katmanlardan biri karar verme sürecinin *bilişsel* boyutuna, diğeri *davranışsal* boyutuna, bir diğeri ise karar verme sürecinin *sosyal* boyutuna odaklanmaktadır.

Şekil 4.2 bireysel yatırımcıların bu çalışma kapsamında araştırılan karar modelini temsil etmektedir. Şekil 4.2'nin zemin rengi bağlamı, şeritler bireysel yatırımcının kararına etki eden boyutları, şeritlerin sarmal yapısı ise karar süreçlerinin bütünlük yapısını sembolize etmektedir. Ayrıca karar modelinde yer alan sosyal, bilişsel ve davranışsal boyutların sınırlarının her zaman kesin çizgilerle ayıramayabileceğini, yani iç içe geçmiş alanların olduğu gözden uzak tutulmamalıdır.

Şekil 4.2: Bireysel Yatırımcı Karar Modeli



Bilişsel katmanda fikri ve duygusal içerik barındıran metinsel verilerin, yani tweetlerin içeriğinin doğal dil işleme ve metin madenciliği yöntemleri kullanılarak anlamlandırılması ve bu anlamdan yola çıkarak etkileşimlerin finansal piyasalarla ilişkisi çözümlenmektedir. Diğer bir ifadeyle bilişsel katmanda bağlam, duygu ve fikir sınıflandırması yapılarak, her bir tweetin bağlamını ve duygusal yönelimini gösteren iki temel öznitelik elde edilmekte ve bu iki temel öz niteliğin ve bunlardan türetilen özniteliklerin borsa ile ilişkisi incelenmektedir. Davranışsal katmanda Türkçe paylaşım yapan Twitter kullanıcılarının paylaştığı tweetlere diğer kullanıcıların verdiği tepkiler, yani retweet sayısı, beğeni sayısı gibi tepkisel öznitelikler analize dâhil edilmektedir. Başka bir deyişle davranışsal katmanda tweetlerin içeriğine bakılmaksızın, tweetlere verilen tepkilerden nesnel bir öznitelik elde edilmektedir. Sosyal katmanda ise Twitter platformundan elde edilen kullanıcı etkileşimleri yer almaktadır. Başka bir deyişle bu katmanda Twitter kullanıcılarının etkileşim dinamikleri aracılığıyla bireysel yatırımcıların bilgi paylaşma ve arama davranışları ile karar verme süreçlerindeki sosyal etki modellenmektedir. Araştırmanın son aşamasında karar vermenin bilişsel, davranışsal ve sosyal katmanları bütünleşik hale getirilerek borsa endeksleriyle ilişkisi analiz edilmektedir.

4.5.1 Borsa Endeksleri ile İlgili Değişkenler

Analiz kapsamına XU030, XU100, XUTUM, XUMAL, XUSIN, XUTEK ve XUHIZ endeksleri dâhil edilmiştir. Analizin ilk aşamasında davranışsal, bilişsel ve sosyal öznitelikler ile her bir endeksin endeks açılış değeri, işlem hacmi, volatilité ve getiri ile ilişkisi Pearson korelasyon analizi yöntemi ile çözümlenmiştir. Korelasyon analizine çok boyutlu büyük veri çalışmalarında analizin nereye odaklanması gerektiği konusunda ön bilgi edinmek amacıyla sıklıkla başvurulmaktadır. Ayrıca korelasyon sonuçları, lineer regresyon modeli kurarken filtreleme yaklaşımlarına dayalı öznitelik seçiminde de kullanılmaktadır.

Çalışmada kullanılan değişkenlerden biri endekslerin açılış değerleridir. Önceki günün açılış saatinden aynı günün açılış saatine kadar ki kolektif duygu ve fikir yöneliminin açılış fiyatlarına etkisini araştırabilmek amacıyla bu değişken analize dâhil edilmiştir. Ayrıca

literatürdeki bazı bulgular (Cui ve diğerleri, 2016) özellikle açılıştan açılışa hesaplanan duygu ve fikir skorları ile açılış fiyatları arasındaki ilişkinin istatistiksel olarak daha anlamlı olduğuna işaret etmektedir. Bu nedenle bu çalışma kapsamında çözümlemeye kapanış fiyatları yerine açılış fiyatları dâhil edilmiştir.

Çalışmada kullanılan bir diğer değişken işlem hacmidir. İşlem hacmi, işleme konu olan menkul kıymetin adedinin işlem fiyatı ile çarpılması suretiyle hesaplanmaktadır. Başka bir ifadeyle işlem hacmi, belirli bir zaman diliminde yapılan alım satım işlemlerinin toplam değeridir. Bu bağlamda işlem hacmi önemli likidite göstergelerinden biridir ve aynı zamanda bir menkul kıymetin, endeksin ya da piyasanın genel faaliyetini temsil etmektedir.

Çalışmada kullanılan değişkenlerden bir diğeri volatilitedir. Volatilite bir menkul kıymetin fiyatının belirli bir zaman aralığındaki dalgalanma düzeyidir. Bu çalışma kapsamında borsa endekslerinin volatilitesi benzer çalışmalarda (Li ve diğerleri, 2018b) tercih edilen ve Eşitlik 4.1’de yer alan Parkinson (1980) yaklaşımına göre hesaplanmıştır. Eşitlikte h_i sembolü en yüksek fiyatı, l_i sembolü en düşük fiyatı temsil etmektedir. Ayrıca bunun bir yıldaki işlem dönemi sayısının kareköküyle çarpılarak yıllıklaştırılması gerekmektedir.

$$Volatilite_p = \sqrt{\frac{1}{4 \ln(2)} \sum_{i=1}^N \left(\ln \frac{h_i}{l_i} \right)^2} \quad (4.1)$$

Bu çalışmada piyasa karakteristiğini temsilen kullanılan bir diğer değişken ise getiridir. Eşitlik 4.2 getirinin nasıl hesaplandığını formüle etmektedir. Eşitlikte p_t sembolü belirli bir zaman noktasındaki kapanış fiyatını, p_{t-1} bir önceki zaman noktasındaki kapanış fiyatını temsil etmektedir.

$$Getiri = \left(\frac{p_t - p_{t-1}}{p_{t-1}} \right) \quad (4.2)$$

4.5.2 Bilişsel Katman: Doğal Dil İşleme ve Derin Öğrenme Yöntemleri ile Finansal Bağlam, Fikir ve Duygu Sınıflandırması

Ekonomi ve finansla ilgili paylaşılan sayısal ve yapılandırılmış verilerin modern finans kuramları çerçevesinde anlamlandırılması, isabetli kararların verilmesinde oldukça önemli bir yer tutmaktadır. Özellikle gelişen bilgisayar ve iletişim teknolojileriyle birlikte bu veriler hızla hesaplanarak elde edilen sonuçlar karar süreçlerine kolaylıkla dâhil edilebilmektedir. Ancak bu sayısal veriler kadar önemli olan bir diğer konu, bu verilerin nasıl yorumlandığı, bu göstergelerle ilgili olarak özelde uzmanların, genelde ise insanların ne düşündüğü hususudur. Zira sayıların ya da verilerin bilgiye dönüşmesi ve bu bilginin alınan kararlara ışık tutması, bunların nasıl yorumlandığı, anlaşıldığı ya da algılandığıyla yakından ilgilidir. Bu bağlamda ekonomi ve finansla ilgili metinler yorumlama ya da anlamlandırma sürecinin hem çıktısı hem girdisi olarak işlev görmektedir. Dolayısıyla ekonomi ve finansla ilgili metinlerin makineler tarafından anlamlandırılması ve buradan elde edilen bilgilerin karar süreçlerine dâhil edilmesi oldukça önemlidir. Bu başlıkta finansal metinlerin makineler tarafından sayısallaştırılarak anlamlandırılması için kullanılan doğal dil işleme, metin madenciliği, sözcük temsilleri (kelime vektörleri) ve derin öğrenme yöntemlerinden bahsedilmekte daha sonra araştırma kapsamında bu yöntemler kullanılarak eğitilen *finansal bağlam sınıflandırıcısı* ile *finansal fikir ve duygu sınıflandırıcısı*na ilişkin ayrıntılar açıklanmaktadır.

4.5.2.1 Doğal Dil İşleme ve Sözcük Temsilleri

Diller genellikle makine dilleri ve doğal diller olmak üzere iki sınıfa ayrılmaktadırlar. Makine dilleri⁵¹ denildiğinde günlük kullanımda Java, Python, C++ gibi programlama dilleri ifade edilmektedir. Bu diller insanlar tarafından kurallarının bilinçli bir şekilde ve açıkça belirlendiği, yapılandırılmış dillerdir. Doğal diller ise insanların günlük yaşamında bilinçli planlama ve kasıt olmadan ortaya çıkan, tekrarlar yoluyla kendiliğinden gelişen ve insanların

⁵¹ Esasen makine dili, programla dillerinin en alt seviyesini oluşturan ikilik tabandaki sayılardan ibarettir. Ancak günlük kullanımda *makine dili* ifadesiyle genelde programlama dilleri kastedilmektedir.

iletiřim kurmalarını saęlayan Trke, Arapa, ince, İngilizce gibi dilleri ifade etmektedir. Bu diller olduka kompleks, deęiřken ve evrilmeye aıktır. Bu baęlamda insanlar dil retme, ayrıntılı ve nanslı anlamları ifade etme, anlama ve yorumlama yeteneęine sahipken, dilin kurallarını anlama ve aıklamada ok bařarılı deęillerdir (Goldberg, 2017). Doęal diller, makine dillerinden farklı olarak yapılandırılmamıř dillerdir. Sembolik ve soyuttur. Bařka bir syleyiřle, dilin en temel gesi olan karakterler ya da sesler bir araya gelerek nesneleri, kavramları, olayları, eylemleri ve fikirleri ifade etmektedirler. Bu baęlamda hem karakterler (ya da sesler) hem szckler soyut sembollerdir. Zira hem genel bir temsili aęrıřtırırlar hem de her bir zihnin kendi yorumuna bırakılan soyut bir sembole tekabl ederler. Ayrıca doęal diller bileřimsel bir zellięe de sahiptir. Dięer bir ifadeyle, harfler veya sesler szckleri, szckler szck beklerini ve cmleleri oluřtururlar. Bu nedenle bir ifadenin anlamı bir takım girift kuralları takip ederek tek tek onu oluřturan szcklerin anlamından farklı bir řeyi ifade edebilmektedir. Dolayısıyla anlam bazen karakterlerde, bazen szcklerde, bazen szck beklerinde, bazen cmlelerde ve aslında btn bunların diziliminde tezahr etmektedir.

Doęal dil iřleme, ontolojik baęlamda “anlam”⁵² kavramına, alıřma kapsamına indirgendięinde ise *dilsel ifadeleri anlamlı kılan řeyin ne olduęuna* odaklanmaktadır. znde tmeller sorunuyla iliřkili bu sorulara yanıt arayan pek ok kuram geliřtirilmiřtir. Bunlardan *znellik*, bir szcęn anlamının insan zihninde yer aldıęını, insanların zihinlerinin dıřında anlamların olduęu bir dnyanın sz konusu olamayacaęını savunmaktadır. Dolayısıyla anlam denilen řey zneldir. John Locke’nin nderlięini yaptıęı bu grř, gnmzde Noam Chomsky’nin anlamın kafanın ierisinde olduęu savıyla srdrlmektedir. stelik biliřsel bilim ve yapay zekâ alıřmalarında da yaygın olarak kabul grmeye devam etmektedir. *Gerekilik*, anlamın dilden ve insan zihninden baęımsız, soyut bir varlık olduęu ve bu sayede insanların iletiřim kurabildięi, anlamı dilden dile nesilden nesile aktarabildięi grřndedir. Dolayısıyla szckler ancak bir araya gelip bir tmce oluřturduklarında anlam kazanmaktadırlar. Dięer bir ifadeyle bir tmcenin dile getirdięi anlam, o tmcenin

⁵² *Anlam szcęnn szlkteki aıklamasında bir kelimeden, bir szden, bir davranıř veya olgudan anlařılan řey, bunların hatırlattıęı dřnce veya nesne olduęu ifade edilmektedir(TDK, 2009).*

parçalarının anlamlarından oluşmaktadır. Bu kuramlara ilave olarak *bütüncülük*, sözcüklerin anlamlarının tek tek değil, bir dilin kavramsal çerçevesi içinde bir bütün olarak kurulduğunu; *doğrulamacılık*, bir tümcenin anlamının o tümcenin doğruluk koşullarında yattığını; *dışsalcılık*, sözcüklerin anlamlarının zihinlerde değil dış dünya tarafından belirlendiğini; *doğalcılık*, her şeyin doğa içinde bilimsel bir açıklaması olduğunu savunmakta ve tümellerin varlığını tümenden reddederek her şeyin tikel olduğunu iddia etmektedir. Son olarak, yirminci yüzyıla iz bırakan pozitivistizmin bir uzantısı olan davranışçılık öğretisinden türeyen ve aynı isimle anılan *davranışçılık* kuramı ise, anlamın sözcükler duyulduğunda gösterilen gözlenebilir davranışlarda yattığını savunan anlam kuramıdır. Bu görüşe göre bir sözcüğün anlamı, o sözcüğün kullanımına yönelik davranışsal eğilimlerde gizlidir. Davranışsal olmasının nedeni, zihnin içerisinde olup bitenlerden ziyade dışarıdan gözlemlenebilir olana işaret etmesiyle, eğilim olmasının nedeni ise hiç gerçekleştirilmiş olmasa bile belirli bir şekilde davranmaya mütemayil olmasıyla ilgilidir (İnan, 2012).

Doğal dil işleme dildeki anlamı anlamaya dönük görüşlerden yola çıkarak metinlerin, yani karakter veya sözcük dizilerinin makinelerce nasıl anlaşılabilir hale getirileceğiyle ilgilenmektedir. Başka bir deyişle dildeki anlamın nasıl sayısallaştırılabileceği ve neyle/nasıl temsil edilebileceği problemine odaklanmaktadır. Bu amaçla geliştirilen üç farklı yaklaşımdan söz edilebilir. Bu yaklaşımlardan birincisi *dilbilimsel* (linguistik) yaklaşımdır. Dil bilgisini kullanarak cümle yapısına göre analiz yapmaktadır. Temelde dilin yapısını oluşturan şekilbilim [morphology], sözdizim [syntax], anlambilim [semantics] ve kullanımbilim (pragmatics) konularına dayanarak dilin özellik ve kurallarına dayalı yaklaşımdır (Şeker, 2015a). İkinci yaklaşım *istatistiksel* yaklaşımdır. İstatistiksel yaklaşımlarda metinlerin bilgisayarlar tarafından anlamlı olabilmesi için karakterlerin ya da kelimelerin sayısal olarak temsil edilmesi gerekmektedir. Bu amaçla geliştiren yöntemlerden bazıları anlamı karakter düzeyinde, bazıları sözcük düzeyinde sembolize etmektedir. Metnin, kelime, karakter ya da n-gram⁵³ gibi farklı birimlere ayrılması sonucu elde edilen yapılara

⁵³ N-gram bir kelime ya da cümleden çıkartılabilen ve birbirini takip eden, n tane karakter ya da kelime gruplarıdır. Örneğin kelime düzeyinde “Borsa yükselişini sürdürdü.” cümlesinin 2-gram kümesi {“Borsa”, “Borsa yükselişini”, “yükselişini”, “yükselişini sürdürdü”, “sürdürü”}; 3-gram kümesi {“Borsa”, “Borsa yükselişini”, “Borsa yükselişini sürdürdü”} şeklinde olmaktadır.

dizgecik [token] denir. Doğal dil işlemede öncelikle karakter ya da kelimelerin sayısal olarak ifade edilebilmesi için metin dizgeciklere [tokens] ayrılarak her bir dizgecik vektörlere dönüştürülmektedir. Bu anlamda metni vektör haline getirme işlemi, metni sayısal tensörlere dönüştürme işlemidir (Chollet, 2019). Diğer bir ifadeyle metnin sayısallaştırılması, dizgeciklere ayrılıp her bir dizgeciğin sayısal tensörlerle ilişkilendirilmesinden ibarettir. Ayrıca TF-IDF, kelime çantası gibi yöntemler de kullanılmaktadır. Son olarak hem dilbilimsel hem istatistiksel yaklaşımların birlikte kullanıldığı üçüncü gruptaki yaklaşımlar ise *hibrit* yaklaşımlar olarak ifade edilmektedir.

Bu araştırma kapsamında kelime düzeyindeki istatistiksel yöntemlerden bahsedilmekte ve tweetlerin sınıflandırılmasında bu yöntemler kullanılmaktadır. Bu yaklaşımda kelimeler dilsel anlamın temel birimleri olarak kabul edilmektedir. Kelimeleri sayısal olarak ifade edebilmek amacıyla her bir kelime dizgeciklere [word tokens] ayrılarak kelime uzayında bir noktayı temsil eden vektörlere kodlanmaktadır. Orijinal uzayda her bir kelime, Şekil 4.3'te ki gibi, *one-hot encoding* ile gösterildiğinden her bir kelime çiftinin birbirine öklid uzaklığı $\sqrt{2}$ 'dir (Goodfellow ve diğerleri, 2018). Başka bir deyişle *one-hot* vektör uzayında kelimeler birbirine eşit uzaklıkta olduğundan, herhangi bir kelime hakkında herhangi bir komşusundan bilgi edinilmesi mümkün değildir.

Şekil 4.3: One-Hot Encoding Örneği

$$w^{altın} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, w^{borsa} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, w^{dolar} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix},$$

Kelimenin anlamı pek çok durumda bağlama göre, yani birlikte bulunduğu kelimelere göre ortaya çıkmaktadır. Dolayısıyla bir kelimenin anlamını ortaya çıkarabilmek için diğer kelimelerle bağlantılarını ya da ilişkilerini gösteren vektörlere ihtiyaç duyulmaktadır. Bu yaklaşımlarda benzer bağlamlarda sıklıkla görülen kelimeler birbirlerine yaklaştırılarak, benzer anlamlara sahip olanların birbirlerine komşu olması sağlanmaktadır. Model

geliştirilirken her bir kelimenin bir vektörle temsil edildiği gömülme matrisi [embedding matrix] oluşturulmaktadır. Model eğitildikten sonra bu matristeki her bir vektör temsil ettiği kelimenin diğer kelimelerle ilişkisini, yani anlamını temsil etmektedir. Kelime bağlantılarının bu şekilde vektörel olarak gösterilmesine kelime gömümleri denilmektedir. Başka bir deyişle kelime gömümleri [word embedding] insan dilinin geometrik uzaya eşlenmesidir (Goodfellow ve diğerleri, 2018).

Kelime vektörü oluşturmak için kullanılan çok farklı yöntemler vardır. Bunlar genel olarak frekans tabanlı [frequency based] ve tahmin tabanlı [prediction based] olmak üzere iki kategoride toplanmaktadır. Frekans tabanlı yöntemlerin sıklık vektörü [count vector], TF-IDF vektörü [TF-IDF vectorization] ve birlikte bulunma matrisi [co-occurrence matrix] olmak üzere üç farklı yaklaşıma sahip olduğu söylenebilir. Sıklık vektöründe [count vector] model tüm dokümanlardan bir sözcük dağarcığı oluşturmak suretiyle öğrenmektedir. Her bir dokümanda her bir kelimenin kaç kez geçtiği hesaplanarak satırlar dokümanlara, sütunlar dizgeciklere [token] karşılık gelecek şekilde sıklık matrisi oluşturulmaktadır. Örneğin 3 dokümanda toplam 7 farklı sözcük içeren bir sözcük dağarcığı oluşturulduğu varsayıldığında 3×7 boyutunda bir matris elde edilmiş olur. Ancak derlem büyüdükçe bu matrisin boyutu da büyümektedir.

Kelimelerin sadece sıklığına bakıldığında, büyük derlemlerde (ve, veya, ama gibi) bazı kelimelerin çok fazla tekrar etmesinden dolayı belgenin içeriği hakkında çok az bilgi sahibi olunmaktadır. Bu sorunu aşmak için TF-IDF yöntemi ortaya atılmıştır (Sparck Jones, 1972; Salton ve Buckley, 1988). Bu yöntem bir sözcüğün sadece tek dokümanda ortaya çıkmasından ziyade, bütün derlemdeki durumu hesaba katmaktadır. Böylece yüksek frekansa sahip, ancak metin hakkında çok fazla bilgi vermeyen kelimelerin etkisi azaltılmış olmaktadır. TF ve IDF'nin hesaplanması ve aralarındaki ilişki aşağıdaki gibidir.

$$TF = \frac{\text{Bir Kelimenin Bir Dökümanda Görünme Sıklığı}}{\text{Dökümandaki Toplam Kelime Sayısı}}$$

$$IDF = \log \left(\frac{\text{Toplam Döküman Sayısı}}{t \text{ Kelimesinin Görüldüğü Dökümanların Sayısı}} \right)$$

$$TF - IDF = TF \times IDF$$

Bunlara ilave olarak tekil değer ayrışımını [singular value decomposition (SVD)] kullanan iki yaklaşım da söz konusudur. Bunlardan kelime-doküman matrisi [word-document matrix] birbiriyle ilişkili kelimelerin aynı dokümanlarda daha sık geçeceğini varsaymaktadır. Örneğin “politika” konusundaki metinlerin “cumhurbaşkanı”, “meclis”, “hükümet” gibi sözcüklerle; “ekonomi” konusundaki metinlerin “borsa”, “altın”, “döviz”, “enflasyon” gibi sözcüklerle inşa edilme ihtimali daha yüksektir. Bu yaklaşıma göre bir metinde birlikte geçen kelimelerin sıklığına bakarak metnin konusu belirlenebilmektedir. Buna Latent Semantic Analysis (LSA) denilmektedir (Deerwester ve diğerleri, 1990). Ancak doküman sayısı arttıkça matris oldukça büyümekte, bu da bilgisayarlarda hafıza problemlerine yol açmaktadır (Rosario, 2000).

Tekil değer ayrışımını kullanan bir diğer yaklaşım ise birlikte bulunma matrisidir [co-occurrence matrix]. Tablo 4.4’te gösterilen bu yöntemde, derlemdeki her kelimenin, ilgili kelime etrafında belirli bir boyuttaki bir pencerede kaç kez birlikte bulunduğu sayılmaktadır. Bu yöntemi diğerlerinden ayıran özellik ise kelimelerin belirli bir pencere aralığında birlikte bulunmalarının olasılığının hesaplanmasıdır. Bu hesaplama tüm doküman üzerinde yapılmaktadır. Böylece kelimelerin birlikte bulunma olasılıklarını temsil eden bir matris oluşturulur (Manning, 2020a). Ancak bu yöntemle vektör oluşturulduğunda da elde edilen matris çok büyük olmaktadır. Bu matrisin boyutunu düşürmek için tekil değer ayrışımı kullanılmaktadır. Böylece birlikte bulunma matrisi, kelime vektörlerini temsilen tekil değer ayrışımı ile hesaplanan faktörlerin kombinasyonuna dönüşmektedir (Manning, 2020b).

Örneğin elimizde aşağıdaki cümlelerden oluşan bir derlem olsun.

- Borsa günü düşüşle tamamladı.
- Borsa, altın ve dolar kazandırdı.
- Borsa günü yükselişle tamamladı.

Tablo 4.4: Birlikte Bulunma Matrisi

	altın	borsa	dolar	düşüşle	günü	kazandırdı	tamamladı	ve	yükselişle
altın	0	1	1	0	0	1	0	1	0
borsa	1	0	1	1	2	1	2	1	1
dolar	1	1	0	0	0	1	0	1	0
düşüşle	0	1	0	0	1	0	1	0	0
günü	0	2	0	1	0	0	2	0	1
kazandırdı	1	1	1	0	0	0	0	1	0
tamamladı	0	2	0	1	2	0	0	0	1
ve	1	1	1	0	0	1	0	0	0
yükselişle	0	1	0	0	1	0	1	0	0

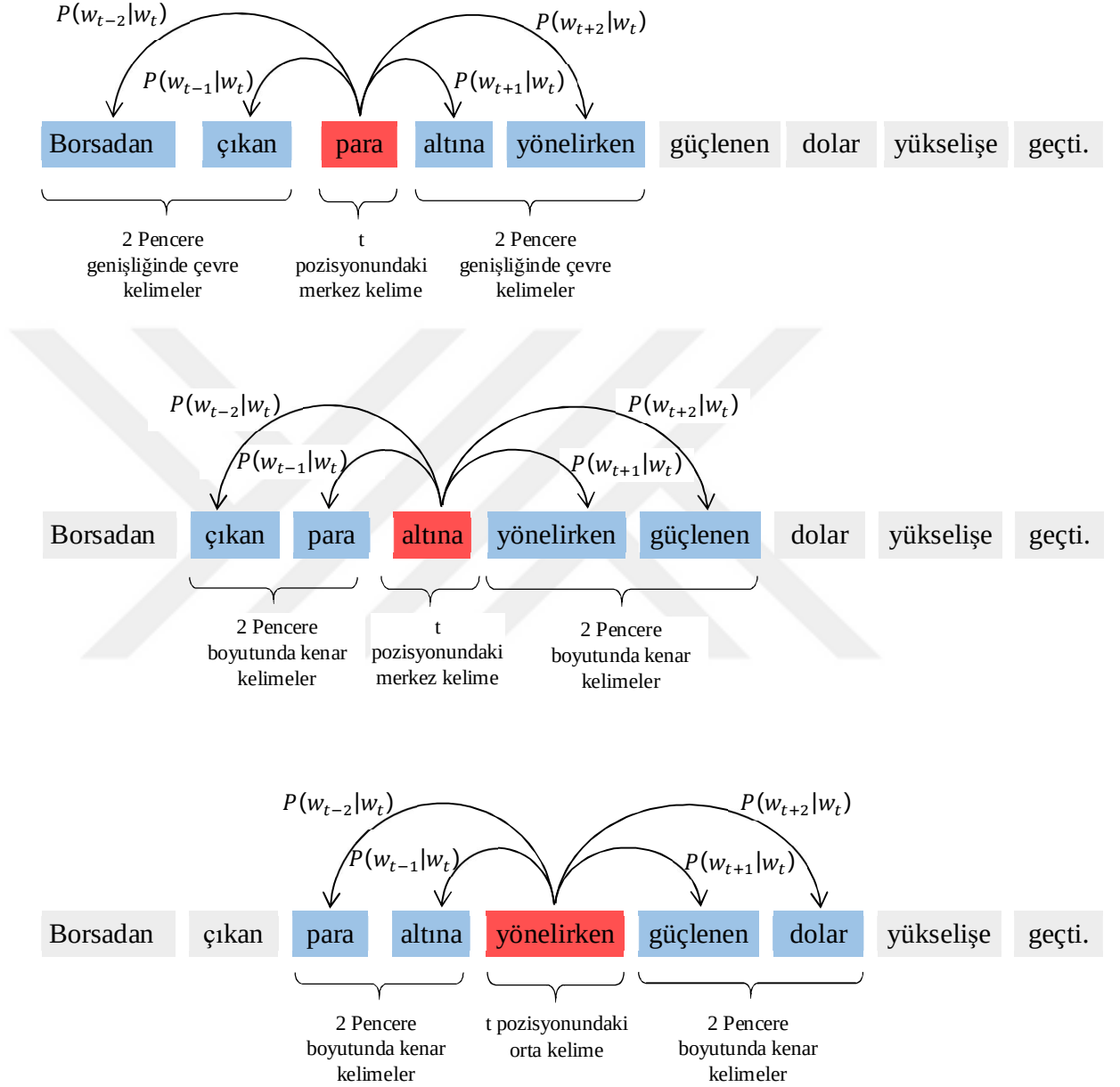
Birlikte bulunma matrisi daha önce bahsedilen yöntemlere kıyasla sözcükler arasındaki ilişkiyi korumaktadır. Ancak bu yaklaşımda da sözcük dağarcığına sıklıkla yeni kelimeler eklendiğinden matrisin boyutu da sıklıkla değişmektedir. Bu durumda da oldukça büyük bir hafızaya ihtiyaç vardır. Bu sorunu aşmak için de küçük bir boyut belirlenip en önemli bilgiler saklanmaktadır (Manning, 2020b). Bu yöntemde karşılaşılan bir diğer problem ise çoğu kelimenin birlikte bulunmamasından dolayı elde edilecek matrisin oldukça seyrek [sparse matrix] olmasıdır. Bunu aşmak için ise etkisiz kelimeler [stopwords] metinden çıkarılmaktadır. Ancak bu sorunların pek çoğunu yinelemeye dayalı [iteration based] ya da tahmin tabanlı [prediction based] yöntemler çok daha başarılı bir şekilde çözmektedir (Roy, 2023). Bu yöntemler Word2Vec, GloVe, fastText, ELMO, BERT olarak sayılabilir.

Word2Vec yöntemi, birlikte bulunma matrisine benzer şekilde, bir kelimenin anlamının etrafındaki kelimelerle şekilleneceği fikrine dayanan gözetimsiz [unsupervised] bir öğrenme

yöntemidir (Mikolov ve diğerleri, 2013). Yöntem her bir kelimenin birlikte kullanıldığı diğer kelimelerle ilişkisine bakarak o kelime hakkında bir özellik vektörü oluşturmaktadır. Bunun için herhangi bir kelime esas alınarak etrafındaki kelimeler ya da etrafındaki kelimeler esas alınarak hedef kelime tahmin edilmeye çalışılmaktadır. Bu olasılıkların hesaplanıp modelin eğitilmesi, loss fonksiyonunun minimize edilirken olasılık değerlerinin maksimize edilmesi ve böylece benzer kelime vektörlerinin birbirine yaklaştırılması şeklinde olmaktadır. Böylece kelime vektörleri etrafındaki diğer vektörler tahmin edebilir hale gelmektedir.

Word2Vec yaklaşımında kelime vektörü oluşturmak için Skip-Gram ve CBOW olmak üzere iki farklı yöntem kullanılmaktadır. Skip-Gram belirli bir kelimenin belirli bir pencere genişliğinde diğer kelimelerle bir arada bulunma olasılığını hesaplamaktadır. Pencere genişliği [windows size] merkezde bulunan kelimenin [center word] öncesindeki ve sonrasındaki kaç sözcüğün hesaplanmaya dâhil edileceğini belirlemektedir. Örneğin pencere genişliği 2 olduğunda algoritma hesaplamanın merkezindeki kelimenin sağında ve solunda bulunan 2, toplamda 4 kelimeyi tahmin etmek için olasılık dağılımları hesaplanmaktadır. Daha sonra bu olasılıklar optimizasyon ile artırılmaktadır.

Şekil 4.4: Skip-Gram Yaklaşımı



Örneğin Şekil 4.4'te görüldüğü gibi *para* kelimesi merkez kelime olarak alındığında, (pencere büyüklüğü 2 alınmışsa) *para* kelimesinin 2 pencere öncesinde ve sonrasında bulunan kelimeler tahmin edilmeye çalışılmaktadır. Skip-Gram algoritması ilk sözcükten

başlayarak kelime kelime kaydırılmak suretiyle metindeki bütün sözcüklerin belirlenen pencere büyüklüğünde, yani aynı pencere içerisinde bulunma olasılıklarını hesaplamaktadır. Bu yöntemle tüm derlem üzerinden elde edilen vektörler kelimelerin birbirleriyle bağlantısını başarılı bir şekilde temsil etmektedir.

Skip-Gram yönteminde amaç bir kelimenin bağlamını tahmin etmektir. Bu nedenle modelin eğitilmesi amaç fonksiyonunu maksimize eden sözcük kümesini bulmak anlamına gelmektedir. Eşitlik 4.3'te m pencere büyüklüğünü, t kelimenin pozisyonunu gösterirken, amaç, t pozisyonundaki bir kelimenin m kadar sağındaki ve solundaki kelimeleri tahmin etmek ve bu tahmin olasılıklarını maksimize etmektir. Olasılıkları maksimize edebilmek için logaritması alınmaktadır. Bu durumda amaç fonksiyonu loss ya da cost fonksiyonunun minimizasyonu eşdeğer hale gelir. Dolayısıyla denklem loss fonksiyonu şekilde ifade edilebilir (Manning, 2020a; Aydoğan ve Karcı, 2019).

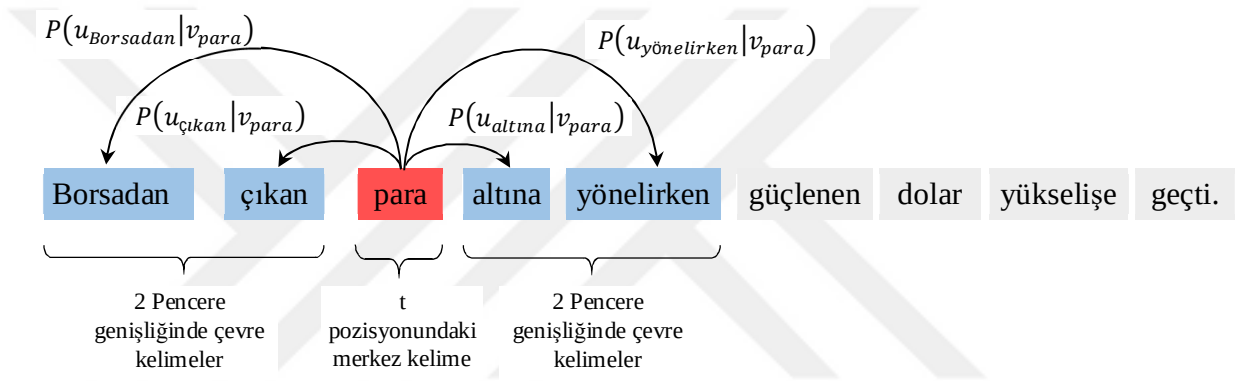
$$J = -\frac{1}{T} \sum_{t=1}^T \sum_{\substack{-m \leq j \leq m \\ j \neq 0}} \log P(w_{t+j}|w_t) \quad 4.3$$

Skip-Gram modelinde amaç fonksiyonu, yani *loss* ya da *cost* fonksiyonu, kaybı minimize ederek tahmin olasılığını artırmayı hedeflemektedir. Eşitlikteki $\frac{1}{T}$ ise olasılıkları tüm derlem üzerinde hesaplamak yerine her bir pozisyonun ortalamasının alınmasını sağlamaktadır. Eşitlikteki $P(w_{t+j}|w_t)$ 'yi hesaplayabilmek amacıyla her bir kelime için iki vektör kullanılmaktadır.

$$P(o|c) = \frac{\exp(u_o^T v_c)}{\sum_{j=1}^n \exp(u_w^T v_c)} \quad 4.4$$

Bu vektörlerden v_w , w merkez kelime iken olasılığını, diğeri yani u_w , w bağlam kelimesi iken olasılığını göstermektedir. Örneğin $P(u_{Borsadan}|v_{para})$ 'nın açılımı $P(Borsadan|para; u_{Borsadan}, v_{para})$ şeklinde olmaktadır. Eşitlik 4.4'teki t kelimenin pozisyonunu, o ve c ise kelimenin sözcük dağarcığındaki indeksini göstermektedir. Bu iki vektörden elde edilen sonuçları nokta çarpımıyla hesaplanarak olasılık dağılımına dönüştürmek için ise softmax kullanılmaktadır (Manning, 2020a).

Şekil 4.5: Skip-Gram Örneği



Skip-Gram algoritmasının bir modifikasyonundan da kısaca bahsetmek gerekmektedir (Mikolov ve diğerleri, 2013). Negatif örnekleme [negative sampling] olarak bilinen bu modifikasyon, eğitimin derlemdeki tüm kelimeler üzerine uygulanmasından ziyade rastgele seçilen birkaç kelimeye uygulamayı hedeflemektedir. Eşitlik 4.5'in birinci kısmı iki kelimenin aynı pencerede bulunma olasılığının logaritmasının softmax yerine sigmoid ile hesaplanmasını ifade etmektedir. Böylece elde edilen sayı 0 ile 1 arasına sıkıştırılmaktadır. Eşitliğin ikinci kısmı ise rastgele bir kelimenin merkez kelime ile birlikte bulunma olasılığını maksimize etmektedir. Böylece bu işlemin tüm derleme uygulanmasına gerek kalmamaktadır.

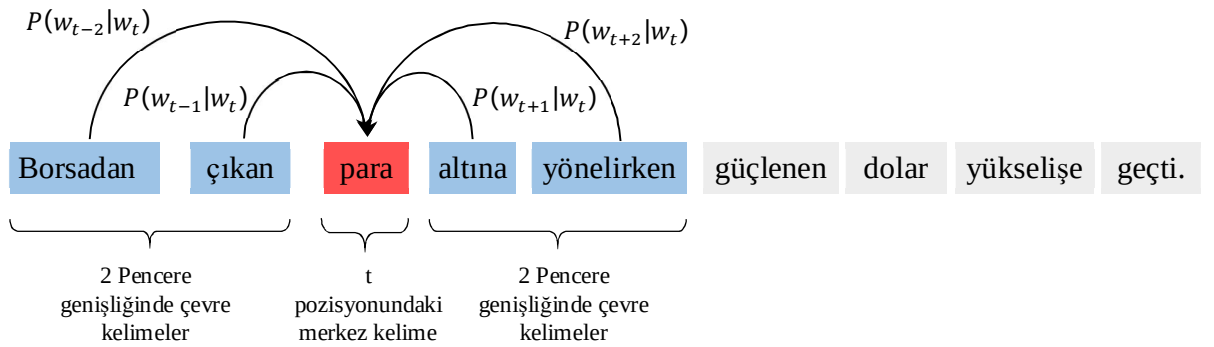
$$J_t = \log \sigma(u_o^T v_c) + \sum_{j \sim P(w)} [\log \sigma(-u_j^T v_c)] \quad 4.5$$

Rastgele alınan kelimeler unigram dağılımıyla belirlenmektedir. Sadece en sık geçen kelimeler alınmasın diye dağılımın $\frac{3}{4}$ üssü alınmaktadır. Böylece örnekleme Eşitlik 4.6'daki şekilde olmaktadır. Neticede bu modifikasyon eğitimin çok hızlı bir şekilde gerçekleşmesini sağlamaktadır (Manning, 2020a).

$$P(w) = U(w)^{\frac{3}{4}}/Z \quad 4.6$$

Word2Vec yaklaşımının bir diğer tekniği ise Continuous Bag of Words (CBOW) yöntemidir. CBOW yönteminde, Şekil 4.6'da görüldüğü gibi Skip-Gram yönteminin aksine, belirli bir pencere genişliğinde yer alan kelimelere bakarak bu kelimelerin ortasında hangi kelimelerin olabileceğine ilişkin olasılıklar hesaplanmaktadır. CBOW modelinde amaç, bir metin dizisinde kendisinden önceki ve sonraki bağlam sözcüklerine dayanarak hedef sözcüğün tahmin edilmesidir. Başka bir deyişle amaç, bağlamı verilen sözcüğe yani içeriğe göre hedef kelimeyi tahmin etmektir. Bu modelde bağlam, merkez sözcüğün etrafındaki sabit boyutlu bir pencerede bulunan sözcükler olarak temsil edilmektedir.

Şekil 4.6: CBOW Yaklaşımı



Kelime temsilleri ile ilgili yaklaşımlardan bir diğeri GloVe [Global Vectors for Word Representation] yaklaşımıdır. GloVe yaklaşımı, birlikte bulunma matrisi ile Skip-Gram ve CBOW yöntemlerinin avantajlarını bir araya getirmeyi hedeflemektedir. Pennington ve diğerlerine (2014) göre LSA gibi yöntemler istatistiksel bilgileri verimli bir şekilde

kullanırken, kelime benzetiminde nispeten daha zayıf kalmakta ve bu da yetersiz bir vektör uzay yapısına neden olmaktadır. Skip-Gram gibi yöntemler ise kelime benzetiminde daha iyi olmakla birlikte, derlemin geneline ilişkin istatistikler yerine yerel yani pencereler düzeyinde çalıştıkları için, topluluğun istatistiklerini kullanmada yetersiz kalmaktadırlar. Buradan hareketle GloVe, yerel ölçek yerine genel kelime birlikteliği istatistiklerini kullanan bir model olarak ortaya çıkmaktadır. Tablo 4.5’te bir örneği sunulan bu yöntemde ham birliktelik olasılıkları yerine, bu birliktelik olasılıklarının oranlarına odaklanılmaktadır.

Tablo 4.5: GloVe Örneği

w_k	“solid”	“gas”	“water”	“fashion”
$p_1 = P(w_k \text{"ice"})$	0,00019	0,000066	0,003	0,000017
$p_2 = P(w_k \text{"steam"})$	0,000022	0,00078	0,0022	0,000018
p_1/p_2	8,9	0,085	1,36	0,96

GloVe yönteminin amaç fonksiyonu Eşitlik 4.7’de verilmiştir (Aydoğan ve Karcı, 2019). Eşitlikteki u ve v simgeleri kelime sayılarının elde edildiği matrisin herhangi bir satır ve sütun değerlerini gösterirken $f(P_{ij})$ ağırlıklandırma fonksiyonu ifade etmektedir. Amaç her i ve j kelimeleri için u ve v vektörleri ile toplam kelime sayısının logaritmasını minimize etmektir. Kelimenin bütün derlem içerisinde kaç kez bir araya geldiğine bakarak kelime gömülmelemlerini doğrudan optimize etmektedir. Başka bir deyişle her bir kelimenin derlem içerisinde birlikte bulunma istatistiğini çıkardıktan sonra optimizasyon yapılmaktadır. Böylece GloVe belirli bir kelimenin başka kelimelerle birlikte bulunma olasılığı tahmin edilerek genel bir birlikte bulunma matrisi oluşturulmaktadır.

$$J = \frac{1}{2} \sum_{i,j=1}^W f(P_{ij}) (u_i^T v_j - \log P_{ij})^2 \quad 4.7$$

Kelime temsilleriyle ilgili bir diğerk yaklaşıım Word2Vec yönteminin bir uzantısı olan fastText'tir (Bojanowski ve diğerkleri, 2017). Doğrudan kelime vektörlerini öğrenmek yerine, her bir kelime, hem kelime olarak hem de bir n-gram karakter olarak temsil edilmektedir. Örneğinin açılı parantezler [angular brackets] kelimenin başlangıcını ve sonunu göstermek üzere “ekonomi” kelimesi $n=3$ iken, $\langle$ “ek”, “eko”, “kon”, “ono”, “nom”, “omi”, “mi” $\rangle$ şeklinde temsil edilmektedir. Bu yaklaşım kelime vektörlerinin, ön ekleri ve son ekleri temsil etmelerini sağlamaktadır. Böylece anlam sözcükaltı [subword] düzeyde de temsil edilebilir hale gelmektedir. Kelimeler n-gram karakter kullanılarak temsil edildikten sonra kelime gömülmeleri Skip-Gram yöntemiyle eğitilmektedir. Bu yöntem kelimelerin içyapısını göz ardı ettiğinden kelime üzerinde kayan pencereli bir kelime çantası modeli olarak kabul edilmektedir. Başka bir deyişle bu yöntemde önemli olan n-gramların sırası değil aynı pencerede yer almalarıdır. Word2Vec ve GloVe'dan farklı olarak sözcük dağarcığında olmayan kelimelere de temsil edilme imkânı sunmaktadır (Bojanowski ve diğerkleri, 2017). Ayrıca daha seyrek geçen kelimelerin temsillerinde de başarılı sonuçlar ortaya koymaktadır.

Bu aşamaya kadar bahsedilen yöntemler kelimelerin anlamlarının bağlama göre ortaya çıkacağını kabul ediyor, ancak bağlamı ya bir doküman çerçevesinde ya da belirli bir pencere genişliğinde birlikte bulunma olarak nitelendiriyordu. Gerçekte doğal diller çokanlamlılık [polysemy] ve kompleks anlamlılık gibi özelliklere sahip olduğundan bağlam sadece kelimelerin birlikte bulunmasına bağlı olarak değil, fakat aynı zamanda dilin yapısal özelliklerine bağlı olarak da tezahür etmektedir. Örneğinin aşağıdaki cümlelerde geçen “altında” kelimesi her iki cümlede de farklı anlamları ifade etmektedir.

- “Dolardaki yükseliş beklentilerin **altında** kaldı.”
- “**Altında** yükseliş sürüyor.”

Birinci cümlede “altında” kelimesi “alt”, “aşağıda olmak”; “altında kalmak” ifadesi ise “ezilmek”, mecazen “karşılığını verememek” anlamlarını ifade etmektedir. İkinci cümlede geçen “altında” kelimesi ise “altın” madenini işaret etmektedir. Bu örneğe biraz daha çeşitlendirmek için üçüncü bir cümle daha eklenebilir.

- “Yatırımcılar için on **altın** kural”

Bu cümlede geçen “altın” kelimesi ise, ikinci cümledeki maden anlamını değil, “üstün nitelikli”, “değerli” anlamını ifade etmektedir. Dolayısıyla makinenin bütün bu anlam farklarını göz önünde bulundurabilmesi için sadece birlikte bulunma istatistiklerini değil, fakat aynı zamanda dil bilgisi kurallarını da göz önünde bulundurması gerekmektedir. Kelimelerin bu şekilde bağlama bağlı olarak anlam kazanması, bağlama duyarlı kelime temsillerinin geliştirilmesini teşvik etmektedir.

Bağlama duyarlı temsiller arasında TagLM [language-model-augmented sequence tagger] (Peters ve diğerleri, 2017), CoVe [Context Vectors] (McCann ve diğerleri, 2017) ve ELMo [Embeddings from Language Models] (Peters ve diğerleri, 2018) sayılabilir. Bu yaklaşımların ortak özelliği dil modellerini [language model] algoritmaya dâhil etmesidir. Bu araştırma kapsamında sadece ELMo yaklaşımından bahsedilmektedir.

ELMo, her bir kelime için sabit bir temsil kullanmak yerine, temsil kelimelerini, kelimenin içerisinde geçtiği cümlelere bakarak atamaktadır. Örneğin Word2Vec, GloVe ve fastText yaklaşımlarında, “altında” kelimesi için tek bir vektör oluşturulurken, ELMo aynı kelimenin bağlama göre farklı temsillerini oluşturmaktadır. ELMo çift yönlü [bi-directional] ve iki katmanlı LSTM dil modeli olarak geliştirilmiştir. Algoritmanın çift yönlü olması, dil modelinin cümleleri hem ileriye doğru hem de geriye doğru okuyarak eğitilmesi anlamına gelmektedir. Yani, özünde iki dil modeli vardır, biri geçmiş kelimelerle bir sonraki kelimeyi tahmin etmeyi öğrenirken, diğeri gelecek kelimelerle birlikte geçmiş kelimeleri tahmin etmeyi öğrenmektedir. Algoritmanın iki katmanlı olması ise bir katmanda sözcük türü etiketleme [part of speec tagging] ve söz dizimsel görevleri icra ederken diğerkatmanda sözcük anlamını açıklığa kavuşturma gibi soyut görevlere odaklanmasını sağlamaktadır. Bu iki katman birleştirildiğinde neticede iyi bir performans elde edilmektedir (Peters ve diğerleri, 2018).

Eşitlik 4.8’de γ^{task} skaler parametresi görev modelinin tüm ELMo vektörünü ölçeklendirmesini sağlamaktadır. s_j^{task} softmax ile normalize edilmiş ağırlıkları temsil etmektedir. k ve j indeksleri, sözcüğün dizinine ve gizli durumun çıkarıldığı katmanın dizinine karşılık gelmektedir. Örneğin $h_{k,j}$ k kelimesi için LSTM’in j ’inci çıktısıdır. s_j k ’nın temsili için hesaplanan $h_{k,j}$ ’nin ağırlığıdır.

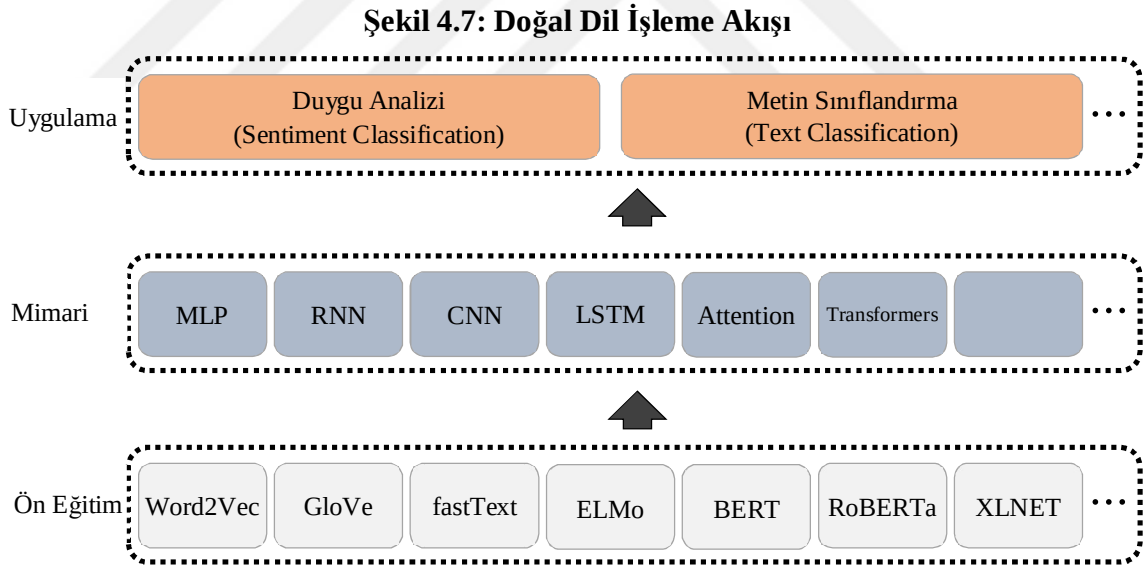
$$ELMo_k^{task} = \gamma^{task} \sum_j^L s_j^{task} h_{k,j}^{LM} \quad 4.8$$

Özetle ELMo, çok katmanlı, çift yönlü, LSTM tabanlı bir dil modelidir. Kelimelerin giriş dizisine bağlı olarak her katmanın gizli durumunu çıkarmaktadır. Daha sonra her bir kelimenin temsili elde etmek için gizli durumların ağırlıklarını hesaplamaktadır. Her durumun ağırlığı modelin icra ettiği görevin mimarisine bağlıdır ve bu ağırlıklar nihai görevin eğitimi esnasında öğrenilmektedir. Ancak ELMo’nun doğal dil işleme problemlerine yönelik çözümleri, duygu analizi, semantik etiketleme, varlık ismi tanıma [named entity recognition] gibi görevlere özgü mimarilere dayanmaktadır (Peters ve diğerleri, 2018).

Kelime temsillerini hem çift yönlü bağlam düzeyinde hem de daha az göreve bağımlı olarak hesaplayan yaklaşım ise BERT’tir (Devlin ve diğerleri, 2018). BERT [Bidirectional Encoder Representations from Transformers] yöntemi ELMo’da olduğu gibi metni hem sağdan sola doğru, hem de soldan sağa doğru değerlendirmekte, böylelikle kelimelerin birbirleriyle olan ilişkilerini daha isabetli bir şekilde temsil edebilmeyi hedeflemektedir. Bunun için iki yönetime başvurmaktadır. Bu yöntemlerden birincisi kelimeler arası ilişkilere odaklanan *maskelenmiş dil modeli* [Masked LM (MLM)] olarak ifade edilmektedir. MLM yönteminde bir cümledeki kelimelerin %15’inde maskeleyme tekniği kullanılmaktadır. Diğer bir deyişle girdi kelimelerinin %15’i maskelenerek, maskelenen kelimeler tahmin edilmeye çalışılmaktadır (Devlin, 2022). Çok fazla kelime maskelendiğinde (yani oran %15’ten büyük olduğunda) modelin eğitimi oldukça zorlaşmakta, daha az kelime maskelendiğinde ise bu kez model bağlamı kavrayamamaktadır. Yöntem, kelimelerin %80’ini [MASK] tokenla,

%10'unu rassal bir kelimeyle değiştirmektedir. Kalan %10'unu ise değiştirmeden sabit bırakmaktadır. Cümleler arası ilişkilere odaklanan modelin ikinci yöntemi ise *bir sonraki cümlenin tahminine* [Next Sentence Prediction (NSP)] odaklanmaktadır. NSP yönteminde birbirini takip eden cümle ikililerinde, ikinci cümlelerin %50'si değiştirilerek tahminler optimize edilmeye çalışılmaktadır (Devlin ve diğerleri, 2018).

Şekil 4.7'de doğal dil işlemenin işleyiş akışını göstermektedir. Başka bir ifadeyle MLP, CNN ve RNN gibi farklı derin öğrenme mimarileri kullanarak doğal dil işleme modelleri tasarlanmanın temel fikirlerini özetlemektedir. İlk basamakta makinelerin dillerin temsilini öğrenmesine ilişkin yaklaşımların bir kısmı yer almaktadır. İkinci basamak bu yöntemlerin geliştirilmesinde kullanılan derin öğrenme mimarilerinin genel bir gösterimini sunmaktadır. Son basamak ise doğal dil işleme yöntemlerinin odaklandığı dilsel çözümlemelerden bazılarını göstermektedir.



Zhang, A., Zachary C. L., Li, M., & Smola, A. J. (2021). Dive into Deep Learning. https://d2l.ai/chapter_natural-language-processing-applications/index.html

Özellikle derin öğrenme yöntemlerindeki gelişmelerle birlikte tümüyle kural tabanlı sistemlerin yerini alan istatistiksel temelli yaklaşımlar önemli başarılar elde etmiştir. Bu

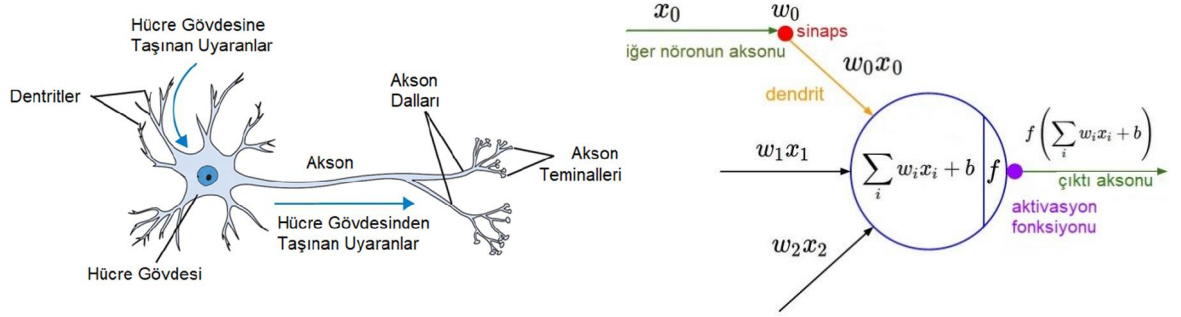
çerçevede doğal dil işleme, metnin insan seviyesinde anlaşılmasından ziyade metinsel verilerin karakter, sözcük ya da cümle düzeylerinde istatistiksel yapısına bakılarak ilişki örüntüsünün ortaya konulmasına (Chaubard ve diğerleri, 2016), bir diğer ifadeyle derin öğrenme yöntemleri ile örüntü tanımının kelimelere, cümlelere ve paragraflara uyarlanmasına dönüşmüştür (Chollet, 2019).

Sonuç olarak yöntemlerin avantaj ve dezavantajları değerlendirildiğinde, araştırma kapsamında fikir ve duygu sınıflandırması yapacak modelin BERT yaklaşımıyla eğitilmesine karar verilmiştir. Bir tweetin borsa ile ilgili olup olmadığını belirlemek için eğitilecek bağlam sınıflandırıcı için yine derin öğrenmeye dayalı farklı bir yöntem tercih edilmiştir. Bu yöntemin ayrıntılarına geçmeden önce, gerek bağlam sınıflandırmasında gerekse duygu ve fikir sınıflandırmasında kullanılan derin öğrenmenin temelleri ve modellerine kısaca değinilmektedir. Ardından finansal bağlam, duygu ve fikir sınıflandırması için kullanılan derin öğrenme modelleri ve elde edilen sonuçlar değerlendirilmektedir.

4.5.2.2 Derin Öğrenme ve Yinelemeli Sinir Ağları

Derin öğrenme, insan beyninden, yani biyolojik nöronlardan esinlenen çok katmanlı yapay sinir ağları algoritmalarının büyük ölçekli verilerle beslendiği bir makine öğrenmesi alt dalıdır. Şekil 4.8’de görüldüğü üzere, biyolojik sinir hücrelerinden esinlenerek geliştirilmiştir. Bilindiği üzere biyolojik sinir hücreleri tek akson ve çok sayıda dentrite sahiptir. Etrafındaki sinir hücrelerinden sinyal almak için dentritlerini, bu hücrelere sinyal göndermek için aksonlarını kullanmaktadırlar. Bu iletişim *sinaps* adı verilen kavşaklarda gerçekleşmektedir. Bu kavşaklarda hücreler, hücreye ulaşan elektrik sinyalinin bir kimyasal madde salgılatması (nörotransmitter) ve bu kimyasal maddenin de karşı hücrede tekrar bir elektriksel etkiye yol açması suretiyle iletişim kurarlar (Kandel, 2011). Yapay sinir ağları da benzer şekilde sinir hücrelerine dışarıdan gelen bilgileri bir toplama fonksiyonu ile toplar. Daha sonra bir aktivasyon fonksiyonundan geçirerek ürettiği çıktıyı ağıın bağlantıları üzerinden diğer hücrelere gönderir (Temel, 2012).

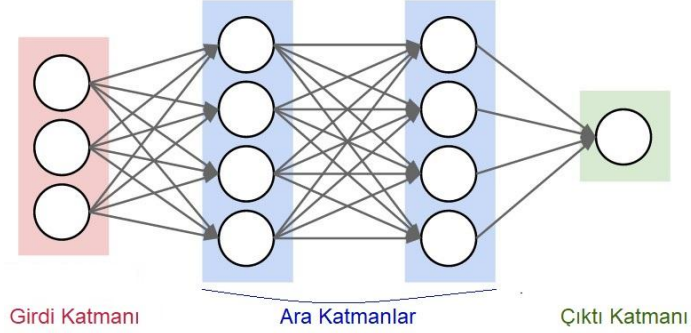
Şekil 4.8: Biyolojik Sinir Hücresi-Yapay Sinir Hücresi Karşılaştırması



Neural networks. (n.d.). CS231n convolutional neural networks for visual recognition. CS231n @Stanford. <https://cs231n.github.io/neural-networks-1>

Yapay sinir ağı, Şekil 4.9’da görüldüğü üzere, girdi katmanı, ara katmanlar [hidden layers] ve çıktı katmanı olmak üzere temelde üç katmanlı bir yapıya sahiptir. Girdi katmanındaki nöronlar ara katmanlardaki nöronlara, ara katmanlardaki nöronlar da çıktı katmanındaki nöronlara bağlıdır. Ağda bilgi akışı da girdi katmanından çıktı katmanına doğru bu silsileyi takip ederek ilerlemektedir. Başka bir ifadeyle nöronların her biri kendisinden önceki nörondan gelen çıktıyı kendi içerisinde işlemiden geçirdikten sonra bir sonraki nörona girdi olarak iletmektedir. Böylece ağa girdi katmanından iletilen bilgiler ara katmanlarda işlenerek çıktı katmanına gönderilmektedir. Ara katmanlar zincirinin tamamının uzunluğu modelin derinliğini vermektedir. Derin öğrenme adı da esasen sinir ağlarının öğrenmeyi sağlayan çok sayıda ara katmana sahip olmasından kaynaklanmaktadır (Goodfellow ve diğerleri, 2018). Nihayetinde hem girdiler hem de o girdilere karşılık ağın üretmesi beklenen çıktılar ağa gösterilmiş olur. Bu çerçevede ağdan beklenen kritik husus, belirli girdiler için belirli çıktıları üretebilmesi, yani öğrenebilmesidir. Öğrenme sürecinde bir dizi veri tekrar tekrar ileriye doğru beslenir [feedforward] ve nöronların ağırlıkları, çıktı gerçek değerlere yaklaşıncaya kadar değiştirilir. Aslında insanın deneyimlerle beceri kazanmasına benzer şekilde, derin öğrenme algoritmaları da bir görevi tekrar tekrar ve her seferinde sonucu biraz daha iyileştirerek gerçekleştirmektedir. Ağdaki ağırlıklar ($w_1, w_2 \dots w_i$), hücreye gelen girdinin hücre üzerindeki etkisini temsil etmektedir. Toplamda bir sabit terim de [bias] yer almaktadır. Dolayısıyla yapay sinir ağlarında öğrenme, girdiyi çıktıya ulaştıran ağın tüm katmanlarının ağırlıklarının bulunması işlemidir (Chollet, 2019).

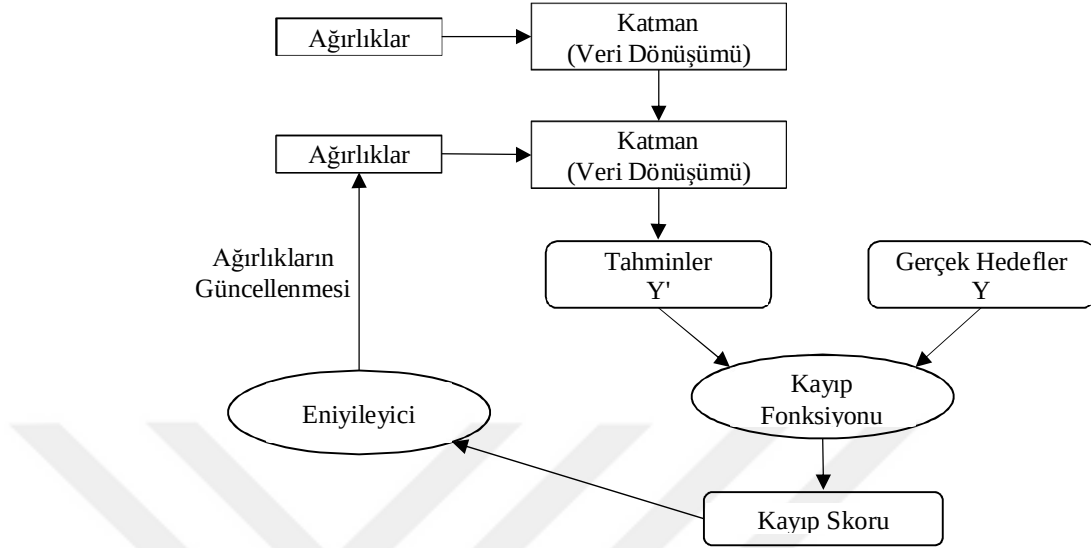
Şekil 4.9: Yapay Sinir Ağı



Neural networks. (n.d.). CS231n convolutional neural networks for visual recognition. CS231n @Stanford. <https://cs231n.github.io/neural-networks-1>

Ağırlık ve bias değerleri eğitimin başlangıcında rastgele atanmakta ve her bir örnek ağı gösterildikten sonra ağı öğrenme kuralları çerçevesinde güncellenmektedir. Daha öncede ifade edildiği gibi çok katmanlı ileri beslemeli sinir ağlarında, tahmin edilen ve gözlenen çıktılar arasındaki farkın mümkün olduğunca küçük olabilmesi için bağlantı ağırlıklarının tekrar tekrar ayarlanması gerekmektedir. Bu fark modelde, Şekil 4.10'da görüldüğü üzere, kayıp fonksiyonu [loss function] ile ölçülmektedir. Diğer bir deyişle kayıp fonksiyonu ağı çıktısı ile beklenen çıktı arasındaki farkı belirleyerek bir girdi için hesaplama başarımının ne kadar iyi olduğunu ortaya koymaktadır (Chollet, 2019).

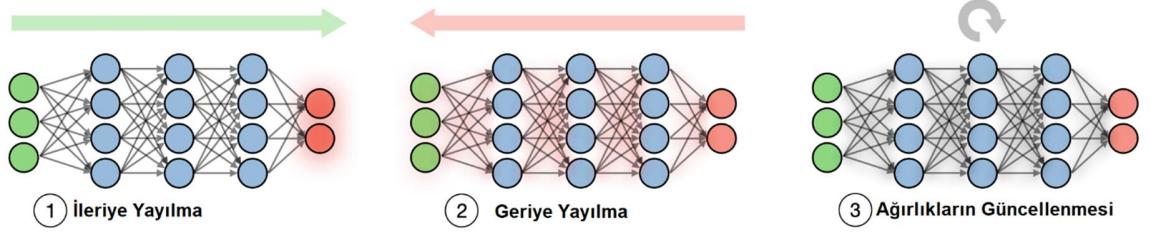
Şekil 4.10: Kayıp Fonksiyonu ile Ağın Çıktısı Arasındaki İlişki



Kaynak: Chollet, F., *Python ile derin öğrenme* (1. baskı). Buzdağı Yayınevi, 2019.

Derin öğrenmede kritik husus, kayıp fonksiyonuyla ölçülen değerin, Şekil 4.11 'de görüldüğü üzere, geri bildirim sinyali şeklinde kullanılarak kaybı azaltacak biçimde ağırlık matrisini güncellemektir. Diğer bir ifadeyle üretilen çıktı ile beklenen çıktı arasındaki farkın geriye doğru dağıtılmasıdır. Geri yayılım [backpropagation] adı verilen bu güncelleme işlemini yapmak için eniyileme [optimization] algoritmalarına başvurulmaktadır. Veri kümesindeki her bir verinin bir kez eğitimden geçmesine, ya da modelin tüm veri kümesini bir kez görmesine epok [epoch] denir. Bu süreç eğitimde kullanılan örneklerin tamamı için başarılı sonuçlar üretilinceye kadar tekrarlanmaktadır (Maas ve Tandon, 2013; Mishra ve Datta-Gupta, 2018; Temel, 2012).

Şekil 4.11: Yapay Sinir Ağı Öğrenme Süreci

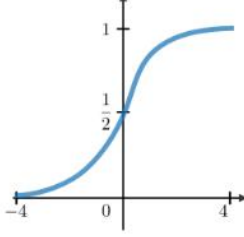
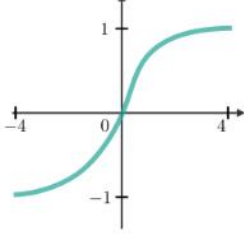
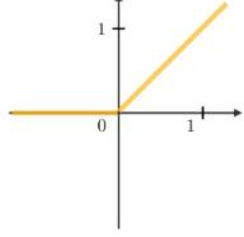
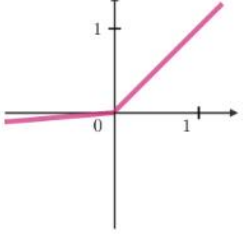


Kaynak: Amidi, A., & Amidi, S. (n.d.-a). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-deep-learning-tips-and-tricks>

Girdi ve çıktı katmanlarındaki nöron sayısı çözülmek istenen problemin özelliklerine göre belirlenmektedir. Ara katmanların sayısını ve ara katmanlardaki nöron sayısını belirlemek için ise ideal bir yöntem bulunmamaktadır. Daha çok tecrübeyle, deneme yanılma yoluyla belirlenmektedir. Dolayısıyla yapay sinir ağlarında kontrol edilmesi gereken en önemli parametreler, gizli katmanların sayısı, gizli katmanlardaki nöronların sayısı, öğrenme hızı, sönümlleme katsayısı veya momentumu ve daha iyi optimizasyon için yineleme sayısı olarak ifade edilebilir.

Toplama fonksiyonu bir nörona gelen net girdiyi, aktivasyon fonksiyonu hücreye gelen net girdiyi işleyerek bu girdiye karşılık üreteceği çıktıyı belirlemektedir. Aktivasyon fonksiyonu doğrusallığı ortadan kaldırarak daha zengin bir hipotez uzayı elde etmeyi sağlamaktadır. Bunun için, Şekil 4.12’de görüldüğü üzere, farklı aktivasyon fonksiyonları kullanılabilir.

Şekil 4.12: Aktivasyon Fonksiyonları

Sigmoid	Tanh	ReLU	Leaky ReLU
$g(z) = \frac{1}{1 + e^{-z}}$	$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	$g(z) = \max(0, z)$	$g(z) = \max(\epsilon z, z)$ with $\epsilon \ll 1$
			

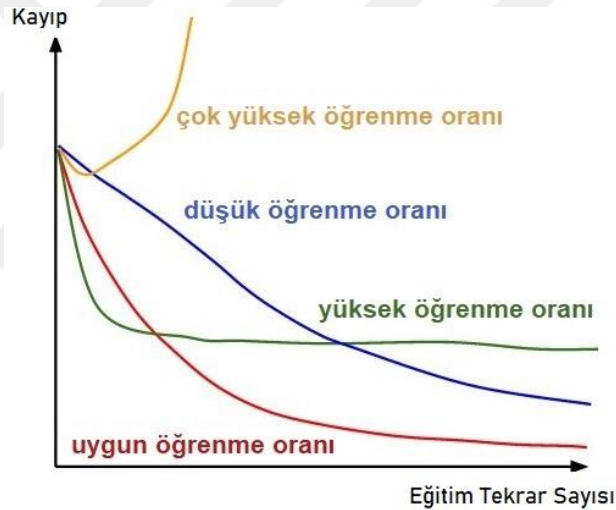
Kaynak: Amidi, A., & Amidi, S. (n.d.-a). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-deep-learning-tips-and-tricks>

Bu fonksiyonlardan sigmoid, aldığı değer ile $[0,1]$ arasında bir değer üretmektedir. Yüksek bir değer geldiğinde 1'e, düşük bir değer geldiğinde 0'a yaklaşmaktadır. Çok küçük negatif ya da çok büyük pozitif bir sayı olduğunda doygunluğa ulaşmaktadır. Başka bir deyişle fonksiyon düzleşerek girdisindeki küçük değişikliklere duyarsız hale gelmektedir. Hiperbolik tanjant fonksiyonu (Tanh) gelen değerleri $[-1,1]$ arasına sıkıştırmaktadır. ReLU [rectified linear unit] girdilerin negatif değerleri için 0 çıktısı verirken, girdi 0'dan büyükse bu değere bir sıkıştırma uygulamamaktadır. Yani $[0, +\infty)$ aralığında değerler almaktadır. Sigmoid ve Tanh fonksiyonlarına kıyasla çok daha hızlı hesaplama yapmaktadır. Sızıntılı ReLU [Leaky ReLU] ise 0,01 gibi bir sızıntı değeri vererek öğrenmeyi negatif bölgedeki değerler için de sağlamış olmaktadır (Goodfellow ve diğerleri, 2018).

Derin öğrenme yönteminde kayıp [loss | cost] fonksiyonu, eğitim sırasındaki başarıyı ölçmektedir. Hataların hesaplanmasını bir eniyileme [optimization] problemine dönüştürerek modelin ürettiği tahminin gerçek değerden ne kadar farklı olduğunu hesaplamaktadır. Modelin eğitimi esnasında kayıp fonksiyonunun 0'a yaklaşması beklenmektedir. Eniyileme algoritması ise kayıp fonksiyonuna göre ağı nasıl güncelleneceğini belirlemektedir (Chollet, 2019). Eniyileme algoritmasıyla kayıp değerini azaltmak için eniyileme yapılmaktadır.

Eniyileme esnasında minimuma ulaşmak için atılacak adımların büyüklüğüne öğrenme oranı [learning rate] denilmektedir. Diğer bir deyişle öğrenme oranı, ağırlıkların hangi hızda güncellendiğini göstermektedir. Sabitlenebilir veya uyarlanabilir bir orandır. Bu oran doğrudan modelin performansı üzerinde etkili olan hiperparametrelerden biridir (Goodfellow ve diğerleri, 2018). Grafik 4.1’de görüldüğü üzere öğrenme oranının küçük belirlenmesi durumunda model yavaş öğrenirken, büyük belirlenmesi durumunda minimuma ulaşamamaktadır. İdeal bir öğrenme oranı ise kayıp fonksiyonuna bakarak belirlenebilmektedir. Mevcut durumda öğrenme hızını adapte edebilen en popüler yöntem Adam’dır. Buna ilave olarak AdaGrad, RMSProp ve benzeri algoritmalar da vardır.

Grafik 4.1: Öğrenme Oranının Belirlenmesi



Neural networks. (n.d.). CS231n convolutional neural networks for visual recognition. CS231n @Stanford. <https://cs231n.github.io/neural-networks-1>

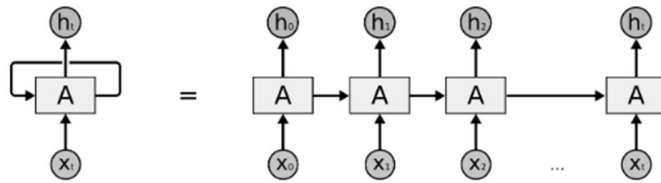
Bu aşamaya kadar anlatılan geleneksel sinir ağları, tüm girdilerin ve çıktıların birbirinden bağımsız olduğunu varsaymaktadır. Ancak bu varsayım sıralı verileri işlemek için sinir ağlarının geleneksel yaklaşımını işlevsiz hale getirmektedir. Örneğin bir metindeki kelimelerin sırası gözetilmeden ağa verilmesi durumunda, sözdiziminin anlamı belirlemedeki rolü göz ardı edilmekte ve büyük oranda anlam kaybı yaşanmaktadır. Dolayısıyla bu ve benzeri görevler için özelleştirilmiş sinir ağlarına ihtiyaç duyulmaktadır. Bu amaçla geliştirilen yinelemeli sinir ağları [recurrent neural networks (RNN)] diziler ya da

zamana bağılı sıralı verileri ardışıklık bilgisini koruyarak işleme imkânı sunmaktadır. Ardışıklık bilgisinin korunması hesaplamaların saklanacağı bir belleği gerekli kılmaktadır. Yinelemeli sinir ağlarını diğer ağlardan farklı kılan en önemli husus hafızalarının olmasıdır. Hafızaları sayesinde sıralı verilerin önceki gözlemlere bağımlılığı göz önünde bulundurulabilmektedir. Dolayısıyla yinelenen sinir ağları, zaman serileri veya doğal dil gibi dizi verilerinin modellenmesi için oldukça güçlü bir sinir ağı mimarisi sunmaktadır.

Sıralı öğrenme insanın biyolojik zekâsının işleyişine benzetilebilir. Zira insan zekâsı bilgiyi, daha önce öğrenilen içsel bir modeli muhafaza ederken aynı zamanda gelen yeni bilgilerle güncelleyerek artırımı olarak işlemektedir. Örneğin bir metni okurken zihin kelimelerin anlamlarını önceki kelimelerin anlamlarına göre inşa eder. Yinelemeli sinir ağları da düşüncelerdeki bu süreklilik prensibini modellemektedir. Sıralı verilerin her bir elemanı üzerinde ilerlerken o ana kadar öğrendiği bilgiyi durum bilgisi olarak muhafaza etmektedirler. Sonuç olarak kendi iç döngüsü bulunan RNN’de tüm veri bir hamlede değil iç döngüde eleman eleman işlenmektedir (Chollet, 2019).

Örneğin Şekil 4.13’te görülen x_t girdisini A sinir ağı istifi [chunk] işleyerek h_t değerini üretmektedir. Yinelenen sinir ağı istifinin açılımındaki döngü, bilginin saklanarak ağın bir adımından diğerine geçirilmesine izin vermektedir. Diğer bir deyişle yinelenen sinir ağı, aynı ağın her biri bir ardılına bilgiyi ileten birden fazla kopyası olarak da düşünülebilir (Olah, 2015). Bunun için yineleme formülü kullanılmaktadır (Li, 2020).

Şekil 4.13: Yinelemeli Ağ Mimarisi



Kaynak: Olah, C. (2015, August 27). Understanding LSTM Networks. Colah's Blog. <https://colah.github.io/posts/2015-08-Understanding-LSTMs>

Eşitlik 4.9’da x_t belirli bir zamndaki girdi vektörünü h_{t-1} önceki durumu f_w w parametrelili bir fonksiyonu h_t ise yeni durumu göstermektedir. Şematik olarak bir RNN katmanı, o ana kadar karşılaştığı zamansal adımlar hakkında kodladığı bilgilerden oluşan içsel durumu [internal state] korurken ardışıklığın zamansal adımları üzerinde bir *for döngüsü* [for loop] döndürmektedir.

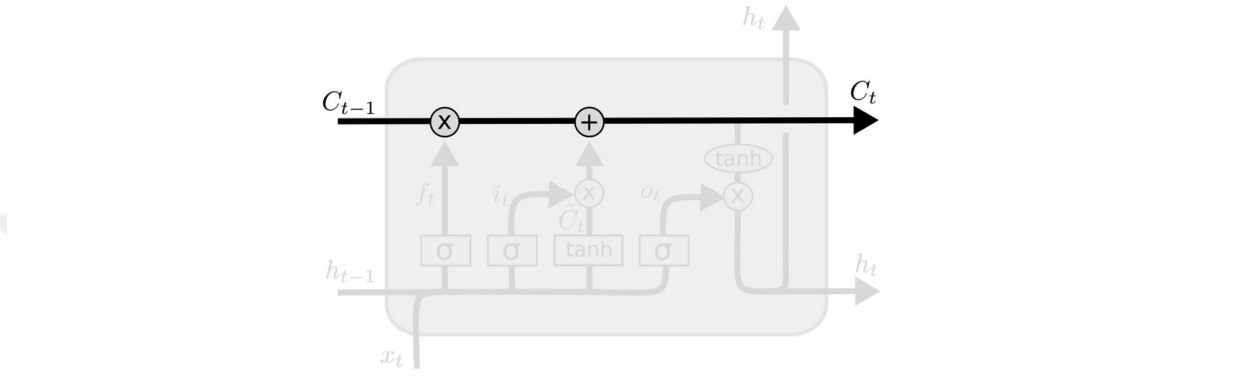
$$h_t = f_w(h_{t-1}, x_t) \quad 4.9$$

Ancak basit yinelemeli sinir ağları da [simpleRNN] çok temel bir problemle karşı karşıya kalmaktadır. Patlayan gradient [exploding gradient] ve kaybolan gradient [vanishing gradient] olarak ifade edilen bu problemler ileri beslemeli bir ağda ağırlıkların güncellenmesi için uygulanan çarpma işleminden kaynaklanmaktadır. Zira hata geriye dağıtılırken [backpropagation] w matrisi ile çarpma işlemi yapılmakta, her bir çarpma işleminde aynı w matrisi kullanıldığı için, w matrisi sürekli kendisiyle çarpılmış olmaktadır. Bu durum büyük bir ağda başlangıç katmanlarına doğru ilerledikçe eğer bu sayı 1’den küçükse *kaybolan gradyan problemine*, eğer bu sayı 1’den büyükse *patlayan gradyan problemine* yol açmaktadır (Li, 2020). Basit RNN’lerdeki bu sorun nedeniyle bilgi, ya bir yerde kaybolmakta ya da aşırı büyümekte ve sonuçta ağı eğitmek zorlaşmaktadır. Yinelenen sinir ağlarında bu sorunların üstesinden gelerek oldukça başarılı bir performans gösteren ve alanyazınında LSTM ve GRU olarak bilinen iki farklı ağ mimarisi geliştirilmiştir.

LSTM yaklaşımının ana fikri, Şekil 4.14’te görüldüğü üzere, zaman içerisinde bilgi taşıyan hücre durumu [cell state] ile hücre modülünde tekrar eden dört etkileşimli katmana dayanmaktadır (Hochreiter ve Schmidhuber, 1997). Dizi zincirinde görelili bilgiyi taşıyan hücre durumu c_t (taşıma anlamında [carry]) girdi ile yinelemeli bağlantıları bileştirerek aktivasyon fonksiyonları ya da matris çarpımıyla sonraki adıma gönderilecek durumu etkilemektedir. Böylece, önceki zaman adımlarından gelen bilgiler kısa vadeli belleğin etkilerini azaltarak daha sonraki zaman adımlarına taşınabilir hale gelmektedir. Hücre durumu yolculuğa devam ederken, kapılar yoluyla hücre durumuna bilgi eklenilmekte veya kaldırılmaktadır. Kapılar, eğitim sırasında hangi bilgilerin saklanması veya unutulması

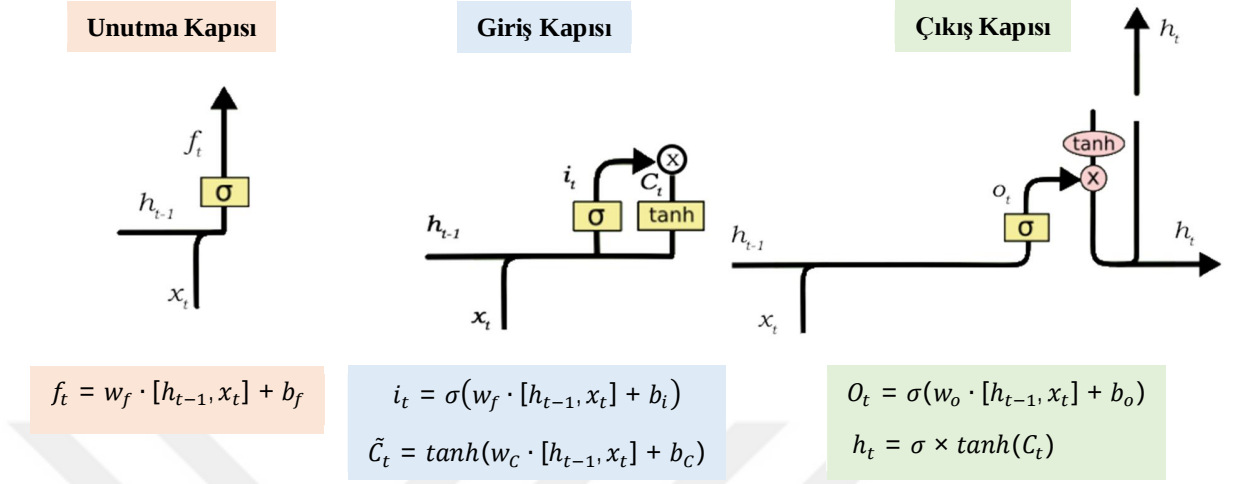
gerektiğini öğrenebilen ve hücre durumunda hangi bilgilere izin verildiğine karar veren farklı sinir ağı katmanlarıdır. LSTM'in en önemli özelliklerinden biri de hücre durumuna bilgi ekleme ya da çıkarma yeteneğidir (Olah, 2015).

Şekil 4.14: LSTM Durum Hücresi



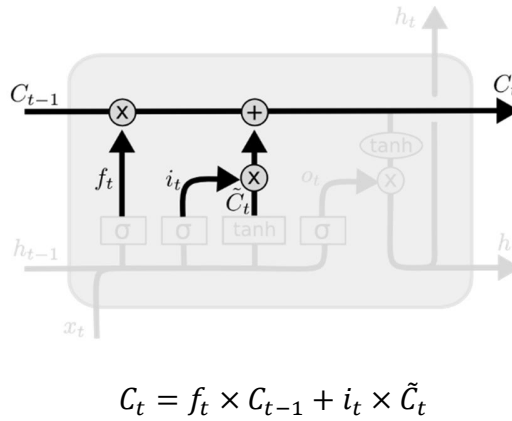
Şekil 4.15'te görüldüğü üzere, bir LSTM hücresindeki bilgi akışını düzenleyen üç farklı kapı vardır. Bunlar unutma kapısı, giriş kapısı ve çıkış kapısıdır. Unutma kapısında LSTM hangi bilgilerin unutulması ve hangi bilgilerin saklanması gerektiğine karar vermektedir. Önceki hücreden gelen gizli durum [hidden state] bilgisi ile kendisine gelen girdi birleştirilerek sigmoid fonksiyonundan geçirilmektedir. Böylece 1'e yakın bilgiler hafızada tutulurken 0'a yakın bilgiler unutulmaktadır. Giriş kapısında, önceki gizli katmandan gelen bilgi ile girdi sigmoid fonksiyonundan geçirilerek hangi değerlerin güncelleneceğine karar verilmektedir. Tanh hücresinde ise durum hücresine eklenmeye aday olan yeni değerlerden bir vektör oluşturulmakta, daha sonra bu hücreler birleştirilmektedir.

Şekil 4.15: LSTM Sinir Hücresini Düzenleyen Kapılar



Durum hücresinde ise öncelikle unutma kapısında elde edilen değer ile önceki hücre durumu çarpılmaktadır. Daha sonra giriş kapısından elde edilen veriler durum hücresinin yeni durumuna eklenir (Şekil 4.16'ye bakınız). Böylece durum hücresinin [cell state] yeni hali elde edilmiş olur.

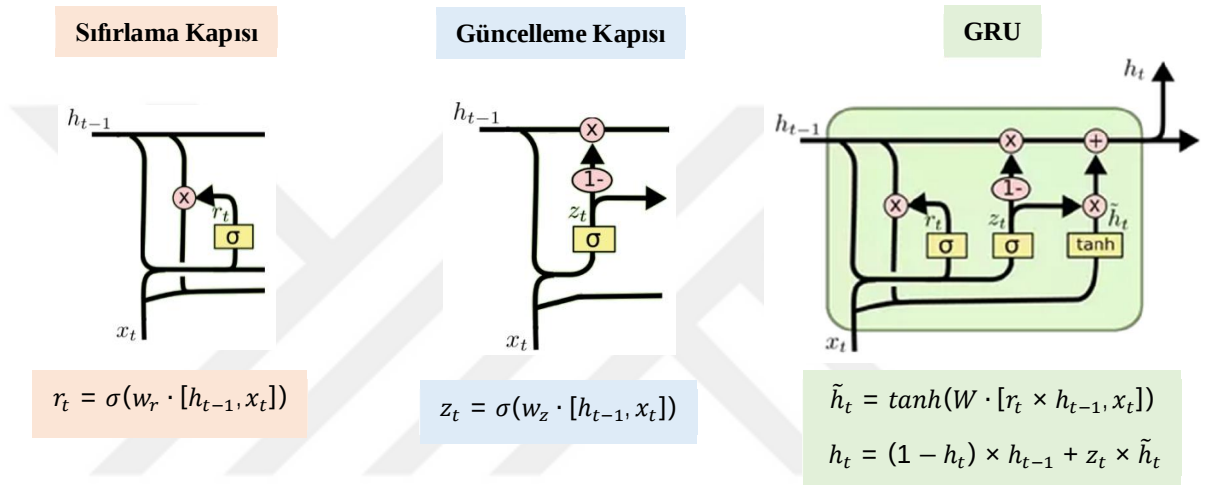
Şekil 4.16: LSTM Durum Hücresin Yeni Durumu



Son olarak LSTM'de bir de çıkış kapısı bulunmaktadır. Çıkış kapısında çıktının ne olması gerektiğine karar verilmektedir. Bu, aynı zamanda bir sonraki gizli durumun [hidden state] ne olacağına karar verilmesi anlamına da gelmektedir. İlk olarak, önceki gizli durum ve

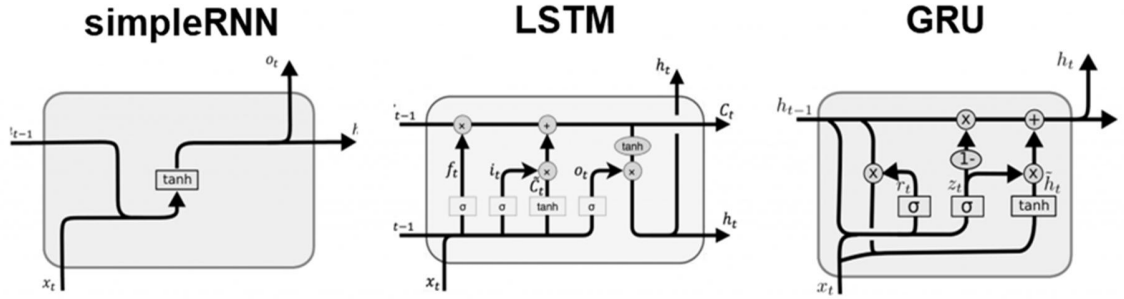
mevcut girdi sigmoid fonksiyonundan geçirilmektedir. Sonra yeni güncellenmiş durum hücresinin değeri tanh fonksiyonundan geçirilmektedir. Yeni gizli durumun hangi bilgileri taşıması gerektiğine karar vermek için bu kez tanh fonksiyonunun çıktısı ile sigmoid fonksiyonunun çıktısı çarpılmaktadır. Elde edilen yeni gizli durum ve yeni durum hücresi bir sonraki zaman adımına taşınmaktadır.

Şekil 4.17: GRU Sinir Hücresini Düzenleyen Kapılar



Geleneksel RNN'lerin karşılaştığı kaybolan ve patlayan gradyan problemini ele alan bir diğer ağ mimarisi ise GRU [Gated Recurrent Units] yaklaşımıdır (Cho, 2014). Şekil 4.17'de görüldüğü üzere, GRU'ların LSTM'den farklı olarak sadece iki kapısı bulunmaktadır. Bunlar güncelleme kapısı ve sıfırlama kapısıdır. Güncelleme kapısı [update gate] LSTM'nin unutma ve giriş kapısına benzer şekilde hangi bilgilerin atılacağına ve hangi bilgilerin ekleneceğine karar vermektedir. Sıfırlama kapısı [reset gate] ise, geçmiş bilgisinin ne kadarının unutulacağına karar vermek için kullanılan başka bir kapıdır. GRU'da LSTM'den farklı olarak hücre durumu bulunmamaktadır (Şekil 4.18'e bakınız). Bunun yerine sadece gizli durum bulunur. GRU'ların daha az tensör işlemi olduğundan, LSTM'leri eğitmek için harcanan zamana kıyasla biraz daha hızlıdır.

Şekil 4.18: RNN Mimarilerinin Karşılaştırması

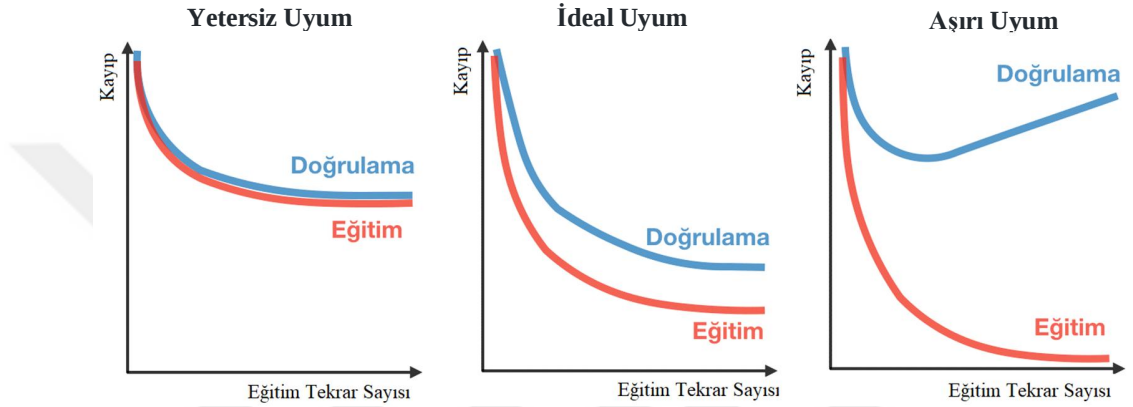


Derin öğrenmeyle ilgili bir diğer husus ise modelin performansının nasıl değerlendirileceğidir. Bunun için eğitime başlamadan önce veri kümesinden belirli bir olasılık dağılımı ile eğitim ve test verileri seçilmektedir. Bu kümelerin özdeşçe dağılmış ve aynı olasılık dağılımı içerisinde seçilmiş ve birbirinden bağımsız örnekler olduğu varsayılmaktadır. Bu varsayımlar dolayısıyla modelin beklenen eğitim hatasının aynı modelin beklenen test hatasına eşit olduğu kabul edildiğinden eğitim ve test hataları kıyaslanabilir hale gelmektedir. Makine öğrenmesinde önce eğitim kümesinin hatasını en aza indireyecek şekilde parametreler seçilmektedir (Goodfellow ve diğerleri, 2018). Eğitim setinde belirli bir başarıya ulaşıldıktan sonra test setindeki örnekler üzerinde modelin performansı test edilmektedir. Şayet test verileri üzerinde eğitim verilerine benzer iyi bir performans yakalanabilirse model eğitilmiş olmaktadır. Bir diğer ifadeyle ağıın bu şekilde *genelleştirme* yapabilme yeteneği kazanmasına ağıın öğrenmesi denilmektedir (Temel, 2012).

Ancak model eğitilirken genellikle beklenen test hatası, beklenen eğitim hatasından daha büyük veya eşit olmaktadır. Bir makine öğrenmesi algoritmasının ne kadar iyi çalışacağını belirleyen faktörler modelin eğitim hatasını küçültme ve eğitim ve test hataları arasındaki farkı azaltma yeteneklerine göre ölçülmektedir. Bu iki faktör makine öğrenmesinin iki temel zorluğuna işaret etmektedir: Yetersiz uyum [underfitting] ve aşırı uyum [overfitting] (Géron, 2021). Şekil 4.19’da görüldüğü üzere, yetersiz uyum model eğitim kümesi üzerinde yeterince düşük hata oranını yakalayamadığında olmakta; aşırı uyum ise eğitim ve test hataları

arasındaki farkın büyüdüğü durumlarda oluşmaktadır. Aşırı uyum ve yetersiz uyum problemlerinin üstesinden gelmek için modelin kapasitesini değiştirmek, ağırlıkları düzenlileştirmek [weight regularization], iletim sönümü [dropout] uygulamak gibi yöntemlere başvurulabilmektedir (Goodfellow ve diğerleri, 2018).

Şekil 4.19: Öğrenme Başarımı



Kaynak: Amidi, A., & Amidi, S. (n.d.-b). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/l/tr/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks>

Derin öğrenmeyle ilgili son olarak eğitimin performansını değerlendirmede kullanılan metriklerden de kısaca bahsetmek gerekir. Bu metriklerin başında, Tablo 4.6'ta yer alan ayırt etme matrisi [confusion matrix] gelmektedir. Gerçek değerler ile tahmin değerlerinin bir karşılaştırmasını yapan hata matrisi aslında sınıflandırma problemine ilişkin tahmin sonuçlarının bir özetini de sunmaktadır. Diğer bir ifadeyle, doğru ya da yanlış tahminlerin sınıf bazında tasnif edilmiş sayısını özetlemektedir. Sadece sınıflandırma modelinin yaptığı hatalar hakkında değil, (1. Tip Hata veya 2. Tip Hata gibi) yapılan hata türleri hakkında da bilgi vermektedir.

Tablo 4.6: Ayırt Etme Matrisi

		TAHMİNİ SINIF	
		+	-
GERÇEK SINIF	-	FP [False Positives] Yanlış Pozitif 1. Tip Hata	TN [True Negatives] Doğru Negatif
	+	TP [True Positives] Doğru Pozitif	FN [False Negatives] Yanlış Negatif 2. Tip Hata

Sınıflandırma modelinin performansını değerlendirmek için ayırt etme matrisinden türetilen Tablo 4.7’teki metrikler yaygın olarak kullanılmaktadır.

Tablo 4.7: Sınıflandırma Performansı Değerlendirme Metrikleri

Metrik	Formül	Açıklama
Doğruluk	$\frac{TP + TN}{TP + TN + FP + FN}$	Modelin genel performansını göstermektedir. Sınıflandırma problemlerinde genellikle modelin performansı, doğruluk [accuracy] metriği ile ölçülmektedir.
Kesinlik	$\frac{TP}{TP + FP}$	Pozitif tahminlerin ne kadar kesin olduğunu göstermektedir. Gerçek pozitiflerin tahmin edilen pozitiflere oranıdır. Diğer bir deyişle pozitif olarak tahmin edilenlerin yüzde kaçının gerçekten pozitif olduğunu göstermektedir.
Duyarlılık Doğru Pozitif Oran	$\frac{TP}{TP + FN}$	Doğru sınıflandırılan pozitif örneklerin toplam pozitiflere oranıdır. Başka bir ifadeyle gerçek pozitifler içinde geliştirilen testin pozitifleri ayırt edebilme yeteneğini göstermektedir. Gerçek pozitif oran, bir ROC eğrisindeki y eksenine karşılık gelmektedir.

Seçicilik	$\frac{TN}{TN + FP}$	Doğru tahmin edilen negatif örneklerin toplam negatiflere oranıdır. Başka bir ifadeyle gerçek negatifler içinde geliştirilen testin negatifleri ayırt edebilme yeteneğini göstermektedir.
Yanlış Pozitif Oran	$\frac{FP}{TN + FP}$	Gerçek negatiflerin yüzde kaçının yanlışlıkla pozitif olarak sınıflandırıldığını göstermektedir.
Yanlış Negatif Oran	$\frac{FN}{TP + FN}$	Gerçek pozitiflerin yüzde kaçının yanlışlıkla negatif olarak sınıflandırıldığını göstermektedir.
F1 Skoru	$\frac{2TP}{2TP + FP + FN}$	Dengesiz sınıflar için yararlı hibrit bir metrik sunmaktadır. Kesinlik ve duyarlılık ölçütlerinin harmonik ortalamasıdır. Her iki ölçütü birlikte değerlendirme imkânı vermektedir. Sınıflandırıcının eşik değeri belirlenirken F1 değeri göz önünde bulundurulmaktadır.
MCC	$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) + (TP + FN) + (TN + FP) + (TN + FN)}}$	Dengesiz dağılan veri setlerinde gerçek sınıflar ile tahmin edilen sınıflar arasındaki korelasyon ilişkisini (phi-coefficient) hesaplamaktadır. MCC değeri 1'e yaklaştıkça tahmin ilişkisinin güçlü olduğu sonucuna varılmaktadır.

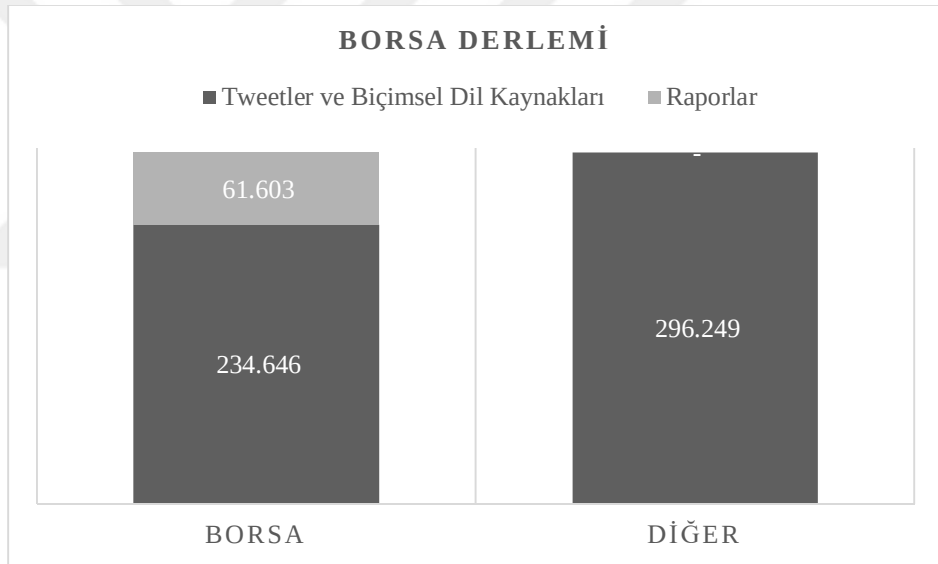
4.5.2.3 Finansal Bağlam Sınıflandırıcı

Doğal dil işleme açısından bağlam, bir dil birimini çevreleyen, ondan önce veya sonra gelen, birçok durumda söz konusu birimi etkileyen, onun anlamını, değerini belirleyen birim veya birimler bütünü olarak ifade edilmektedir (TDK, 2019). Dolayısıyla bağlamın doğru bir şekilde belirlenebilmesi, anlamın başarılı bir şekilde belirlenebilmesinin en önemli adımlarından biridir. Bir diğer ifadeyle anlam, bağlamda tezahür ettiği için, öncelikle bağlamın belirlenebilmesi gerekmektedir. Araştırma kapsamında geliştirilen *finansal bağlam sınıflandırıcı* [financial context classifier], anlamın doğru bir şekilde tespit edilebilmesi için,

borsayla ilgili olan metinlerin belirlenerek borsayla ilgili olmayanlardan ayrıştırılmasını ve tweetlerin bu ölçüte göre sınıflandırılmasını hedeflemektedir.

Metinsel verilerin borsa ile ilgisini belirleyebilmek için, etiketli 592.498 belgeden oluşan iki sınıflı (“borsa” ve “diğer” sınıflarını temsil eden) bir derlem oluşturulmuştur. Grafik 4.2’de görüldüğü üzere *borsa* etiketli 296.249 belgenin 61.603 tanesi formel dille yazılmış metinler, 234.646 tanesi informal dille yazılmış tweetler ve diğer biçimsel dil kaynaklarından oluşmaktadır⁵⁴. *Diğer* etiketli 296.249 belgenin tümü, politika, teknoloji, spor, magazin gibi konularda yazılmış, yani borsa ile ilgili olmayan tweetlerden oluşmaktadır.

Grafik 4.2: Borsa Derlemi



⁵⁴ Bu çalışma borsaya odaklandığı için çeşitli biçimsel dil kaynaklarından elde edilen 739.614 belgeden genel ekonomi, altın, döviz vb. konu etiketli olanlar ayrıştırılmış, sınıflandırıcının eğitiminde sadece borsa ile ilgili olanlar kullanılmıştır. Bu da finansal derlemin yaklaşık %20’sine tekabül etmektedir.

Borsa derleminin %80'i eğitim için; %20'si test için ayrıldıktan sonra, bağlam sınıflandırıcı, yöntemler arasında karşılaştırma yapabilmek amacıyla RNN-GRU ve RNN-LSTM yöntemleriyle eğitilmiştir. RNN-GRU ile geliştirilen model özeti Şekil 4.8'da sunulmaktadır.

Tablo 4.8: RNN-GRU Bağlam Sınıflandırıcı Model Özeti

Model: "sequential"		
Layer (type)	Output Shape	Param #
embedding_layer (Embedding)	(None, 57, 50)	1000000
cu_dnngru (CuDNNGRU)	(None, 57, 16)	3264
cu_dnngru_1 (CuDNNGRU)	(None, 57, 8)	624
cu_dnngru_2 (CuDNNGRU)	(None, 4)	168
dense (Dense)	(None, 1)	5
Total params:1.004.061		
Trainable params:1.004.061		
Non-trainable params:0		

Modelde her kelimeye karşılık gelen kelime vektörünün uzunluğu [embedding size] 50 olarak belirlenmiştir. Gömülme matrisi [embedding matrix] oluşturulurken kelime vektörleri herhangi bir öneğitimli kelime gömülmesine başvurmaksızın oluşturulmuştur. Bu yöntemde başlangıçta rassal olarak oluşturulan kelime vektörleri model eğitilirken eğitilmektedir. Ancak istenildiğinde word2vec ya da GloVe ve benzeri bir yöntemle eğitilmiş hazır vektörler gömülme katmanına [embedding layer] eklenebilmektedir. Matrisin boyutu kelime sayısı ve gömülme büyüklüğü kadar, yani 20.000x50 boyutlarında oluşmaktadır. Metinler dizgeciklere [token] ayrıldıktan sonra sözde dolgu [padding] eklenerek girdi [input] boyutu 65 olarak belirlenmiştir. Sözde dolgu [padding], girdi boyutunu, yani metinlerin uzunluğunu standardize etmek için boşlukları tamamlama işlemidir. Böylece 57 dizgeciğe karşılık gelen 50 uzunluğunda vektörler oluşturmuştur. Her dizgecik vektörü bir sonraki katman için girdi olurken ilk GRU katmanı 57'e 16 boyutunda çıktı vermektedir. Model özetindeki 16 sayısı katmandaki nöron sayısını göstermektedir. Üçüncü katmandan sonra sonuçlar tek nöronlu

dense layera gönderilmektedir. Dense layerda sigmoid fonksiyonu 0 ile 1 arasına sıkıştırılmış tek bir sonuç üretmektedir. Ayrıca model özetindeki “none”, yığın boyutu [batch size] kadar olmaktadır. Yığın boyutu, verileri modele dâhil ederken bir defada alınan yığındaki örnek sayısını göstermektedir. Eğitim verisindeki verilerin bir defa eğitimden geçmesine ya da modelin tüm veri kümesini bir kez görmesine epok [epoch] denmektedir.

Borsa derlemi ayrıca RNN-LSTM mimarisiyle oluşturulan model ile de eğitilmiştir. RNN-LSTM modelin özeti Tablo 4.9’da yer almaktadır.

Tablo 4.9: RNN-LSTM ile Bağlam Sınıflandırıcı Model Özeti

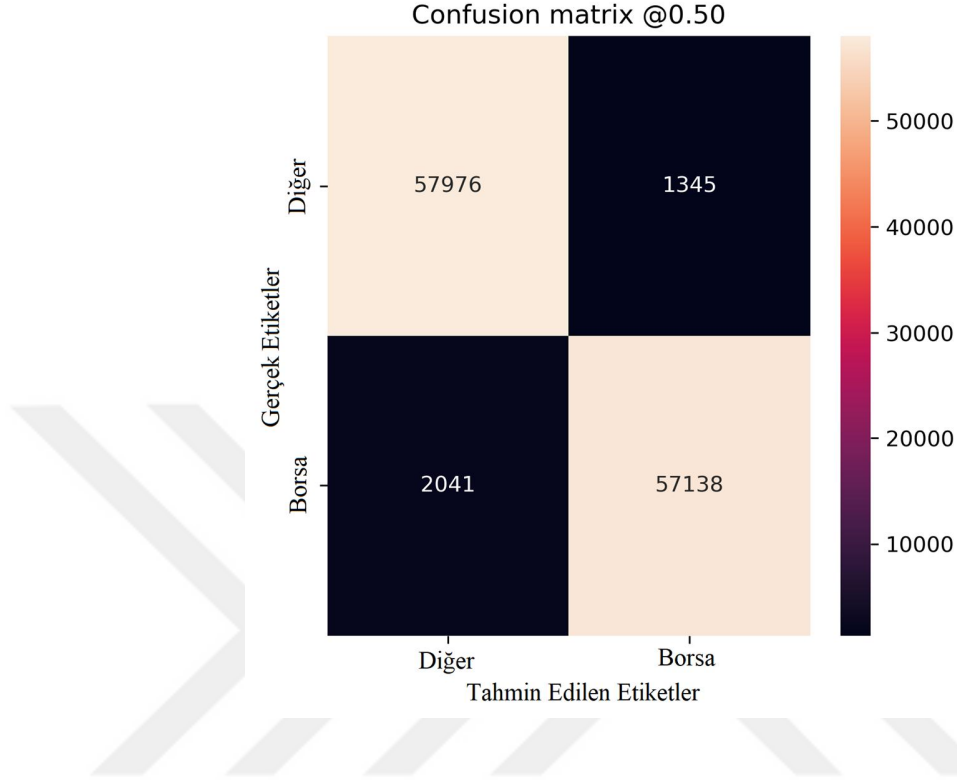
Model: “sequential”		
Layer (type)	Output Shape	Param #
embedding_layer (Embedding)	(None, 65, 50)	1000000
cu_dnngru (CuDNNLSTM)	(None, 65, 16)	4352
cu_dnngru_1 (CuDNNLSTM)	(None, 65, 8)	832
cu_dnngru_2 (CuDNNLSTM)	(None, 4)	224
dense (Dense)	(None, 1)	5
Total params:1.005.413		
Trainable params:1.005.413		
Non-trainable params:0		

Her iki RNN mimarisi ile elde edilen eğitim sonuçları Tablo 4.11’de verilmiştir. Modelin eğitimi esnasında aşırı uyum [overfitting] sorunuyla karşılaşmış, bu problemi aşmak için ağırlıklar düzenlenileştirilmek [weight regularization] suretiyle modeller yeniden eğitilmiştir. ℓ_1 düzenlenileştirmesinde maliyet, ağırlıkların katsayılarının mutlak değerine oransal olarak eklenirken, ℓ_2 düzenlenileştirmesinde maliyet, ağırlıkların katsayılarının karelerine oransal olarak eklenmektedir. Modelin eğitiminde, sinir ağırları bağlamında ağırlık azaltımı [weight decay] olarak bilinen ℓ_2 düzenlenileştirmesi “kernel_regularizer=’l2” ve “bias_regularizer=’l2” şeklinde eklenmiştir. Eniyileme yöntemi olarak “Adam” tercih edilmiş, ayrıca aşırı uyum problemine karşı erken durdurma [early stopping] parametresi

kullanılmıştır. Erken durdurma parametresi, eğitim esnasında elde edilen en iyi model korunurken doğrulama kaybının [val_loss] belirlenen sayıda epokta iyileşmemesi durumunda eğitimi durdurmayı sağlamaktadır. Modelde erken durdurma parametresi, eğitim esnasında eğitim kaybı ile test kaybı arasındaki farkın açılması ve doğrulama kaybının 5 kez minimumdan uzaklaşması durumunda eğitimi durduracak şekilde oluşturulmuştur. Her iki modelde de öğrenme oranı [learning rate] “1e-5” olarak belirlenmiş, her iki mimari için de eğitim, hem düzenlenleştirilmiş ağırlıklar kullanılarak hem de düzenlenleştirilmiş ağırlıklar kullanılmadan gerçekleştirilmiştir.

Grafik 4.3'te modelin test veri kümesindeki doğru ve hatalı tahminleri yer almaktadır. Ayırt etme matrisi (doğruluk-hata matrisi) gerçek değerler ile tahmin değerlerinin bir karşılaştırmasını yaparken esasında sınıflandırma problemine ilişkin tahmin sonuçlarının bir özetini sunmaktadır. Ayrıca ayırt etme matrisi sadece sınıflandırma modelinin yaptığı hatalar hakkında değil, yapılan hata türleri hakkında da bilgi vermektedir. Bu matrise göre tahmin modeli, gerçekte “Borsa” etiketli metinlerin 57.138 tanesini doğru bir şekilde “Borsa” olarak sınıflandırırken, 2.041 tanesini hatalı bir şekilde “Diğer” olarak sınıflandırmaktadır. Yine aynı matrise göre tahmin modeli gerçekte “Diğer” sınıfında yer alan metinlerin 57.976 tanesini doğru bir şekilde “Diğer” olarak sınıflandırırken, 1.345 tanesini hatalı bir şekilde “Borsa” olarak sınıflandırmaktadır.

Grafik 4.3: Test Veri Kümesi Doğruluk-Hata Matrisi



Tablo 4.10’da sınıflandırma sonucunda elde edilen doğru ve hatalı tahminlerin rakamsal değerleri yer almaktadır. Nitekim diğer performans metrikleri bu değerler esas alınarak hesaplandığında, modelin performansını ham bir şekilde ortaya koymaktadır.

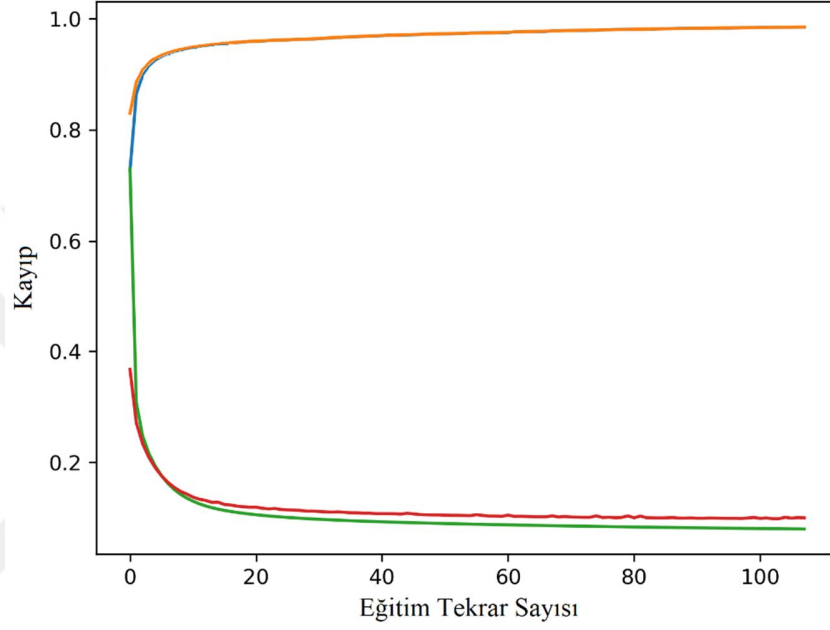
Tablo 4.10: Tahminlerin Doğruluk Durumuna Göre Dağılımı

İsim			Değer
Doğru Pozitif	[True Positive]	[tp]	57138
Yanlış Pozitif	[False Positive]	[fp]	1345
Doğru Negatif	[True Negative]	[tn]	57976
Yanlış Negatif	[False Negative]	[fn]	2041

Grafik 4.4 ağırlıkların düzenlenmesi ve erken durdurma parametresiyle kullanılan RNN-GRU mimarisinin eğitim ve doğrulama kaybı ile eğitim ve doğrulama başarımını göstermektedir. Buna göre modelin sınıflandırma başarısı 108’inci epokta %98 oranına

ulaşmıştır. Diğer bir ifadeyle sınıflandırma metriklerinden biri olan ve doğru tahminlerin tahminler toplamına oranını gösteren doğruluk [accuracy] oranı, modelin hem eğitim hem test verisinde %98 oranında doğru sınıflandırma başarısı gösterdiğini ortaya koymaktadır.

Grafik 4.4: RNN GRU Mimarisi ile Eğitim ve Doğrulama Başarımı

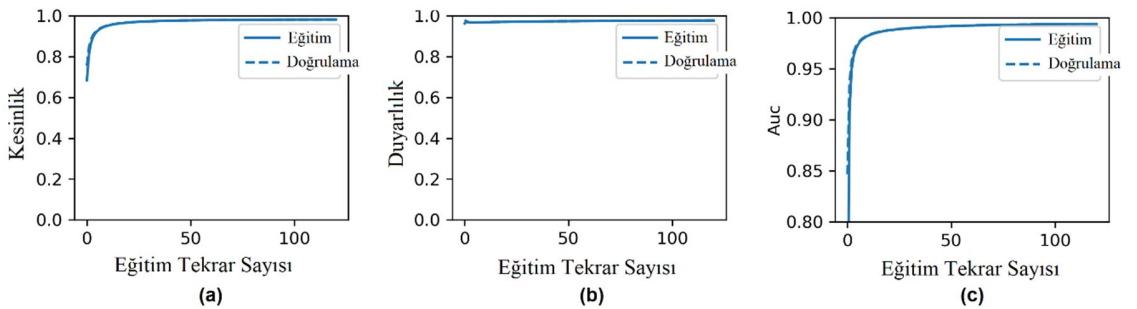


Tablo 4.11 ve Grafik 4.5'te modele ilişkin performans sonuçları bulunmaktadır. Tablo 4.11 ayrıca GRU ve LSTM mimarileri ile eğitilen modelin bir karşılaştırmasını da sunmaktadır. Düzenleştirilmiş ağırlıklar parametresi kullanmanın modelin performansına sadece %1'lik bir katkı sağladığı ancak esasen modelin aşırı uyum probleminin önüne geçtiği görülmektedir. Ayrıca GRU mimarisinin performans metriklerinde LSTM mimarisine kıyasla çok küçük bir avantajı olduğu görülmektedir. Ancak bu mimariler arasında asıl fark eğitim süresinden kaynaklanmaktadır. Zira LSTM mimarisiyle eğitim Nvidia GPU ile 40 çekirdekli cuda kuyruğunda yaklaşık 8 saat 48 dakikada, GRU mimarisiyle eğitim ise 5 saat 27 dakikada tamamlanmıştır. Dolayısıyla sınıflandırma için zamansal avantajından dolayı GRU mimarisiyle eğitilen model tercih edilmiştir.

Tablo 4.11: Bağlam Sınıflandırıcısının Eğitim Performans Sonuçları

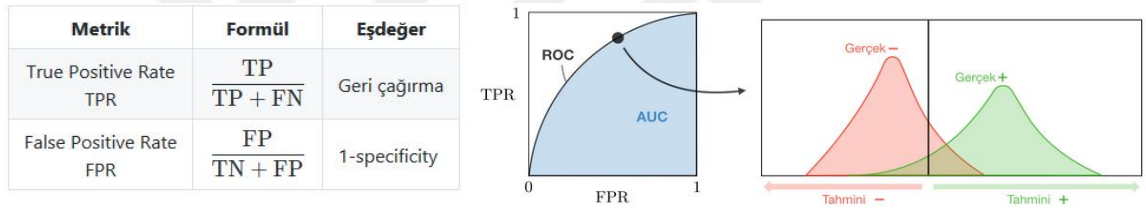
	Düzenleştirilmiş Ağırlıklar		Düzenleştirilmemiş Ağırlıklar	
	GRU	LSTM	GRU	LSTM
loss	0,07	0,08	0,06	0,06
accuracy	0,98	0,98	0,97	0,96
precision	0,98	0,98	0,98	0,96
recall	0,98	0,98	0,98	0,98
auc	0,99	0,99	0,99	0,99
val_loss	0,11	0,12	0,09	0,09
val_accuracy	0,98	0,98	0,97	0,96
val_precision	0,98	0,98	0,98	0,97
val_recall	0,98	0,98	0,98	0,98
val_auc	0,99	0,99	0,99	0,99

Kesinlik [precision] pozitif tahminlerin ne kadar kesin olduğunu göstermektedir. Diğer bir deyişle doğru sınıflandırılan pozitif örneklerin toplam pozitif örnek sayısına (TP+FP) oranıdır. Bu oran hem eğitim verisi, hem test verisi için %98'dir. Bir diğer performans metriği ise duyarlılık [recall] oranıdır. Duyarlılık oranı doğru sınıflandırılan pozitif örneklerin toplam pozitif tahminlere (TP+FN) oranıdır. Bu oran da eğitim ve test verileri için %96 olarak gerçekleşmiştir. F1 skoru ise dengesiz sınıflar için kesinlik ve duyarlılık ölçütlerinin harmonik ortalamasından oluşan yararlı hibrit bir metrik sunmaktadır. Modelde her iki ölçütü birlikte değerlendirme imkânı veren F1 skoru da %98 olarak hesaplanmıştır.

Grafik 4.5: RNN GRU Kesinlik, Duyarlılık ve AUC Oranları

Sınıflandırma performansını gösteren ölçütlerden bir tanesi de, Şekil 4.20’de görülen işlem karakteristik eğrisi [receiver operating curve] (ROC) ve eğri altında kalan alandır [area under curves] (AUC). ROC eğrisi, iki grubun ne kadar iyi sınıflandırılabilirdiğinin bir ölçütüdür. ROC eğrisinin dikey ekseninde *doğru pozitif*, yatay ekseninde *yanlış pozitif* değerleri yer almaktadır. ROC eğrisi sol üst köşeye ne kadar yakın ise, eğrinin altında kalan alan 1’e o kadar yaklaşmaktadır. Eğri altında kalan alanın alabileceği en büyük değer 1’dir. Bu durum modelin sınıflandırma başarısının kusursuz olduğunu göstermektedir (Alpar, 2022). Başka bir deyişle AUC değerinin 1’e yaklaşması testin ayırt ediciliğinin arttığına, 0,5 olması modelin sınıflandırma başarısına sahip olmadığına işaret etmektedir. Grafik 4.5’te görülen AUC değeri eğitim ve test verisi için %99 olarak hesaplanmıştır.

Şekil 4.20: Alıcı İşletim Karakteristik Eğrisi [ROC]

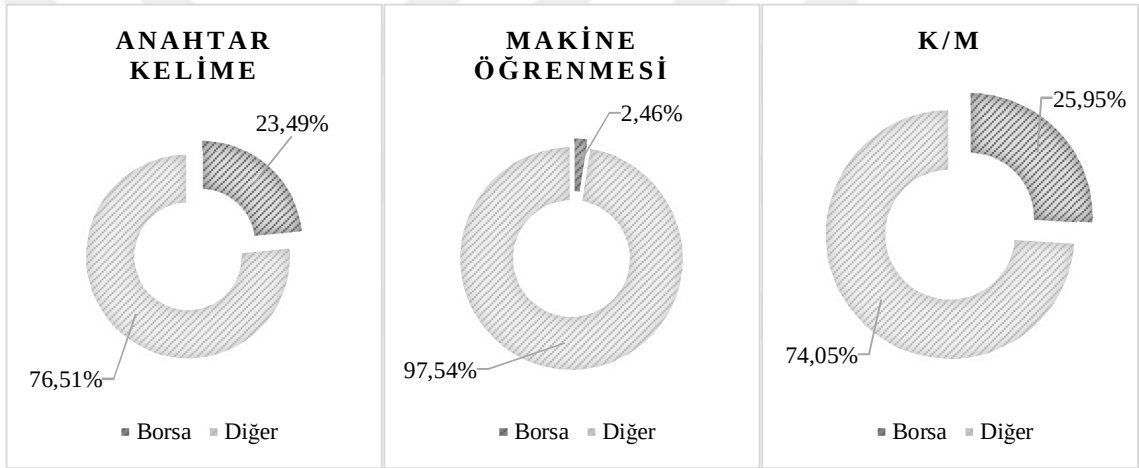


Kaynak: Amidi, A., & Amidi, S. (n.d.-b). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/l/tr/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks>

Finansal bağlam sınıflandırması, derin öğrenme yöntemlerinin yanı sıra anahtar kelimeler yardımıyla da yapılmıştır. Bu yaklaşımda ise borsa ile ilgili olması kesin olan sözcükler, ifadeler, şirket kodları ve şirket isimlerinden oluşan bir anahtar kelimeler listesi oluşturulmuştur. Daha sonra oldukça basit bir mantık yürütülerek, bu sözcükler #hashtag’lerde aratılmış, anahtar sözcüklerin içerisinde geçtiği tweetler borsa ile ilgili tweetler olarak sınıflandırılmıştır. Bu yöntemin dezavantajı KONYA, ADANA, AFYON, FENER gibi bazı hisse senedi kodlarının borsa ile ilgisi olmayan pek çok tweeti de borsa bağlamında sınıflandırmasıdır. Nitekim derin öğrenme yöntemiyle sınıflandırma modeli de esasen bu sorunu aşmak için geliştirilmiştir. Böylece makine öğrenmesi tabanlı bağlam sınıflandırıcı ile anahtar kelime tabanlı bağlam sınıflandırıcının performanslarını, tweetlerin borsa ile ilişkisini araştıran analizlerde karşılaştırma imkânı doğmuştur.

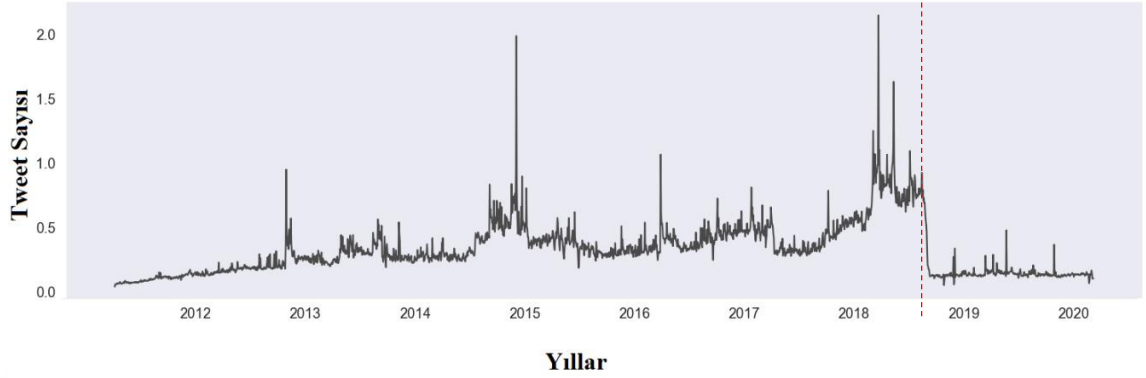
Analiz kapsamında 01.01.2012 tarihinden 28.02.2020 tarihine kadar toplanan 4.257.094.438 tweet bağlam sınıflandırıcıları yardımıyla sınıflandırılmıştır. Grafik 4.6 bağlamı belirlenen tweetlerin dağılımını göstermektedir. Buna göre anahtar kelime tabanlı yaklaşım toplam tweetlerin %24'ünü (953.965.658), makine öğrenmesi tabanlı yaklaşım ise toplam tweetlerin %2,5'ini (90.344.468) borsa ile ilgili olarak sınıflandırmıştır. Ayrıca anahtar kelime veya makine öğrenmesi tabanlı yaklaşımlardan (k/m) herhangi birinde borsa ile ilgili olarak sınıflandırılan tweetler ise toplam tweetlerin %26'sına tekabül etmektedir.

Grafik 4.6: Tweetlerin Borsa İle İlgisine Göre Bağlam Sınıflandırıcı Sonuçları



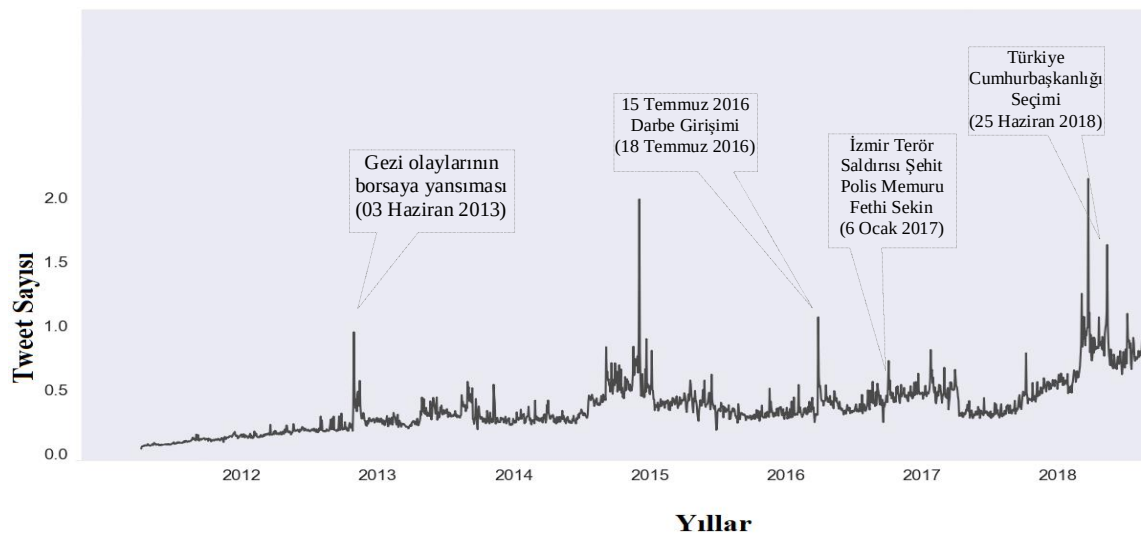
01.01.2012 tarihinden 28.02.2020 tarihine kadar, günlük olarak paylaşılan ve makine öğrenmesi (m) ya da anahtar kelime (k) yöntemlerinden herhangi birinin borsa ile ilgili olarak sınıflandırdığı tweetlerin yıllara göre dağılımı Grafik 4.7'de gösterilmektedir. Çalışma kapsamında, Twitter'ın veri paylaşımıyla ilgili yaptığı değişiklikler nedeniyle kullanılan tweet toplama yöntemi değiştiği için 10.10.2018 tarihinde tweet sayısında belirgin bir düşüş göze çarpmaktadır. Analiz sonuçlarını etkileyeceği düşüncesiyle 10.10.2018 tarihinden sonrası kapsam dışında bırakılmıştır.

Grafik 4.7: 2012-2020 Tarihleri Arasında Borsa Bağlamında Günlük Tweetler



01.01.2012-10.10.2018 tarihleri arasındaki günlük tweet sayılarını ve dikkat çekici olaylar Grafik 4.8’de yer almaktadır. Buna göre 03 Haziran 2013’te gezi olaylarının, 18 Temmuz 2016’da 15 Temmuz 2016 darbe girişiminin, 6 Ocak 2017’de İzmir’deki terör saldırısının ve 25 Haziran 2018’de Türkiye Cumhurbaşkanlığı seçiminin tweet sayılarına etkisi görülmektedir.

Grafik 4.8: 2012-2018 Tarihleri Arasında Borsa Bağlamında Günlük Tweetler ve Dikkat Çekici Artışlar



4.5.2.4 Finansal Fikir ve Duygu Sınıflandırıcı

Finansal fikir ve duygu sınıflandırıcı [opinion and sentiment classifier], finansal bir metnin içerdği duygusal tonun ve fikri yönelimin tespit edilerek sınıflandırılmasını amaçlamaktadır. Esasen alanyazınında duygu analizi [sentiment analysis] ve fikir madenciliği [opining mining] başlıkları altında ayrı ayrı ele alınan bu husus, finansal metinlerin doğası gereği hem fikir hem de duygu barındırabilmesinden dolayı çalışma kapsamında bütünlük bir yaklaşımla ele alınmaktadır.

Fikir analiziyle, borsa, endeks ya da borsada işlem gören hisse senedi hakkında, fiyatı ya da endeks değerini etkileme ihtimali bulunan, beklenti içeren ancak herhangi bir duygusal ifade içermeyen metinlerin belirlenerek sınıflandırılması kastedilmektedir. Bu yaklaşımla örneğin “Şirket beklentilerin üzerinde kâr açıkladı.” ifadesi, şirketin beklenen kârının piyasada fiyatlandığı, ancak beklentinin üzerinde elde edilen kârın fiyatlara henüz yansımadığı için fiyatın olumlu etkileneceği düşüncesiyle “pozitif” olarak etiketlenmektedir. Benzer şekilde “Şirket beklentilerle uyumlu kâr açıkladı.” ifadesi “nötr”, “Şirket beklentilerin altında kâr açıkladı.” ifadesi “negatif”, “Teknik göstergelere göre BIST’te yükseliş beklentisi korunuyor.” ifadesi “pozitif” ve “Hisse senedi, haftalık bazda genel olarak Al ve hacim dolgunluk olarak Güçlü Al sinyali vermektedir.” ifadesi “pozitif” olarak etiketlenmektedir. Dikkat edileceği üzere bu ifadeler, içerisinde herhangi bir duygu barındırmadan fiyatların yönüne işaret eden fikirler içermektedir.

Duygu analizinde ise bir metnin içerisinde geçen duygu kelimelerine odaklanılmaktadır. Metnin duygu durumu, ya tek tek kelimelerin bir kombinasyonu olarak ya da tek tek kelimelerin duygu içeriklerinin toplamı olarak ele alınmaktadır. Örneğin “*güzel kâr*”, “*bist mutluluğun resmi*”, “*Daha sana itimat eder miyim.*”, “*bir heyecan yaratmadın bizde*” gibi ifadelerde geçen duygu kelimeleri, daha önceden hazırlanan kelime sözlükleriyle belirlenerek metnin duygusal yönelimine göre sınıflandırılmaya çalışılmaktadır. Bu amaçla İngilizce için geliştirilmiş pek çok duygu kelimeleri sözlüğü vardır (AFINN, 2011; Mohammad ve Turney, 2011; BING, n.d.). Bunlardan bazıları sözcükleri *pozitif*, *negatif* ve

nötr olarak etiketlerken bazıları, *mutluluk, korku, güven, üzüntü, tikslenme* gibi duygu durumlarına göre etiketlemektedir. Bazı sözlüklerde ise kelimelere belirli aralıklarda (örneğin -5 ile 5 arasında) değişen puanlar atanmaktadır. Bu tür sözlüklerde negatif değerler olumsuz, pozitif değerler olumluyu işaret etmektedir. Bunlara ek olarak bazı duygu sözlükleri ise belirli alana özgü duyguları analiz etmek amacıyla hazırlanmıştır. Örneğin finans alanında İngilizce metinlerin duygusal yönelimini belirleyebilmek amacıyla hazırlanan Harvard IV-4 (2018) ya da Loughran gibi sözlükler alanyazınında önemli bir yere sahiptir (Loughran ve McDonald, 2011).

Türkçe metinlerin duygusal yöneliminin belirlenmesinde ise genellikle İngilizce için geliştirilen bu sözlükler Türkçeye tercüme edilerek kullanılmaktadır. Ancak bu yaklaşımda tercümeden kaynaklanan anlam kayıpları yaşanabilmektedir. Bunun dışında Türkçe için geliştirilmiş sözlükler (Dehkharghani ve diğerleri, 2016) ve daha geniş bir muhtevaya sahip dil kaynakları da mevcuttur (<http://ddi.itu.edu.tr/araclarkaynaklar>). Bu tür doğal dil işleme çalışmaları dilin doğasını anlamak ve bunu makinelere öğretmek açısından oldukça büyük önem taşımaktadır.

Sınıflandırma problemlerinin çözümünde doğal dil işleme ve derin öğrenme yöntemlerindeki gelişmeler, doğrudan bu tür sözlüklere başvurmadan, etiketli bir veri seti kullanılarak bir metnin fikri ve duygusal yöneliminin başarılı bir şekilde belirlenebilmesine imkân tanımaktadır. Dolayısıyla bu tür yaklaşımlarda önemli olan öncelikle ilgili alandaki metinleri temsil düzeyi yüksek ve başarılı bir şekilde etiketlenmiş bir veri seti oluşturmaktır. Bu amaçla araştırma kapsamında fikir ve duygu sınıflandırıcısını eğitmek için borsa, endeks ya da borsada işlem gören şirketler hakkındaki metinlerden oluşan bir veri kümesi (derlem) oluşturulmuştur. Veri kümesi oluşturulurken iki farklı kaynağa başvurulmuştur. Bunlardan birincisi yatırım kuruluşlarının raporlarıdır. Bu raporlar alanın uzmanlarınca hazırlanan oldukça zengin ve formel bilgi kaynaklarıdır. Bu raporların veri kümesine dâhil edilmesinin iki temel sebebi vardır. Bunlardan birincisi, bu raporlarda duru bir finans terminolojisi kullanılması diğeri ise, bu raporlardan bazılarında şirketlerle ilgili haberlerin uzmanlarca etiketlenerek yatırımcıların istifadesine sunulmuş olmasıdır. Diğer bir ifadeyle bu raporlar

Tablo 4.12’de görüldüğü üzere, doğrudan borsada işlem gören bir şirketle ilgili herhangi bir haberin fiyatlara muhtemel etkisinin nesnel bir şekilde etiketlendiği veriler de sunduğundan fikir ve duygu etiketçisinin eğitiminde kullanılmak üzere veri kümesine dahil edilmiştir.

Tablo 4.12: Yatırım Kuruluşları Uzmanlarınca Etiketlemiş Veri Örneği

Yatırım Kuruluşu	Rapor Tarihi	Yorum Yapılan Firma	Firma Kodu	Yorum	Beklenti
Garanti Yatırım	11.10.2017	Aselsan	ASELS	Aselsan 44m ABD\$ değerinde bir sözleşme imzaladı.	Olumlu
Halk Yatırım	08.03.2017	Ülker Bisküvi	ULKER	2016 yılın son çeyreğinde şirket, 51,2 milyon TL net kâr olan ortalama piyasa beklentisinin aksine 1,8 milyon TL net zarar açıkladı.	Olumsuz
qnbfi.com	15.06.2015	Türk Hava Yolları	THYAO	Türk Hava Yolları’nın yolcu trafiği Mayıs ayında bir önceki yıla göre %13 arttı, Ocak-Nisan dönemindeki büyüme %7 idi. Mayıs ayında bir önceki yıla göre iç hatlar %20, dış hatlar %9 arttı, doluluk faktörü bir önceki aya göre 90 baz puan düşerek %78,6 oldu. Arz edilen koltuk km %10 arttı ve ücretli yolcu km %9 yükseldi.	Sınırlı Olumlu
Ziraat Yatırım	14.05.2019	Karsan Otomotiv	KARSN	Şirket ile Industria Italiana Autobus S.p.A. (IIA) arasındaki Tesis Kullanım Sözleşmesi çerçevesinde, IIA, 2019 yılı içinde farklı müşteriler için ayrı ayrı zamanlarda şirkete 61,3m Euro tutarında yeni tam ve yarı mamul otobüs üretim siparişinde bulunmuş olup, söz konusu siparişlerin sipariş programına göre 2019 yılı içinde üretilerek teslim edilmesi planlanmaktadır.	Sınırlı Pozitif
Vakıf Yatırım	04.02.2019	Yapı Kredi Bankası	YKBNK	4Ç18 Finansal Sonuçları: Beklentilerin üzerinde net kar	(+)
Vakıf Yatırım	04.02.2019	Şişe Cam	SISE	4Ç18 Finansal Sonuçları: Net kar piyasa beklentisinin hafif üzerinde	(=)
Ziraat Yatırım	23.12.2019	Güneş Sigorta	GUSGR	Şirket ortaklarından Türkiye Vakıflar Bankası T.A.O. Memur ve Hizmetlileri Emekli ve Sağlık Yardım Sandığı Vakfı ile Vakıfbank Personeli Özel Sosyal Güvenlik Hizmetleri Vakfı, Şirket üzerinde sahip oldukları sırasıyla 27.054.000 TL ve 54.000.000 TL nominal değerli pay senetlerinin Borsada satışa konu edilebilmesi amacıyla Borsada işlem gören statüye çevrilmesi için Merkezi Kayıt Kuruluşu’na başvuruda bulunmuşlardır.	Negatif
Ziraat Yatırım	06.02.2020	Pegasus	PGSUS	İzmir-İstanbul seferini yapan Pegasus yolcu uçağı, inişten sonra Sabiha Gökçen Havalimanında pistten çıktı. Uçakta 175 yolcu, 2 bebek, 6 mürettebat olmak üzere toplam 183 yolcu bulunuyordu. Kazada 3 kişi hayatını kaybetti, 180 kişi yaralandı. Kaynak: Dünya	Sınırlı Negatif

Vakıf Yatırım	16.01.2017	Bimeks	BMEKS	Şirket bankalar ve leasing şirketleri nezdindeki finansal borçlarını, 1 yılı ödemesiz 6 yıl vadeli olarak yapılandırdı. Refinansman paketi ile finansal borçların uzun vadeli bir döneme taşınması ve daha düşük bir finansal maliyet ile yeniden finansmanının sağlanarak, şirketin borçluluk seviyesinin azaltılması ile nakit akışı ve kar üretme kapasitesinin artması planlanıyor.	nötr
Vakıf Yatırım	10.02.2020	Arçelik	ARCLK	Şirket 4Ç19'da beklentilere paralel yıllık bazda %14 düşüşle 240 mn TL net kâr açıkladı. Ciro ve brüt kârlılıkta beklentimize paralel gerçekleşmeye rağmen, daha yüksek faaliyet gider baskısı nedeniyle operasyonel kârlılık beklentimizin bir miktar altında kaldı. Buna göre, Şirketin 4Ç19'da FAVÖK'ü beklentimizin %10, piyasa beklentisinin %9 altında yıllık bazda %7 düşüşle 819 mn TL, FAVÖK marjı da yıllık bazda 2,1 puan düşüşle %9,8 düzeyinde gerçekleşti. Yorum: Beklentilere paralel net kâr performansına karşın, operasyonel kârlılığın tahminlerin altında kalması nedeniyle sonuçların hisse fiyatına sınırlı olumsuz yansıtacağını düşünüyoruz.	(-)
Halk Yatırım	07.01.2020	Alkim Kimya	ALKIM	Şirketin Acıgöl - Koralkim Tesislerinde kurulacak olan Potasyum Sülfat Yatırımı için; T.C. Sanayi ve Teknoloji Bakanlığı Teşvik Uygulama ve Yabancı Sermaye Genel Müdürlüğü'ne yapmış olduğu başvuru sonucunda teşvik belgesi düzenlendi. Teşviğe konu yatırım tutarı 106.992.565 TL'dir.	Pozitif
Halk Yatırım	27.04.2017	İş GYO	ISGYO	Şirket, 1Ç17'de ortalama piyasa beklentisinin %3,7 üzerinde 62 milyon TL satış geliri açıkladı. FAVÖK piyasa beklentisinin %4,7 altında 36,7 milyon TL gerçekleşirken, net kâr piyasa beklentisinin %16,9 altında yıllık %59 azalarak 20,6 milyon TL seviyesinde gerçekleşti.	Negatif

Yatırım kuruluşlarının etiketlemede ortak bir terminoloji kullanmadıkları göz önünde bulundurulmalıdır. Diğer bir deyişle, yatırım kuruluşlarının =, +, -, *olumlu*, *olumsuz*, *sınırlı olumlu*, *sınırlı olumsuz*, *hafif olumlu*, *hafif olumsuz*, *sınırlı pozitif*, *sınırlı negatif*, *hafif negatif*, *nötr* gibi farklı sözcükleri, karakterleri veya sembolleri tercih ettikleri görülmektedir. Dikkat edilirse her ne kadar sözcük ve sembol olarak birbirinden farklı olsa da yöneldiği anlam birbiriyle örtüşmektedir. Dolayısıyla analizde terminoloji birliğini sağlamak amacıyla tüm etiketler “pozitif”, “negatif” ve “nötr” sözcükleri ile temsil edilir hale getirilmiştir.

Ayrıca veri setine, Tablo 4.13'te bir örneği yer alan ve Twitter'dan toplanan, borsa, endeks ya da borsada işlem gören firmalarla ilgili tweetler de iki uzman tarafından elle etiketlenerek

dâhil edilmiştir. Twitter’den toplanan verilerin temel özelliği genellikle informal dilin kullanılmış olmasıdır. Bu veriler, iğneleme, ima, argo⁵⁵ gibi, biçimsel dille yazılmış raporlarda karşılaşılmayan özellikleri de kısmen de olsa modelleyebilme imkânı sunmaktadır. Böylece modele, biçimsel ya da biçimsel olmayan dille yazılmış metinleri sınıflandırabilme yeteneği kazandırılması amaçlanmıştır.

Tablo 4.13: Duygu-Fikir Etiketli Tweet Örnekleri

Olumlu [Pozitive]	Olumsuz [Negative]
kfein gıbraştı.....☺☺	XU100 HOŞ DURMUYOR.
HerŞeyGüzelOlacak DOAS	BIST'te aşağı eğilim belirginleşiyor.
gubrf hoş	SODA gözü aşağıda
isgsy güzel...	acsel şerefsizlik yapıyon
heks istikrarlı	banvt uyuzlanıp duruyor.
THYAO güzelleşiyor	Sıkıntı büyük xu100
BIST'te alımlar kuvvetli.	kRDMD olumsuz görüntü
nibas iyi haber	garan sebest düşüş
isdmr güzel kar	Aman dikkat XU100
prkab olumlu bilanço	avgyo satıldı,yollar ayrıldı
utpya mutluluğun resmi	BIST'te zayıflama belirginleşiyor.
kuyas iyi duruyor.	nibas düşüşe geçti
NIBAS GÜZEL OLACAK.	yesil olumsuz bilanço
odas yükselişi sağlıklı	birileri yine THYAO 'yu piste indirme çalışıyor, bakalım hangi havalimanına inecek bu sefer
halkb ve KRDMD güzeller	(yürümeyecek yol verdim buna ☺) allah bin bereket ☺☺☺
gozde güzel gitti masallah.	fmızp bugün de tavana dokundu 25,95 . yorulma belirtileri başladı. dikkatli olun .iz süren stop kullanın.
krdmd güzel göründü gözüme.	BNTAS Hisse acayip yoruldu kademeci oyuncular girişor günlükte. Elimde hala bir miktar var benim. 2.12 -2.19 üzeri vedalaşılırım. 1.86 Kar al Stop. YTD ...
msgyo Çok güzel çokkkk	gsray teknik göstergelerde 60 dakkalıkta yorulma belirtelire hakim artık dikkat etmek gerek
Baya iyiymiş KATMR ...	bist Yalanci bahar bunlar asil soguk gelince cicek acanlarin hepsi dokulecek.
TSPOR alemin kralı sensin	Bugün sağlam dayak var bist'te mevizleri terk edin :(
KRONT bana hoşgeldin dedi.Eyvallahhh.☺☺☺	GÖRÜNÜM ZAYIF, NEGATİF AÇILIŞ EKLENİYOR
asuzu emin adımlarla ha!	TRGYO, BİLANÇOSUNDA KÖTÜ SONUÇLAR GELDİ
Güzel bilanço ... ASELS ...	karsn 1.47 de analiz ettik seni ama bir heyecan yaratmadın bize henüz ...

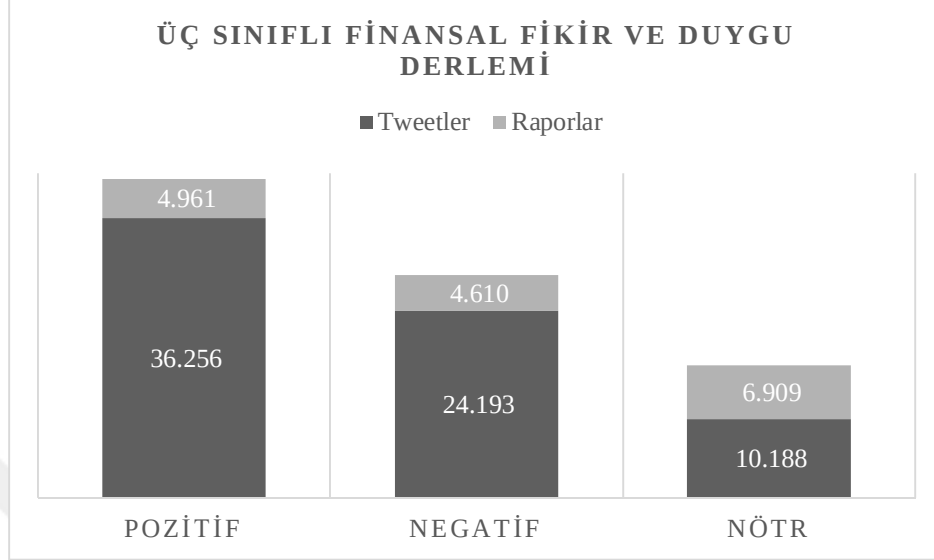
⁵⁵ Türkçe’de kullanılan argo sözcükler, Sezer’den (2015) elde edilerek, tez kapsamında oluşturulan finansal fikir ve duygu derlemine dahil edilmiştir.

"Türk Hava Yolları yeni bir take off'a geçecek"	Yüksekten atılan kedi bile düştüğünde zıplar.bist kedi den beter yapıştı resmen yere.
tire çok hoş	Acımadıki... Acımadıki..😞 bist100
gerel alım fırsatı olabilir	iyi silkelemişler
idgyo ver coşkuyu oğlum	Bjkas spegi tam bir tecavüzcü
kozal Baba büyüün ...	Ykbnk yaklaşık 1 yıllık amele sümüklüğü statüsünü devam ettiriyor.
tknsa yine uçtu 🙌🙌🙌🙌😄	BIST'tir git ya bu nedir bir yukarı bir aşağıya yorulmadınız mı arkadaş.
BIST'te iyimserlik devam ediyor.	kırılırsa diplerden dip beğen
BIST'te yukarı eğilim kuvvetleniyor	PETKM matriks ... 62 62 zorladım doce kızdı bulaşm lan tahtama dedi tırtım
Hastasıyız dedee! petkm	tektu anasını sikeyim bütün hisseler uçtu gitti amk hissesi hala aynıyerinde bu ne yavşak hisse be amk
desa günün roketi	otkar bı s. Git. Daha sana itimat edermiyim. dolandırıcı.
# KENT-İSTK-RAMAZAN ŞEKERİ OLABİLİRİ...:D	İş yatırım biz senden mal almanı beklerken sen tahtanın AMK'dun yuh sana SISE matriks ...

Grafik 4.9 verilerin edinildiği kaynağa ve duygu durumuna göre dağılımını göstermektedir. Üç sınıflı finansal fikir ve duygu derlemindeki 87.102 adet metnin 70.622 tanesi Twitter'dan, 16.480 tanesi yatırım kuruluşlarının raporları ve diğer kaynaklardan⁵⁶ elde edilmiştir. Ayrıca üç farklı sınıflı temsil eden veri kümesindeki verilerin 41.212 tanesi *pozitif*, 28.793 tanesi *negatif* ve 17.097 tanesi *nötr* olarak etiketlidir. Sonuç olarak finansal fikir ve duygu derlemi, ağırlıklı olarak iki farklı kaynaktan derlenmiş üç sınıflı bir veri kümesi elde edilmiştir.

⁵⁶ 16.480 metnin 13.810 tanesi yatırım kuruluşlarının raporlarından, 2670 tanesi ise Arslan (n.d.) kaynağında paylaşılan, içeriğindeki ürün ve kargo yorumları arındırılarak seçilen duygu etiketli metinlerden oluşmaktadır.

Grafik 4.9: Üç Sınıflı Finansal Fikir ve Duygu Derleminin Duyguya ve Edinildiği Kaynağa Göre Dağılımı



Türkçe dilinde yazılmış, borsa ile ilgili fikir ve duygu etiketli veri setlerinin çok sınırlı sayıda olması nedeniyle modelin başarımını artırmak için iki farklı mekanizmaya başvurulmuştur. Bunlardan biri “Attention” mekanizması diğeri öneğitimli [pretrained] kelime gömülmeleridir. RNN’in “Encoder” ve “Decoder” mimarisinde önemli bir başarı elde edilse de girdi dizisi uzun olduğunda dizinin ilk elemanlarının bütün içerisindeki nispi önemi azalmaktadır. Sıralı hesaplamaların temel kısıtı olan bu sorunu aşabilmek amacıyla, unutma ve hatırlamayı kısmen de olsa modelleyebilen LSTM, GRU gibi yaklaşımlar geliştirilmiştir. Bu yaklaşımlar hesaplama verimliliğini ve model performansını önemli ölçüde artırsa da sıralı hesaplamaların temel kısıtına tam bir çözüm getirememiştir. Bu bağlamda Attention mekanizması, girdi veya çıktı dizileri arasındaki mesafeye bakılmaksızın, yani en son oluşan gizli katmanı [hidden layer] Decoder’a göndermek yerine bütün gizli katmanları Decoder’a göndermektedir. Böylece dizi şeklindeki verinin ilk değerlerinin nispi önemi bir dereceye kadar korunarak bilgi bütünlüğü sağlanmış olmaktadır. “Attention” mekanizmasını kullanan “Transformer” teknolojisi ise paralel hesaplama yapabilme ve tek bir dikkat mekanizması yerine değişkenlere dizi veriyi soldan sağa ve sağdan sola değerlendirerek farklı açılardan

bakılmasını sağlayan multi head attetion mekanizmasına da sahiptir (Vaswani ve diğerleri, 2017).

Ayrıca modelin genelleme başarısı, modelin kelime temsillerini öğrenmesi için yeterli büyüklükte veri kümesine sahip olup olmamakla da ilgilidir. Dolayısıyla mevcut veri kümesinin kelime gömülmelerini öğrenmeye yetecek büyüklükte olmadığı ya da belirli bir düzeyde genelleme yapma başarısına ulaşamadığı durumlarda, öğrenilen niteliklerin genelleme yapabilmesi istendiğinde ön eğitimden geçmiş [pretrained] kelime temsillerine başvurulmaktadır. Alanyazınında öğrenme aktarımı [transfer learning] olarak kavramlaştırılan bu durum, makine öğrenmesi yöntemlerinin bir problemi çözerken edindiği bilgiyi saklayıp başka bir problemin çözümünde o bilgiyi kullanması olarak ifade edilmektedir (Torrey ve Shavlik, 2009). Öneğitimli kelime temsilleri, kelime kullanım istatistiklerini, yani metinlerde hangi kelimelerin birlikte görüldüğüne dair gözlemleri çeşitli tekniklerle hesaplayarak oluşturulmaktadır (Chollet, 2019). Böylece modelin kelime temsillerini mevcut veri kümesinden öğrenmesi yerine, dilin jenerik yapısını kavramış, kullanışlı özellikler sunan, önceden hesaplanmış kelime temsil vektörleriyle genelleme yapma başarısı artırılmaktadır. Çeşitli doğal diller için daha önce ayrıntılı şekilde bahsedilen Word2Vec, GloVe, FastText, ELMo, BERT gibi yöntemlerle önceden hesaplanmış çok çeşitli kelime gömülmeleri verisi elde edilebilir. Bu yöntemlerle öneğitimden geçmiş kelime gömülmeleri, araştırma kapsamında Embedding katmanıyla kullanılarak her seferinde sıfırdan öğrenmek yerine geri yayılımı sürdürerek önceden eğitilmiş ağırlıklarına ince ayar yapılmaktadır (Transfer Learning., n.d.). Böylece önceden edinilen bilgilerle daha küçük eğitim veri kümeleri daha yüksek tahmin başarısı gösteren modeller şeklinde eğitilebilmektedir. Bir diğer ifadeyle gözetimsiz ve gözetimli öğrenme fazlarını birleştiren bu yaklaşımda gözetimli öğrenme fazı, gözetimsiz öneğitme aşamasında öğrenilmiş olan özniteliklerin üzerine basit bir sınıflandırıcı eğitmeyi ya da öneğitme fazında öğrenilmiş olan ağda gözetimli hassas ayar yapmayı gerektirmektedir (Goodfellow ve diğerleri, 2018). Neticede bu yaklaşım daha az veriyle ve daha iyi performans gösteren modelleri daha hızlı bir şekilde eğitilebilme imkânı sunmaktadır.

Araştırma kapsamında fikir ve duygu sınıflandırıcısını eğitmek amacıyla BERT ile öneğitimli Türkçe dil modeli (MDZ Digital Library. (n.d.)) kullanılmış, ince ayar yapılarak [fine-tuning] mevcut veri kümesiyle bütünleşik hale getirilmiştir (McCormick, 2019) ⁵⁷. Veri kümesinin %80'i eğitim için %20'i test için ayrılmıştır. Modelin özeti Tablo 4.14'te yer almaktadır.



⁵⁷ Kodun oluşturulmasında aşağıdaki linkte verilen bilgilerden yararlanılmıştır.
<https://mccormickml.com/2019/07/22/BERT-fine-tuning/#sentence-length--attention-mask>, (Erişim Tarihi: 07.07.2020).

Tablo 4.14: BERT Fikir ve Duygu Sınıflandırıcı Model Özeti

GPU: Tesla V100-SXM2-32GB

Training: 69682

Test: 17341

55,745 training samples

13,937 validation samples

The BERT model has 201 different named parameters.

==== Embedding Layer====

bert.embeddings.word_embeddings.weight	(128000,768)
bert.embeddings.position_embeddings.weight	(512, 768)
bert.embeddings.token_type_embeddings.weight	(2,768)
bert.embeddings.LayerNorm.weight	(768,)
bert.embeddings.LayerNorm.bias	(768,)

==== First Transformer ====

bert.encoder.layer.0.attention.self.query.weight	(768,768)
bert.encoder.layer.0.attention.self.query.bias	(768,)
bert.encoder.layer.0.attention.self.key.weight	(768,768)
bert.encoder.layer.0.attention.self.key.bias	(768,)
bert.encoder.layer.0.attention.self.value.weight	(768,768)
bert.encoder.layer.0.attention.self.value.bias	(768,)
bert.encoder.layer.0.attention.output.dense.weight	(768,768)
bert.encoder.layer.0.attention.output.dense.bias	(768,)
bert.encoder.layer.0.attention.output.LayerNorm.weight	(768,)
bert.encoder.layer.0.attention.output.LayerNorm.bias	(768,)
bert.encoder.layer.0.intermediate.dense.weight	(3072, 768)
bert.encoder.layer.0.intermediate.dense.bias	(3072,)
bert.encoder.layer.0.output.dense.weight	(768, 3072)
bert.encoder.layer.0.output.dense.bias	(768,)
bert.encoder.layer.0.output.LayerNorm.weight	(768,)
bert.encoder.layer.0.output.LayerNorm.bias	(768,)

==== Output Layer ====

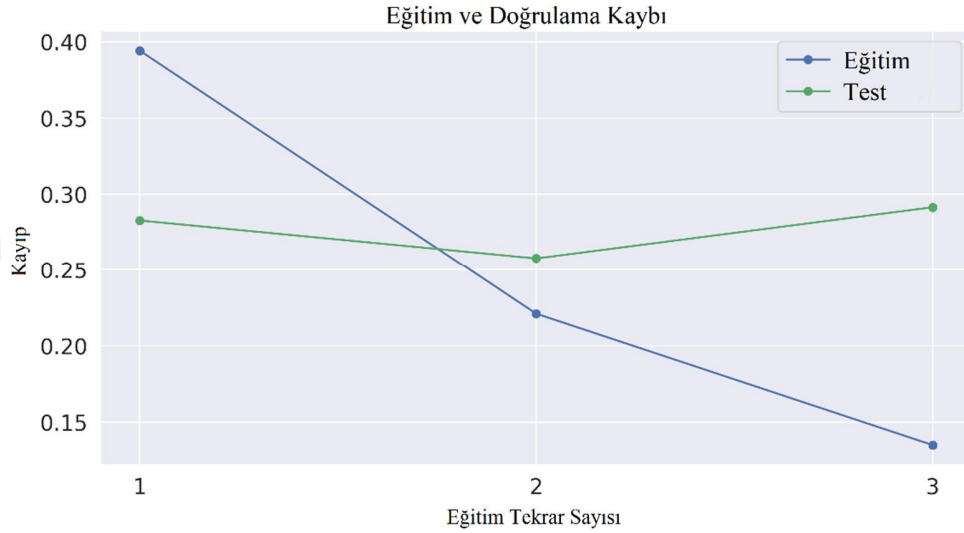
bert.pooler.dense.weight	(768,768)
bert.pooler.dense.bias	(768,)
classifier.weight	(3,768)
classifier.bias	(3,)

Öneğitimli modellerin kullanıldığı durumlarda eğitim genellikle 2 ya da 5 arasında değişen epoklarda gerçekleştirilmesi yeterlidir. Tablo 4.15 ve Grafik 4.10’da, 3 epokta ulaşılan başarı performansını göstermektedir. Üçüncü tekrardan sonra model aşırı uyum problemiyle karşılaştığından, eğitim 3 tekrarda tamamlanmıştır.

Tablo 4.15: Ön Eğitilmiş BERT Mimarisi ile Üç Kategorili Sınıflandırma Performansı

Sınıflar	Eğitim Kaybı	Doğrulama Kaybı	Doğrulama Başarımı	Eğitim Süresi	Doğrulama Süresi
Nötr	0,39	0,28	0,89	00:04:18	00:00:18
Pozitif	0,22	0,25	0,91	00:04:18	00:00:18
Negatif	0,13	0,29	0,91	00:04:18	00:00:18

Grafik 4.10: Ön Eğitilmiş BERT Mimarisi ile Üç Sınıflı Eğitim ve Doğrulama Başarımı



Fikir ve duygu sınıflandırıcısı için eğitilen model, Tablo 4.16’da görüldüğü üzere, test verisinde %91 düzeyinde bir sınıflandırma başarısına ulaşmıştır. Kesinlik [precision], duyarlılık [recall] ve F skorları da %90 civarında bir performans göstermiştir. Makine öğrenmesinde ikili ya da çoklu sınıflandırmaların performansını gösteren metriklerden biri de Matthews korelasyon katsayısıdır. Farklı boyutlara sahip sınıflar için dengeli bir ölçüt sunan Matthews korelasyon katsayısı (MCC) istatistikte phi katsayısı olarak da bilinmektedir. MCC özünde -1 ile +1 arasında bir korelasyon katsayısı değerleri alırken +1, mükemmel bir tahmini, 0 ortalama bir tahmini ve -1 ise son derece başarısız bir tahmini

temsil etmektedir (Baldi ve diğ erleri, 2000; Gorodkin, 2004; Jurman ve diğ erleri, 2012). Bu  er evede modelin MCC'ye g re %86'lık bir tahmin ba arısı sergilediđi g r lmektedir. Oranın 1'e yakla tı a b t n sınıfların tahmin ba arısının y ksek olduđu anlamına geldiđinden finansal fikir ve duygu sınıflandırıcısının sınıflandırma ba arısı yeterli g r lm   t r.

Tablo 4.16:    Sınıflı Fikir ve Duygu Sınıflandırıcısının Eđitim Performans Sonu ları

Dođrulama Veri Seti Sonu�ları	
F-Score	0,903
Recall	0,905
Precision	0,900
Accuracy	0,913
Total MCC	0,862

Modelin performansı sınıflar bazında deđerlendirildiđinde ise pozitif ve negatif fikir ve duygu y nelimlerini sınıflandırmada %91'ler d zeyinde ba arı g sterdiđi, ancak n tr metinlerin sınıflandırmasında ba arı oranının %2 gibi k   k bir fark ile %89'larda kaldıđı g r lmektedir (Bakınız Tablo 4.15). Bu durum fikir ve duygu sınıflandırıcının pozitif ve negatif metinleri n tr metinlere kıyasla daha ba arılı bir  ekilde ayırt edebildiđi anlamına gelmektedir. Esasen bir insan i in de bazı durumlarda metnin n tr olduđuna karar vermek, pozitif ya da negatif olduđuna karar vermekten daha zor olabilmektedir.  rneđin kimileri i in, Tablo 4.12'de yer alan, “Net k r piyasa beklentisinin hafif  zerinde” c mlesi pozitif olarak algılanabilirken, beklenti piyasada fiyatlandıđı d   ncesiyle uzmanlar tarafından n tr olarak etiketlenmi tir. Bu c mle s zc k d zeyinde deđerlendirildiđinde, *k r*, *beklenti*, * zerinde* gibi olumlu  ađrışım yapan s zc klerle in a edildiđi g r lmektedir. Dolayısıyla bu c mlenin duygusal y nelimini sınıflandırmak, yani n tr veya pozitif sınıflardan hangisine dahil olduđuna karar vermek, a ık bir  ekilde pozitif ya da negatif s zc klerle in a edilen c mleleri sınıflandırmaya kıyasla daha zordur. Bu c mlenin n tr olarak sınıflandırılmasında ayırt edici unsur *hafif* s zc đ d r. Ba ka bir deyi le bu c mle, *hafif* s zc đ  olmadan “Net k r piyasa beklentisinin  zerinde”  eklinde kurulmu  olsaydı, pozitif olarak sınıflandırılması gerekirdi. İ te sınıflandırıcının bu ayrımı  đrenebilmesi, pozitif ya da negatif sınıflara g re

daha zordur. Bu nedenle %2’lik kaybın bu tür pozitif ya da negatif ifadelerle yakın cümlelerden kaynaklandığı düşünülmektedir.

Tablo 4.17: Üç Sınıflı Fikir ve Duygu Sınıflandırıcısının Kategorilere Göre Eğitim Performans Sonuçları

	nötr	pozitif	negatif	accuracy	macro avg	weighted avg
Precision	0,833	0,932	0,934	0,913	0,900	0,913
Recall	0,865	0,918	0,934	0,913	0,905	0,913
F1-score	0,849	0,925	0,934	0,913	0,903	0,913
support	3408	8206	5727	0,913	17341	17341

Tablo 4.16’da ortalama olarak verilen üç sınıflı fikir ve duygu sınıflandırıcısının eğitim performans sonuçları, Tablo 4.17’de her bir sınıf düzeyi için ayrıntılı olarak verilmiştir. Buna göre Tablo 4.17’de yer alan üç sınıflı modelin eğitim sonuçları incelendiğinde, *nötr* kategorisindeki sınıflandırma başarısının, *pozitif* ve *negatif* kategorilerindeki başarıya kıyasla daha düşük kaldığı gözlemlenmiştir. Bu durumun temel nedeninin, *nötr* etiketli cümlelerin ayırt edici özelliklerinin, *pozitif* ya da *negatif* cümlelerin ayırt edici özellikleri kadar belirgin olmamasından kaynaklanabileceği düşünülmektedir. Sonuç olarak bu farkın, analizde büyük veri kullanıldığı için, sonuçlar üzerinde önemli bir etkisinin olmayacağı varsayılmaktadır.

4.5.2.5 Bilişsel Özniteliklere İlişkinin Analizi

Bilişsel etki temelli analizde, borsa ile ilgili Twitter kullanıcılarının paylaştıkları tweetlerin fikri ve duygusal yönelimlerinin Borsa İstanbul pay piyasası endeksleri ile ilişkisine odaklanılmaktadır. Bilişsel öznitelikler hazırlanırken, belirli bir günde borsa bağlamında paylaşılan tweetlerin duygusal yönelimleri esas alınmıştır. Tweetlerin duygusal yönelimleri önceki gün saat 10:00’dan işlem günü saat 10:00’a kadar olan t-1 zamanı için hesaplanmış, yani açılıştan açılışa pozitif, negatif ve nötr tweetlerin sayısı elde edilmiş, borsa göstergeleri ise t günündeki borsanın açılış değerlerinden elde edilmiştir. Tatil günleri için duygu ve fikir yönelimleri, tatil gün sayısının aritmetik ortalaması alınarak hesaplanmıştır.

Bilişsel etkiyi ortaya koymak amacıyla, başvuru göstergelerden biri, mesaj içeriklerindeki duygusal sinyallerle ilgili olan yükseliş potansiyeli [bullishness] indeksidir. Yükseliş potansiyeli indeksi Antweiler ve Frank'ın (2004) çalışması baz alınarak Eşitlik 4.10'da yer aldığı şekliyle hesaplanmıştır. Bu eşitlikte B_t sembolü yükseliş potansiyelini, M_t sembolü belirli bir zamandaki tweet sayısını, p sembolü pozitif tweetleri, n sembolü negatif tweetleri temsil etmektedir. Yükseliş potansiyelinin artması piyasada pozitif beklentilerin hâkim olduğu anlamına gelmektedir. Yükseliş potansiyeli indeksi arttığında endeks değerinin ve işlem hacminin artması, volatilitenin düşmesi beklenir.

$$B_t = \ln \left(\frac{1 + M_t^p}{1 + M_t^n} \right) \quad (4.10)$$

Yükseliş potansiyeli indeksi hem tweetler için hem retweetler için hem de beğeni tweetler için hesaplanmıştır. Analizde kullanılan yükseliş potansiyeli indekslerinin kodları, açıklamaları ve hangi katmanı temsil ettiklerine ilişkin ayrıntılı bilgi Tablo 4.18'de yer almaktadır. Buna göre tweetler için hesaplanan yükseliş potansiyeli indeksi (bullgt, bullkt) bilişsel katmanda, retweetler (bullRTgt, bullRTkt) ve beğeni tweetler (bullFAVgt, bullFAVkt) için hesaplanan yükseliş potansiyeli indeksleri hem bilişsel hem davranışsal hem sosyal katmanda yer almaktadır.

Tablo 4.18: Yükseliş Potansiyeli İndeksleri

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
bullgt	bullkt	Yükseliş potansiyeli TW	Bilişsel
bullRTgt	bullRTkt	Yükseliş potansiyeli RT	Bilişsel-Davranışsal-Sosyal
bullFAVgt	bullFAVkt	Yükseliş potansiyeli FAV	Bilişsel-Davranışsal-Sosyal

Bilişsel etkiye ilişkin bir diğer ölçüt sözbirliği etkisidir. Söz birliği indeksi [agreement consistent] (Antweiler ve Frank, 2004) belirli bir platformda paylaşım yapan hesaplar arasındaki hâkim fikir birliği düzeyini yansıtan bir ölçüttür. Eşitlik 4.11'de sözbirliği etkisini

nasıl hesaplandığı gösterilmektedir. Bu eşitlikte A sembolü sözbirliği düzeyini, M_t^p sembolü belirli bir zamandaki pozitif tweet sayısını ve M_t^n sembolü belirli bir zamandaki negatif tweet sayısını temsil etmektedir.

$$A_t = 1 - \sqrt{1 - \left(\frac{M_t^p - M_t^n}{M_t^p + M_t^n} \right)} \quad (4.11)$$

Söz birliği indeksinin artması duygu ve fikirlerin homojenleştiği anlamına gelmekte ve ***tam fikir birliği olduğunda*** işlem hacminin ve volatilitenin azalması beklenmektedir. Başka bir ifadeyle alım yönünde tam fikir birliği olduğunda daha az satıcı olacağından, satım yönünde tam fikir birliği olduğunda daha az alıcı olacağından işlem hacminin ve volatilitenin düşmesi beklenmektedir. Fikirlerin ayrışması durumunda ise işlem hacmi ve volatilitenin artması beklenilmektedir.

Sözbirliği indeksi de hem tweetler için hem retweetler için hem de beğeni tweetler için hesaplanmıştır. Analizde kullanılan sözbirliği indekslerinin kodları, açıklamaları ve hangi katmanı temsil ettiklerine ilişkin ayrıntılı bilgi Tablo 4.19’da yer almaktadır. Buna göre tweetler için hesaplanan sözbirliği indeksi (aggregt, aggregt) bilişsel katmanda, retweetler (aggreRTgt, aggreRTkt) ve beğeni tweetler (aggreFAVgt, aggreFAVkt) için hesaplanan sözbirliği indeksleri hem bilişsel hem davranışsal hem sosyal katmanda yer almaktadır.

Tablo 4.19: Sözbirliği İndeksleri

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
aggregt	aggrekt	Söz birliği TW	Bilişsel
aggreRTgt	aggreRTkt	Söz birliği RT	Bilişsel-Davranışsal-Sosyal
aggreFAVgt	aggreFAVkt	Söz birliği FAV	Bilişsel-Davranışsal-Sosyal

Bilişsel katmanda yükseliş potansiyeli ve söz birliği skoru hesaplanırken nötr tweetler göz ardı edilmektedir. Bu eksikliği gidermek amacıyla nötr tweetlerin pozitif ve negatif tweetlere göre durumunu göz önünde bulunduran öznitelikler de hesaplanmıştır. Tablo 4.20’de görüldüğü üzere, nötr tweetlerin pozitif ve negatif tweetlerin toplamına oranı (neutralTWgt, neutralTWkt) bilişsel katmanda, nötr retweetlerin pozitif ve negatif retweetlerin toplamına oranı (neutralRTgt, neutralRTkt) ve nötr beğeni tweetlerin pozitif ve negatif beğeni tweetlerin toplamına oranı (neutralFAVgt, neutralFAVkt) bilişsel, davranışsal ve sosyal katmanda yer almaktadır.

Tablo 4.20: Nötr Tweet Öznitelikleri

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
neutralTWgt	neutralTWkt	nötr / (p+n) TW	Bilişsel
neutralRTgt	neutralRTkt	rtnötr / (rtp+rtn) RT	Bilişsel-Davranışsal-Sosyal
neutralFAVgt	neutralFAVkt	favnötr / (rtp+rtn) FAV	Bilişsel-Davranışsal-Sosyal

Bilişsel katmanda yer alan bir diğer grup öznitelik ise pozitif ve negatif tweetlerin toplam tweetlere oranıyla ilgilidir. Bu grupta yer alan öznitelikler, belirli bir günde borsa bağlamında paylaşılan tweetlerin içeriklerindeki pozitif ve negatif yönelimleri göz önünde bulundurmaktadır.

Tablo 4.21: Pozitif-Negatif Tweet Öznitelikleri

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
bgpts_gt	bkpts_kt	Toplam pozitif tweet / toplam tweet	Bilişsel
bgnts_gt	bknts_kt	Toplam negatif tweet / toplam tweet	Bilişsel

Tablo 4.21’de de görüldüğü üzere, belirli bir gündeki pozitif tweetlerin toplam tweetlere oranı (bgpts_gt, bkpts_kt) ve negatif tweetlerin toplam tweetlere oranı (bgnts_gt, bknts_kt)

hesaplanarak elde edilen öznitelikler, sadece içeriğin duygusal yönelimi ile ilişkili olduğu için bilişsel katmanın temsilcisi olarak kabul edilmiştir.

4.5.3 Davranışsal Katman: Etki Tepki Analizi

Davranışsal katmanda, Twitter kullanıcılarının tweetlerinin içeriğinden ziyade tweet atma, retweet etme ve beğenme gibi sözlü ya da yazılı olmayan ve daha çok tepkisel bir nitelikte olan davranışlarına odaklanılmaktadır. Bu katmanda, bağlam sınıflandırması sonucunda borsa ile ilgili olarak sınıflandırılan tweet, retweet ve beğeni tweet sayıları ile borsa göstergeleri arasındaki ilişki analiz edilmektedir.

Bu çalışma kapsamında davranışsal katmanda yer alan öznitelikler Tablo 4.22’de sunulmuştur. Buna göre davranışsal katmanda, belirli bir günde borsa bağlamında paylaşılan toplam tweet sayısı (gt_norm, kt_norm), retweetlerin toplam tweetlere oranı (RTgtt, RTktt), beğeni tweetlerin toplam tweetlere oranı (FAVgtt, FAVktt), toplam retweet sayısı (bggRTnorm, bkgRTnorm) ve toplam beğeni tweet sayısı (bggFAVnorm, bkgFAVnorm) yer almaktadır.

Tablo 4.22: Davranışsal Öznitelikler

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
gt_norm	kt_norm	Toplam tweet sayısı	Davranışsal
RTgtt	RTktt	Retweetlerin toplam tweetlere oranı	Davranışsal-Sosyal
FAVgtt	FAVktt	Beğeni tweetlerin toplam tweetlere oranı	Davranışsal-Sosyal
bggRTnorm	bkgRTnorm	Toplam genel RT	Davranışsal-Sosyal
bggFAVnorm	bkgFAVnorm	Toplam beğeni tweet	Davranışsal-Sosyal

4.5.4 Sosyal Katman: Sosyal Değişkenlerin Finansal Kararlar ile İlişkisi

Bireysel yatırımcı karar modelinin sosyal katmanında, bireysel yatırımcıların bilgi arayışlarında sosyal çevrenin etkisine ve sosyal çevre ile etkileşimlere odaklanılmaktadır. Bu çalışma kapsamındaki sosyal öznitelikler Tablo 4.23'te yer almaktadır. Bu özniteliklerden biri toplam takipçi sayısının toplam arkadaş sayısına oranıdır (FolFrigtt, FolFrikt). Bu öznitelik, kanaat önderlerinin takipçi sayılarının arkadaş sayılarından daha fazla olacağı varsayımına dayanarak, belirli bir günde borsa bağlamında tweet atan kullanıcıların kanaat önderi olma potansiyelini analize dâhil etmektedir. Bir diğer öznitelik toplam takipçi sayısıdır(bggfols_norm, bkgfols_norm). Belirli bir günde borsa bağlamında tweet atan kullanıcıların takipçi sayıları, tweetlerin kaç kişiyi etkileyebileceğine ilişkin bir gösterge olarak kullanılmaktadır. Bir diğer öznitelik ise arkadaş sayısıdır (bggfris_norm, bkgfris_norm). Belirli bir günde borsa bağlamında tweet atan kullanıcıların kaç kişi takip ettiğine ilişkin bir göstergedir.

Tablo 4.23: Sosyal Öznitelikler

Öznitelik Kodu		Öznitelik Açıklaması	Öznitelik Türü
Makine Öğrenmesi Tabanlı [m]	Anahtar Kelime Tabanlı [k]		
FolFrigtt	FolFrikt	Toplam takipçi sayısı / toplam arkadaş sayısı	Sosyal
bggfols_norm	bkgfols_norm	Takipçi sayısı	Sosyal
bggfris_norm	bkgfris_norm	Arkadaş sayısı	Sosyal
interactionG	interactionK	Etkileşim skoru	Sosyal

Sosyal katmanda yer alan bir diğer öznitelik etkileşim skorudur(interactionG, interactionK). Etkileşim skoru, takipçi sayısı, tweet sayısı, retweet sayısı ve beğeni tweet sayısını birbirleri ile ilişkilendiren bir skordur. Etkileşim skoru, Eşitlik 4.12'de yer aldığı şekliyle hesaplanmıştır. Etkileşim skoru, t günündeki tweetlerin ne kadar retweet edildiğini ve beğenildiğini, t gününde etkileşime dâhil olanların takipçi sayıları ile ilişkilendiren bir skordur (Tardy, 2023). Eşitlikte RT_t sembolü belirli bir gündeki retweet sayısını, FAV_t sembolü belirli bir gündeki beğeni tweetlerin sayısını, TTW_t sembolü belirli bir gündeki

toplam tweet sayısını, FOL_t sembolü ise belirli bir günde tweet atan kullanıcıların takipçi sayılarını temsil etmektedir.

$$Etkileşim\ Skoru = \frac{\left(\frac{RT_t + FAV_t}{TTW_t}\right)}{FOL_t} \times 100 \quad (4.12)$$

4.5.5 Bütünleşik Karar Modeli ve Borsa İstanbul Pay Piyasası Endeksleri ile İlişkisi

Analizin bu aşamasında anahtar kelime ve makine öğrenmesi sınıflandırması varyasyonları dâhil, 20 öznitelik elde edilmiştir. Twitter verilerinden elde edilen bu öznitelikler, endeks açılış değeri, işlem hacmi, getiri ve volatilité gibi borsa göstergesi olarak belirlenen bağımlı değişkenleri tahmin başarısı araştırılmaktadır. Ancak öznitelik sayısının çok fazla olması nedeniyle, ilgili ve gerekli öznitelikler çeşitli yöntemlerle seçilmiş, tahmin modeli seçilen bu öznitelikler ile oluşturulmuştur. Lineer regresyon için öznitelik seçimi korelasyon değerlerine göre yapılırken lasso regresyon, rassal orman ve XGBoost yöntemleri için öznitelik seçimi öğrenme algoritmasının bir parçası olarak öğrenme aşamasında gömülü olarak yapılmaktadır.

4.5.5.1 Tweet Öznitelikleri ile BIST Pay Endekslerinin Genel İlişkisi

Tweet öznitelikleri ile BIST Pay Endekslerinin genel ilişkisi ilk aşamada korelasyon analizi ile ortaya konulmaktadır. Korelasyon analizine çok boyutlu büyük veri çalışmalarında çözümlemenin nereye odaklanması gerektiği konusunda ön bilgi edinmek amacıyla sıklıkla başvurulmaktadır. Bu yaklaşım bütünleşik modele geçmeden önce ilişkileri öznitelik düzeyinde gözlemlemek açısından oldukça önemlidir. Korelasyon katsayıları, değişkenler arasındaki ilişkinin yönü ve kuvveti hakkında bilgi vermektedir. Genellikle -1 ile +1 arasında değişen değerler alırlar. Korelasyon katsayılarının -1 ve +1'e yaklaşması ilişkinin kuvvetinin arttığı, 0'a yaklaşması ilişkinin kuvvetinin azaldığı anlamına gelmektedir. Buna göre +1 pozitif tam ilişkiyi, -1 negatif tam ilişkiyi, 0 ise ilişki olmadığını ifade etmektedir.

Değişkenler arasındaki ilişkinin anlamlı olup olmadığını belirlemek için anlamlılığın da test edilmesi gerekmektedir (Alpar, 2022).

Ölçümle belirtilen iki değişken arasındaki doğrusal ilişkinin kuvveti ve yönü Pearson korelasyon katsayısı (r) yöntemi ile belirlenebilmektedir (Alpar, 2022). Pearson korelasyon katsayısının hesaplanması Eşitlik 4.13'te gösterilmektedir.

$$r = \frac{\sum_{i=1}^n x_i y_i - \frac{\sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n}}{\sqrt{\left(\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}\right) \left(\sum_{i=1}^n y_i^2 - \frac{(\sum_{i=1}^n y_i)^2}{n}\right)}} \quad (4.13)$$

Korelasyon katsayıları, büyüklüğü dikkate alınarak nitelendirilmektedirler. Tablo 4.24 Pearson korelasyon katsayısının mutlak değerine göre niteliğini göstermektedir. Buna göre 0,00 ile 0,19 arası korelasyon skoru, düşük düzeyde ilişki olduğu; 0,20-0,39 arası korelasyon skoru, zayıf ilişki olduğu; 0,40-0,69 arası korelasyon skoru, orta düzeyde ilişki olduğu; 0,70-0,89 arası korelasyon skoru, kuvvetli ilişki olduğu; 0,90-1,00 arası korelasyon skoru ise çok kuvvetli ilişki olduğu anlamına gelmektedir.

Tablo 4.24: Pearson Korelasyon Katsayısının Nitelendirilmesi

r	Nitelendirme
0,00-0,19	İlişki yok ya da önemsiz düzeyde düşük ilişki
0,20-0,39	Zayıf (düşük düzeyde) ilişki
0,40-0,69	Orta düzeyde ilişki
0,70-0,89	Kuvvetli (yüksek düzeyde) ilişki
0,90-1,00	Çok kuvvetli düzeyde ilişki

Öznitelik seçimi ve bütünlük karar modeline geçmeden önce tweetlerden elde edilen özniteklere BIST Pay Endekslerinin açılış değeri, volatilité, işlem hacmi ve getiri ile ilişkisi korelasyon analizi ile ortaya konulmuştur. Bu genel değerlendirmelerden sonra öznitelik seçimi ve bütünlük karar modeli analizlerine geçilmiştir.

4.5.5.1.1 Açılış Değeri

Bu çalışma kapsamında Twitter verilerinden elde edilen özniteliklerin BIST Pay Endekslerinin açılış değerleri ile ilişkisini gösteren ikili korelasyonlar Tablo 4.25'te yer almaktadır. Tweetlerden elde edilen özniteliklerden biri yükseliş potansiyeli indeksidir (bullkt, bullgt). Yükseliş potansiyeli indeksinin artması piyasada pozitif beklentilerin arttığına, indeksin azalması piyasada negatif beklentilerin hâkim olmaya başladığına işaret etmektedir. Dolayısıyla yükseliş potansiyeli indeksinin artması endeks değerlerinin yukarı yönlü hareket etmesine, indeksin azalması endeks değerlerinin aşağı yönlü hareket etmesine yol açması beklenmektedir. Bu çalışma kapsamında belirli bir günde borsa bağlamında atılan tweetlerden elde edilen yükseliş potansiyeli indeksi ile BIST Pay Endekslerinin açılış değeri arasında pozitif yönlü zayıf bir korelasyon olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,25$, $p < 0,001$; XU100 [k] için: $r = 0,17$, $p < 0,001$). Bu durum yükseliş potansiyeli indeksi arttığında endeks açılış değerlerinin artacağı, azaldığında endeks açılış değerinin azalacağı anlamına gelmektedir. Beğeni tweetler için hesaplanan yükseliş potansiyeli indeksi (bullFAVkt, bullFAVgt) ile de benzer bir ilişki söz konusudur (XU100 [m] için: $r = 0,21$, $p < 0,001$; XU100 [k] için: $r = 0,18$, $p < 0,001$). Ancak retweetlerden oluşan yükseliş potansiyeli indeksi (bullRTkt, bullRTgt) ile endeks açılış değerleri arasında anlamlı bir korelasyon olmadığı görülmektedir. Dolayısıyla endekslerin açılış değerleri retweetlerden ziyade beğeni tweetlerin ilişkili olduğu sonucuna varılabilir. Borsa İstanbul için yükseliş potansiyeli ile endeks açılış değeri arasındaki ilişkiyi ortaya koyan bu bulgular, Antweiler ve Frank'ın (2004) DIA (Dow Jones Industrial Average) ve XLK (Dow Jones Internet Commerce Index) endeksleri ile Sprenger ve diğerlerinin (2014) S&P 100 indeksi için elde ettikleri bulgular ile örtüşmektedir. Bu durum yatırımcıların pozitif tweetler ya da beklentiler arttığında pişmanlıktan kaçınma eğiliminin etkisiyle, fırsatları kaçırmamak için işlem yapıyor olabileceğine işaret etmektedir.

Bir diğer gösterge söz birliği indeksidir (aggrekt, aggregt). Söz birliği indeksi belirli bir günde borsa bağlamında paylaşılan tweetlerin fikri ve duygusal içeriklerinin ne düzeyde homojen olduğunu göstermektedir. Bu çalışma kapsamında Borsa İstanbul pay endekslerinin

açılış değerleri ile söz birliği indeksi arasında pozitif yönlü zayıf bir korelasyon olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,27$, $p < 0,001$; XU100 [k] için: $r = 0,18$, $p < 0,001$). Bu korelasyon beğeni tweetlerle hesaplanan sözbirliği indeksi (aggreFAVkt, aggreFAVgt) için de geçerlidir (XU100 [m] için: $r = 0,17$, $p < 0,001$; XU100 [k] için: $r = 0,24$, $p < 0,001$). Ancak retweetlerle hesaplanan söz birliği indeksi (aggreRTkt, aggreRTgt) ile endekslerin açılış değerleri arasında bir ilişki bulunamamıştır. Başka bir ifadeyle söz birliği indeksi retweetler ve beğeni tweetler için de hesaplanmış, endekslerin açılış değerlerinin beğeni tweetlerden hesaplanan sözbirliği indeksi ile ilişkili, retweetlerden hesaplanan söz birliği indeksi ile ilişkisiz olduğu sonucuna varılmıştır. Söz birliği indeksi ile endekslerin açılış değeri arasındaki ilişkiyi ortaya koyan bu bulgular Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulguları ile örtüşmektedir. Bu bulgular söz birliği indeksi arttığında endeks açılış değerlerinin artacağı, azaldığında endeks açılış değerlerinin azalacağı anlamına gelmektedir. Söz birliği indeksinin yükselmesi genel anlamda fikir birlikteliğine işaret ederken aynı zamanda sosyal medyanın yankı odaları ile ilgili bir boyuta da sahip olabilir. Bir diğer ifadeyle söz birliği benzer fikirlerin benzer sosyal çevrelerde dolaşmaya başlamasından kaynaklanıyor da olabilir. Böyle bir durumda sürü psikolojisi, hazırda bulunma eğilimi ve (gayri iradi olarak) doğrulama (farklı kaynaklardan gelen benzer fikirler) eğilimlerinin etkisinden söz etmek mümkündür. Ancak yankı odalarının etkisinin tam olarak ortaya çıkartılabilmesi için analizin mikro ve mezo ölçekte, yani kullanıcılar ve sosyal gruplar ölçeğinde yapılması daha somut bulgular elde edilmesini sağlayabilir.

İkili korelasyonlar nötr tweetler ile endekslerin açılış değerleri arasındaki ilişkileri de ortaya koymaktadır. Nötr tweetlerde pozitif ya da negatif bir yönelim olmadığı için, yöneldimsel belirsizliğin bir temsilcisi olarak düşünülmektedir. Belirli bir günde borsa bağlamında paylaşılan nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt) ile endeks açılış değerleri arasında orta düzeyde pozitif bir ilişki vardır (XU100 [m] için: $r = 0,38$, $p < 0,001$; XU100 [k] için: $r = 0,53$, $p < 0,001$). Bu pozitif yönlü ilişki hem nötr retweetler (neutralRTkt, neutralRTgt) için (XU100 [m] için: $r = 0,15$, $p < 0,001$; XU100 [k] için: $r = 0,40$, $p < 0,001$) hem de nötr beğeni tweetler (neutralFAVkt, neutralFAVgt) için (XU100 [m]: $r = 0,32$, $p < 0,001$; XU100 [k] için: $r = 0,32$, $p < 0,001$) geçerlidir. Bu bulgu pozitif tweetlerden

ziyade negatif tweetlerin etkili olduğu bulgusu ile birlikte değerlendirildiğinde, nötr tweetlerin artmasının yatırımcılar açısından genellikle olumlu bir sinyal olarak algılandığına işaret ediyor olabilir.

Tablo 4.25: Korelasyon Tablosu (Açılış Değeri)

		Açılış								
Özmitelik Kodu	Özmitelik Açıklaması	Özmitelik Türü	C	XU030	XU100	XUTUM	XUSIN	XUMAL	XUTEK	XUHZ
bullkt	Yükseliş potansiyeli TW	Bilişsel	k	0.16***	0.17***	0.16***	0.13***	0.19***	0.08**	0.17***
bullgt			m	0.24***	0.25***	0.25***	0.23***	0.2***	0.15***	0.28***
bullRTkt	Yükseliş potansiyeli RT	Bilişsel-Davranışsal-Sosyal	k	0.06*	0.06*	0.06*	0.03	0.09***	0.01	0.04
bullRTgt			m	-0.0	-0.0	-0.0	-0.02	0.02	-0.04	0.0
bullFAVkt	Yükseliş potansiyeli FAV	Bilişsel-Davranışsal-Sosyal	k	0.18***	0.18***	0.18***	0.18***	0.15***	0.14***	0.19***
bullFAVgt			m	0.2***	0.21***	0.21***	0.21***	0.15***	0.15***	0.23***
aggrekt	Söz birliği TW	Bilişsel	k	0.18***	0.18***	0.18***	0.14***	0.2***	0.09***	0.18***
aggregt			m	0.26***	0.27***	0.27***	0.25***	0.22***	0.17***	0.31***
aggreRTkt	Söz birliği RT	Bilişsel-Davranışsal-Sosyal	k	0.03	0.03	0.03	-0.01	0.09***	-0.02	0.01
aggreRTgt			m	-0.0	-0.0	-0.0	-0.01	0.01	-0.02	0.0
aggreFAVkt	Söz birliği FAV	Bilişsel-Davranışsal-Sosyal	k	0.24***	0.24***	0.24***	0.24***	0.19***	0.2***	0.23***
aggreFAVgt			m	0.17***	0.17***	0.17***	0.19***	0.11***	0.16***	0.18***
neutralTWkt	nötr / (p+n) TW	Bilişsel	k	0.52***	0.53***	0.55***	0.66***	0.28***	0.67***	0.57***
neutralTWgt			m	0.38***	0.38***	0.4***	0.48***	0.2***	0.54***	0.37***
neutralRTkt	rt+nötr / (rtp+rt+n) RT	Bilişsel-Davranışsal-Sosyal	k	0.39***	0.4***	0.41***	0.52***	0.16***	0.53***	0.45***
neutralRTgt			m	0.15***	0.15***	0.16***	0.2***	0.05*	0.19***	0.18***
neutralFAVkt	fav+nötr / (rtp+rt+n) FAV	Bilişsel-Davranışsal-Sosyal	k	0.32***	0.32***	0.34***	0.44***	0.12***	0.48***	0.33***
neutralFAVgt			m	0.32***	0.32***	0.34***	0.44***	0.12***	0.48***	0.33***
bkpts_kt	Toplam pozitif / toplam tweet	Bilişsel	k	-0.14***	-0.14***	-0.15***	-0.22***	-0.02	-0.25***	-0.16***
bgnts_gt			m	-0.04	-0.04	-0.05	-0.09***	0.01	-0.16***	-0.01
bknts_kt	Toplam negatif / toplam tweet	Bilişsel	k	-0.32***	-0.33***	-0.33***	-0.32***	-0.28***	-0.26***	-0.34***
bgnts_gt			m	-0.33***	-0.33***	-0.34***	-0.34***	-0.26***	-0.26***	-0.37***
kt_norm	Toplam tweet sayısı	Davranışsal	k	0.56***	0.56***	0.58***	0.68***	0.31***	0.57***	0.63***
gt_norm			m	0.45***	0.45***	0.46***	0.52***	0.25***	0.43***	0.51***
RTkt	retweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	0.54***	0.54***	0.57***	0.66***	0.35***	0.63***	0.56***
RTgt			m	0.7***	0.7***	0.72***	0.82***	0.45***	0.81***	0.7***
FAVkt	Beğeni tweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	0.77***	0.78***	0.79***	0.91***	0.46***	0.86***	0.81***
FAVgt			m	0.72***	0.73***	0.74***	0.84***	0.45***	0.8***	0.75***
bkgRTnorm	Toplam genel RT	Davranışsal-Sosyal	k	0.56***	0.56***	0.58***	0.7***	0.3***	0.63***	0.6***
bggRTnorm			m	0.22***	0.23***	0.23***	0.27***	0.12***	0.24***	0.25***
bkgFAVnorm	Toplam beğeni tweet	Davranışsal-Sosyal	k	0.62***	0.62***	0.64***	0.78***	0.3***	0.7***	0.68***
bggFAVnorm			m	0.59***	0.59***	0.61***	0.74***	0.29***	0.65***	0.66***
FolFrikt	Toplam takipçi sayısı / toplam arkadaş sayısı	Sosyal	k	0.14***	0.15***	0.16***	0.15***	0.19***	0.26***	0.05
FolFrigt			m	0.09***	0.09***	0.1***	0.09***	0.12***	0.13***	0.05*
bkgfols_norm	Takipçi sayısı	Sosyal	k	0.62***	0.63***	0.64***	0.62***	0.54***	0.64***	0.62***
bggfol_norm			m	0.39***	0.4***	0.41***	0.39***	0.34***	0.4***	0.41***
bkgfris_norm	Arkadaş sayısı	Sosyal	k	0.35***	0.35***	0.36***	0.35***	0.28***	0.27***	0.4***
bggfris_norm			m	0.22***	0.22***	0.22***	0.21***	0.18***	0.15***	0.26***
interactionK	Etkileşim skoru	Sosyal	k	0.07**	0.06**	0.07**	0.21***	-0.15***	0.17***	0.1***
interactionG			m	-0.02	-0.02	-0.02	0.12***	-0.22***	0.1***	-0.0

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.001$. k: Anahtar Kelime; m: Makine Öğrenmesi

Bir diğer gösterge pozitif tweetlerin toplam tweetlere oranıdır. Bu çalışma kapsamında özellikle makine öğrenmesi yöntemiyle sınıflandırılan pozitif tweetlerin toplam tweetlere oranı ile endekslerin açılış değerleri arasında anlamlı ve güçlü bir ilişki bulunamamış, ancak negatif tweetlerin pozitif tweetlere oranı (bknts_kt, bgnts_gt) ile açılış değerleri arasında

negatif yönlü zayıf bir korelasyon bulunmuştur (XU100 [m] için: $r = -0,33$, $p < 0,001$; XU100 [k] için: $r = -0,33$, $p < 0,001$). Bu bulgu He ve diğerlerinin (2016), benzer şekilde piyasada olumsuz fikirlerin arttığında endeks değerinde düşüş olacağı bulgusu ile örtüşmektedir.

Bir diğer gösterge, belirli bir günde borsa bağlamında paylaşılan tweetlerin sayısıdır. Belirli bir günde borsa ile ilgili paylaşılan tweet sayısı (kt_norm, gt_norm) ile endeks açılış değerleri arasında orta düzeyde pozitif bir ilişki vardır (XU100 [m] için: $r = 0,45$, $p < 0,001$; XU100 [k] için: $r = 0,56$, $p < 0,001$). Belirli bir günde borsa ile ilgili tweetlerdeki artış ya da azalışlarla endeks değerlerindeki artış ya da azalışların aynı yönde olduğunu göstermektedir. Bu bulgu, belirli bir günde borsa bağlamında paylaşılan tweet sayısının yatırımcı ilgisini temsil ettiği varsayımı altında beklentilerle örtüşmektedir. Benzer bulgular Mao ve diğerleri (2012), Wei ve diğerleri (2016) ve Alostad ve Davulcu (2017) tarafından da raporlanmıştır.

Belirli bir gündeki retweetlerin ya da beğeni tweetlerin aynı gün atılan tweetlerin toplamına oranı bir çeşit etkileşim göstergesi olarak kullanılabilir. Bu nedenle retweetlerin toplam tweetlere oranı (RTktt, RTgtt) ile endekslerin açılış değerleri arasındaki ilişki araştırılmış, pozitif yönde kuvvetli ilişki olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,70$, $p < 0,001$; XU100 [k] için: $r = 0,54$, $p < 0,001$). Beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt) ile endekslerin açılış değerleri arasındaki ilişki de araştırılmış ve benzer şekilde pozitif yönde kuvvetli ilişki olduğu sonucuna varılmıştır (XU100 [m] için: $r = 0,73$, $p < 0,001$; XU100 [k] için: $r = 0,78$, $p < 0,001$). Retweet (bkgRTnorm, bggRTnorm) ve beğeni tweet sayısının (bkgFAVnorm, bggFAVnorm) artması tweetlerin, daha fazla yayıldığı anlamına gelmektedir. Bu durum Twitter’da borsa ile ilgili bilgi arayan yatırımcıların bazı tweetleri tekrar tekrar göreceği anlamına da gelmektedir. Mesajların bu sıklığı, içeriğinden bağımsız bir şekilde, yatırımcıların aşinalık eğilimi nedeniyle bilişsel bir rahatlık hissine kapılması ve piyasalarda örtük pozitif bir etki oluşturmaya neden olabilir.

Takipçi sayısı genellikle bir popülerlik ölçüsü olarak kabul edilmektedir. Zira bir kullanıcının ne kadar çok takipçisi olursa, o kadar çok kişiyi etkilemesi beklendiğinden, takipçi sayıları, etki düzeyinin bir göstergesi niteliğindedir. Belirli bir günde borsa ile ilgili tweet atan

kullanıcıların takipçi sayıları (bkgfols_norm, bggfols_norm) ile Borsa İstanbul pay piyasası endeks açılış değerleri arasında orta düzeyde pozitif bir ilişki vardır (XU100 [m] için: $r = 0,40$, $p < 0,001$; XU100 [k] için: $r = 0,63$, $p < 0,001$). Bir kullanıcının etkili bir kişi (influencer) olabilmesi için takipçi sayısının yanı sıra, takipçilerinin arkadaşlarından çok olması beklenir. Belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayısının arkadaş sayısına oranı (FolFrikt, FolFrigt) ile endeks açılış değerleri arasında pozitif yönlü zayıf bir korelasyon vardır (XU100 [m] için: $r = 0,09$, $p < 0,001$; XU100 [k] için: $r = 0,15$, $p < 0,001$). Başka bir ifadeyle belirli bir günde borsa bağlamında paylaşımlar yapan kullanıcıların takipçi sayıları (yani etki düzeyleri) ile endeks değerleri aynı yönde hareket etmektedir.

Etkileşim skoru belirli bir günde tweet atan kullanıcıların takipçi sayılarını retweet, beğeni tweet ve takipçi sayılarıyla ilişkilendiren bir başka skordur. Belirli bir günde borsa ile ilgili tweet atan kullanıcıların etkileşim skoru (interactionK, interactionG) ile endeks açılış değerleri arasında pozitif yönlü zayıf bir korelasyon vardır (XU100 [k] için: $r = 0,06$, $p < 0,01$). Bu bulgu belirli bir günde borsa ile ilgili tweet atanlar ve diğer kullanıcılar arasındaki etkileşim düzeyi ile BIST Pay Endekslerinin açılış değerleri arasında ilişki olduğunu göstermektedir.

4.5.5.1.2 İşlem Hacmi

Bu çalışma kapsamında Twitter verilerinden elde edilen özniteliklerin Borsa İstanbul pay piyasası endekslerinin işlem hacimleri ile ilişkisini gösteren ikili korelasyonlar Tablo 4.26'da yer almaktadır. Tweetlerden elde edilen özniteliklerden biri yükseliş potansiyeli indeksidir. Bu indeks arttığında işlem hacminin artması beklenir. Bu çalışma kapsamında yükseliş potansiyeli indeksi (bullkt, bullgt) ile işlem hacmi arasında pozitif yönlü zayıf bir korelasyon olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,26$, $p < 0,001$; XU100 [k] için: $r = 0,14$, $p < 0,001$). Beğeni tweetler için hesaplanan yükseliş potansiyeli indeksi (bullFAVkt, bullFAVgt) ile de benzer bir ilişki söz konusudur (XU100 [m] için: $r = 0,22$, $p < 0,001$; XU100 [k] için: $r = 0,14$, $p < 0,001$). Ancak retweetlerden oluşan yükseliş potansiyeli indeksi (bullRTkt, bullRTgt) ile işlem hacmi arasında anlamlı bir korelasyon olmadığı

görülmektedir. Yükseliş potansiyeli ile işlem hacmi arasındaki ilişkiyi ortaya koyan bu bulgular, Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulguları ile örtüşmektedir. Bu bulgu piyasada pozitif beklentiler arttığında, yatırımcıların pişmanlıktan kaçınma eğilimiyle pozisyon almaya ve dolayısıyla daha fazla işlem yapmaya sevk edilmiş olabileceğine işaret etmektedir.

Sözbirliği skoru borsa ile ilgili tweet atan kullanıcıların belirli bir günde ne düzeyde fikir birliği içerisinde olduklarının bir göstergesi olarak kullanılmaktadır. Bu çerçevede sözbirliği arttığında, fikirler daha homojen olacağı için işlem hacminin azalması beklenmektedir. Araştırma kapsamında endekslerin işlem hacimleri ile söz birliği indeksi (aggrekt, aggregt) arasında, beklentinin aksine, pozitif yönlü zayıf bir korelasyon olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,28$, $p < 0,001$; XU100 [k] için: $r = 0,16$, $p < 0,001$). Bu durum söz birliği indeksi arttığında işlem hacminin artacağı, azaldığında işlem hacminin azalacağı anlamına gelmektedir. Bu korelasyon beğeni tweetlerle hesaplanan sözbirliği indeksi (aggreFAVkt, aggreFAVgt) için de geçerlidir (XU100 [m] için: $r = 0,16$, $p < 0,001$; XU100 [k] için: $r = 0,19$, $p < 0,001$). Ancak retweetlerle hesaplanan söz birliği indeksi (aggreRTkt, aggreRTgt) ile işlem hacimleri arasında bir ilişki bulunamamıştır. Genel olarak söz birliği indeksi ile işlem hacmi arasındaki ilişkiyi ortaya koyan bu bulgular Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulguları ile örtüşmemektedir. Ancak burada dikkat edilmesi gereken husus esasında tam bir fikir birliği olup olmaması durumudur. Zira tam bir fikir birliği oluşmadan önce, fikirlerin homojenleşmeye başlamasıyla birlikte, yatırımcılar kayıptan kaçınmak ya da pişmanlıktan kaçınmak ve diğer yatırımcılardan daha hızlı pozisyon alabilmek için daha fazla işlem yapmaya sevk edilmiş olabilir. Bu bağlamda değerlendirildiğinde bulgular beklentilerle örtüşmektedir.

Tablo 4.26: Korelasyon Tablosu (İşlem Hacmi)

İşlem Hacmi										
Öznitelik Kodu	Öznitelik Açıklaması	Öznitelik Türü	C	XU030	XU100	XUTUM	XUSIN	XUMAL	XUTEK	XUHIZ
bullkt	Yükseliş potansiyeli TW	Bilişsel	k	0.13***	0.14***	0.14***	0.1***	0.16***	0.02	0.13***
bullgt			m	0.24***	0.26***	0.25***	0.21***	0.23***	0.13***	0.24***
bullRTkt	Yükseliş potansiyeli RT	Bilişsel-Davranışsal-Sosyal	k	0.01	0.03	0.03	0.02	0.05	-0.04	0.03
bullRTgt			m	0.02	0.02	0.02	-0.01	0.05*	-0.04	0.01
bullFAVkt	Yükseliş potansiyeli FAV	Bilişsel-Davranışsal-Sosyal	k	0.13***	0.14***	0.14***	0.14***	0.09***	0.06*	0.16***
bullFAVgt			m	0.21***	0.22***	0.22***	0.17***	0.18***	0.13***	0.23***
aggrekt	Söz birliği TW	Bilişsel	k	0.14***	0.16***	0.15***	0.12***	0.17***	0.02	0.14***
aggrekt			m	0.26***	0.28***	0.27***	0.23***	0.25***	0.14***	0.26***
aggreRTkt	Söz birliği RT	Bilişsel-Davranışsal-Sosyal	k	-0.02	0.0	0.0	-0.0	0.04	-0.07**	-0.01
aggreRTgt			m	0.02	0.02	0.02	-0.01	0.05*	-0.03	0.01
aggreFAVkt	Söz birliği FAV	Bilişsel-Davranışsal-Sosyal	k	0.18***	0.19***	0.19***	0.19***	0.12***	0.1***	0.21***
aggreFAVgt			m	0.15***	0.16***	0.16***	0.15***	0.1***	0.13***	0.19***
neutralTWkt	nötr / (p+n) TW	Bilişsel	k	0.54***	0.58***	0.59***	0.58***	0.37***	0.5***	0.6***
neutralTWgt			m	0.33***	0.37***	0.4***	0.45***	0.18***	0.45***	0.39***
neutralRTkt	rtnötr / (rt+p+rt+n) RT	Bilişsel-Davranışsal-Sosyal	k	0.46***	0.49***	0.5***	0.47***	0.33***	0.4***	0.48***
neutralRTgt			m	0.19***	0.19***	0.2***	0.18***	0.15***	0.16***	0.18***
neutralFAVkt	favnötr / (rt+p+rt+n) FAV	Bilişsel-Davranışsal-Sosyal	k	0.35***	0.37***	0.39***	0.42***	0.19***	0.38***	0.38***
neutralFAVgt			m	0.35***	0.37***	0.39***	0.42***	0.19***	0.38***	0.38***
bkpts_kt	Toplam pozitif / toplam tweet	Bilişsel	k	-0.17***	-0.18***	-0.19***	-0.2***	-0.08**	-0.21***	-0.2***
bgpts_gt			m	-0.02	-0.03	-0.04	-0.1***	0.04	-0.12***	-0.05
bknts_kt	Toplam negatif / toplam tweet	Bilişsel	k	-0.29***	-0.31***	-0.31***	-0.27***	-0.27***	-0.15***	-0.31***
bgnts_gt			m	-0.32***	-0.35***	-0.35***	-0.31***	-0.29***	-0.22***	-0.33***
kt_norm	Toplam tweet sayısı	Davranışsal	k	0.62***	0.63***	0.63***	0.59***	0.46***	0.48***	0.62***
gt_norm			m	0.46***	0.46***	0.46***	0.46***	0.31***	0.36***	0.45***
RTkt	retweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	0.56***	0.6***	0.59***	0.52***	0.44***	0.42***	0.66***
RTgt			m	0.59***	0.66***	0.68***	0.72***	0.37***	0.63***	0.71***
FAVkt	Beğeni tweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	0.68***	0.74***	0.76***	0.81***	0.42***	0.73***	0.76***
FAVgt			m	0.65***	0.71***	0.73***	0.74***	0.44***	0.67***	0.73***
bkgrTnorm	Toplam genel RT	Davranışsal-Sosyal	k	0.62***	0.64***	0.64***	0.61***	0.44***	0.54***	0.64***
bggrTnorm			m	0.2***	0.21***	0.22***	0.24***	0.12***	0.18***	0.22***
bkgrFAVnorm	Toplam beğeni tweet	Davranışsal-Sosyal	k	0.64***	0.67***	0.68***	0.71***	0.39***	0.66***	0.65***
bggrFAVnorm			m	0.62***	0.64***	0.65***	0.67***	0.39***	0.62***	0.62***
FolFrikt	Toplam takipçi sayısı / toplam arkadaş sayısı	Sosyal	k	0.03	0.08**	0.09***	0.1***	0.02	0.06**	0.17***
FolFrikt			m	0.01	0.04	0.05*	0.06**	0.01	0.02	0.08**
bkgrfols_norm	Takipçi sayısı	Sosyal	k	0.38***	0.45***	0.48***	0.52***	0.26***	0.38***	0.52***
bggrfols_norm			m	0.25***	0.3***	0.32***	0.34***	0.18***	0.25***	0.33***
bkgrfris_norm	Arkadaş sayısı	Sosyal	k	0.3***	0.33***	0.33***	0.31***	0.27***	0.26***	0.31***
bggrfris_norm			m	0.21***	0.22***	0.22***	0.19***	0.2***	0.15***	0.19***
interactionK	Etkileşim skoru	Sosyal	k	0.24***	0.21***	0.21***	0.23***	0.12***	0.32***	0.16***
interactionG			m	0.13***	0.11***	0.12***	0.16***	0.03	0.29***	0.07**

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.001$. k: Anahtar Kelime; m: Makine Öğrenmesi

İkili korelasyonlar nötr tweetler ile endekslerin işlem hacimleri arasındaki ilişkileri de ortaya koymaktadır. Nötr tweetlerde pozitif ya da negatif bir yönelim olmadığı için, yönelimsel belirsizliğin bir temsilcisi olarak kabul edilmektedir. Bu çerçevede belirli bir gündeki borsa bağlamında paylaşılan nötr tweetlerin toplamının toplam tweetlere oranı (neutralTWkt, neutralTWgt) ile işlem hacimleri arasında orta düzeyde pozitif bir ilişki vardır (XU100 [m] için: $r = 0,37$, $p < 0,001$; XU100 [k] için: $r = 0,58$, $p < 0,001$). Bu pozitif yönlü ilişki hem nötr retweetler (neutralRTkt, neutralRTgt) için (XU100 [m] için: $r = 0,19$, $p < 0,001$; XU100 [k]

için: $r = 0,49$, $p < 0,001$) hem de nötr beğeni tweetler (neutralFAVkt, neutralFAVgt) için (XU100 [m]: $r = 0,37$, $p < 0,001$; XU100 [k] için: $r = 0,37$, $p < 0,001$) geçerlidir. Bu bulgular yönelimsel belirsizlik arttığında işlem hacminin arttığını göstermektedir. Nötr tweetlerin toplam tweetlere oranı ile işlem hacimleri arasındaki pozitif yönlü bu ilişki yatırımcıları, pozisyonlarını gözden geçirmeye sevk etmesiyle ilgili olduğu düşünülmektedir. Bunun altında yatan neden, yatırımcıların belirsizlik sonrasında geleceğe dair olası pozitif ya da negatif beklentisine göre kayıptan kaçınma, belirsizlikten kaçınma ya da pişmanlıktan kaçınma eğilimleri nedeniyle işlem yapması olabilir.

Tweetlerin fikri ve duygusal yönelimiyle ilgili bir başka gösterge pozitif tweetlerin toplam tweetlere oranıdır. Bu çalışma kapsamında özellikle makine öğrenmesi yöntemiyle sınıflandırılan pozitif tweetlerin toplam tweetlere oranı (bgpts_gt) ile endekslerin işlem hacmi arasında anlamlı ve güçlü bir ilişki bulunamamıştır. Ancak negatif tweetlerin pozitif tweetlere oranı (bknts_kt, bgnts_gt) ile işlem hacimleri arasında negatif yönlü zayıf bir korelasyon bulunmuştur (XU100 [m] için: $r = -0,35$, $p < 0,001$; XU100 [k] için: $r = -0,31$, $p < 0,001$). Bu bulgu piyasada olumsuz fikirlerin yaygınlaştığında işlem hacminde düşüş olacağı beklentisi ile örtüşmektedir.

Araştırmaya dâhil edilen bir diğer gösterge borsa bağlamında paylaşılan toplam tweet sayısıdır. Borsa bağlamında paylaşılan tweet sayılarındaki değişim ile işlem hacmindeki değişiminin aynı yönde olması beklenir. Zira belirli bir günde borsa bağlamında paylaşılan tweet sayısı, yatırımcıların borsaya ilgisinin bir göstergesi olarak düşünülebilir. Dolayısıyla kullanıcıların belirli bir günde borsa bağlamında daha fazla tweet paylaşımları, tweetlerin içeriğinden bağımsız bir şekilde işlem hacmini etkileme potansiyeline sahiptir. Bu çalışma kapsamında belirli bir günde borsa ile ilgili paylaşılan tweet sayısı (kt_norm, gt_norm) ile endeks işlem hacimleri arasında orta düzeyde pozitif bir ilişki olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,46$, $p < 0,001$; XU100 [k] için: $r = 0,63$, $p < 0,001$). Belirli bir günde borsa ile ilgili tweetlerdeki artış ya da azalışlarla endeks işlem hacimlerinde artış ya da azalışların aynı yönde olduğunu göstermektedir. Benzer sonuçlar Mao ve diğerleri (2012),

Garcia-Lopez ve diğçerleri (2018) ve Sprenger ve diğçerleri (2014) tarafından da raporlanmıřtır.

Belirli bir günde ki retweetlerin ya da beğeni tweetlerin aynı gün atılan tweetlerin toplamına oranı bir çeřit etkileřim göstergesi olarak kullanılabilir. Bu nedenle retweetlerin toplam tweetlere oranı (RTktt, RTggt) ile endekslerin iřlem hacimleri arasındaki iliřki arařtırılmıř, pozitif yönde kuvvetli iliřki olduđu bulgusuna ulařılmıřtır (XU100 [m] için: $r = 0,66$, $p < 0,001$; XU100 [k] için: $r = 0,60$, $p < 0,001$). Beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVggt) ile endekslerin iřlem hacimleri arasındaki iliřki de arařtırılmıř ve benzer řekilde pozitif yönde kuvvetli iliřki olduđu sonucuna varılmıřtır (XU100 [m] için: $r = 0,71$, $p < 0,001$; XU100 [k] için: $r = 0,74$, $p < 0,001$). Bu durum kullanıcılar arasındaki sosyal etkileřim arttıkça iřlem hacminin de artacađı, etkileřimin azalmasının iřlem hacmini azaltacađı anlamına gelmektedir.

Analizde bir popülerlik ölçüsü olarak belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayılarının iřlem hacmi ile iliřkisi de arařtırılmıřtır. Belirli bir günde borsa ile ilgili tweet atanların takipçi sayılarının fazla olması, etki düzeyinin bir göstergesi olarak düşünülebilir. Bu çalıřma kapsamında, belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayıları (bkgfols_norm, bggfols_norm) ile iřlem hacimleri arasında orta düzeyde pozitif bir iliřki olduđu bulgusuna ulařılmıřtır (XU100 [m] için: $r = 0,30$, $p < 0,001$; XU100 [k] için: $r = 0,45$, $p < 0,001$). Bir kullanıcının etkili bir kiři olabilmesi için takipçi sayısının yanı sıra, kendi takipçilerinin arkadaşlarından çok olması muhtemeldir. Belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayısının arkadaş sayısına oranı (FolFrikt, FolFrigt) ile iřlem hacimleri arasında pozitif yönlü zayıf bir korelasyon vardır (XU100 [k] için: $r = 0,08$, $p < 0,001$). Ayrıca belirli bir günde tweet atan kullanıcıların takipçi sayılarını retweet ve beğeni tweet sayılarıyla iliřkilendiren etkileřim skoru (interactionK, interactionG) ile iřlem hacimleri arasında pozitif yönlü zayıf bir korelasyon vardır (XU100 [m] için: $r = 0,11$, $p < 0,001$; XU100 [k] için: $r = 0,21$, $p < 0,001$). Bu bulgu belirli bir günde borsa ile ilgili tweet atan kullanıcılar arasındaki etkileřim skorundaki artış ya da azalıřların iřlem hacimlerini aynı yönde etkilediđi anlamına gelmektedir. Genel olarak takipçi sayısı

kaynaklı ve etkileşime dayalı bu bulgular takipçi sayısının fazla olmasının, retweet ve beğeni şeklinde etkileşim almasının belirli bir fikrin yayılmasında sosyal faktörlerin etkin bir rol oynadığının göstergesidir. Aynı zamanda bu durum, takipçi rolüyle bireysel yatırımcıların bilgi arayışında olduğunu, fikirlerini paylaşan kullanıcılardan gelecek etkiye açık olduklarının bir işareti de olabilir. Ayrıca işlem hacmi ile ilişkisi açısından retweet ve beğeni tweetlerin sayısının takipçi sayısından daha yüksek korelasyona sahip olması, etkileşimin daha iyi bir gösterge olduğuna işaret etmektedir. Nitekim bu bulgu Zhang ve diğerlerinin (2011), retweet sayının takipçi sayısından daha iyi bir dayanak olduğu bulgusuyla da örtüşmektedir.

4.5.5.1.3 Volatilite

Twitter verilerinden elde edilen öznitelikler ile volatilite ilişkisi de araştırılmış, korelasyon tablosu Tablo 4.27’de sunulmuştur. Yükseliş potansiyeli indeksi (bullkt, bullgt) ile volatilite arasında negatif yönlü önemsizlenecek kadar düşük bir korelasyon olduğu bulgusuna ulaşılmıştır (XU100 [k] için: $r = -0,08$, $p < 0,05$). Retweetler için hesaplanan yükseliş potansiyeli indeksi (bullRTkt, bullRTgt) ile de benzer bir ilişki söz konusudur (XU100 [m] için: $r = -0,06$, $p < 0,1$; XU100 [k] için: $r = -0,06$, $p < 0,1$). Ancak beğeni tweetlerden oluşan yükseliş potansiyeli indeksi (bullFAVkt, bullFAVgt) ile volatilite arasında bir korelasyon olmadığı görülmektedir. Yükseliş potansiyeli ile endekslerin volatilitesi arasındaki ilişkiyi ortaya koyan bu bulgular Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulguları ile örtüşmektedir.

Volatilite ile ilişkisi araştırılan bir başka gösterge söz birliği indeksidir. Bu çerçevede sözbirliği arttığında, piyasa hakkındaki fikirler daha homojen hale geleceği için volatilitenin azalması beklenir. Endekslerin volatilitesi ile söz birliği indeksi (aggrekt, aggregt) arasında negatif yönlü çok düşük düzeyde ilişki vardır (XU100 [k] için: $r = -0,08$, $p < 0,001$). Bu durum söz birliği indeksi arttığında volatilitenin azalacağı, söz birliği indeksi azaldığında artacağı anlamına gelmektedir. Bu korelasyon retweetlerle hesaplanan sözbirliği indeksi (aggreRTkt, aggreRTgt) için de geçerlidir (XU100 [m] için: $r = -0,05$, $p < 0,1$; XU100 [k] için: $r = -0,11$,

$p < 0,001$). Ancak beğeni tweetlerle hesaplanan söz birliği indeksi (aggrelFAVkt, aggreFAVgt) ile endekslerin volatilitesi arasında bir ilişki bulunamamıştır. Söz birliği indeksi ile volatilité arasındaki ilişkiyi ortaya koyan bu bulgular Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulguları ile örtüşmektedir.

Tablo 4.27: Korelasyon Tablosu (Volatilité)

Volatilité										
Öznitelik Kodu	Öznitelik Açıklaması	Öznitelik Türü	C	XU030	XU100	XUTUM	XUSIN	XUMAL	XUTEK	XUHIZ
bullkt	Yükseliş potansiyeli TW	Bilişsel	k	-0.08**	-0.08**	-0.08**	-0.06*	-0.08***	-0.07**	-0.09***
bullgt			m	-0.03	-0.03	-0.03	-0.01	-0.03	-0.04	-0.01
bullRTkt	Yükseliş potansiyeli RT	Bilişsel-Davranışsal-Sosyal	k	-0.06*	-0.06*	-0.06*	-0.06*	-0.05*	-0.08**	-0.08***
bullRTgt			m	-0.06*	-0.06*	-0.06*	-0.06*	-0.06*	-0.07**	-0.07**
bullFAVkt	Yükseliş potansiyeli FAV	Bilişsel-Davranışsal-Sosyal	k	0.01	0.01	0.01	0.03	0.02	-0.01	0.0
bullFAVgt			m	-0.01	-0.01	-0.01	0.02	-0.01	-0.02	0.01
aggrekt	Söz birliği TW	Bilişsel	k	-0.08***	-0.08***	-0.08***	-0.06**	-0.08***	-0.08**	-0.09***
aggrekt			m	-0.03	-0.03	-0.03	-0.01	-0.04	-0.04	-0.01
aggreRTkt	Söz birliği RT	Bilişsel-Davranışsal-Sosyal	k	-0.1***	-0.11***	-0.1***	-0.11***	-0.09***	-0.12***	-0.14***
aggreRTgt			m	-0.05	-0.05*	-0.05*	-0.05*	-0.05*	-0.06*	-0.05*
aggrelFAVkt	Söz birliği FAV	Bilişsel-Davranışsal-Sosyal	k	-0.03	-0.02	-0.02	0.01	-0.02	-0.03	-0.02
aggreFAVgt			m	-0.03	-0.02	-0.02	0.01	-0.03	-0.01	0.0
neutralTWkt	nötr / (p+n) TW	Bilişsel	k	0.12***	0.15***	0.15***	0.26***	0.1***	0.2***	0.21***
neutralTWgt			m	0.04	0.06*	0.06*	0.15***	0.03	0.2***	0.12***
neutralRTkt	rt nötr / (rt+p+rt)n RT	Bilişsel-Davranışsal-Sosyal	k	0.18***	0.2***	0.2***	0.28***	0.17***	0.22***	0.25***
neutralRTgt			m	0.06**	0.07**	0.07**	0.09***	0.07**	0.08***	0.1***
neutralFAVkt	fav nötr / (rt+p+rt)n FAV	Bilişsel-Davranışsal-Sosyal	k	0.17***	0.19***	0.18***	0.25***	0.16***	0.21***	0.23***
neutralFAVgt			m	0.17***	0.19***	0.18***	0.25***	0.16***	0.21***	0.23***
bkpts_kt	Toplam pozitif / toplam tweet	Bilişsel	k	-0.12***	-0.13***	-0.13***	-0.17***	-0.11***	-0.15***	-0.16***
bgpts_gt			m	-0.05*	-0.06*	-0.06*	-0.09***	-0.05*	-0.13***	-0.08**
bknts_kt	Toplam negatif / toplam tweet	Bilişsel	k	0.05	0.04	0.04	-0.01	0.05*	0.02	0.03
bgnts_gt			m	0.02	0.01	0.02	-0.02	0.02	0.0	-0.01
kt_norm	Toplam tweet sayısı	Davranışsal	k	0.19***	0.22***	0.21***	0.32***	0.16***	0.2***	0.26***
gt_norm			m	0.19***	0.21***	0.2***	0.28***	0.18***	0.17***	0.23***
RTkt	retweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	-0.03	0.01	0.0	0.15***	-0.07**	0.04	0.07**
RTgt			m	0.02	0.05*	0.05*	0.21***	-0.01	0.18***	0.12**
FAVkt	Beğeni tweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	k	0.05*	0.08**	0.07**	0.22***	0.04	0.21***	0.16***
FAVgt			m	0.02	0.05	0.04	0.18***	-0.0	0.16***	0.12***
bkgRTnorm	Toplam genel RT	Davranışsal-Sosyal	k	0.14***	0.17***	0.17***	0.3***	0.11***	0.18***	0.23***
bggRTnorm			m	0.08**	0.09***	0.09***	0.13***	0.08**	0.09***	0.1***
bkgFAVnorm	Toplam beğeni tweet	Davranışsal-Sosyal	k	0.17***	0.2***	0.19***	0.31***	0.16***	0.24***	0.26***
bggFAVnorm			m	0.18***	0.2***	0.19***	0.3***	0.16***	0.24***	0.26***
FolFrikt	Toplam takipçi sayısı / toplam arkadaş sayısı	Sosyal	k	-0.19***	-0.17***	-0.17***	-0.1***	-0.2***	-0.08***	-0.16***
FolFrikt			m	-0.08**	-0.07**	-0.07**	-0.04	-0.1***	-0.04	-0.06*
bkgfols_norm	Takipçi sayısı	Sosyal	k	-0.1***	-0.08**	-0.08**	0.03	-0.11***	0.07**	-0.03
bggfol_norm			m	-0.04	-0.03	-0.03	0.04	-0.06**	0.06*	0.01
bkgfris_norm	Arkadaş sayısı	Sosyal	k	0.03	0.03	0.03	0.07**	0.02	0.07**	0.05*
bggfris_norm			m	0.02	0.03	0.03	0.04	0.02	0.03	0.03
interactionK	Etkileşim skoru	Sosyal	k	0.18***	0.18***	0.17***	0.17***	0.21***	0.2***	0.2***
interactionG			m	0.12***	0.12***	0.11***	0.11***	0.15***	0.17***	0.15***

* $p < 0.1$. ** $p < 0.05$. *** $p < 0.001$. k: Anahtar Kelime; m: Makine Öğrenmesi

İkili korelasyonlar nötr tweetler ile endekslerin volatiliteleri arasındaki ilişkileri de ortaya koymaktadır. Belirli bir gündeki borsa bağlamında atılan nötr tweetlerin toplamının toplam

tweetlere oranı (neutralTWkt, neutralTWgt) ile volatilité arasında orta düzeyde pozitif bir ilişki vardır (XU100 [m] için: $r = 0,06$, $p < 0,1$; XU100 [k] için: $r = 0,15$, $p < 0,001$). Bu pozitif yönlü ilişki hem nötr retweetler (neutralRTkt, neutralRTgt) için (XU100 [m] için: $r = 0,07$, $p < 0,001$; XU100 [k] için: $r = 0,20$, $p < 0,001$) hem de nötr beğeni tweetler (neutralFAVkt, neutralFAVgt) için (XU100 [m]: $r = 0,19$, $p < 0,001$; XU100 [k] için: $r = 0,19$, $p < 0,001$) geçerlidir. Nötr tweetler piyasadaki belirsizliğin bir temsilcisi olarak düşünülürse, belirsizliğin artmasının volatilitéyi artırması beklenen bir durumdur. Başka bir ifadeyle belirsizlik durumunda farklı fikirler ve beklentiler ortaya çıkmakta, fikirlerin değişkenliği fiyatların değişkenliğine yansımaktadır. Bu bulgu, böyle durumlarda yatırımcıların belirsizlikten kaçınma, kayıptan kaçınma ya da pişmanlıktan kaçınma eğilimlerinin etkisiyle işlem yaptığı anlamına gelebilir.

Tweetlerin fikri ve duygusal yönelimiyle ilgili bir başka gösterge pozitif tweetlerin toplam tweetlere oranıdır. Bu çalışma kapsamında pozitif tweetlerin toplam tweetlere oranı (bkpts_kt, bgpts_gt) ile endekslerin volatilitesi arasında negatif yönlü zayıf bir ilişki bulunmuştur (XU100 [m] için: $r = -0,06$, $p < 0,1$; XU100 [k] için: $r = -0,13$, $p < 0,001$). Pozitif tweetler arttığında volatilitenin azalacağı anlamına gelmektedir. Ancak negatif tweetlerin pozitif tweetlere oranı (bknts_kt, bgnts_gt) ile volatilité arasında anlamlı bir ilişki bulunamamıştır.

Kullanıcıların belirli bir günde borsa bağlamında paylaştığı tweetlerin artması tweetlerin içeriğinden bağımsız bir şekilde volatilitéyi etkileme potansiyeline sahiptir. Bu nedenle belirli bir günde borsa ile ilgili paylaşılan tweet sayısı (kt_norm, gt_norm) ile volatilité arasında ilişki olup olmadığı araştırılmış, zayıf düzeyde pozitif bir ilişki olduğu bulgusuna ulaşılmıştır (XU100 [m] için: $r = 0,21$, $p < 0,001$; XU100 [k] için: $r = 0,22$, $p < 0,001$). Benzer bulgular Wei ve diğerleri (2016), Alostad ve Davulcu (2017), Garcia-Lopez ve diğerleri (2018) tarafından da raporlanmıştır. Belirli bir günde borsa ile ilgili tweet sayısında artış ya da azalış olduğunda volatilitenin de aynı yönde artış ya da azalış gösterdiği anlamına gelmektedir. Bir kullanıcının borsa bağlamında tweet atmasının temelinde, metnin içeriğinden bağımsız bir şekilde duygu ve düşüncesini dışa vurma isteği vardır. Bu bağlamda

tweet sayısı ile volatilité arasındaki pozitif yönlü bu bulgu, daha çok tweet paylaşılmasının fikri ve duygusal anlamda daha çok çeşitlilik anlamına gelebileceği, fikri ve duygusal çeşitliliğin artmasının ise volatilitéyi artırması beklenen bir durumdur.

Belirli bir gündeki tweetlerin ne kadar retweetlendiği ya da beğenildiği bir çeşit etkileşim göstergesi olarak kullanılabilir. Belirli bir günde borsa bağlamında atılan tweetlerle ilgili etkileşimin artmasının volatilitéyi artırması beklenir. Bu nedenle retweetlerin toplam tweetlere oranı (RTktt, RTggt) ile volatilité arasındaki ilişki araştırılmış, pozitif yönde zayıf ilişki olduğu bulgusuna ulaşılmıştır (XUSIN [m] için: $r = 0,21$, $p < 0,001$; XUSIN [k] için: $r = 0,15$, $p < 0,001$). Beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVggt) ile endekslerin volatilité arasındaki ilişki de araştırılmış ve benzer şekilde pozitif yönde zayıf ilişki olduğu sonucuna varılmıştır (XUSIN [m] için: $r = 0,18$, $p < 0,001$; XUSIN [k] için: $r = 0,22$, $p < 0,001$). Bu bulgu sosyal etkileşim düzeyi ile volatilité arasında pozitif yönlü bir ilişki olduğu anlamına gelmektedir.

Bir popülerlik ölçüsü olarak kullanıcıların takipçi sayısı arttıkça, daha çok kişiyi etkilemesi beklenir. Bu çalışma kapsamında belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayıları (bkgfols_norm, bggfols_norm) ile volatilité arasında ilişki olup olmadığı araştırılmış, neticede makine öğrenmesi yaklaşımı ile sınıflandırılan tweetler için anlamlı bir ilişki bulunamamıştır. Bu bulgu paylaşılan mesajların kullanıcılara ulaşmış olmasından ziyade, paylaşımların ne kadar etkileşim aldığı, yani retweet edildiğinin ve beğenildiğinin yatırımcı davranışlarını tahmin etmede daha önemli bir gösterge olduğuna işaret etmektedir. Bu bulgu ayrıca Zhang ve diğerlerinin (2011), retweet sayının takipçi sayısından daha iyi bir dayanak olduğu bulgusuyla da örtüşmektedir. Diğer taraftan anahtar kelime yaklaşımı ile sınıflandırılan tweetlere odaklanıldığında, belirli bir günde borsa ile ilgili tweet atan kullanıcıların takipçi sayıları ile volatilité arasında ilişki olup olmadığı araştırılmış negatif ve önemsiz düzeyde zayıf bir ilişki olduğu bulgusuna ulaşılmıştır (XU100 [k] için: $r = -0,08$, $p < 0,05$). Bu bulgunun geçerli olması halinde, belirli bir fikirden daha çok kişinin etkilenmesinin volatilitéyi azaltacağı şeklinde yorumlanabilir. Başka bir ifadeyle daha çok

takipçi sayısı daha çok kişinin etkilenmesi anlamına geleceğinden, takipçi sayısının artması söz birliğinin artmasına ve dolayısıyla volatilitenin azalmasına yol açıyor olabilir.

Belirli bir günde tweet atan kullanıcıların takipçi sayılarını retweet ve beğeni tweet sayılarıyla ilişkilendiren etkileşim skoru (interactionK, interactionG) ile volatilité arasında pozitif yönlü zayıf bir korelasyon vardır (XU100 [m] için: $r = 0,12$, $p < 0,001$; XU100 [k] için: $r = 0,18$, $p < 0,001$). Bu bulgu belirli bir günde borsa ile ilgili tweet atan kullanıcılar arasındaki etkileşim skorundaki artış ya da azalışların volatilitéyi aynı yönde etkilediği anlamına gelmektedir.

4.5.5.1.4 Getiri

Twitter verilerinden elde edilen öznitelikler ile endeks getirileri arasında ilişki olup olmadığı da araştırılmış, korelasyon sonuçları Tablo 4.28’de sunulmuştur. Bu çalışma kapsamında tweetlerden elde edilen öznitelikler ile piyasa getirileri arasında anlamlı bir ilişki bulunamamıştır. Bu çalışma kapsamında belirli bir günde borsa bağlamında paylaşılan tweetlerden elde edilen fikir ve duygu skorları ile getiri arasında ilişki bulunamamasının nedeninin, analizin hisse senedi düzeyinde değil de, endeks düzeyinde yapılmış olması ile ilgili olabileceği düşünülmektedir. Zira yatırımların hisse senedi bazında olması nedeniyle, bazı hisse senetlerinin fiyatları yükselirken bazı hisse senetlerinin fiyatları düşebildiği için, kullanılan öznitelikler ile endeks getirisi arasında (orta/yüksek pozitif/negatif herhangi) bir ilişki bulunamamış olabilir.

Tweetlerden elde edilen öznitelikler ile getiri arasındaki ilişkiyi araştıran çalışmalarda elde edilen bulgular da bu tespiti destekler niteliktedir. Örneğin Sprenger ve diğerleri (2014) gibi bazı çalışmalar getiri ile anlamlı fakat çok zayıf bir ilişki olduğunu raporlamışlardır. Ranco ve diğerleri (2015) ise duygu kutupluluğu endeksi ile getiri arasında anlamlı bir ilişki bulunamamış, ancak olay çalışması yöntemiyle analiz ettiklerinde getiri ilişkisi olduğu sonucuna ulaşmışlardır. Diğer taraftan Liew ve Budavari (2017), Xun ve Guo (2017), Sul ve

diğerleri (2017), Tabari ve diğerleri (2018) ve Rakowski ve diğerleri (2021) gibi hisse senedi düzeyinde yapılan çalışmalarda getiri ilişkisi olduğu bulgusuna ulaşılmıştır.

Tablo 4.28: Korelasyon Tablosu (Getiri)

Getiri										
Öznitelik Kodu	Öznitelik Açıklaması	Öznitelik Türü	C	XU030	XU100	XUTUM	XUSIN	XUMAL	XUTEK	XUHIZ
bullkt	Yükseliş potansiyeli TW	Bilişsel	m	0.01	0.01	0.01	0.01	0.01	0.04	0.03
bullgt			k	-0.01	-0.01	-0.01	-0.02	-0.0	0.0	-0.0
bullRTkt	Yükseliş potansiyeli RT	Bilişsel-Davranışsal-Sosyal	m	0.04	0.04	0.04	0.01	0.04	0.03	0.05*
bullRTgt			k	-0.0	-0.0	-0.0	-0.01	0.0	-0.01	-0.01
bullFAVkt	Yükseliş potansiyeli FAV	Bilişsel-Davranışsal-Sosyal	m	-0.0	0.0	0.0	0.01	-0.0	0.01	-0.0
bullFAVgt			k	-0.02	-0.02	-0.02	-0.01	-0.02	-0.01	-0.01
aggrekt	Söz birliği TW	Bilişsel	m	0.02	0.02	0.02	0.01	0.01	0.04	0.04
aggrekt			k	-0.01	-0.01	-0.01	-0.02	-0.0	0.01	-0.0
aggreRTkt	Söz birliği RT	Bilişsel-Davranışsal-Sosyal	m	0.03	0.03	0.03	0.02	0.04	0.03	0.04
aggreRTgt			k	-0.0	-0.0	-0.0	-0.01	0.01	0.0	-0.01
aggreFAVkt	Söz birliği FAV	Bilişsel-Davranışsal-Sosyal	m	-0.01	-0.01	-0.01	0.01	-0.02	-0.01	-0.0
aggreFAVgt			k	-0.03	-0.03	-0.03	-0.02	-0.03	-0.01	-0.01
neutralTWkt	nötr / (p+n) TW	Bilişsel	m	-0.01	-0.01	-0.01	-0.01	-0.02	-0.02	-0.02
neutralTWgt			k	-0.02	-0.02	-0.02	-0.0	-0.02	-0.02	-0.03
neutralRTkt	rtnötr / (rtp+rtm) RT	Bilişsel-Davranışsal-Sosyal	m	-0.02	-0.02	-0.02	-0.0	-0.02	-0.02	-0.03
neutralRTgt			k	-0.01	-0.01	-0.01	-0.0	-0.01	-0.01	-0.02
neutralFAVkt	favnötr / (rtp+rtm) FAV	Bilişsel-Davranışsal-Sosyal	m	-0.01	-0.01	-0.02	-0.01	-0.02	-0.03	-0.02
neutralFAVgt			k	-0.01	-0.01	-0.02	-0.01	-0.02	-0.03	-0.02
bkpts_kt	Toplam pozitif / toplam tweet	Bilişsel	m	0.02	0.02	0.02	0.01	0.02	0.05	0.04
bkpts_gt			k	0.01	0.01	0.01	-0.01	0.01	0.02	0.01
bknts_kt	Toplam negatif / toplam tweet	Bilişsel	m	-0.01	-0.01	-0.01	-0.0	-0.0	-0.03	-0.03
bknts_gt			k	0.02	0.02	0.02	0.03	0.01	0.0	0.01
kt_norm	Toplam tweet sayısı	Davranışsal	m	-0.05*	-0.06*	-0.06*	-0.04	-0.06*	-0.06*	-0.06*
gt_norm			k	-0.07**	-0.08**	-0.08**	-0.07**	-0.08**	-0.06**	-0.09***
RTkt	retweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	m	-0.02	-0.02	-0.02	-0.0	-0.03	-0.01	-0.02
RTgt			k	-0.03	-0.03	-0.03	-0.01	-0.04	-0.0	-0.03
FAVkt	Beğeni tweetlerin toplam tweetlere oranı	Davranışsal-Sosyal	m	-0.03	-0.03	-0.03	-0.01	-0.04	-0.0	-0.03
FAVgt			k	-0.02	-0.02	-0.02	-0.0	-0.03	-0.0	-0.01
bkgrThorm	Toplam genel RT	Davranışsal-Sosyal	m	-0.01	-0.02	-0.01	-0.0	-0.02	-0.03	-0.02
bggrThorm			k	-0.06*	-0.06*	-0.07**	-0.05*	-0.06**	-0.05*	-0.07**
bkgrFAVnorm	Toplam beğeni tweet	Davranışsal-Sosyal	m	-0.04	-0.05	-0.05	-0.03	-0.05*	-0.03	-0.05
bggrFAVnorm			k	-0.03	-0.04	-0.04	-0.03	-0.04	-0.03	-0.04
FolFrikt	Toplam takipçi sayısı / toplam arkadaş sayısı	Sosyal	m	0.04	0.04	0.04	0.03	0.04	0.03	0.03
FolFrikt			k	0.01	0.01	0.01	0.01	0.01	0.02	0.01
bkgrFols_norm	Takipçi sayısı	Sosyal	m	-0.0	0.0	0.0	0.0	-0.0	0.01	0.01
bggrFols_norm			k	-0.02	-0.02	-0.02	-0.01	-0.02	-0.01	-0.02
bkgrFris_norm	Arkadaş sayısı	Sosyal	m	-0.04	-0.04	-0.04	-0.03	-0.04	-0.01	-0.03
bggrFris_norm			k	-0.04	-0.03	-0.03	-0.02	-0.03	-0.01	-0.04
interactionK	Etkileşim skoru	Sosyal	m	0.02	0.01	0.01	0.01	0.02	0.01	-0.0
interactionG			k	0.02	0.01	0.01	0.01	0.01	0.01	-0.01
*p < 0.1. **p < 0.05. ***p < 0.001. k: Anahtar Kelime; m: Makine Öğrenmesi										

*p < 0.1. **p < 0.05. ***p < 0.001. k: Anahtar Kelime; m: Makine Öğrenmesi

4.5.5.2 Öznitelik Seçimi

Özellikle büyük veri analitiğinde karşılaşılan yüksek boyutlu veriler, sınıflandırıcıların ya da tahmincilerin performansını olumsuz yönde etkileyebilmektedir (Shi, 2022). Başka bir

deyişle bir taraftan hızla artan yüksek boyutlu veriler hem depolama hem de hesaplama maliyetlerini beraberinde getirirken, diğer taraftan sınıflandırıcıların ya da tahmincilerin gereksiz öznitelikler kullanılarak karmaşıklığının artmasına ve genelleme kabiliyetinin düşmesine neden olabilmektedir (Zheng ve Casari, 2018). Makine öğrenmesinde boyut laneti [curse of dimensionality] olarak da ifade edilen bu durum hem gözlem sayılarının, hem de her bir birimle ilgili özniteliklerin sayısının çok fazla olmasından kaynaklanmaktadır (Çağlayan Akay, 2020). Bu bağlamda öznitelik seçimi [feature selection], ilgisiz [irrelevant], tahmin modeline katkısı olmayan [gereksiz: redundant] ve gürültülü öznitelikleri dışarıda tutarak ilgili, gerekli ve gürültüsüz özniteliklerin küçük bir alt kümesinin seçilmesi işlemidir. Başka bir deyişle öznitelik seçimi, büyük bir veri kümesindeki gereksiz ve ilgisiz değişkenlerin modelden çıkartılırken, veri setini iyi bir şekilde temsil edebilen değişkenlerin tahmininde kullanılmak üzere modelde kalmasını sağlayan boyut azaltma tekniklerini ifade etmektedir (Wang ve diğerleri, 2016). Dolayısıyla başarılı bir öznitelik seçimi, genellikle daha iyi öğrenme performansı, daha düşük hesaplama maliyeti ve daha iyi model yorumlanabilirliği sağlamaktadır (Kuhn ve Johnson, 2020).

Makine öğreniminde kullanılan öznitelik seçme yöntemleri genel olarak üç yaklaşıma dayanmaktadır. Bunlardan biri filtreleme yöntemleridir. Filtreleme yöntemleri [filter methods], herhangi bir makine öğrenimi algoritması kullanmadan verinin karakteristiğine göre seçim yapan ve genellikle ön işleme aşamasında başvuru yaklaşımıdır. Bu yaklaşımların ilk adımında öznitelikler belirli kriterlere göre sıralanmakta, ikinci adımında ise en yüksek sıralamaya sahip olanlar seçilmektedir. Hesaplama açısından hızlı ve maliyetsizdirler. Bu yaklaşıma dayalı tekniklerde öznitelik seçimi her bir özellik ve ilgili değişken arasındaki ilişkiye göre ayrı ayrı değerlendirildiğinden, çoklu bağlantı [multicollinearity] problemine karşı bağışık değillerdir. Filtreleme yaklaşımının bir başka dezavantajı ise, seçilen öznitelik altkümesinin tahmin veya sınıflandırma algoritmasının performansı üzerindeki etkilerini tamamen göz ardı etmesidir. Filtreleme yöntemlerinde, özniteliklerin farklı özelliklerini ölçen çeşitli sıralama kriterleri vardır. Bunların başlıcaları arasında bilgi kazanımı [information gain], ki-kare testi [chi-square test], Fisher skoru

[Fisher's score], korelasyon katsayısı [Pearson's correlation coefficient] ve varyans eşiği [variance threshold] teknikleri sayılabilir (Cherrington ve diğerleri, 2019).

Öznitelik seçiminde başvuru bir diğer yaklaşım kombinleme yöntemleridir. Kombinleme yöntemleri [wrapper models] olası tüm öznitelik altkümelerini, tahmine dayalı bir makine öğrenimi modeli ile değerlendirmektedir. Açgözlü algoritmalar olarak da adlandırılan bu yaklaşım teknikleri özyinelemeli bir şekilde öznitelik altkümelerini kullanarak algoritmayı eğitmekte ve eğitimden elde edilen sonuçlara dayanarak özniteliklerin eklenmesi ya da çıkartılmasına karar verilmektedir. Ancak n değişkenden oluşan bir veri kümesinin olası alt küme sayısı 2^n olacağından, pek çok durumda tüm olası öznitelik kombinasyonlarını eğitmek yerine, en uygun altküme seçmeye yönelik bazı kriterler belirlenmekte (örneğin modelin performansının düşmesi, belirli sayıda özelliğe ulaşılması ya da tepe tırmanışı [hill-climbing], en iyi ilk [best-first], dal-sınır [branch-and-bound] ve genetik algoritmalar gibi) ve bu kriterler sağlandığında seçim işlemi sonlandırılmaktadır (Wang ve diğerleri, 2016). Kombinleme yöntemlerinin başlıca avantajı, tahmin modelinin eğitiminde en uygun öznitelikler setini oluşturarak filtreleme yöntemlerine göre daha başarılı sonuçlar sağlamaları, dezavantajı ise hesaplama açısından maliyetli olmalarıdır. Ardışık ileri doğru seçim, ardışık geriye doğru eleme, çift yönlü eleme ve özyinelemeli seçim ise kombinleme yaklaşımında başvuru başlıca öznitelik seçimi teknikleridir (Liu, 2011).

Öznitelik seçimindeki bir diğer yaklaşım gömülü yöntemlerdir. Gömülü yöntemler [embedded methods] öznitelik seçimini öğrenme algoritmasının bir parçası olarak öğrenme sürecine dâhil etmektedirler (Liu, 2011). Gömülü modeller, öznitelik seçimini model yapısına yerleştirirken hem filtreleme modellerinden hem de kombinleme modellerinden yararlanmakta ve bu modellerin avantajlarını birleştirmektedirler. Zira gömülü yöntemler, filtreleme yöntemleri gibi daha hızlı sonuçlar ortaya koyarken kombinleme yöntemleri gibi öznitelik kombinasyonlarını da dikkate almaktadırlar (Wang ve diğerleri, 2016). Başlıca teknikleri arasında LASSO (ℓ_1 regularization) ve RIDGE (ℓ_2 regularization) regresyon, elastik ağlar [elastic nets] (ℓ_1 ve ℓ_2 regularization), rassal orman [random forest] ve gradyan artırma [gradient boosting] sayılabilir.

Bu çalışma kapsamında bütünleşik tahmin modellerinde farklı öznitelik seçim tekniklerine dayanan makine öğrenmesi yöntemleri kullanılmış, böylece arka planda bu tekniklerin kıyaslanması da amaçlanmıştır.

4.5.5.3 Makine Öğrenmesi Yöntemleri ile Bütünleşik Karar Analizi

Bu aşamada Twitter verilerinden elde edilen öznitelikler ile Borsa İstanbul pay endekslerinin açılış değerleri, işlem hacimleri, getirileri ve volatilitelerinin tahmin edilip edilemeyeceği araştırılmaktadır. Analizde makine öğrenmesi yöntemlerinden lineer regresyon, lasso regresyon, rassal orman ve XGBoost yöntemleri kullanılmaktadır. Farklı makine öğrenmesi yöntemleri kullanmak suretiyle hem bazı yöntemlerin dezavantajlarını ortadan kaldırmak, hem de farklı yöntemlerin performanslarını karşılaştırmak amaçlanmaktadır. Ayrıca borsayla ilgili tweetlerin sınıflandırılması için kullanılan anahtar kelime (k) tabanlı yaklaşım ile makine öğrenmesi (m) temelli yaklaşım arasında performans açısından önemli bir fark olup olmadığı da araştırılmaktadır.

4.5.5.3.1 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini

Analizin ilk aşamasında, tweetlerden elde edilen özniteliklerin BIST Pay Endekslerinin açılış değerlerinin tahmininde kullanılıp kullanılamayacağı araştırılmıştır. Tablo 4.29’da, tweetlerden elde edilen öznitelikler ile BIST Pay Endekslerinin açılış değerlerinin tahmin sonuçları yer almaktadır. Tabloda yer alan performans ölçüm metriklerinden biri olan R^2 , bağımlı değişkendeki değişimin ne kadarının model tarafından açıklanabileceğini ölçmektedir. Korelasyon katsayısının karesidir. R^2 değeri, 0 ile 1 arasında değerler alır ve daha büyük bir değer, tahmin ile gerçek değer arasında daha iyi bir uyum olduğunu göstermektedir. Ancak tek başına R^2 değeri aşırı uyum sorununu dikkate almamaktadır. Bu bağlamda bir tahmin fonksiyonunun parametrelerini öğrenmek ve bunu aynı veriler üzerinde test etmek metodolojik bir hata olacağından eğitim ve test verilerinin birbirinden farklı olması gerekmektedir. Böylece modelin, eğitildiği veri setinden farklı bir veri seti ile test

edilerek benzer bir performans gösterip gösteremediği, yani aşırı uyum sorunu olup olmadığı ortaya konulabilmektedir. Özellikle modelde çok sayıda açıklayıcı değişkenin olduğu durumlarda eğitim verilerinde çok iyi sonuçlar alınabilirken test verilerinde kötü bir performans ile karşılaşılabilir. Bu durum literatürde aşırı uyum sorunu olarak bilinmektedir. Bu çalışma kapsamında verilerin bir kısmı eğitim için (%67) bir kısmı test için (%33) ayrılmış, aşırı uyum problemi olup olmadığını görebilmek amacıyla tabloda eğitim ve test sonuçları birlikte verilmiştir.

Makine öğrenmesinde tahmin modelinin performansını artırmak için hiperparametre ayarları yapılarak test için ayrılan veri setlerinde çok daha iyi performans sağlanabilmektedir. Bununla birlikte, parametrelerin test kümesine göre optimize edilmesi modelin, hem eğitim hem test verilerinde başarıyla hiç karşılaşmadığı veriler üzerinde kötü performans göstermesine neden olabilmektedir. Başka bir ifadeyle hiperparametre ayarı, test veri setinde aşırı uyum riskine yol açabilmektedir. Bu sorunu aşmak için başvurulan yöntemlerden biri, veri setinin eğitim ve test verilerinden farklı başka bir parçasının doğrulama seti olarak ayrılmasıdır. Bu yaklaşımda eğitim setinde eğitim yapıldıktan sonra performans değerlendirmesi doğrulama setinde yapılmaktadır. Burada belirli bir başarıya ulaşıldığında son değerlendirme test seti üzerinde gerçekleştirilmektedir. Ancak mevcut verilerin eğitim, doğrulama ve test şeklinde üç kümeye ayrılması, modeli öğrenmek için kullanılabilecek örnek sayısını büyük ölçüde azaltmaktadır. Bu sorunu aşmak için çapraz doğrulama [cross validation] adı verilen bir yöntem ortaya konulmuştur. K katmanlı çapraz doğrulama yaklaşımında eğitim seti k küçük kümeye bölünmekte, eğitim verisi olarak katmanların k-1'i kullanılarak model eğitilmekte ve test verisi ile test edilmektedir. K-katmanlı çapraz doğrulama sonucunda ortaya konulan performans ölçüsü, döngüde hesaplanan değerlerin ortalamasıdır (Pedregosa ve diğerleri, 2011). Bu çalışma kapsamında kullanılan makine öğrenmesi yöntemlerinde 10 katmanlı doğrulama yapılmış, elde edilen performans sonuçları (cvs) ve bu sonuçların standart sapması (std) Tablo 4.29'da raporlanmıştır.

Analiz sonuçlarının değerlendirilmesinde kullanılan bir diğer metrik, Eşitlik 4.14'te yer alan ortalama karesel hatadır. R^2 , modelin bağımlı değişkene ne kadar iyi uyduğunu gösterirken

ortalama karesel hata [mean square error - MSE], tahmin sonuçlarının gerçek sayıdan ne kadar saptığına dair mutlak bir sayı vermektedir. Burada y_i i'inci gözlemlenen değer, $\hat{y}_i$, y_i için karşılık gelen tahmin değeri ve n gözlem sayısıdır.

$$MSE = \frac{\sum (y_i - \hat{y}_i)^2}{n} \quad (4.14)$$

Bu metriğin karekökü alınarak hesaplanan bir diğer metrik ise Eşitlik 4.15'te yer alan, hataların ortalama kareköküdür. Hataların ortalama karekökü [root mean square error - RMSE] çıktı değişkeni ile aynı birim olduğu ve yorumlamayı kolaylaştırdığı için ortalama karesel hata metriğine göre daha sık kullanılmaktadır.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4.15)$$

Bir diğer performans metriği, Eşitlik 4.16'da yer alan ortalama mutlak hatadır. Ortalama mutlak hata [mean absolute error - MAE] MSE'de hatanın karelerinin toplamı yerine hatanın mutlak değerinin toplamını almaktadır. Çıktı değişkeni ile aynı birimdir ve aykırı değerlere karşı dirençlidir. Öngörülen değerler ile gerçek değerler arasındaki ortalama mutlak fark olarak ölçülür ve bir regresyon modelinin etkinliğini değerlendirmek için kullanılmaktadır. Genel olarak ifade etmek gerekirse R^2 modelin uyumunu açıklamak için tercih edilirken, RMSE ve MAE farklı regresyon modelleri arasındaki performans karşılaştırmak için kullanılmaktadır.

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (4.16)$$

Makine öğrenmesi yöntemlerinin performansını gösteren bir diğer metrik ise Eşitlik 4.17’de yer alan ortalama mutlak yüzde hata [mean absolute percentage error - MAPE] değeridir. MAPE istatistiği, özellikle farklı değere sahip modellerin karşılaştırılmasında kullanılmaktadır. Diğer kriterlerden üstünlüğü öngörü hatalarını yüzde olarak ifade etmesidir. MAPE değeri %10’un altında olan tahmin modelleri *yüksek doğruluk* derecesine sahip modeller olarak kabul edilirken MAPE değeri %10 - %20 arasında olan modeller *doğru* tahmin modeli olarak kabul edilmektedir. Başka bir ifadeyle MAPE < %10 ise *çok iyi*, %10 < MAPE < %20 ise *iyi*, %20 < MAPE < %50 ise *kabul edilebilir*, %50 < MAPE ise *hatalı* model olarak sınıflandırılmaktadır.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100 \quad (4.17)$$

Bu nedenle Tablo 4.29’da analizler sonucunda elde edilen RMSE ve MAE ve MAPE değerleri de raporlanmıştır.

Tablo 4.29: BIST Pay Endekslerinin Açılış Değerleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları

Y	P	C	Endeks Açılış													
			XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Lineer Regresyon	cvs (R ²)	k	0,71	0,73	0,71	0,73	0,74	0,75	0,88	0,89	0,86	0,86	0,80	0,82	0,36	0,30
		m	0,67	0,69	0,68	0,70	0,71	0,73	0,86	0,88	0,84	0,86	0,71	0,76	0,25	0,24
	std	k	0,06	0,09	0,06	0,09	0,05	0,09	0,03	0,07	0,03	0,08	0,06	0,05	0,09	0,16
		m	0,05	0,09	0,05	0,09	0,04	0,08	0,02	0,03	0,02	0,03	0,05	0,05	0,09	0,12
	RMSE	k	70,95	71,47	70,95	71,47	69,85	70,19	69,71	68,97	121,31	115,26	43,72	43,02	114,59	120,68
		m	96,98	99,95	76,88	79,12	75,61	77,72	79,26	83,23	130,94	129,15	53,57	51,67	125,31	126,45
	MAE	k	55,92	55,87	55,92	55,87	55,00	54,60	52,98	51,67	84,91	80,62	34,32	33,15	92,87	98,59
		m	76,77	78,22	60,48	61,73	59,11	60,26	58,97	60,15	94,43	90,97	41,24	38,90	101,15	102,10
	MAPE	k	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,06	0,18	0,17	0,06	0,06	0,09	0,09
		m	0,08	0,08	0,07	0,07	0,07	0,07	0,07	0,07	0,19	0,18	0,07	0,07	0,09	0,09
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
Lasso Regresyon	cvs (R ²)	k	0,74	0,75	0,75	0,76	0,78	0,78	0,91	0,91	0,88	0,90	0,84	0,83	0,49	0,49
		m	0,68	0,70	0,69	0,71	0,72	0,73	0,86	0,88	0,84	0,87	0,70	0,75	0,42	0,42
	std	k	0,05	0,07	0,05	0,08	0,05	0,07	0,03	0,04	0,03	0,04	0,05	0,05	0,07	0,13
		m	0,07	0,10	0,07	0,10	0,06	0,09	0,05	0,04	0,05	0,06	0,08	0,07	0,10	0,13
	RMSE	k	83,09	85,00	65,81	67,61	64,40	66,04	63,96	66,87	108,76	106,45	38,81	41,47	98,35	100,62
		m	91,98	99,46	72,68	78,77	71,31	77,22	73,58	79,82	122,21	125,98	50,32	51,19	104,23	111,17
	MAE	k	65,82	66,61	52,14	53,16	50,99	51,71	48,92	49,64	76,92	74,26	30,62	31,92	77,55	78,64
		m	73,43	77,82	57,72	61,36	56,31	59,73	55,38	58,49	87,84	86,43	39,05	38,74	82,86	86,72
	MAPE	k	0,07	0,07	0,07	0,07	0,06	0,06	0,06	0,06	0,16	0,16	0,06	0,06	0,07	0,07
		m	0,07	0,08	0,07	0,07	0,07	0,07	0,07	0,07	0,18	0,16	0,07	0,07	0,08	0,08
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
Rassal Orman	cvs (R ²)	k	0,92	0,91	0,92	0,91	0,93	0,92	0,96	0,97	0,96	0,95	0,93	0,92	0,85	0,81
		m	0,91	0,89	0,91	0,90	0,92	0,91	0,96	0,96	0,96	0,95	0,91	0,90	0,82	0,79
	std	k	0,02	0,02	0,02	0,03	0,02	0,02	0,01	0,01	0,01	0,01	0,03	0,02	0,03	0,05
		m	0,02	0,03	0,02	0,03	0,02	0,03	0,01	0,02	0,01	0,01	0,03	0,04	0,04	0,06
	RMSE	k	30,58	48,28	23,47	37,40	22,25	35,97	21,71	38,54	30,81	63,67	16,57	27,19	34,00	53,98
		m	31,05	51,35	24,73	39,42	23,50	38,82	22,67	37,84	33,53	65,21	18,38	28,39	38,08	58,77
	MAE	k	23,37	36,82	17,87	28,51	16,91	27,45	15,27	26,75	18,97	38,23	12,43	20,48	26,44	41,46
		m	23,86	39,14	19,05	30,16	18,02	29,48	17,10	27,69	22,10	39,50	13,96	21,82	29,27	44,52
	MAPE	k	0,02	0,04	0,02	0,04	0,02	0,03	0,02	0,03	0,04	0,07	0,02	0,04	0,02	0,04
		m	0,02	0,04	0,02	0,04	0,02	0,04	0,02	0,03	0,04	0,07	0,02	0,04	0,03	0,04
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
XGBoost	cvs (R ²)	k	0,90	0,90	0,91	0,91	0,92	0,91	0,95	0,95	0,96	0,96	0,92	0,92	0,82	0,80
		m	0,89	0,88	0,89	0,88	0,90	0,89	0,96	0,95	0,95	0,94	0,90	0,88	0,78	0,76
	std	k	0,03	0,02	0,02	0,03	0,02	0,03	0,02	0,03	0,01	0,02	0,03	0,02	0,04	0,06
		m	0,03	0,03	0,03	0,03	0,03	0,03	0,01	0,03	0,02	0,01	0,03	0,03	0,05	0,08
	RMSE	k	13,71	52,01	10,21	40,11	10,32	39,38	9,38	41,47	11,72	69,74	6,55	28,21	16,33	60,26
		m	12,66	55,01	10,60	42,32	9,88	42,52	9,55	42,37	11,38	74,42	7,29	31,15	15,18	65,22
	MAE	k	10,53	39,83	7,84	30,08	7,93	30,04	7,16	29,08	8,67	42,33	5,04	21,46	12,52	45,73
		m	9,66	42,26	7,96	32,65	7,57	32,88	7,17	30,62	8,50	46,99	5,58	23,74	11,51	49,73
	MAPE	k	0,01	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,02	0,07	0,01	0,04	0,01	0,04
		m	0,01	0,04	0,01	0,04	0,01	0,04	0,01	0,04	0,02	0,08	0,01	0,04	0,01	0,05

Y: Yöntem; P: Performans, C: Bağlam; k: Anahtar Kelime; m: Makine Öğrenmesi; cvs: Çapraz Doğrulama Skoru; std: Standart Sapma

4.5.5.3.1.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini

Bu çalışma kapsamında tweetlerden elde edilen özniteliklerin BIST Pay Endekslerinin açılış değerini tahmin etmek için başvuru yöntemlerinden biri lineer regresyon yöntemidir. Lineer regresyon yönteminde öznitelikler seçilirken korelasyona dayalı filtreleme yöntemine başvurulmuştur. Twitter verilerinden elde edilen değişkenler ile BIST Pay Endeksleri

arasındaki ilişkiyi gösteren Pearson korelasyon katsayıları hesaplanmış ve sonuçlar Tablo 4.25, Tablo 4.26, Tablo 4.27 ve Tablo 4.28’de sunulmuştur.

Lineer regresyon yönteminde korelasyon sonuçları filtreleme yaklaşımı benimsenerek modele dahil olacak özniteliklerin seçiminde kullanılmıştır. Bu nedenle literatürdeki ağırlıklı yaklaşıma uygun olarak (Cunningham ve diğerleri, 2021) bağımlı değişken ile mutlak %50’nin üzerinde korelasyona sahip olan öznitelikler modele dâhil edilmiştir. Ancak filtrelemede tercih edilen korelasyon değeri için bağlayıcı bir kural bulunmamaktadır. Ayrıca lineer regresyon yöntemi çoklu bağlantı sorununa karşı bağışık değildir. Bunu kısmen azaltabilmek için filtreleme yöntemi, açıklayıcı değişkenlerin kendi aralarında yüksek korelasyona sahip olanları elemek suretiyle de yapılabilmektedir. Ancak bu şekilde bir filtreleme yaklaşımı da nihayetinde, belirli bir eşiğin altındaki korelasyon ilişkisine sahip öznitelikleri modelde tutacağı için çoklu bağlantı sorununu tümüyle ortadan kaldıramamaktadır. Ayrıca bu çalışma kapsamında kullanılan diğer makine öğrenmesi yöntemleri gömülü yöntemlerle öznitelik seçimi yaptığı için lineer regresyon modellerinde bu tekniğe başvurulmamıştır.

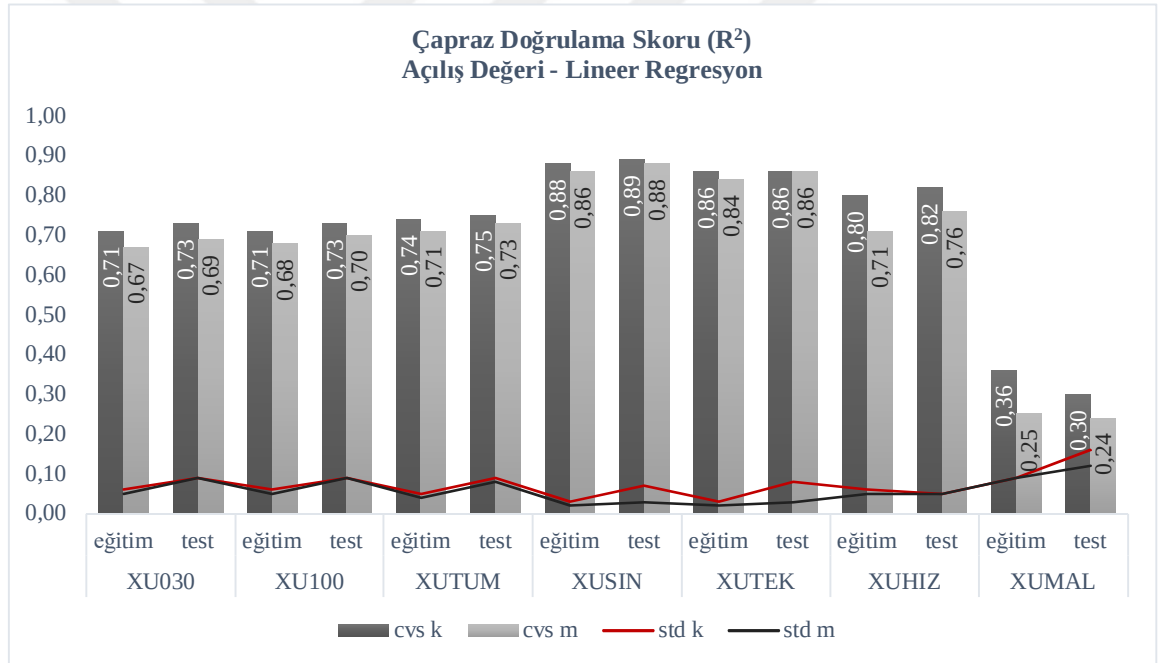
Tablo 4.29’da yer alan ve lineer regresyon analizinde elde edilen bulgular yer almakta, lineer regresyon modelinin çapraz doğrulama, standart sapma, eğitim ve test sonuçları Grafik 4.11’de sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 ve XU100 endeksindeki değişimin %73’ünü, XUTUM endeksindeki değişimin %75’ini, XUSIN endeksindeki değişimin %89’unu, XUTEK endeksindeki değişimin %86’nı, XUHIZ endeksindeki değişimin %82’ini ve XUMAL endeksindeki değişimin %30’unu açıklayabildiğini göstermektedir.

Bu bulgular Rao ve Srivastava’nın (2012) Twitter’dan gelen işlem sinyalleri kullanılarak piyasanın yönünün, Urolagin’nin (2017) tweetlerin duygusal yönelimi ile hisse senedinin açılış ve kapanış fiyatlarındaki değişimin tahmin edilebileceği bulgularıyla örtüşmektedir. Bu bulgular aynı zamanda Tiwari ve diğerleri’nin (2017) tweetlerden elde edilen duygu

skorları ile hisse senedi açılış fiyatlarının tahmininde elde ettikleri başarı ile benzerlik göstermektedir.

Lineer regresyon analizden elde edilen dikkat çekici bulgulardan bir diğeri tweetlerden elde edilen özniteliklerin XUSIN endeksi için en yüksek, XUMAL endeksi için en düşük açıklayıcılık değerine sahip olmasıdır. Bu farklılığın XUMAL endeksinin analiz döneminde diğer endekslerden negatif ayrışmasından (İlhan ve Bayir, 2021) kaynaklandığı düşünülmektedir.

Grafik 4.11: Lineer Regresyon Çapraz Doğrulama Skoru (Açılış Değeri)



Grafik 4.11’de görüldüğü üzere, eğitim ve test sonuçlarının birbirlerine yakın olması, modelin aşırı uyum probleminin olmadığı anlamına gelmektedir. Grafikte 4.11’de ayrıca eğitim ve test verilerinin 10 katmanlı doğrulamalarının standart sapmaları da yer almaktadır. Bu standart sapmaların düşük olması modelin başarısını, yüksek olması başarısızlığını ortaya koyan bir başka metriktir. XUMAL endeksinin tahmininde diğer endekslerin tahmin başarısına kıyasla elde edilen düşük performans bu metriklerle de yansımaktadır. Nitekim

hem eğitim ve test verileri arasındaki farkın diğerlerine kıyasla daha fazla olması hem de standart sapmanın yükselmesi esasında bu durumun bir göstergesidir.

Lineer regresyon modelinin tahmin başarısı MAPE istatistiği ile ölçülmüştür. Bu sonuçlara göre XUTEK endeksi dışındaki endeksler için MAPE değeri %10'un altında olduğu için model *yüksek doğruluk derecesine* sahiptir. XUTEK endeksi ise MAPE değeri %10 - %20 arasında olduğu için *doğru* tahmin modeli olarak kabul edilmektedir. XUTEK endeksi için her ne kadar MAPE değeri, diğer endekslerin MAPE değerinden farklı olsa da makul sınırlar içerisinde. Ayrıca modelin açıklayıcılığı da XUTEK endeksi için oldukça yüksektir.

Son olarak lineer regresyon analizden elde edilen bir diğer bulgu, anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koymuş olmalarıdır.

4.5.5.3.1.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini

Genellikle lasso regresyonu olarak adlandırılan en küçük mutlak daralma ve seçme operatörü regresyonu, doğrusal regresyonun düzenlenleştirilmiş bir versiyonudur. Eşitlik 4.18'de görüldüğü üzere, lasso regresyonda maliyet fonksiyonuna ℓ_1 düzenlenleştirme terimi eklenmektedir (Géron, 2021).

$$J(\theta) = MSE(\theta) + \alpha \sum_{i=1}^n |\theta_i| \quad (4.18)$$

ℓ_1 düzenlenleştirmesinde maliyet, ağırlıkların katsayılarının mutlak değerine oransal olarak eklenirken, ℓ_2 düzenlenleştirmesinde maliyet, ağırlıkların katsayılarının karelerine oransal olarak eklenmektedir. ℓ_2 düzenlenleştirmesi ile karşılaştırıldığında ℓ_1 düzenlenleştirmesi daha seyrek (yani bazı parametrelerin en iyi değerinin sıfır olması) bir sonuç verdiğinden, öznelik seçimi için yaygın olarak kullanılmaktadır. ℓ_1 cezası ağırlıkların bir alt kümesinin sıfır

olmasına neden olarak bu ağırlıklara karşılık gelen özniteliklerin atılabileceği anlamına gelmektedir (Goodfellow ve diğerleri, 2018). Lasso regresyon yöntemi bu özelliği sayesinde ilgisiz, yani modelin açıklayıcılığına katkısı olmayan değişkenlerin elenerek modelin yorumlanabilirliğini artırmaya yardımcı olmakta ve aşırı uyum problemini de azaltmaktadır (Çağlayan Akay, 2020).

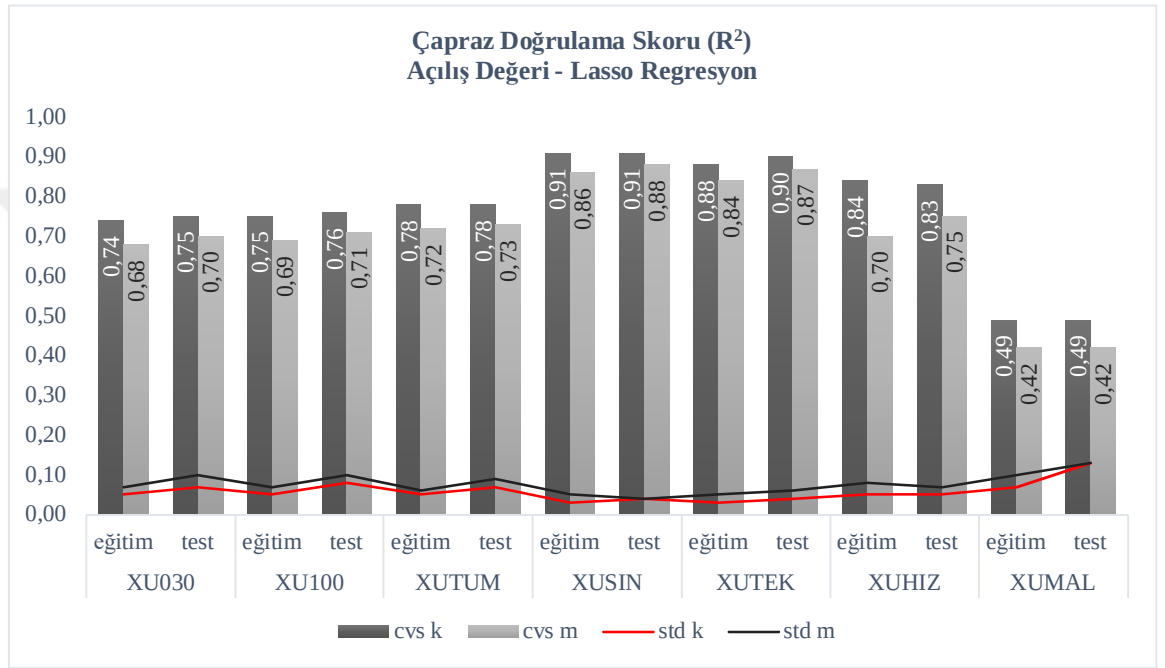
Lasso regresyon modellerinde sapma ve varyans arasındaki dengeyi kontrol etmek amacıyla kullanılan daraltma [shrinkage] veya ceza [penalty] parametresi (α) temelde veri değerinin ortalama gibi merkezi bir noktaya doğru daralma miktarının ne kadar olduğunu göstermektedir. Sapma ve varyans dengesinin belirlenmesinde en uygun parametre değerinin seçimi oldukça önemlidir. Zira bu parametrenin $\alpha = 0$ olarak seçilmesi durumunda, hiçbir öznitelik elenmezken yüksek seçilmesi durumunda daha az özniteliğe sahip bir model ortaya çıkmaktadır. Başka bir deyişle α 'nın küçük değerleri kompleks modelleri seçme eğilimindeyken α 'nın büyük değerleri basit modelleri seçme eğilimindedir (Çağlayan Akay, 2020).

Bu çalışma kapsamında, yukarıda sayılan özellikleri nedeniyle, öznitelik seçiminde başvurulacak yöntemlerden biri de gömülü yöntemler arasında sayılan Lasso regresyon yöntemi olmuştur.

Tablo 4.29'da yer alan lasso regresyon modelinin çapraz doğrulama sonuçları Grafik 4.12'te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksindeki değişimin %75'ini, XU100 endeksindeki değişimin %76'sını, XUTUM endeksindeki değişimin %78'ini, XUSIN endeksindeki değişimin %91'ini, XUTEK endeksindeki değişimin %90'ını, XUHIZ endeksindeki değişimin %83'ünü XUMAL endeksindeki değişimin %49'unu açıklayabildiğini göstermektedir. Lasso regresyon analizinden elde edilen bulgular da lineer regresyon analizinden elde edilen bulgulara benzer şekilde tweetlerden elde edilen özniteliklerin XUSIN endeksi için en yüksek açıklayıcılık değerine, XUMAL endeksi için en düşük açıklayıcılık değerine sahip olmasıdır. Lasso regresyon yönteminde bütün endeksler için açıklayıcılık lineer regresyona kayısla 3-5 puan

daha yüksektir. Bunun temel nedeninin lineer regresyon ve lasso regresyon yöntemlerinin öznitelik seçimindeki yaklaşım farklılığı olduğu düşünülmektedir. Tahmin başarısındaki iyileşme XUMAL endeksi için diğerlerine kıyasla daha fazla olmuştur.

Grafik 4.12: Lasso Regresyon Çapraz Doğrulama Skoru (Açılış Değeri)



Grafik 4.12’te lasso regresyon yönteminde aşırı uyum problemi olmadığı ve standart sapmaların düşük olduğu görülmektedir.

Lasso regresyon yönteminin performansına ilişkin MAPE değerleri Tablo 4.29’da yer almaktadır. Bu sonuçlara göre, lineer regresyon modelinde olduğu gibi XUTEK endeksi dışındaki endeksler için MAPE değeri %10’un altında olduğu için model *yüksek doğruluk derecesine* sahiptir. XUTEK endeksinin MAPE değeri ise, lineer regresyon modelinde elde edilen bulgulara benzer şekilde %10 - %20 arasındadır. Bu nedenle *doğru* tahmin modeli olarak kabul edilmektedir. XUTEK endeksi için modelin açıklayıcılığı, lineer regresyon yönteminden küçük bir farkla da olsa daha başarılıdır.

Lasso regresyon yönteminde de XUTEK endeksi için MAPE değeri, diğer endekslerin MAPE değerinden yüksek olsa da makul sınırlar içerisinde. Lineer regresyon yönteminden elde edilen bulgulara benzer şekilde modelin açıklayıcılığı XUTEK endeksi için oldukça yüksektir.

Lasso regresyon yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koyduğu görülmektedir.

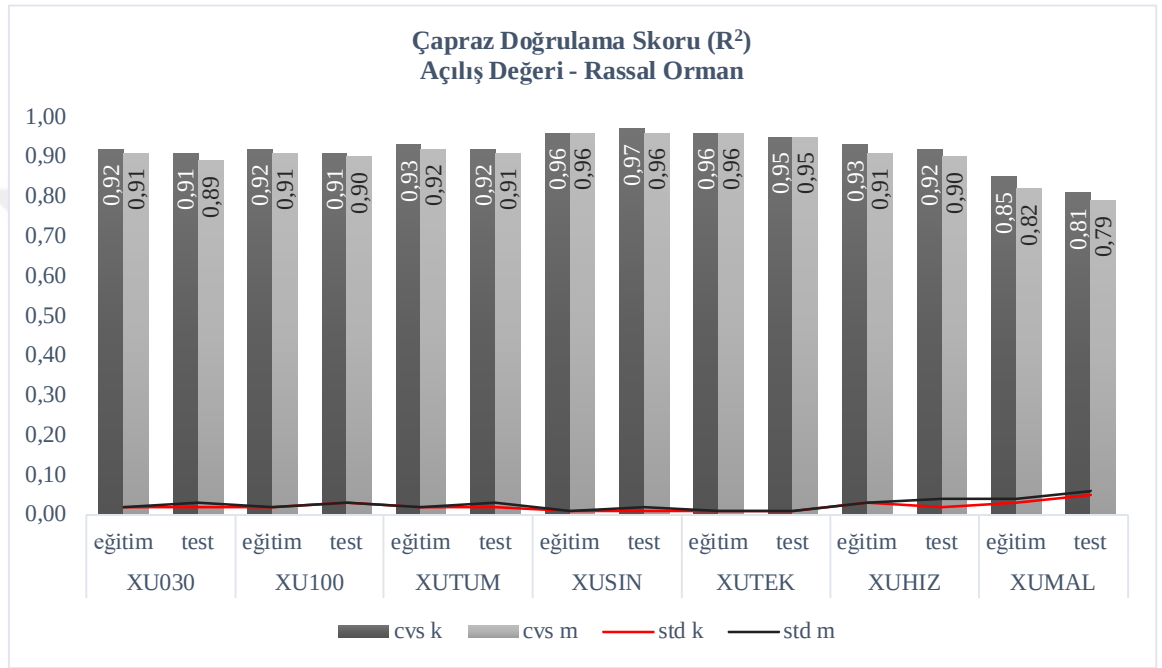
4.5.5.3.1.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini

Rassal orman birden çok karar ağacından oluşan bir topluluk öğrenmesi modelidir. Başka bir ifadeyle topluluk öğrenme yöntemini kullanan rassal orman regresyonu, tek bir modelden daha doğru bir tahmin yapmak için birden çok makine öğrenimi algoritmasından gelen tahminleri birleştiren bir tekniktir. Rassal orman yöntemi “torbalama” [bagging] yaparak, varyansı ve aşırı uyum sorununu en aza indirmektedir.

Tablo 4.29’da yer alan rassal orman modelinin çapraz doğrulama sonuçları Grafik 4.13’te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksindeki değişimin %91’ini, XU100 endeksindeki değişimin %91’ini, XUTUM endeksindeki değişimin %92’sini, XUSIN endeksindeki değişimin %97’sini, XUTEK endeksindeki değişimin %95’ini, XUHIZ endeksindeki değişimin %92’sini XUMAL endeksindeki değişimin %81’ini açıklayabildiğini göstermektedir. Rassal orman modelinde elde edilen dikkat çekici bulgu, tweetlerden elde edilen özniteliklerin XUSIN endeksi için analiz kapsamında ulaşılmış en yüksek açıklayıcılık değeri olmasıdır. Diğer bir deyişle tweetlerden elde edilen öznitelikler, rassal orman yöntemi kullanıldığında, XUSIN endeksinin açılış değerlerinin %97’sini açıklayabilmektedir. Rassal orman modelinde de XUMAL, diğer endekslere kıyasla en düşük açıklayıcılık değerine sahiptir. Ancak XUMAL endeksi için rassal orman yönteminde ulaşılan açıklayıcılık değeri, lineer ve lasso regresyon yöntemlerinde elde edilen açıklayıcılık değerlerinden oldukça yüksektir. Başka bir ifadeyle

lineer regresyonda bu değer %30, lasso regreyonda %49 iken rassal ormanda %81'e çıkmıştır. Bu başarının rassal orman yönteminin doğrusal olmayan (nonlinear) ve topluluk öğrenmesine dayalı yapısından kaynaklandığı düşünülmektedir.

Grafik 4.13: Rassal Orman Çapraz Doğrulama Skoru (Açılış Değeri)



Grafik 4.13'te, görüldüğü üzere, eğitim ve test sonuçlarının birbirine yakın olması rassal orman modelinde de aşırı uyum probleminin olmadığını göstermektedir.

Rassal orman modelinin tahmin performansını gösteren MAPE istatistiği Tablo 4.29'da yer almaktadır. Bu sonuçlara göre, rassal orman modelinde tüm endeksler için MAPE değeri %10'un altında olduğundan model bütün endeks değerleri için *yüksek doğruluk derecesine* sahiptir.

Rassal orman regresyonunda XUTEK endeksi için MAPE değeri, lineer ve lasso regresyonda elde edilen MAPE değerlerinden farklı olarak %10'un altında gerçekleşmiştir. Bu durum

rassal orman yönteminin lineer ve lasso regresyon yöntemlerine kıyasla daha başarılı olduğunu göstermektedir.

Rassal orman regresyonunda da anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koyduğunu göstermektedir.

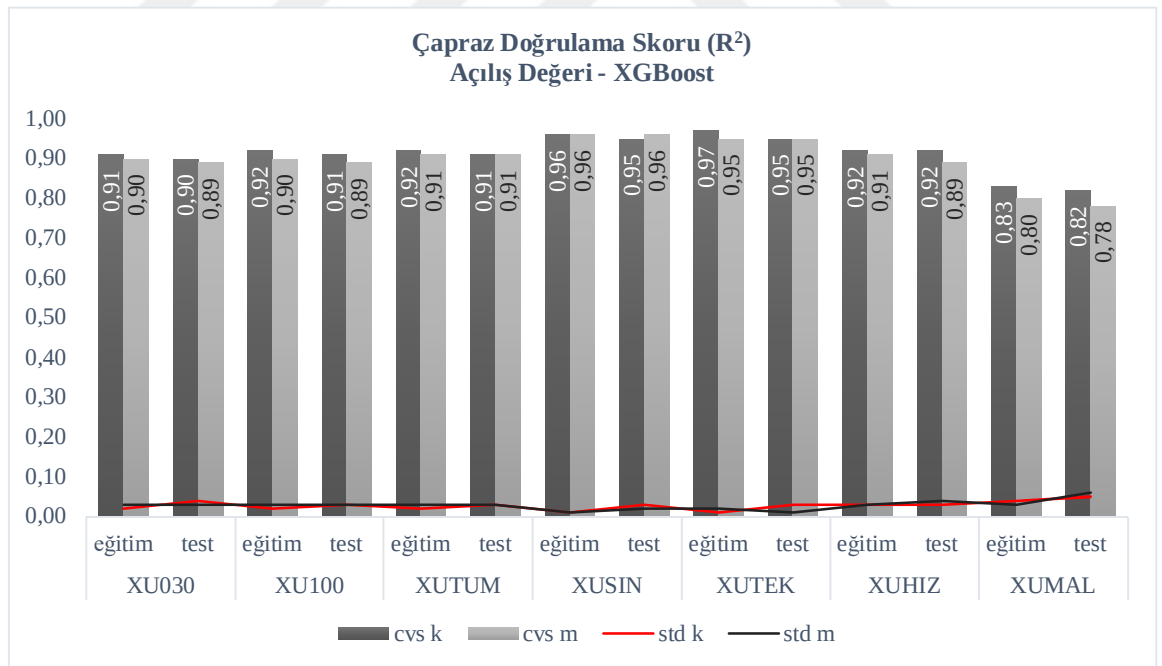
4.5.5.3.1.4 XGBoost Yöntemi ile BIST Pay Endekslerinin Açılış Değerlerinin Tahmini

XGBoost yöntemi de, rassal orman yöntemi gibi bir topluluk öğrenme yöntemidir. XGBoost yöntemi gradyan artırılmış karar ağaçları [gradient boosting decision trees - GBDT] algoritmasının bir uygulamasıdır. Gradyan artırma [gradient boosting], bir dizi daha basit, daha zayıf modelin tahminlerini birleştirerek bir hedef değişkeni doğru bir şekilde tahmin etmeye çalışan gözetimli bir karar ağacı topluluğu öğrenme [ensemble learning] algoritmasıdır. Başka bir deyişle XGBoost yöntemi de (Chen ve Guestrin, 2016), rassal orman yöntemi gibi bir topluluk öğrenme yöntemidir. Topluluk öğrenme algoritmaları, daha iyi bir model elde etmek için çoklu makine öğrenme algoritmalarını birleştirirler (Amazon SageMaker Documentation). Bir diğer ifadeyle kalabalığın bilgeliği fikrine dayanarak tahmincilerin tahminlerini bir araya getirmeyi amaçlarlar (Géron, 2021). Hem rassal orman hem de GBDT, birden çok karar ağacından oluşan bir model oluştururlar. Aralarındaki fark, ağaçların nasıl inşa edildiği ve birleştirildiği ile ilgilidir. XGBoost, konveks bir kayıp fonksiyonunu, öngörülen ve hedef çıktılar arasındaki farka dayalı olarak ve model karmaşıklığı için bir ceza terimini birleştiren düzenleştirilmiş (ℓ_1 ve ℓ_2) bir amaç fonksiyonunu minimize eder. Bu yaklaşımda bir sonraki modele uyması için önceki modelin hata kalıntıları kullanılarak her bir yineleme ile sık karar ağaçlarından oluşan bir topluluğu tekrarlı bir şekilde eğitirler. Nihai tahmin, tüm ağaç tahminlerinin ağırlıklı toplamıdır. Rassal orman yöntemi “torbalama” [bagging] yaparak, varyansı ve aşırı uyum sorununu en aza indirirken, GBDT “artırma” [boosting] yaparak sapmayı ve yetersiz uyumu en aza indirmektedir (XGBoost, n.d.). Yeni modeller eklerken kaybı en aza indirmek için bir gradyan iniş algoritması kullandığından buna gradyan artırma denilmektedir (dmlcXGBoost,

n.d.). Hızlı bir öğrenme eğrisine sahip olan XGBoost yöntemi ayrıca paralel işlem yapmaya uygundur. Ayrıca XGBoost, aşırı uyum problemini kontrol etmek için daha düzenli bir model formalizasyonu kullanmakta ve bu da ona daha iyi performans sağlamaktadır (XGBoost, n.d.).

Tablo 4.29’da yer alan XGBoost modelinin çapraz doğrulama sonuçları Grafik 4.14’te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksindeki değişimin %90’ı, XU100 endeksindeki değişimin %91’ini, XUTUM endeksindeki değişimin %91’ini, XUSIN endeksindeki değişimin %95’ini, XUTEK endeksindeki değişimin %96’sını, XUHIZ endeksindeki değişimin %92’sini XUMAL endeksindeki değişimin %80’ini açıklayabildiğini göstermektedir. XGBoost yöntemiyle elde edilen bulgular, rassal orman yönteminden elde edilen sonuçlar oldukça yakındır.

Grafik 4.14: XGBoost Çapraz Doğrulama Skoru (Açılış Değeri)



Grafik 4.14’te de görüldüğü üzere, XGBoost modeliyle elde edilen eğitim ve test verileri arasındaki yakın sonuçlar, modelin aşırı uyum probleminin olmadığını göstermektedir.

Tablo 4.29’da XGBoost modelinin tahmin başarısını gösteren MAPE istatistiği sonuçları yer almaktadır. Bu sonuçlara göre, XGBoost modelinde, rassal orman modelinde olduğu gibi, tüm endeksler için MAPE değeri %10’un altında olduğundan model *yüksek doğruluk derecesine* sahiptir. XGBoost regresyonunda, tüm endeksler için MAPE değeri, rassal orman regresyonunda olduğu gibi %10’un altında gerçekleşmiştir.

XGBoost yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koyduğu görülmektedir.

4.5.5.3.1.5 Açılış Değerlerinin Tahmininde Özniteliklerin Önem Düzeyleri

Tahmin modellerinden elde edilen ve Tablo 4.29’da sunulan bulgular genel olarak bir açıklayıcılık oranı verse de, hangi özniteliklerin bağımlı değişkeni açıklamada daha önemli bir role sahip olduğu konusunda bilgi vermemektedir. Özellikle öznitelik seçiminde gömülü yöntemleri kullanan lasso regresyon, rassal orman ve XGBoost yöntemlerinde özniteliklerin modele katkısının ortaya konabilmesi gerekmektedir. Böyle durumlarda makine öğrenmesi yöntemlerinin şeffaflığını ve yorumlanabilirliğini artırmak için SHAP (SHapley Additive exPlanations) değerlerine sıklıkla başvurulmaktadır. İşbirlikçi ya da koalisyon oyun teorisine dayalı bir yöntem olan SHAP değerleri, modelin çıktısına her bir özneliğin bireysel katkısını veya önemini göstermek için kullanılmaktadır (Lundberg ve Lee, 2017).

Özniteliğin tüm veri seti için önemini veya katkısını anlayabilmek için genellikle arı sürüsü tipi grafik kullanılmaktadır. Arı sürüsü tipi grafik, özneliğin önemini, özneliğin etkisiyle birleştirmektedir. Grafikteki her bir nokta, bir öznitelik ve bir örnek için SHAP değerine tekabül etmektedir. Y eksenindeki konum öznitelik ve x eksenindeki konum Shapley değeri aracılığıyla belirlenmektedir. Arı sürüsü tipi grafikte öznitelikler en önemliden en aza doğru önem düzeylerine göre sıralanmaktadır. Grafikte kullanılan renk ise, özneliğin değerini düşükten yükseğe doğru temsil etmektedir. Çakışan noktalar y ekseninde yoğunlaşırken [jitter] özniteliklerin Shapley değerlerinin dağılımı hakkında fikir

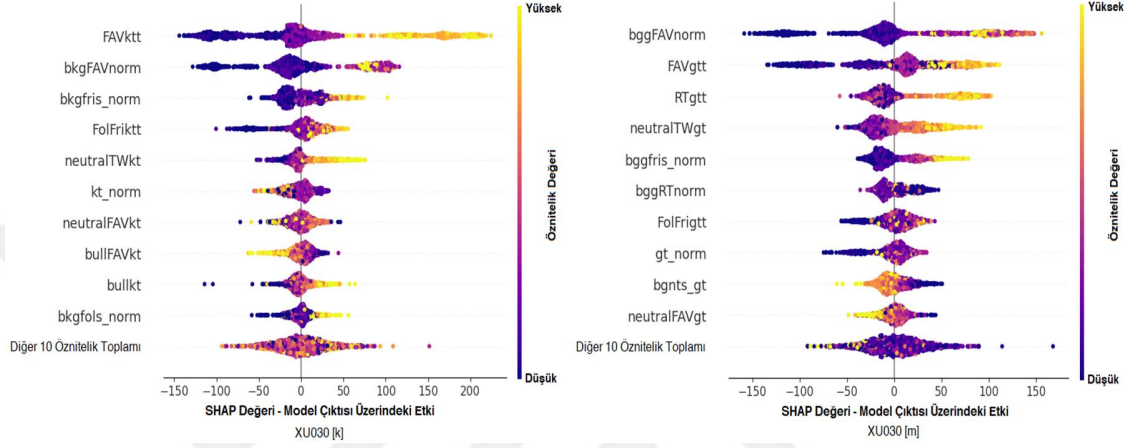
vermektedir. Arı sürüsü tipi grafik, bir özniteliğin değeri ile tahmin üzerindeki etkisi arasındaki ilişkinin göstergelerini ortaya koymaktadır.

Grafik 4.15, XU030 Endeksinin açılış değeri için XGBoost yöntemiyle kurulan tahmin modelinde yer alan özniteliklerin önemini arı sürüsü tipi grafik ile göstermektedir. Grafiğin sol tarafında bağlam sınıflandırması anahtar kelime tabanlı [k] yaklaşımdan elde edilen, sağ tarafında makine öğrenmesi temelli yaklaşımdan [m] elde edilen sonuçlar yer almaktadır. Bu bulgulara göre belirli bir günde borsa ile ilgili anahtar kelime yaklaşımına göre, beğeni tweetlerin toplam tweetlere oranının (FAVkt) ve toplam beğeni tweet sayısının (bkgFAVnorm) XU030 endeksinin açılış değerlerindeki değişimi açıklamada en fazla katkısı olan iki özniteliktir. Beğeni tweet sayısının yüksek değerleri modele pozitif katkı sağlarken, düşük değerlerinin modele negatif katkı sağladığı görülmektedir. Toplam takipçi sayısının toplam arkadaş sayısına oranıyla elde edilen ve bir nevi kanaat önderi olma olasılığına işaret eden bu oranda (FolFrikt) modele en çok katkı sunan oranlar arasındadır. Belirli bir günde anahtar kelime yaklaşımına göre borsa bağlamında tweet paylaşan kullanıcıların takipçi sayılarının arkadaş sayılarından fazla olması modele pozitif yönde katkı sunarken, takipçi sayısının arkadaş sayısına göre daha az olması modele negatif katkı sunmaktadır. Ayrıca belirli bir günde anahtar kelime yaklaşımına göre borsa ile ilgili olduğu belirlenen tweetlerin sayısı (kt_norm), bu tweetler içerisinde yer alan nötr tweet sayısı (neutralTWkt), yükseliş potansiyeli indeksi (bullkt) modelin açıklayıcılığında en çok katkısı olan özniteliklerdir.

Grafik 4.15'in sağ tarafında (XU030 [m]), belirli bir günde makine öğrenmesi yaklaşımına göre borsa ile ilgili olduğu belirlenen tweetlerden elde edilen özniteliklerin modelin açıklayıcılığına katkısı yer almaktadır. Özellikle beğeni tweetlerin (bkgFAVnorm, FAVgtt) modelin açıklayıcılığına katkısı burada da ön plana çıkmaktadır. Bu yaklaşımda, diğer yaklaşımdan farklı olarak öne çıkan özniteliklerden biri retweetlerle ilgilidir. Buna göre borsa ile ilgili makine öğrenmesi yaklaşımıyla belirlenen toplam retweet sayısı (bkgRTnorm) ve retweetlerin toplam tweetlere oranının (RTgtt) yüksek değerlerinin modele pozitif katkı sağladığı, düşük değerlerinin modele negatif katkı sağladığı söylenebilir. Burada öne çıkan bir diğer öznitelik ise negatif tweetlerin toplam tweetlere oranıdır. Buna göre negatif

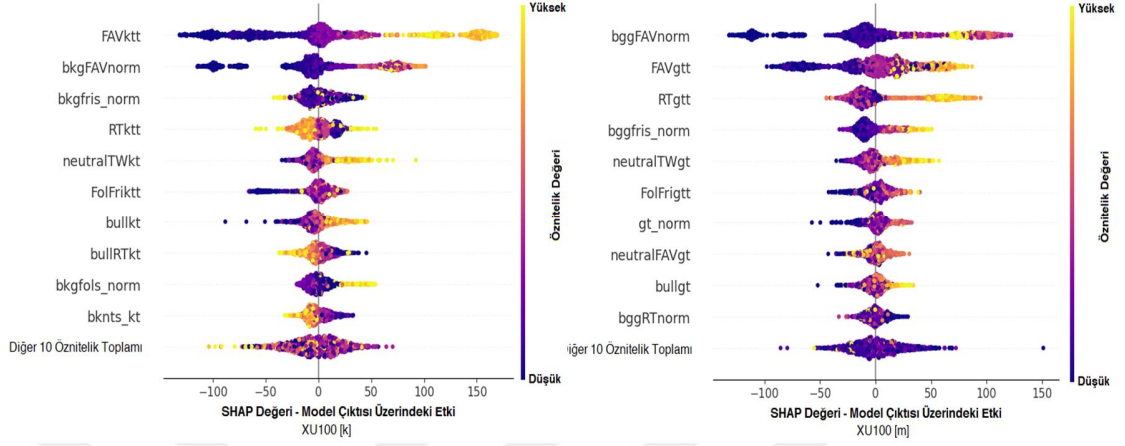
tweetlerin toplam tweetlere oranının (bgnts_gt) yüksek değerlerinin modele negatif katkı sağlarken, düşük değerlerinin modele pozitif katkı sağladığı görülmektedir.

Grafik 4.15: XU030 Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



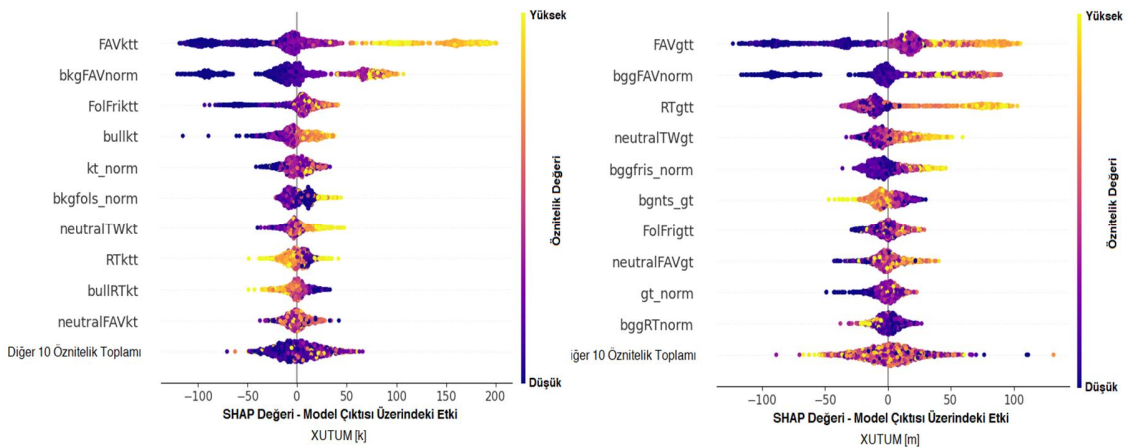
Grafik 4.16’da XU100 endeksi için XGBoost regresyon modelinde Twitter özniteliklerinin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bu endeks için de, modelin açıklayıcılığına en önemli katkıyı sunan öznitelik, beğeni tweetlerle (FAVkt, bkgFAVnorm) ilgili olanlardır. Retweetler, nötr tweetler, yükseliş potansiyeli endeksi gibi öznitelikler de modelin açıklayıcılığına katkı sunan diğer önemli özniteliklerdir.

Grafik 4.16: XU100 Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



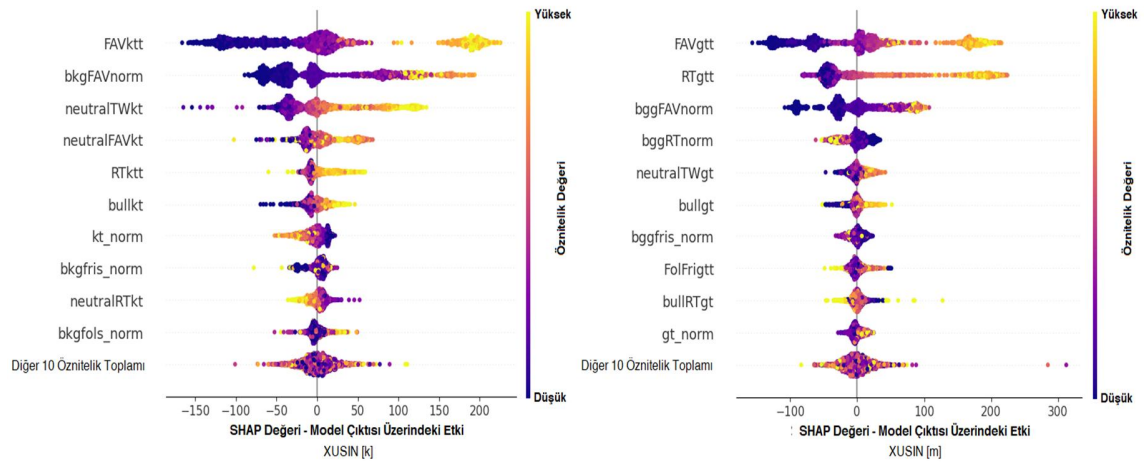
Grafik 4.17, XUTUM endeksinin açılış değerlerini Twitter öznitelikleri ile açıklamaya çalışan modele katkısı olan özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bu endeks için de, modelin açıklayıcılığında beğeni tweetler (FAVkt, bkgFAVnorm) en fazla paya sahiptir. Benzer şekilde tweet sayısı, retweet sayısı, yükseliş indeksi, takipçi sayısı gibi özniteliklerde modelin açıklayıcılığına katkı sunan diğer önemli özniteliklerdir.

Grafik 4.17: XUTUM Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



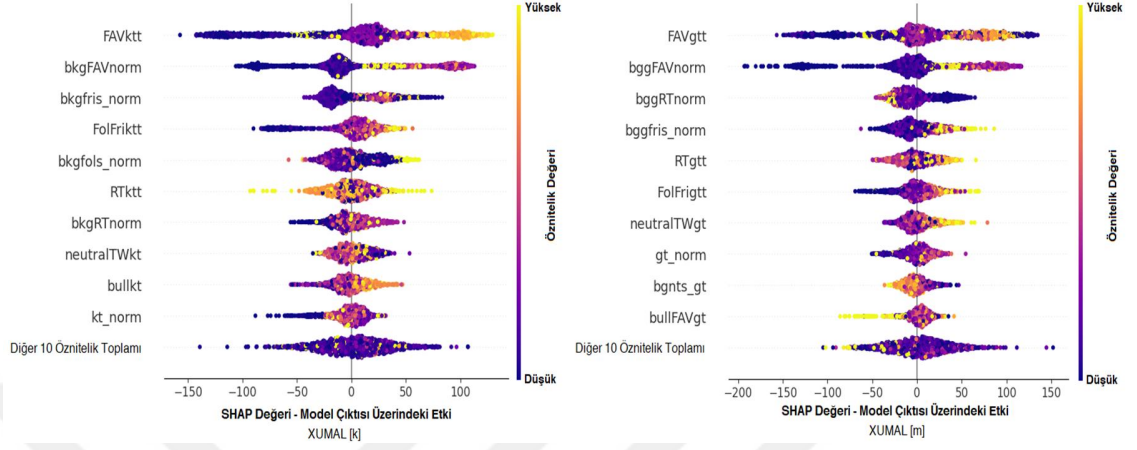
Grafik 4.18’de, XUSIN endeksinin açılış değerlerini borsa ile ilgili tweetlerden elde edilen öznitelikler ile açıklamaya çalışan XGBoost regresyon modelinde özniteliklerin modele katkı düzeyine göre öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular XUSIN endeksi için de, modelin açıklayıcılığında en fazla paya beğeni tweetlerin (FAVkt, bkgFAVnorm) sahip olduğunu göstermektedir. Ayrıca bu endekste nötr tweetlerin modelin açıklayıcılığına daha fazla katkı sunduğu görülmektedir. Benzer şekilde tweet sayısı, retweet sayısı, yükseliş indeksi, takipçi sayısı gibi öznitelikler de modelin açıklayıcılığına önemli katkılar sunan diğer önemli özniteliklerdir.

Grafik 4.18: XUSIN Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Önemini Arı Sürüsü Grafiği ile Gösterimi



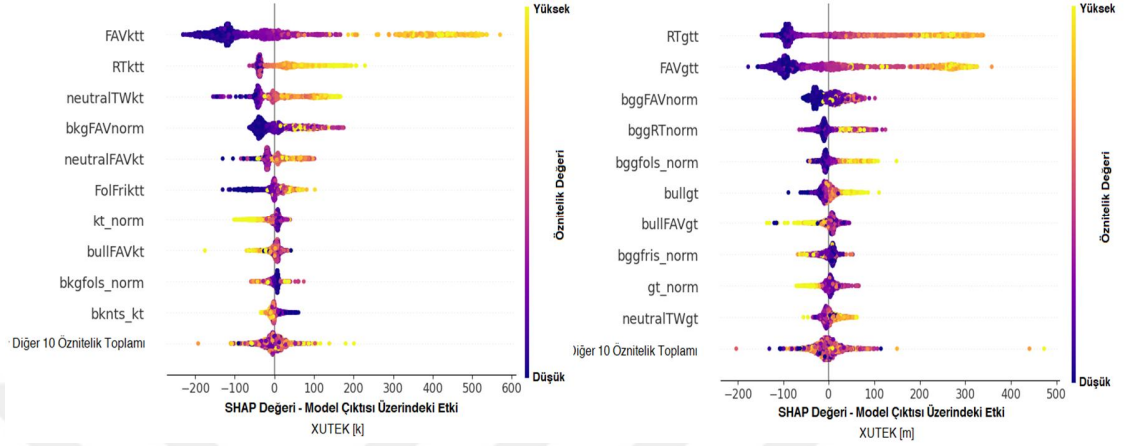
Grafik 4.19’da, XUMAL endeksinin açılış değerlerini borsa ile ilgili tweetlerden elde edilen öznitelikler ile açıklamaya çalışan modelde, özniteliklerin katkı düzeyine göre öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Modelin açıklayıcılığında bulgular, diğer endekslerdekine benzer şekilde, en fazla beğeni tweetlerin (FAVkt, bkgFAVnorm) katkı sağladığını göstermektedir. Ayrıca bu endekste takipçi ve arkadaşlarla ilgili özniteliklerin modelin açıklayıcılığına daha fazla katkı sunduğu görülmektedir. Benzer şekilde tweet sayısı, retweet sayısı, yükseliş indeksi gibi öznitelikler de modelin açıklayıcılığına önemli katkılar sunan diğer önemli özniteliklerdir.

Grafik 4.19: XUMAL Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



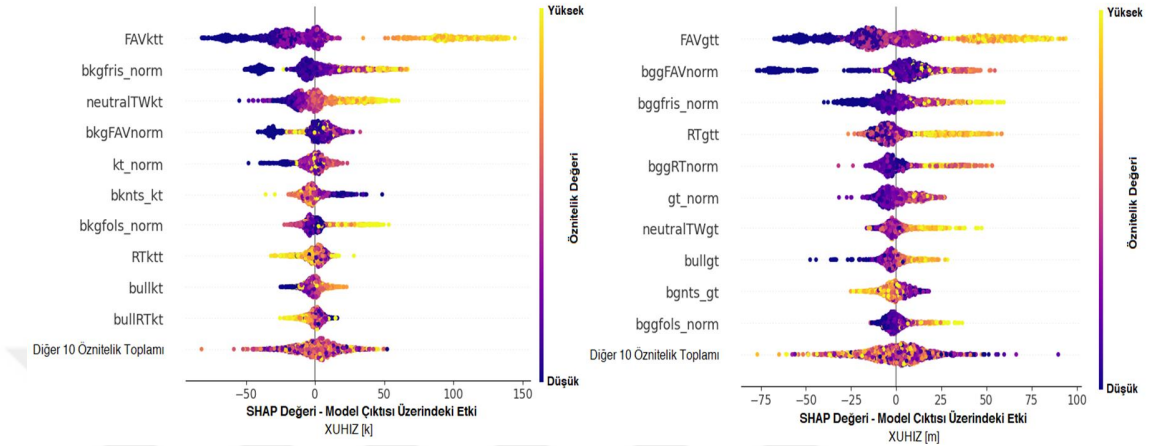
Grafik 4.20, XUTEK endeksinin açılış değerlerini açıklamada, borsa ile ilgili tweetlerden elde edilen özniteliklerin hangilerinin katkısının daha fazla olduğunu arı sürüsü grafiği ile göstermektedir. XUTEK endeksi için bulgular da, beğeni tweetlerin (FAVkt, FAVgt, bkgFAVnorm, bggFAVnorm) ve retweetlerin (RTgtt, RTktt, bggRTnorm, bkgRTnorm) modeli açıklamaya katkısı en fazla olan öznitelikler olduğuna işaret etmektedir. Ayrıca bu endekste retweetler ve nötr tweetlerle ilgili özniteliklerin modelin açıklayıcılığına daha fazla katkı sunduğunu göstermektedir. Benzer şekilde tweet sayısı, yükseliş indeksi gibi öznitelikler de modelin açıklayıcılığına önemli katkılar sunan diğer önemli özniteliklerdir.

Grafik 4.20: XUTEK Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



Grafik 4.21, borsa ile ilgili tweetlerden elde edilen özniteliklerden hangilerinin, XUHIZ endeksinin açılış değerlerini açıklamada, katkısının daha fazla olduğunu arı sürüsü grafiği ile göstermektedir. XUHIZ endeksi için bulgular, önceki bulgulara benzer şekilde beğeni tweetlerin (FAVkt, bkgFAVnorm) modeli açıklamaya katkısı en fazla olan öznitelikler olduğuna işaret etmektedir. Ayrıca bu endekste tweet sayısı, retweetler ve nötr tweetlerle ilgili özniteliklerin modelin açıklayıcılığına daha fazla katkı sunduğunu göstermektedir. Benzer şekilde yükseliş endeksi gibi öznitelikler de modelin açıklayıcılığına önemli katkılar sunmaktadır.

Grafik 4.21: XUHIZ Endeksi Açılış Değeri için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



Genel olarak, belirli bir günde borsa ile ilgili olduğu anahtar kelime tabanlı ve makine öğrenmesi tabanlı yaklaşımlarla belirlenen tweetlerden elde edilen özniteliklerin BIST Pay Endekslerinin açılış değerlerini açıklamada sundukları katkı ile ilgili olarak bulgular, beğeni tweet sayısı, retweet sayısı, takipçi sayısı gibi etkileşime dayalı özniteliklerin önemli katkılar sunduğunu ortaya koymaktadır. Ayrıca yükseliş potansiyeli indeksi, nötr tweetlerden elde edilen öznitelikler de modelin açıklayıcılığına katkı sunan önemli özniteliklerdendir. Bu bulgular, borsa bağlamındaki olduğu belirlenen tweetlerin fikri ve duygusal içeriklerinin, bazı durumlarda içerikten bağımsız olarak sayılarının ve bu tweetlerin kullanıcılar arasında, yani sosyal bir çevrede, sosyal bir etkileşim doğuracak şekilde paylaşımlarının, BIST Pay Endekslerinin açılış değerlerini etkilediğine işaret etmektedir.

4.5.5.3.2 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini

İşlem hacmi, bir menkul kıymetin belirli bir süre içerisindeki alım satım işlemlerinin toplam değeridir. İşlem hacmi bir menkul kıymetin veya piyasanın genel faaliyetini temsil etmektedir. İşlem hacmi ayrıca piyasanın likiditesi hakkında da bilgi vermektedir.

Bu çalışma kapsamında tweetlerden elde edilen özniteliklerin BIST Pay Endekslerinin işlem hacimlerinin ne kadarını açıklayabildiği de araştırılmıştır. Tablo 4.30’da, tweetlerden elde edilen öznitelikler ile BIST Pay Endekslerinin işlem hacimlerinin tahmin sonuçları yer almaktadır. Analiz, lineer regresyon, lasso regresyon, rassal orman ve XGBoost yöntemleri ile gerçekleştirilmiştir.

Tablo 4.30: BIST Pay Endekslerinin İşlem Hacimleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları

İşlem Hacmi (Log)																
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Linear Regresyon	cvs (R ²)	k	0,54	0,52	0,60	0,61	0,63	0,64	0,68	0,70	0,52	0,60	0,66	0,64	0,27	0,23
		m	0,52	0,50	0,58	0,59	0,60	0,62	0,66	0,69	0,53	0,62	0,64	0,65	0,28	0,25
	std	k	0,09	0,10	0,09	0,07	0,09	0,06	0,06	0,05	0,05	0,07	0,07	0,05	0,09	0,09
		m	0,09	0,10	0,08	0,07	0,08	0,07	0,06	0,06	0,07	0,07	0,07	0,04	0,08	0,12
	RMSE	k	0,30	0,32	0,27	0,29	0,26	0,27	0,31	0,32	0,75	0,71	0,35	0,36	0,31	0,33
		m	0,31	0,33	0,28	0,29	0,26	0,28	0,31	0,33	0,73	0,68	0,36	0,37	0,30	0,32
	MAE	k	0,22	0,24	0,20	0,22	0,19	0,20	0,24	0,25	0,59	0,56	0,27	0,29	0,23	0,24
		m	0,23	0,24	0,21	0,22	0,20	0,21	0,24	0,25	0,56	0,53	0,29	0,29	0,22	0,23
	MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,03	0,03	0,01	0,01	0,01	0,01
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,03	0,03	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Lasso Regresyon	cvs (R ²)	k	0,54	0,50	0,61	0,58	0,64	0,61	0,71	0,69	0,62	0,61	0,67	0,64	0,29	0,24
		m	0,52	0,52	0,59	0,52	0,61	0,53	0,67	0,53	0,61	0,54	0,65	0,55	0,27	0,24
	std	k	0,10	0,12	0,04	0,07	0,04	0,07	0,05	0,07	0,07	0,09	0,04	0,09	0,06	0,10
		m	0,10	0,10	0,05	0,15	0,05	0,18	0,03	0,25	0,06	0,17	0,04	0,20	0,08	0,08
	RMSE	k	0,30	0,32	0,27	0,28	0,26	0,26	0,30	0,31	0,64	0,70	0,34	0,35	0,30	0,31
		m	0,30	0,33	0,28	0,28	0,27	0,27	0,31	0,34	0,66	0,72	0,36	0,37	0,30	0,30
	MAE	k	0,22	0,24	0,20	0,21	0,19	0,20	0,23	0,24	0,50	0,53	0,27	0,28	0,22	0,23
		m	0,23	0,24	0,21	0,22	0,20	0,21	0,24	0,26	0,50	0,55	0,28	0,29	0,22	0,23
	MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,03	0,03	0,01	0,01	0,01	0,01
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,03	0,03	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Rassal Orman	cvs (R ²)	k	0,58	0,54	0,64	0,63	0,68	0,65	0,78	0,73	0,73	0,70	0,73	0,70	0,33	0,30
		m	0,57	0,54	0,63	0,60	0,66	0,62	0,77	0,72	0,71	0,69	0,71	0,70	0,31	0,28
	std	k	0,09	0,11	0,05	0,07	0,06	0,08	0,03	0,06	0,06	0,06	0,04	0,08	0,09	0,09
		m	0,09	0,12	0,06	0,08	0,06	0,08	0,04	0,05	0,07	0,05	0,05	0,07	0,08	0,09
	RMSE	k	0,18	0,31	0,16	0,26	0,15	0,24	0,16	0,27	0,31	0,56	0,19	0,32	0,19	0,29
		m	0,18	0,32	0,17	0,27	0,15	0,25	0,16	0,28	0,33	0,58	0,21	0,33	0,20	0,30
	MAE	k	0,13	0,22	0,12	0,20	0,11	0,18	0,12	0,20	0,24	0,43	0,15	0,25	0,14	0,22
		m	0,14	0,23	0,12	0,20	0,11	0,18	0,13	0,21	0,25	0,45	0,16	0,25	0,15	0,22
	MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,01
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
XGBoost	cvs (R ²)	k	0,49	0,41	0,58	0,54	0,60	0,58	0,74	0,66	0,70	0,65	0,68	0,64	0,17	0,13
		m	0,50	0,43	0,56	0,48	0,60	0,48	0,73	0,63	0,62	0,56	0,66	0,59	0,16	0,06
	std	k	0,07	0,10	0,07	0,12	0,08	0,09	0,06	0,07	0,07	0,11	0,05	0,08	0,14	0,11
		m	0,12	0,11	0,08	0,11	0,08	0,08	0,05	0,07	0,08	0,10	0,04	0,11	0,13	0,15
	RMSE	k	0,09	0,34	0,09	0,29	0,08	0,27	0,08	0,28	0,16	0,60	0,10	0,35	0,09	0,32
		m	0,09	0,35	0,08	0,29	0,07	0,26	0,07	0,31	0,17	0,64	0,09	0,35	0,09	0,31
	MAE	k	0,07	0,24	0,06	0,22	0,06	0,21	0,06	0,21	0,12	0,46	0,07	0,27	0,07	0,23
		m	0,07	0,25	0,06	0,22	0,05	0,20	0,06	0,24	0,13	0,50	0,07	0,27	0,07	0,24
	MAPE	k	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,01	0,03	0,00	0,01	0,00	0,01
		m	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,01	0,03	0,00	0,01	0,00	0,01

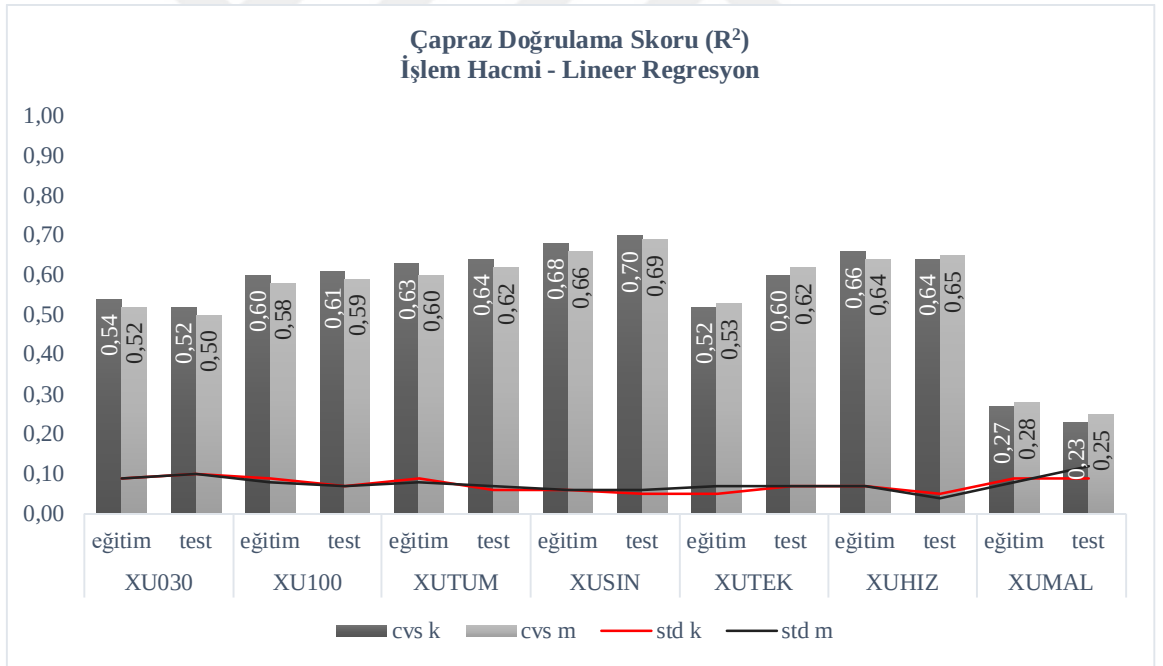
Y: Yöntem; P: Performans, C: Bağlam; k: Anahtar Kelime; m: Makine Öğrenmesi; cvs: Çapraz Doğrulama Skoru; std: Standart Sapma

4.5.5.3.2.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini

Tablo 4.30’da yer alan lineer regresyon modelinin doğal logaritması alınmış işlem hacmi için çapraz doğrulama sonuçları Grafik 4.22’de sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin işlem hacmindeki değişimin %52’sini, XU100 endeksinin işlem hacmindeki değişimin %61’ini, XUTUM endeksinin işlem hacmindeki değişimin

%64'ünü, XUSIN endeksinin işlem hacmindeki değişimin %70'ini, XUTEK endeksinin işlem hacmindeki değişimin %60'ını, XUHIZ endeksinin işlem hacmindeki değişimin %64'ünü XUMAL endeksinin işlem hacmindeki değişimin %23'ünü açıklayabildiğini göstermektedir. Lineer regresyon analizden elde edilen bulgulardan biri tweet özniteliklerin XUSIN endeksinin işlem hacmindeki değişim için en yüksek açıklayıcılık değerine, XUMAL endeksinin işlem hacmindeki değişim için en düşük açıklayıcılık değerine sahip olmasıdır. XUMAL için açıklayıcılık değeri diğerlerine kıyasla dikkat çekici bir şekilde düşüktür. Lineer regresyon analizinden elde edilen bir diğer bulgu, anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koymuş olmalarıdır.

Grafik 4.22: Lineer Regresyon Çapraz Doğrulama Skoru (İşlem Hacmi)



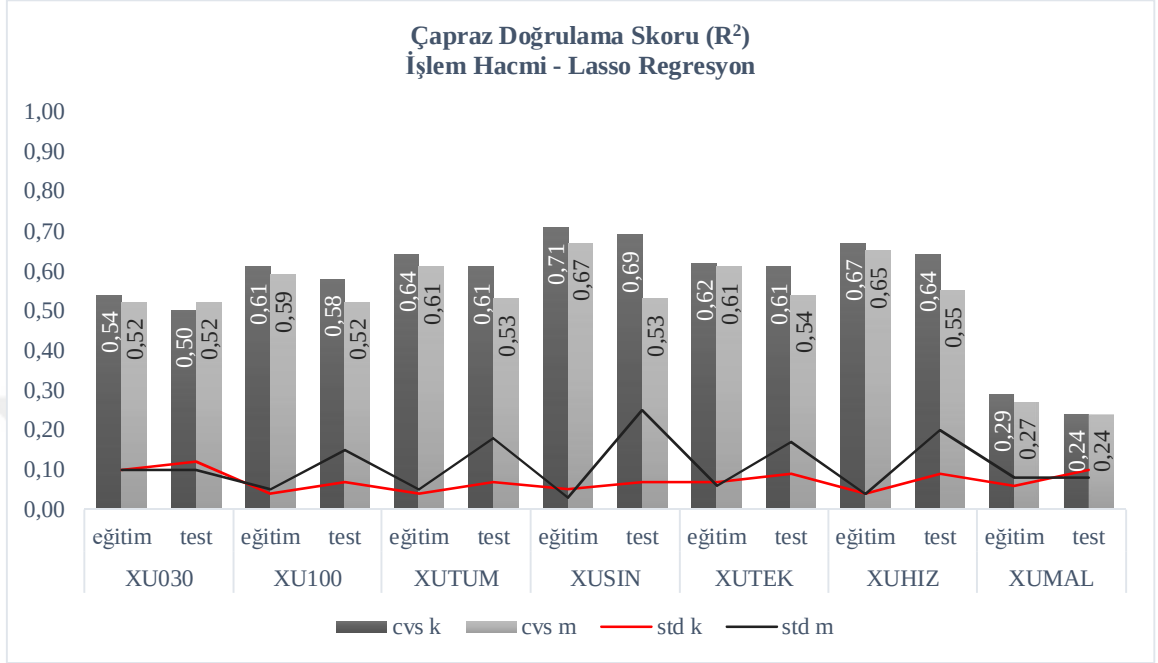
Grafik 4.22’de de görüldüğü üzere, eğitim ve test sonuçları birbirine çok yakındır. Bu durum işlem hacmini tahmin eden modelin aşırı uyum probleminin olmadığı anlamına gelmektedir.

Lineer regresyon modelinin tahmin başarısı ayrıca Tablo 4.30’da görüldüğü üzere MAPE istatistiği ile ölçülmüştür. Bu sonuçlara göre BIST Pay Endeksleri için MAPE değeri %10’un altında olduğu için model *yüksek doğruluk derecesine* sahiptir.

4.5.5.3.2.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini

Tablo 4.30’da yer alan lasso regresyon modelinin çapraz doğrulama sonuçları Grafik 4.23’te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin işlem hacmindeki değişimin %50’sini, XU100 endeksinin işlem hacmindeki değişimin %58’ini, XUTUM endeksinin işlem hacmindeki değişimin %61’ini, XUSIN endeksinin işlem hacmindeki değişimin %69’unu, XUTEK endeksinin işlem hacmindeki değişimin %61’ini, XUHIZ endeksinin işlem hacmindeki değişimin %64’ünü XUMAL endeksinin işlem hacmindeki değişimin %24’ünü açıklayabildiğini göstermektedir. Lasso regresyon analizinde de, lineer regresyonda elde edilen bulgulara yakın bulgular elde edilmiştir. Benzer şekilde tweetlerden elde edilen özniteliklerin XUSIN endeksi için en yüksek açıklayıcılık değerine, XUMAL endeksi için en düşük açıklayıcılık değerine sahip olmasıdır. Lasso regresyon yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı ile makine öğrenmesi tabanlı bağlam sınıflandırıcının modelin tahmin başarısında benzer bir performans ortaya koyduğu görülmektedir.

Grafik 4.23: Lasso Regresyon Çapraz Doğrulama Skoru (İşlem Hacmi)



Grafik 4.23'te görüldüğü üzere, eğitim ve test sonuçları birbirine çok yakın olması işlem hacmini tahmin eden modelin aşırı uyum probleminin olmadığı anlamına gelmektedir.

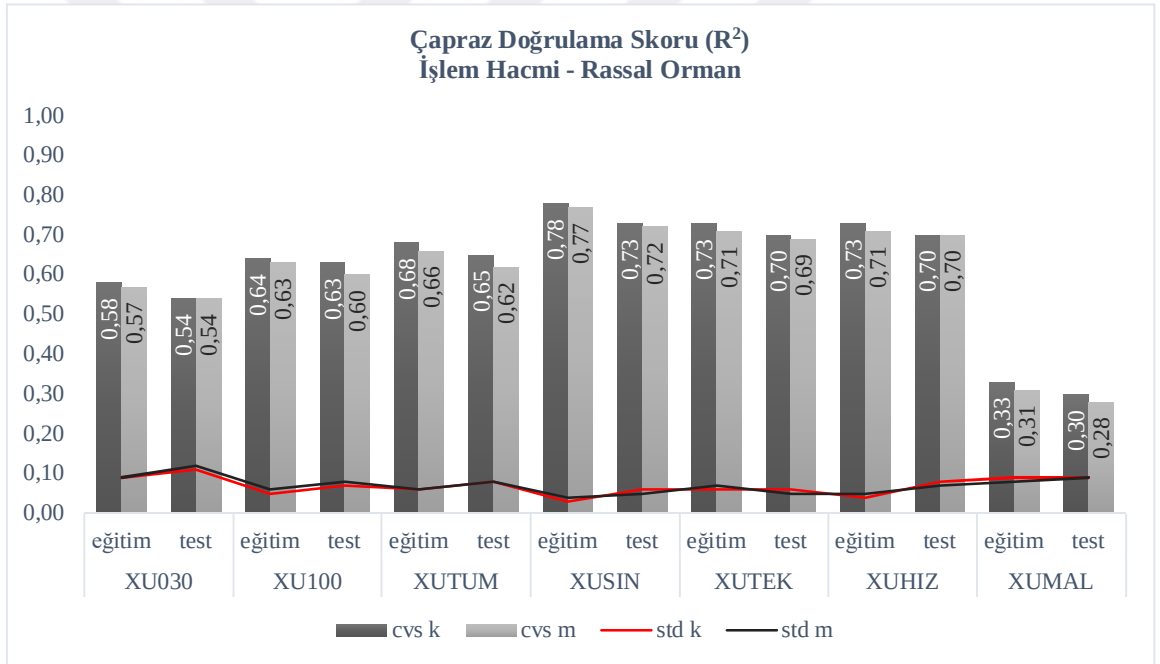
Lasso regresyon modelinin tahmin başarısı ayrıca Tablo 4.30'da görüldüğü üzere MAPE istatistiği ile ölçülmüştür. Bu sonuçlara göre BIST Pay Endeksleri için MAPE değeri %10'un altında olduğu için model *yüksek doğruluk derecesine* sahiptir.

4.5.5.3.2.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini

Tablo 4.30'da yer alan rassal orman modelinin çapraz doğrulama sonuçları Grafik 4.24'te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özneliliklerin XU030 endeksinin işlem hacmindeki değişimin %54'ünü, XU100 endeksinin işlem hacmindeki değişimin %63'ünü, XUTUM endeksinin işlem hacmindeki değişimin %65'ini, XUSIN endeksinin işlem hacmindeki değişimin %73'ünü, XUTEK endeksinin işlem hacmindeki

değişimin %70'ini, XUHIZ endeksinin işlem hacmindeki değişimin %70'ini XUMAL endeksinin işlem hacmindeki değişimin %30'unu açıklayabildiğini göstermektedir. Rassal orman modelinde elde edilen dikkat çekici bulgu, tweet özniteliklerinin XUSIN endeksinin işlem hacmi için analiz kapsamında ulaşılmış en yüksek açıklayıcılık değeri olmasıdır. Diğer bir deyişle tweetlerden elde edilen öznitelikler, rassal orman yöntemi kullanıldığında, XUSIN endeksinin işlem hacmindeki değişimin %73'ünü açıklayabilmektedir. Rassal orman modelinde de XUMAL, diğer endekslere kıyasla en düşük açıklayıcılık değerine sahiptir. Lineer regresyonda bu değer %23, lasso regresyonda %24 iken rassal ormanda %30'dur.

Grafik 4.24: Rassal Orman Çapraz Doğrulama Skoru (İşlem Hacmi)



Grafik 4.24'te, görüldüğü üzere, eğitim ve test sonuçlarının birbirine yakın olması rassal orman modelinde de aşırı uyum probleminin olmadığını göstermektedir.

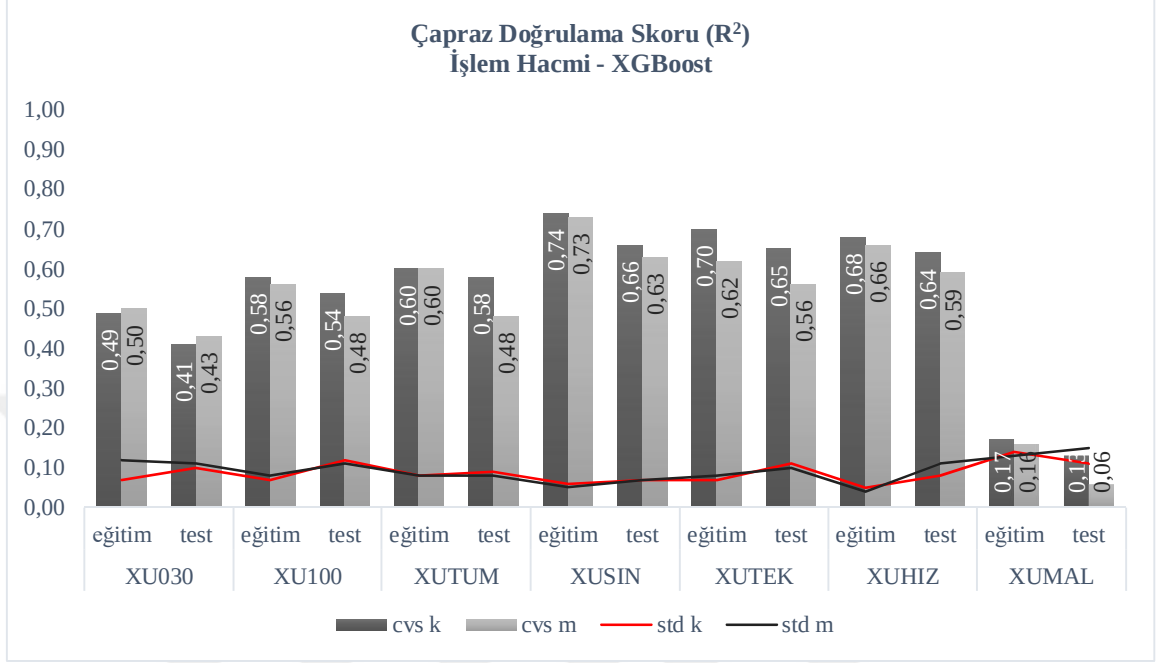
Rassal orman modelinin tahmin performansını gösteren MAPE istatistiği Tablo 4.30'da yer almaktadır. Bu sonuçlara göre, rassal orman modelinde tüm endeksler için MAPE değeri

%10'un altında olduğundan model bütün endeks değerleri için *yüksek doğruluk derecesine* sahiptir.

4.5.5.3.2.4 XGBoost Yöntemi ile BIST Pay Endekslerinin İşlem Hacimlerinin Tahmini

Tablo 4.30'da yer alan rassal orman modelinin çapraz doğrulama sonuçları Grafik 4.25'te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin işlem hacmindeki değişimin %41'ini, XU100 endeksinin işlem hacmindeki değişimin %54'ünü, XUTUM endeksinin işlem hacmindeki değişimin %58'ini, XUSIN endeksinin işlem hacmindeki değişimin %66'sını, XUTEK endeksinin işlem hacmindeki değişimin %65'ini, XUHIZ endeksinin işlem hacmindeki değişimin %64'ünü XUMAL endeksinin işlem hacmindeki değişimin %13'ünü açıklayabildiğini göstermektedir. XGBoost yönteminde model, XUSIN endeksi için en yüksek açıklayıcılık gücüne, XUMAL için en düşük açıklayıcılık gücüne sahiptir. XGBoost yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında önemli bir farklılıklarının olmadığını göstermektedir.

Grafik 4.25: XGBoost Çapraz Doğrulama Skoru (İşlem Hacmi)



Grafik 4.25'te de görüldüğü üzere, XGBoost modeliyle elde edilen eğitim ve test verileri arasındaki yakın sonuçlar, modelin aşırı uyum probleminin olmadığını göstermektedir.

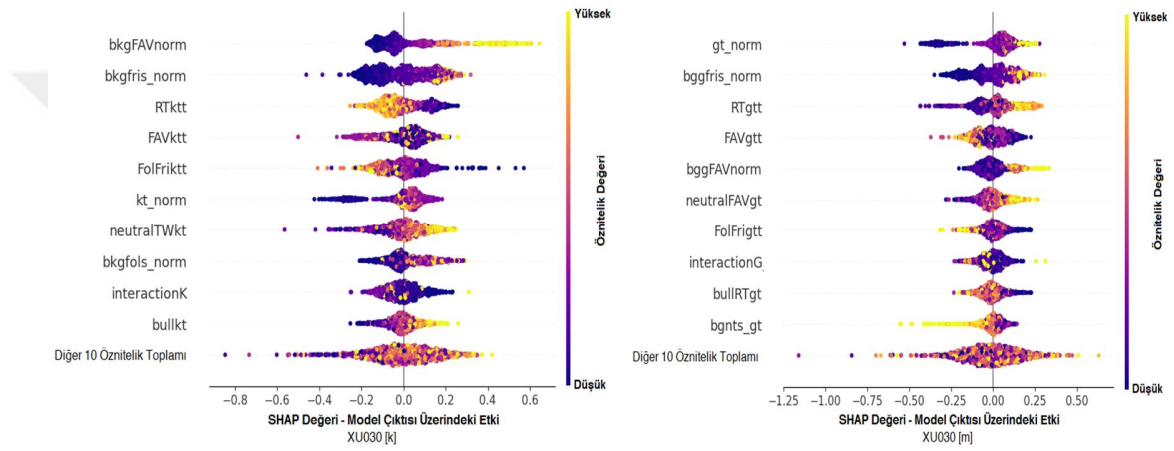
Tablo 4.30'da XGBoost modelinin tahmin başarısını gösteren MAPE istatistiği sonuçları yer almaktadır. Bu sonuçlara göre, XGBoost modelinde, rassal orman modelinde olduğu gibi, tüm endeksler için MAPE değeri %10'un altında olduğundan model *yüksek doğruluk derecesine* sahiptir.

4.5.5.3.2.5 İşlem Hacimlerinin Tahmininde Öznitelik Önem Düzeyleri

Bu çalışma kapsamında tweetlerden elde edilen özniteliklerin, işlem hacmi ile ilgili modelin açıklayıcılığına katkısı da SHAP değerleri yardımıyla araştırılmıştır. Grafik 4.26, XU030 endeksinin işlem hacmi için XGBoost tahmin modelinde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bu çerçevede modele en çok katkısı olan özniteliklerin özellikle etkileşime dayalı öznitelikler olduğu (bkgFAVnorm, bggFAVnorm, bggfris_norm,

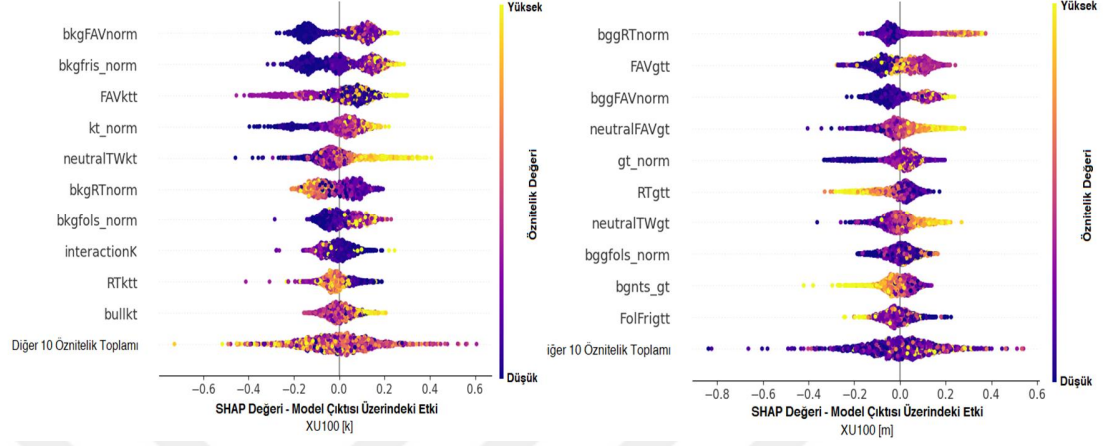
RTgtt, RTktt, FolFrigtt, FolFrikt, FAVktt, FAVgtt, bkgfols_norm, interactionK, interactionG) söylenebilir. Ayrıca yükseliş potansiyeli indeksi de (bullkt) modele katkı sağlamaktadır. Bunlara ilave olarak toplam tweet sayısı (gt_norm, kt_norm) ve nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır.

Grafik 4.26: XU030 Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



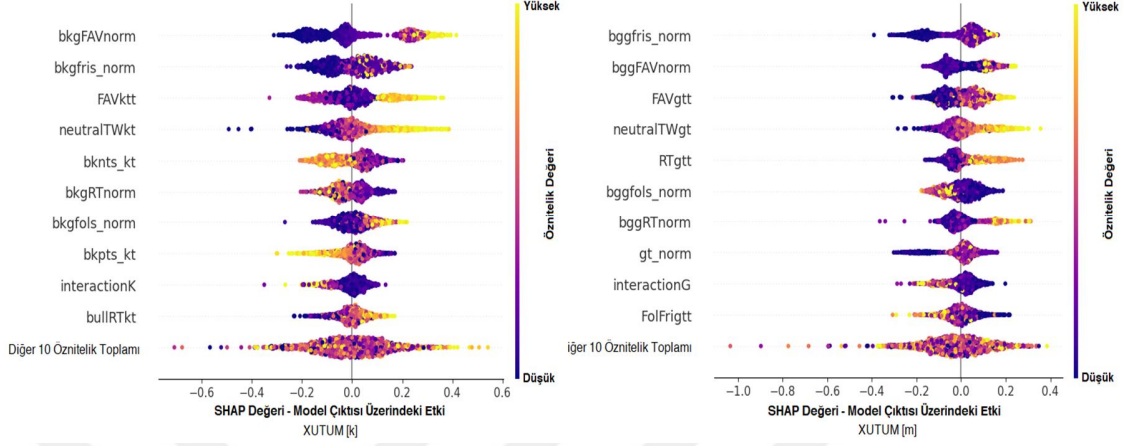
Grafik 4.27, XU100 endeksinin işlem hacminin tahmininde kullanılan XGBoost yönteminde, modele katkısı olan özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular modelin açıklayıcılığında retweetlerin toplam tweetlere oranı (RTgtt, RTktt), retweet sayısı (bkgRTnorm, bggRTnorm), beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), beğeni tweet sayısı (bkgFAVnorm, bggFAVnorm), takipçi sayısı (bkgfols_norm, bggfols_norm) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Ayrıca toplam tweet sayısı (gt_norm, kt_norm) ve nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt) da modelin açıklayıcılığında önemli paya sahip öznitelikler arasındadır.

Grafik 4.27: XU100 Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



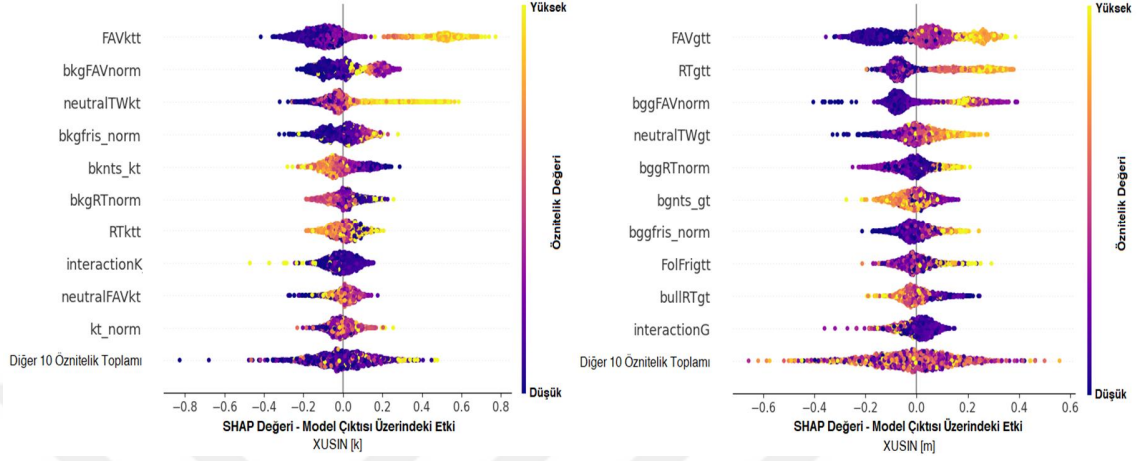
Grafik 4.28, XUTUM endeksinin işlem hacminin XGBoost yöntemi ile tahmininde, modele katkısı olan özniteliklerin önemini göstermektedir. Bulgular modelin açıklayıcılığında retweetlerin toplam tweetlere oranı (RTgtt, RTktt), retweet sayısı (bkgRTnorm, bggRTnorm), beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), beğeni tweet sayısı (bkgFAVnorm, bggFAVnorm), takipçi sayısı (bkgfols_norm, bggfols_norm), arkadaş sayısı (bkgfris_norm, bggfris_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğunu göstermektedir. Buna ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığında önemli paya sahip öznitelikler arasındadır.

Grafik 4.28: XUTUM Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



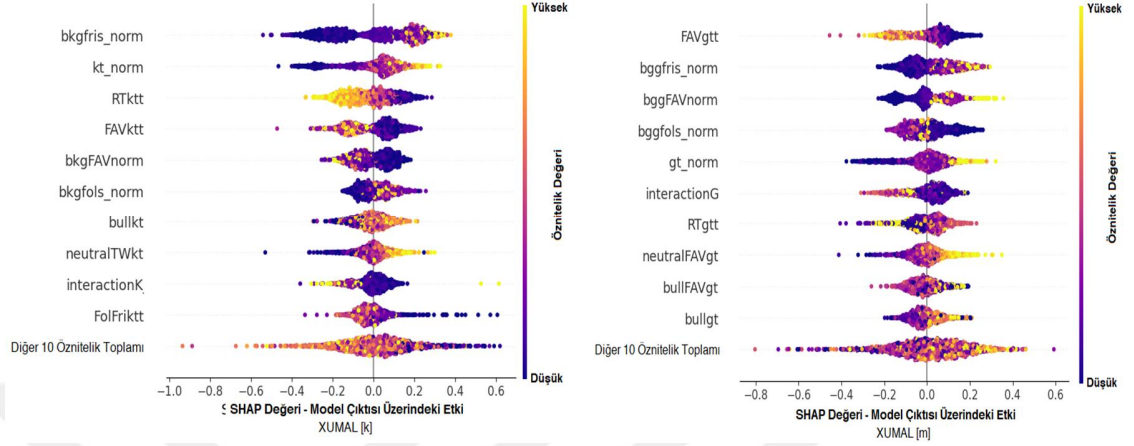
Grafik 4.29, XUSIN endeksi için XGBoost yöntemi ile işlem hacminin tahmininde, modele katkısı olan öznitelikleri göstermektedir. Bulgular modelin açıklayıcılığında beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), beğeni tweet sayısı (bkgFAVnorm, bggFAVnorm), retweetlerin toplam tweetlere oranı (RTgtt, RTktt), retweet sayısı (bkgRTnorm, bggRTnorm), arkadaş sayısı (bkgfris_norm, bggfris_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Ayrıca yükseliş potansiyeli endeksi (bullRTgt), nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığında önemli paya sahip öznitelikler arasında sayılabilir.

Grafik 4.29: XUSIN Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



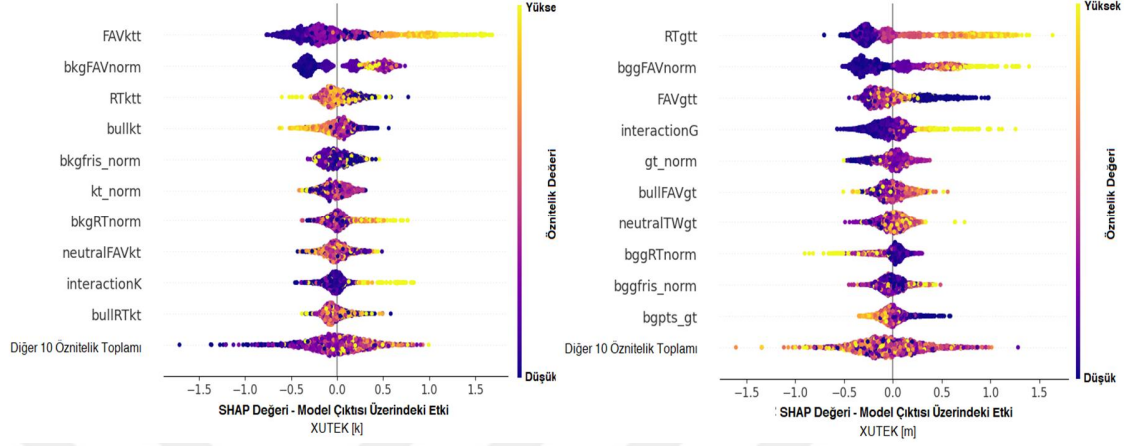
Grafik 4.30, XUMAL endeksi için işlem hacminin tahmininde XGBoost yöntemi ile kurulan modelde en çok katkısı olan öznitelikleri göstermektedir. Bulgular modelin açıklayıcılığında retweetlerin toplam tweetlere oranı (RTgtt, RTktt), retweet sayısı (bkgRTnorm, bggRTnorm), beğeni tweetlerin toplam tweetlere oranı (FAVkt, FAVgtt), beğeni tweet sayısı (bkgFAVnorm, bggFAVnorm), arkadaş sayısı (bkgfris_norm, bggfris_norm), takipçi sayısı (bkgfols_norm, bggfols_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Ayrıca yükseliş potansiyeli indeksi (bullgt, bullkt), beğeni tweetler için yükseliş potansiyeli indeksi (bullFAVgt), nötr tweetlerin toplam tweetlere oranı (neutralTWkt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVgt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığına katkı sunan önemli öznitelikler arasında yer aldığı söylenebilir.

Grafik 4.30: XUMAL Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



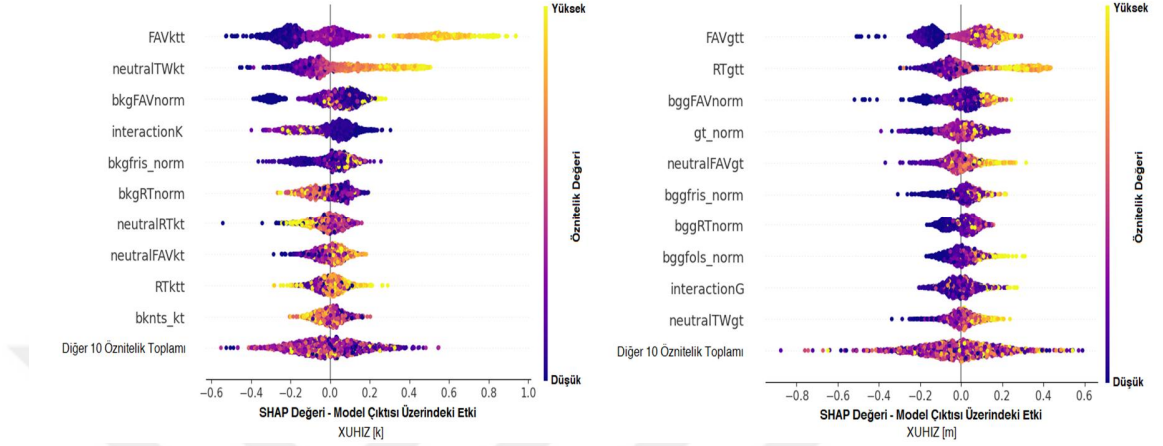
Grafik 4.31, XUTEK endeksinin XGBoost modeliyle işlem hacminin tahmini için kurulan modele en çok katkısı olan öznitelikleri göstermektedir. Bulgular modelin açıklayıcılığında retweetlerin toplam tweetlere oranı (RTgtt, RTktt), beğeni tweetlerin toplam tweetlere oranı (FAVgtt, FAVktt), beğeni tweet sayısı (bkgFAVnorm, bkgFAVnorm), retweet sayısı (bkgRTnorm, bkgRTnorm), arkadaş sayısı (bkgfris_norm, bkgfris_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Ayrıca yükseliş potansiyeli indeksinin (bullkt), beğeni tweetler için hesaplanan yükseliş potansiyeli indeksinin (bullFAVgt) ve retweetler için hesaplanan yükseliş potansiyeli indeksinin (bullRTkt) de modelin açıklayıcılığa katkısının daha üst sıralarda yer aldığı görülmektedir. Bunlara ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır.

Grafik 4.31: XUTEK Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



Grafik 4.32, XUHIZ endeksinin XGBoost modeliyle işlem hacminin tahmini için kurulan modele en çok katkısı olan öznitelikleri göstermektedir. Bulgular modelin açıklayıcılığında beğeni tweetlerin toplam tweetlere oranı (FAVkttt, FAVgtt), retweetlerin toplam tweetlere oranı (RTgtt, RTkttt), beğeni tweet sayısı (bkgFAVnorm, bkgFAVnorm), retweet sayısı (bkgRTnorm, bkgRTnorm), arkadaş sayısı (bkgfris_norm, bkgfris_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Bunlara ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt), nötr retweetlerin toplam tweetlere oranı (neutralRTkt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır.

Grafik 4.32: XUHIZ Endeksi İşlem Hacmi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



İşlem hacimlerinin tahmininde özniteliklerin önem düzeylerine ilişkin bulgular genel olarak değerlendirildiğinde, bazı özniteliklerin hemen hemen bütün endeksler için modelin açıklayıcılığına katkı sunduğu görülmektedir. Bu özniteliklerden bazıları retweetlerin toplam tweetlere oranı, beğeni tweetlerin toplam tweetlere oranı, retweet sayısı, beğeni tweet sayısı, tweet sayısı, takipçi sayısı, arkadaş sayısı, etkileşim skoru gibi etkileşime dayalı, yani karar modelinin sosyal ve davranışsal boyutu ile ilişkili özniteliklerdir. Ayrıca yükseliş potansiyeli indeksi de modelin açıklayıcılığına katkı sunmakta ve karar modelinin bilişsel katmanında yer almaktadır. Son olarak bilişsel, davranışsal ve sosyal boyutları olan nötr tweetlerin de modelin açıklayıcılığında önemli bir paya sahip olduğu görülmektedir. Bu bulgular, borsa bağlamında olduğu belirlenen tweetlerin hem fikri ve duygusal içeriklerinin hem içerikten bağımsız olarak sayılarının hem de bu tweetlerin kullanıcılar arasında, yani sosyal bir çevrede, sosyal bir etkileşim doğuracak şekilde paylaşılmalarının, BIST Pay Endekslerindeki işlem hacimlerini etkilediğine işaret etmektedir.

4.5.5.3 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Getirilerinin Tahmini

Bu çalışma kapsamında tweetlerin BIST Pay Endekslerinin getirilerine etkisi de araştırılmıştır. Tablo 4.31’de, tweetlerden elde edilen öznitelikler aracılığıyla BIST Pay Endekslerinin getirilerinin tahmin sonuçları yer almaktadır. Analiz, açılış değerlerinin tahmininde olduğu gibi, lineer regresyon, lasso regresyon, rassal orman ve XGBoost yöntemleri ile gerçekleştirmiş, elde edilen sonuçlar Tablo 4.31’de yöntem ve endeks bazında ayrı ayrı verilmiştir. Tablo 4.31’de de görüldüğü üzere hiçbir yöntem endeks getirilerinin tahmininde başarılı ve tutarlı bir açıklayıcılık ortaya koyamamıştır. Dikkat edileceği üzere eğitim ve test verileri arasında büyük farklılıklar oluşmakta, başka bir ifadeyle modellerde aşırı uyum sorunu olduğu görülmektedir.

Tablo 4.31: BIST Pay Endekslerinin Getirileri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları

Getiri																
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Linear Regresyon	cvss (R ²)	k	-0,25	-0,56	-0,25	-0,51	-0,25	-0,50	-0,13	-0,46	-0,29	-0,60	-0,25	-0,51	-0,31	-0,46
		m	-0,07	-0,13	-0,06	-0,13	-0,06	-0,13	-0,03	-0,11	-0,05	-0,11	-0,03	-0,09	-0,10	-0,17
	std	k	0,56	0,91	0,58	0,80	0,59	0,76	0,25	0,75	0,61	0,76	0,67	0,67	0,72	0,65
		m	0,10	0,07	0,10	0,08	0,10	0,08	0,07	0,07	0,08	0,12	0,05	0,06	0,13	0,13
	RMSE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,02	0,01	0,01	0,02	0,02
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,02	0,01	0,01	0,02	0,02
MAE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Lasso Regresyon	cvss (R ²)	k	-0,01	-0,05	0,00	-0,05	0,00	-0,05	-0,01	-0,04	-0,01	-0,05	0,00	-0,06	0,00	-0,05
		m	-0,01	-0,05	-0,01	-0,04	-0,01	-0,04	-0,02	-0,03	0,00	-0,04	0,00	-0,04	0,00	-0,04
	std	k	0,03	0,05	0,03	0,05	0,03	0,05	0,03	0,03	0,03	0,04	0,03	0,07	0,03	0,05
		m	0,04	0,04	0,03	0,04	0,03	0,04	0,03	0,02	0,02	0,04	0,04	0,06	0,03	0,05
	RMSE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,02	0,01	0,01	0,02	0,02
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,02	0,01	0,01	0,02	0,02
MAE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Rassal Orman	cvss (R ²)	k	-0,03	-0,09	-0,02	-0,09	-0,04	-0,09	-0,03	-0,10	-0,07	-0,12	-0,04	-0,12	-0,05	-0,10
		m	-0,02	-0,10	-0,02	-0,12	-0,02	-0,09	-0,03	-0,08	-0,03	-0,11	-0,02	-0,09	-0,03	-0,07
	std	k	0,04	0,05	0,04	0,09	0,04	0,08	0,05	0,09	0,05	0,08	0,06	0,12	0,05	0,07
		m	0,04	0,05	0,04	0,07	0,03	0,06	0,05	0,05	0,04	0,05	0,06	0,09	0,04	0,06
	RMSE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,02
		m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,02	0,01	0,01	0,01	0,02
MAE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
MAPE	k	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
	m	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
XGBoost	cvss (R ²)	k	-0,29	-0,42	-0,35	-0,45	-0,27	-0,47	-0,28	-0,41	-0,28	-0,56	-0,33	-0,39	-0,32	-0,38
		m	-0,33	-0,32	-0,31	-0,39	-0,33	-0,38	-0,31	-0,33	-0,28	-0,38	-0,34	-0,43	-0,29	-0,40
	std	k	0,11	0,33	0,09	0,30	0,09	0,29	0,10	0,22	0,08	0,40	0,11	0,24	0,10	0,21
		m	0,10	0,25	0,14	0,20	0,11	0,12	0,11	0,15	0,15	0,16	0,06	0,18	0,09	0,25
	RMSE	k	0,00	0,02	0,00	0,02	0,00	0,01	0,00	0,01	0,01	0,02	0,00	0,01	0,01	0,02
		m	0,00	0,02	0,00	0,02	0,00	0,01	0,00	0,01	0,01	0,02	0,00	0,01	0,01	0,02
MAE	k	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,00	0,01
	m	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,00	0,01
MAPE	k	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,00	0,01
	m	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,01	0,00	0,00	0,01

Y: Yöntem; P: Performans, C: Bağlam; k: Anahtar Kelime; m: Makine Öğrenmesi; cvs: Çapraz Doğrulama Skoru; std: Standart Sapma

4.5.5.3.4 Tweetlerden Elde Edilen Öznitelikler ile BIST Pay Endekslerinin Volatilite Tahmini

Volatilite, bir menkul kıymetin fiyatının zaman içerisindeki dalgalanma oranının bir ölçüsüdür. Başka bir ifadeyle volatilite bir menkul kıymetin fiyat değişiklikleriyle ilişkili risk seviyesini göstermektedir. Fiyat dalgalanmaları ne kadar büyük ve sık olursa, piyasanın o kadar değişken olduğu anlamına gelmektedir.

Bu çalışma kapsamında tweetlerden elde edilen özniteliklerin BIST Pay Endekslerinin volatilitelerinin ne kadarını açıklayabildiğini de araştırılmıştır. Tablo 4.32’de, tweetlerden elde edilen öznitelikler ile BIST Pay Endekslerinin volatilitelerinin tahmin sonuçları yer almaktadır. Analiz, lineer regresyon, lasso regresyon, rassal orman ve XGBoost yöntemleri ile gerçekleştirilmiştir.

Tablo 4.32: BIST Pay Endekslerinin Volatiliteleri ve Twitter Öznitelikleri Makine Öğrenmesi Analiz Sonuçları

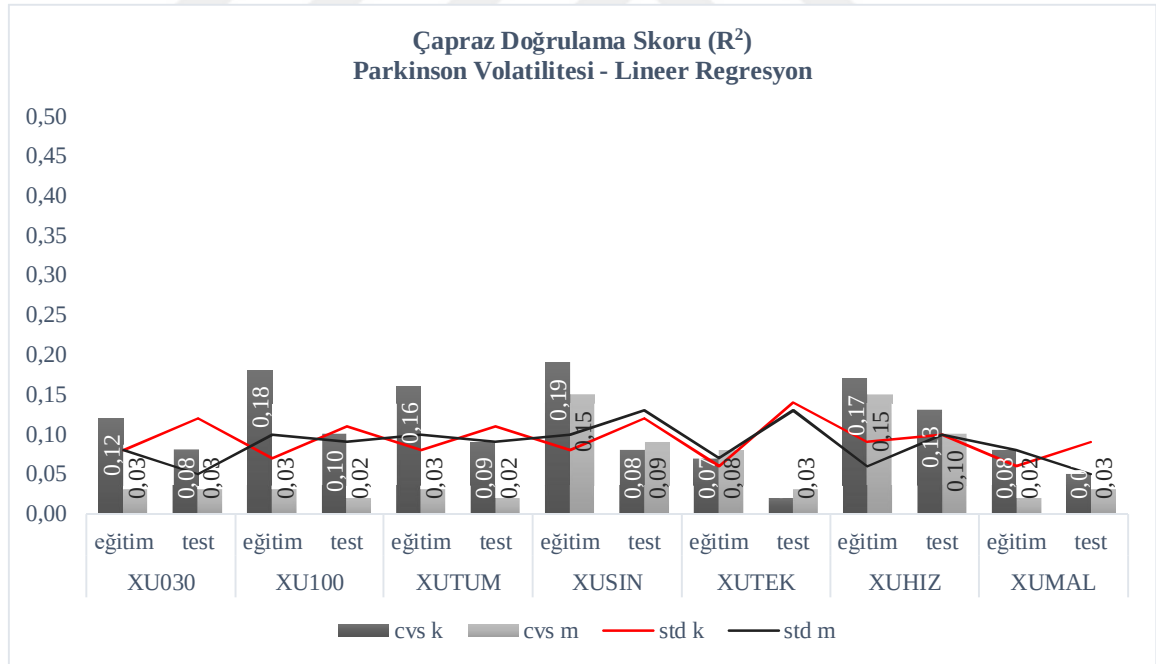
Parkinson Volatilitesi																
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Linear Regresyon	cvs (R ²)	k	0,12	0,08	0,18	0,10	0,16	0,09	0,19	0,08	0,07	0,02	0,17	0,13	0,08	0,05
		m	0,03	0,03	0,03	0,02	0,03	0,02	0,15	0,09	0,08	0,03	0,15	0,10	0,02	0,03
	std	k	0,08	0,12	0,07	0,11	0,08	0,11	0,08	0,12	0,06	0,14	0,09	0,10	0,06	0,09
		m	0,08	0,05	0,10	0,09	0,10	0,09	0,10	0,13	0,07	0,13	0,06	0,10	0,08	0,05
	RMSE	k	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,10	0,05	0,05	0,07	0,06
		m	0,07	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,10	0,06	0,05	0,07	0,07
	MAE	k	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,07	0,07	0,04	0,04	0,05	0,05
		m	0,05	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,07	0,07	0,04	0,04	0,05	0,05
	MAPE	k	0,27	0,27	0,28	0,27	0,28	0,28	0,29	0,30	0,39	0,37	0,27	0,26	0,27	0,27
		m	0,28	0,27	0,28	0,28	0,29	0,29	0,30	0,29	0,39	0,37	0,28	0,27	0,28	0,27
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Lasso Regresyon	cvs (R ²)	k	0,37	0,37	0,36	0,37	0,36	0,36	0,39	0,40	0,24	0,28	0,34	0,36	0,37	0,37
		m	0,29	0,34	0,28	0,33	0,29	0,33	0,31	0,39	0,26	0,38	0,26	0,32	0,31	0,34
	std	k	0,13	0,07	0,12	0,07	0,12	0,07	0,13	0,11	0,07	0,09	0,13	0,08	0,13	0,08
		m	0,17	0,09	0,16	0,09	0,15	0,09	0,14	0,12	0,09	0,08	0,17	0,13	0,15	0,07
	RMSE	k	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,07	0,07	0,03	0,03	0,04	0,04
		m	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,07	0,07	0,04	0,04	0,04	0,04
	MAE	k	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,05	0,05	0,03	0,03	0,03	0,03
		m	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,03	0,05	0,05	0,03	0,03	0,03	0,03
	MAPE	k	0,16	0,16	0,17	0,16	0,18	0,17	0,20	0,20	0,24	0,24	0,17	0,17	0,15	0,15
		m	0,17	0,16	0,18	0,17	0,19	0,17	0,21	0,20	0,24	0,24	0,18	0,18	0,16	0,15
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
Rassal Orman	cvs (R ²)	k	0,74	0,68	0,72	0,67	0,73	0,67	0,74	0,67	0,59	0,55	0,71	0,65	0,71	0,66
		m	0,66	0,67	0,65	0,67	0,65	0,66	0,67	0,66	0,48	0,53	0,63	0,63	0,65	0,66
	std	k	0,06	0,08	0,06	0,08	0,05	0,08	0,07	0,11	0,10	0,09	0,06	0,09	0,07	0,07
		m	0,07	0,06	0,08	0,06	0,07	0,06	0,07	0,09	0,09	0,08	0,08	0,10	0,07	0,06
	RMSE	k	0,01	0,03	0,01	0,03	0,01	0,02	0,02	0,03	0,03	0,05	0,01	0,02	0,02	0,03
		m	0,02	0,03	0,02	0,03	0,02	0,03	0,02	0,03	0,03	0,06	0,02	0,03	0,02	0,03
	MAE	k	0,01	0,02	0,01	0,02	0,01	0,02	0,01	0,02	0,02	0,03	0,01	0,02	0,01	0,02
		m	0,01	0,02	0,01	0,02	0,01	0,02	0,01	0,02	0,02	0,04	0,01	0,02	0,01	0,02
	MAPE	k	0,07	0,10	0,07	0,11	0,07	0,11	0,08	0,12	0,11	0,16	0,08	0,12	0,06	0,10
		m	0,08	0,11	0,08	0,11	0,08	0,12	0,10	0,13	0,12	0,18	0,08	0,12	0,08	0,11
Y	P	C	XU030		XU100		XUTUM		XUSIN		XUTEK		XUHIZ		XUMAL	
			eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test	eğitim	test
XGBoost	cvs (R ²)	k	0,70	0,60	0,68	0,62	0,67	0,58	0,71	0,61	0,59	0,49	0,69	0,57	0,66	0,58
		m	0,61	0,60	0,58	0,62	0,60	0,59	0,62	0,63	0,40	0,45	0,57	0,60	0,60	0,60
	std	k	0,05	0,11	0,07	0,10	0,06	0,09	0,06	0,13	0,10	0,13	0,04	0,12	0,05	0,09
		m	0,08	0,08	0,08	0,08	0,07	0,08	0,09	0,10	0,14	0,13	0,08	0,12	0,07	0,08
	RMSE	k	0,01	0,03	0,01	0,03	0,01	0,03	0,01	0,03	0,01	0,05	0,01	0,03	0,01	0,03
		m	0,01	0,03	0,01	0,03	0,01	0,03	0,01	0,03	0,01	0,06	0,01	0,03	0,01	0,03
	MAE	k	0,00	0,02	0,00	0,02	0,00	0,02	0,00	0,02	0,01	0,04	0,00	0,02	0,01	0,02
		m	0,01	0,02	0,00	0,02	0,00	0,02	0,00	0,02	0,01	0,04	0,00	0,02	0,01	0,02
	MAPE	k	0,03	0,11	0,03	0,11	0,03	0,12	0,03	0,13	0,05	0,17	0,03	0,12	0,03	0,11
		m	0,03	0,12	0,03	0,13	0,03	0,13	0,03	0,14	0,05	0,20	0,03	0,13	0,03	0,12
Y: Yöntem; P: Performans; C: Bağlam; k: Anahtar Kelime; m: Makine Öğrenmesi; cvs: Çapraz Doğrulama Skoru; std: Standart Sapma																

Y: Yöntem; P: Performans, C: Bağlam; k: Anahtar Kelime; m: Makine Öğrenmesi; cvs: Çapraz Doğrulama Skoru; std: Standart Sapma

4.5.5.3.4.1 Lineer Regresyon Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini

Tablo 4.32’de yer alan lineer regresyon modelinin çapraz doğrulama sonuçları Grafik 4.33’te sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin volatilitesindeki değişimin %8’sini, XU100 endeksinin volatilitesindeki değişimin %10’unu, XUTUM endeksinin volatilitesindeki değişimin %9’unu, XUSIN endeksinin volatilitesindeki değişimin %8’ini, XUTEK endeksinin volatilitesindeki değişimin %2’sini, XUHIZ endeksinin volatilitesindeki değişimin %13’ünü XUMAL endeksinin volatilitesindeki değişimin %5’ini açıklayabildiğini göstermektedir. Açıklayıcılıkları oldukça düşük ve tutarsızdır. Eğitim ve test verilerindeki bu çarpıcı farklılık lineer regresyon modelinin aşırı uyum sorunuyla karşılaştığını göstermektedir.

Grafik 4.33: Lineer Regresyon Çapraz Doğrulama Skoru (Parkinson Volatilitesi)

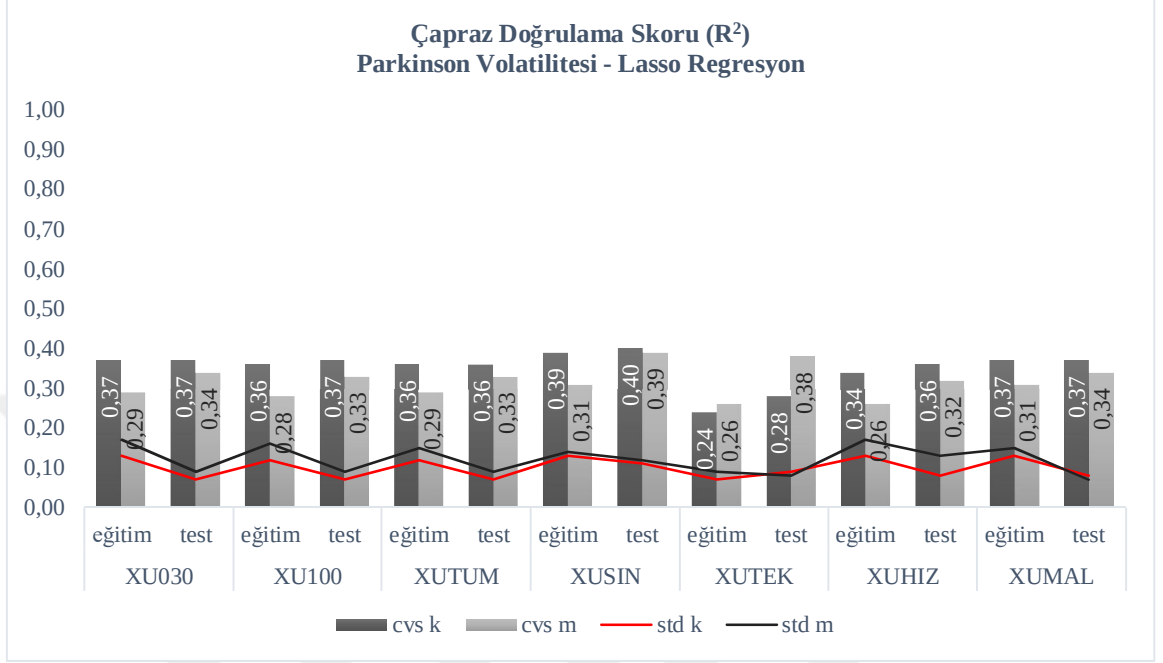


Lineer regresyon modelinin tahmin başarısı ayrıca Tablo 4.32’de görüldüğü üzere MAPE istatistiği ile ölçülmüştür. Bu sonuçlara göre BIST Pay Endeksleri için MAPE değerleri %20’nin üzerindedir. Özellikle XUTEK için bu skor %40’a yaklaşmaktadır. MAPE değerlerinin %50’nin altında olması her ne kadar kabul edilebilir olsa da açıklayıcılık düşük ve model aşırı uyum sorununa sahip olduğu için sonuçlar güvenilir değildir.

4.5.5.3.4.2 Lasso Regresyon Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini

Tablo 4.32’de yer alan lasso regresyon modelinin çapraz doğrulama sonuçları Grafik 4.34’te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin volatilitesindeki değişimin %37’sini, XU100 endeksinin volatilitesindeki değişimin %37’ini, XUTUM endeksinin volatilitesindeki değişimin %36’sını, XUSIN endeksinin volatilitesindeki değişimin %40’ını, XUTEK en endeksinin volatilitesindeki değişimin %28’ini, XUHIZ endeksinin volatilitesindeki değişimin %36’sını XUMAL endeksinin volatilitesindeki değişimin %37’sini açıklayabildiğini göstermektedir. Lasso regresyon analizinde de, lineer regresyonda elde edilen bulgulara yakın bulgular elde edilmiştir. Benzer şekilde tweetlerden elde edilen öznitelikler XUSIN endeksi için en yüksek açıklayıcılık düzeyine sahiptir. Lasso regresyon yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koyduğu görülmektedir.

Grafik 4.34: Lasso Regresyon Çapraz Doğrulama Skoru (Parkinson Volatilitesi)



Grafik 4.34'te görüldüğü üzere, eğitim ve test sonuçları birbirine çok yakın olması endekslerin volatilitelerini tahmin eden modelin aşırı uyum probleminin olmadığı anlamına gelmektedir.

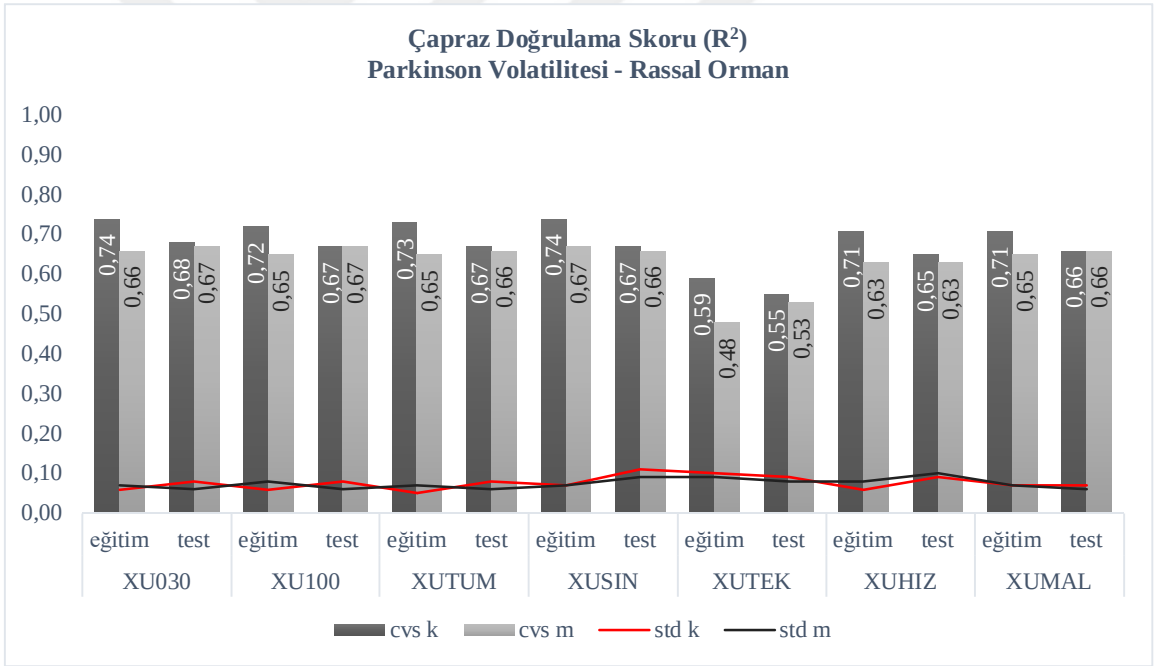
Lasso regresyon modelinin tahmin başarısı ayrıca Tablo 4.32'de görüldüğü üzere MAPE istatistiği ile ölçülmüştür. Bu sonuçlara göre BIST Pay Endeksleri için MAPE değeri %10 ile %20 arasındadır. Modelin doğru olduğu ancak *yüksek doğruluk derecesine* sahip olmadığı anlamına gelmektedir. MAPE değeri XUSIN ve XUTEK için %20 civarındır. Bu skor diğer kadar güçlü olmasa da kabul edilebilir bir değerdir.

4.5.5.3.4.3 Rassal Orman Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini

Tablo 4.32'de yer alan rassal orman modelinin çapraz doğrulama sonuçları Grafik 4.35'te görsel olarak sunulmaktadır. Buna göre tweetlerden elde edilen özneliliklerin XU030

endeksinin volatilitesindeki değişimin %68'ini, XU100 endeksinin volatilitesindeki değişimin %67'sini, XUTUM endeksinin volatilitesindeki değişimin %67'sini, XUSIN endeksinin volatilitesindeki değişimin %67'sini, XUTEK endeksinin volatilitesindeki değişimin %55'ini, XUHIZ endeksinin volatilitesindeki değişimin %65'ini XUMAL endeksinin volatilitesindeki değişimin %66'sını açıklayabildiğini göstermektedir. Analizin dikkat çekici bulgularından biri diğerlerine kıyasla XUTEK endeksinin en düşük açıklayıcılık değerine sahip olmasıdır. Rassal orman yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısında benzer bir performans ortaya koyduğu görülmektedir.

Grafik 4.35: Rassal Orman Çapraz Doğrulama Skoru (Parkinson Volatilitesi)



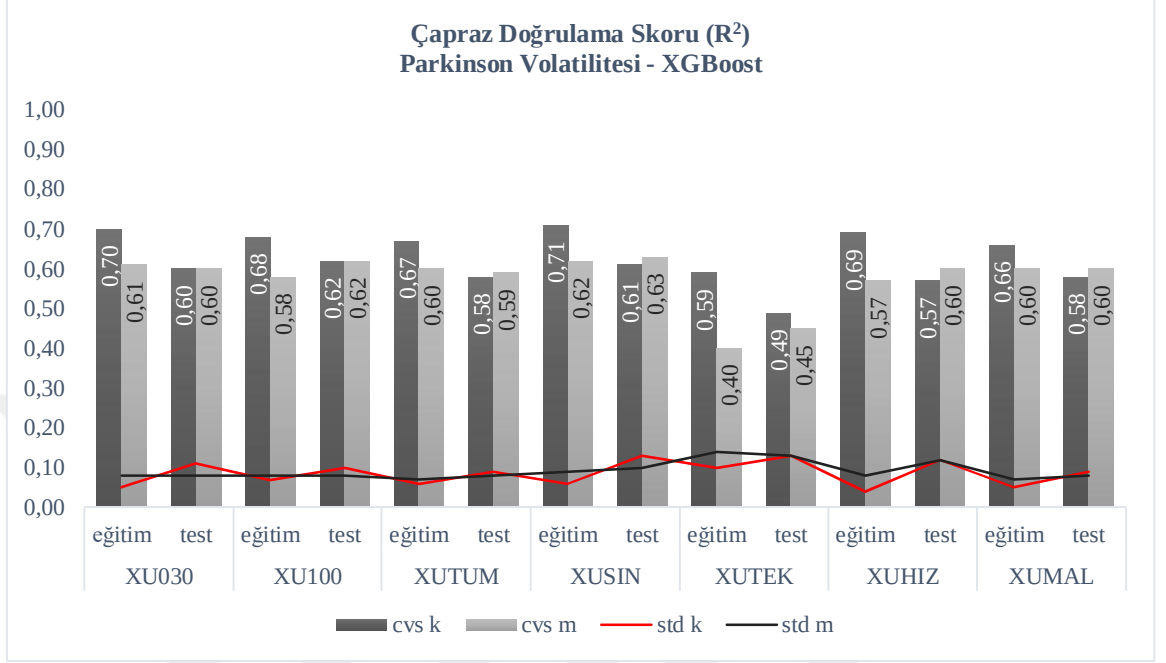
Grafik 4.35'te, görüldüğü üzere, eğitim ve test sonuçlarının birbirine yakın olması rassal orman yöntemi ile endekslerin volatilitelerini tahmin eden modelin aşırı uyum probleminin olmadığını göstermektedir. Özellikle makine öğrenmesi yöntemiyle sınıflandırılan tweetlerle yapılan analizlerin aşırı uyum problemine daha dirençli olduğu görülmektedir.

Rassal orman modelinin tahmin performansını gösteren MAPE istatistiği Tablo 4.32’de yer almaktadır. Bu sonuçlara göre, rassal orman modelinde tüm endeksler için MAPE değeri %10’un hemen üzerindedir. Dolayısıyla model bütün endeks değerleri için *doğruluk derecesine* sahiptir.

4.5.5.3.4.4 XGBoost Yöntemi ile BIST Pay Endekslerinin Volatilitelerinin Tahmini

Tablo 4.32’de yer alan XGBoost modelinin çapraz doğrulama sonuçları Grafik 4.36’da sunulmaktadır. Buna göre tweetlerden elde edilen özniteliklerin XU030 endeksinin volatilitesindeki değişimin %60’ını, XU100 endeksinin volatilitesindeki değişimin %62’sini, XUTUM endeksinin volatilitesindeki değişimin %58’ini, XUSIN endeksinin volatilitesindeki değişimin %61’ini, XUTEK endeksinin volatilitesindeki değişimin %49’unu, XUHIZ endeksinin volatilitesindeki değişimin %57’sini XUMAL endeksinin volatilitesindeki değişimin %58’ini açıklayabildiğini göstermektedir. XGBoost yönteminde model, XUSIN endeksi için en yüksek açıklayıcılık gücüne, XUTEK için en düşük açıklayıcılık gücüne sahiptir. XGBoost yönteminde de anahtar kelime tabanlı bağlam sınıflandırıcı (k) ile makine öğrenmesi tabanlı bağlam sınıflandırıcının (m) modelin tahmin başarısının birbirine yakın olduğu, ancak makine öğrenmesi tabanlı bağlam sınıflandırıcının kullanıldığı modelin aşırı uyum problemine daha dayanıklı olduğu görülmektedir.

Grafik 4.36: XGBoost Çapraz Doğrulama Skoru (Parkinson Volatilitesi)



Grafik 4.36’da görüldüğü üzere, XGBoost modeliyle elde edilen eğitim ve test verileri arasındaki yakın sonuçlar, modelin aşırı uyum probleminin olmadığını göstermektedir. Özellikle makine öğrenmesi tabanlı sınıflandırıcıların kullanıldığı modellerin aşırı uyum problemine çok daha dirençli olduğu görülmektedir.

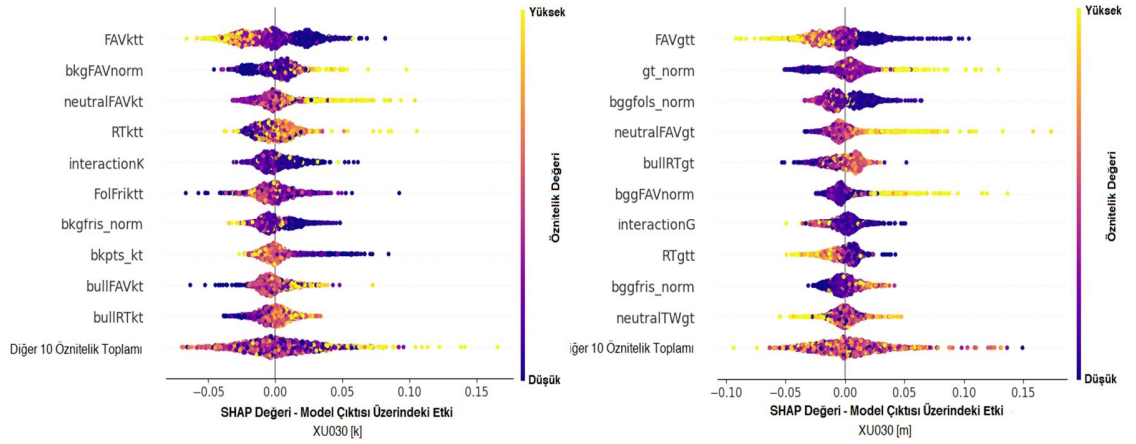
Tablo 4.32’de XGBoost modelinin tahmin başarısını gösteren MAPE istatistiği sonuçları yer almaktadır. Bu sonuçlara göre, XGBoost modelinde tüm endeksler için MAPE değeri %10’un hemen üzerindedir. Dolayısıyla model bütün endeks değerleri için *doğruluk derecesine* sahiptir. MAPE değerlerinin eğitim ve test verileri arasındaki farklılığı da gözden uzak tutulmamalıdır.

4.5.5.3.4.5 Volatilitelerin Tahmininde Öznitelik Önem Düzeyleri

Bu çalışma kapsamında son olarak tweetlerden elde edilen özniteliklerin, endekslerin volatiliteleri ile ilgili modelin açıklayıcılığına katkısı yine SHAP değerleri yardımıyla

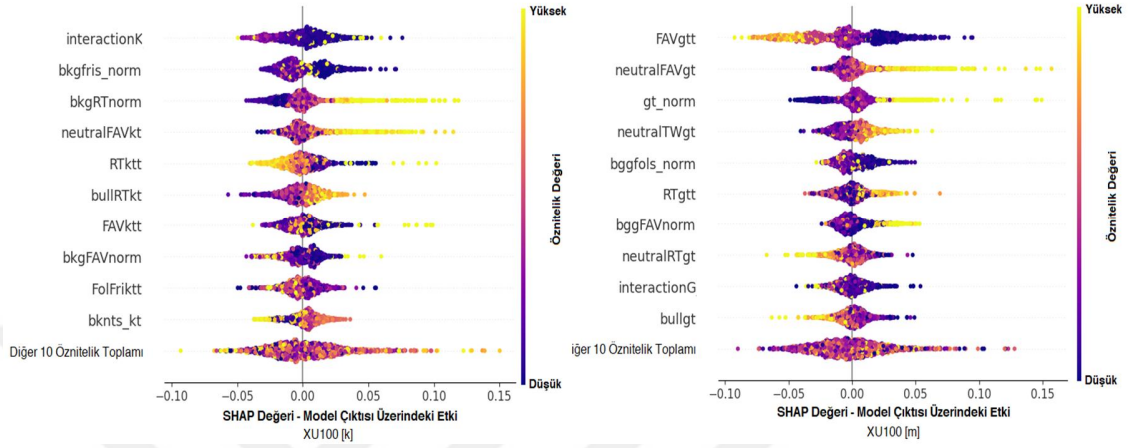
araştırılmıştır. Grafik 4.37’de, XU030 endeksinin volatilitesi için XGBoost tahmin modelinde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bu çerçevede modele en çok katkısı olan özniteliklerin önceki bulgulara benzer şekilde özellikle etkileşime dayalı öznitelikler olduğu (bggFAVnorm, bggfris_norm, RTktt, FolFrigtt, FAVktt, bkgfols_norm, interactionK) söylenebilir. Burada dikkat çekici husus, beğeni tweetlerin ve retweetlerin yüksek değerlerinin modele negatif katkı sağlarken, düşük değerlerinin modele pozitif katkı sağlamasıdır. Bu durum retweet ve beğeni tweetlerin arttığında benzer fikirlerin daha fazla dolaşımda olacağı anlamına geldiği için volatilitenin düşmesi beklenen bir durumdur. Ayrıca yükseliş potansiyeli özniteliklerinin (bullFAVkt, bullRTkt) yüksek değerlerinin modele pozitif katkı sağlarken düşük değerlerinin modele negatif katkı sağladığı görülmektedir.

Grafik 4.37: XU030 Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Önemini Arı Sürüsü Grafiği ile Gösterimi



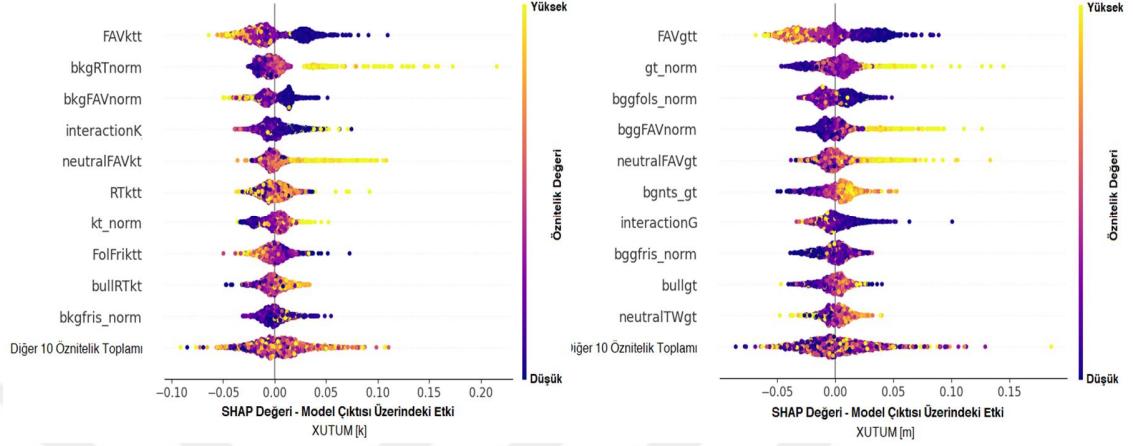
Grafik 4.38’de, XU100 endeksi için XGBoost yöntemiyle volatilitate tahmini modelinde özniteliklerin önemini arı sürüsü grafiği ile gösterimi yer almaktadır. Bu bağlamda modele en fazla katkı sağlayan özniteliklerin önceki bulgulara benzer şekilde özellikle etkileşime dayalı öznitelikler olduğu (bggFAVnorm, bggfris_norm, RTktt, RTgtt, FolFrigtt, FAVktt, bkgfols_norm, interactionK, interactionG) söylenebilir. Ayrıca yükseliş potansiyelinin (bullgt) yüksek değerlerinin modele negatif katkı sağlarken düşük değerlerinin modele pozitif katkı sağladığı görülmektedir.

Grafik 4.38: XU100 Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



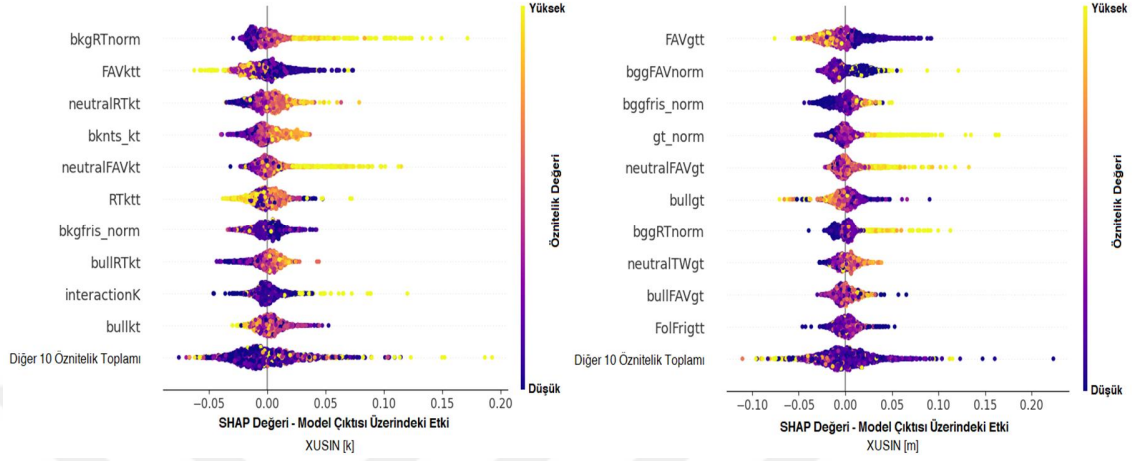
Grafik 4.39’de, XUTUM endeksi için volatilité modelinde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular modelin açıklayıcılığında beğeni tweetlerin toplam tweetlere oranı (FAVkt, FAVgt), retweetlerin toplam tweetlere oranı (RTgt, RTkt), beğeni tweet sayısı (bkgFAVnorm, bggFAVnorm), retweet sayısı (bkgRTnorm), arkadaş sayısı (bkgfris_norm, bggfris_norm), etkileşim skoru (interactionK, interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Bunlara ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt, neutralFAVgt) ve toplam tweet sayısı (gt_norm, kt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır. Ayrıca yükseliş potansiyeli indeksleri (bullgt, bullRTkt) de modelin açıklayıcılığına katkı sunmaktadır.

Grafik 4.39: XUTUM Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



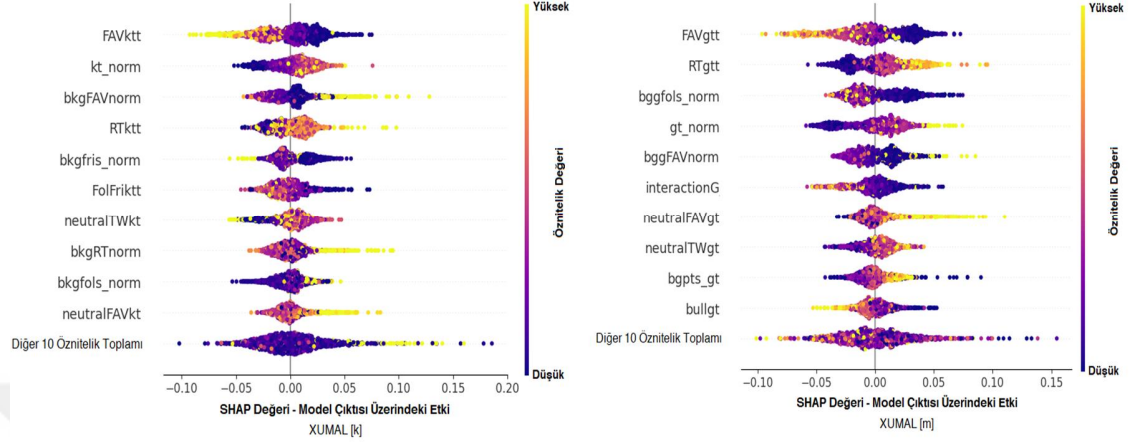
Grafik 4.40’da, XUSIN endeksi için XGBoost yöntemiyle volatilité tahmini modelinde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular modelin açıklayıcılığında retweet sayısı (bkgRTnorm, bggRTnorm), beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), retweetlerin toplam tweetlere oranı (RTktt), beğeni tweet sayısı (bggFAVnorm), arkadaş sayısı (bkgfris_norm, bggfris_norm), etkileşim skoru (interactionK) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Bunlara ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt, neutralFAVgt), nötr retweetlerin toplam tweetlere oranı (neutralRTkt) ve toplam tweet sayısı (gt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır. Ayrıca yükseliş potansiyeli indeksleri (bullgt, bullkt, bullRTkt, bullFAVgt) de modelin açıklayıcılığına katkı sunmaktadır.

Grafik 4.40: XUSIN Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



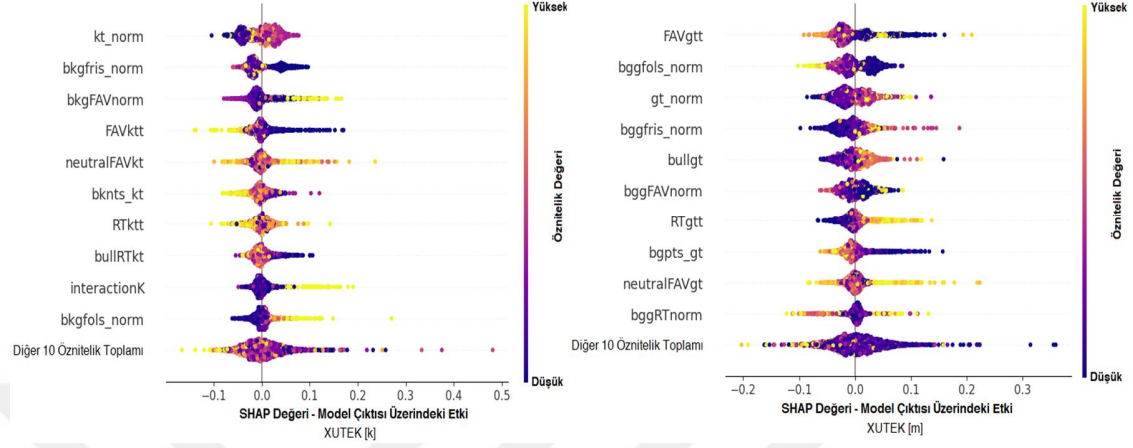
Grafik 4.41’de, XUMAL endeksi için volatilité tahmini modelinde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular modelin açıklayıcılığında beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), retweetlerin toplam tweetlere oranı (RTktt, RTgtt), beğeni tweet sayısı (bkgFAVnorm), arkadaş sayısı (bkgfris_norm), takipçi sayısı (bkgfols_norm, bkgfols_norm), etkileşim skoru (interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Bunlara ek olarak nötr tweetlerin toplam tweetlere oranı (neutralTWkt, neutralTWgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt, neutralFAVgt) ve toplam tweet sayısı (kt_norm, gt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır. Ayrıca yükseliş potansiyeli endeksi (bullgt) de modelin açıklayıcılığına katkı sunmaktadır.

Grafik 4.41: XUMAL Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



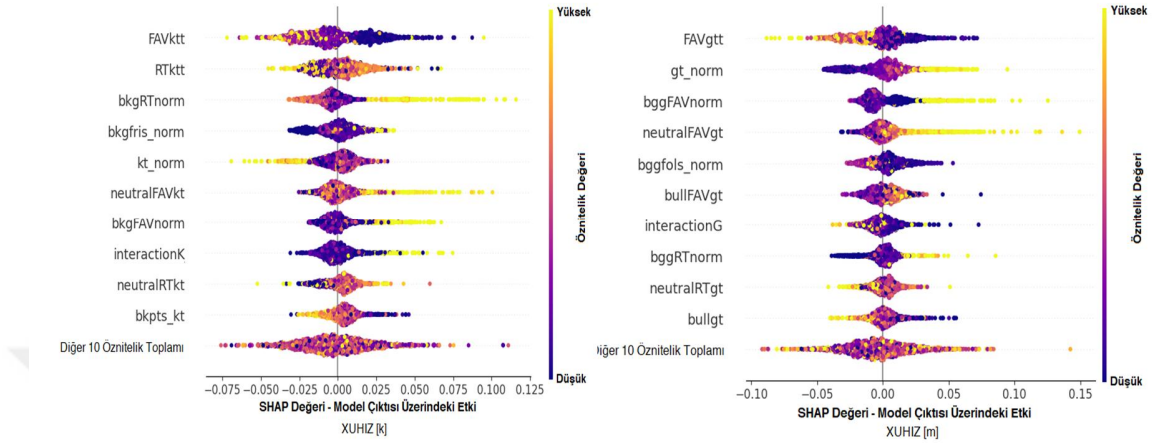
Grafik 4.42’de, XUTEK endeksinin volatilitesinin XGBoost yöntemiyle tahmini için kurulan modelde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bulgular modelin açıklayıcılığında beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), beğeni tweet sayısı (bkgFVnorm, bggFVnorm), retweetlerin toplam tweetlere oranı (RTgtt, RTktt), retweet sayısı (bggRTnorm), arkadaş sayısı (bkgfris_norm, bggfris_norm), etkileşim skoru (interactionK) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Ayrıca yükseliş potansiyeli indeksleri (bullgt, bullRTgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt, neutralFAVgt) ve toplam tweet sayısı (gt_norm, kt_norm) da XUTEK endeksinin volatilitesinin açıklanmasında önemli düzeyde katkısı olan özniteliklerdendir.

Grafik 4.42: XUTEK Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüsü Grafiği ile Gösterimi



Grafik 4.43'te, XUHIZ endeksinin volatilitésinin tahmini için kurulan modelde özniteliklerin öneminin arı sürüsü grafiği ile gösterimi yer almaktadır. Bu bağlamda modele en çok katkı sağlayan öznitelikler beğeni tweetlerin toplam tweetlere oranı (FAVktt, FAVgtt), retweetlerin toplam tweetlere oranı (RTktt), beğeni tweet sayısı (bggFAVnorm), retweet sayısı (bkgRTnorm, bggRTnorm), arkadaş sayısı (bkgfris_norm), takipçi sayısı (bggfol_s_norm), etkileşim skoru (interactionG) ve benzeri etkileşim göstergelerinin önemli bir payı olduğuna işaret etmektedir. Bunlara ek olarak nötr retweetlerin toplam tweetlere oranı (neutralRTkt, neutralRTgt), nötr beğeni tweetlerin toplam tweetlere oranı (neutralFAVkt, neutralFAVgt) ve toplam tweet sayısı (kt_norm, gt_norm) da modelin açıklayıcılığına katkısı olan öznitelikler arasındadır. Ayrıca yükseliş potansiyeli indeksleri (bullgt, bullFAVgt) de modelin açıklayıcılığına katkı sunmaktadır.

Grafik 4.43: XUHIZ Endeksi Parkinson Volatilitesi için XGBoost Tahmin Modelinde Özniteliklerin Öneminin Arı Sürüşü Grafiği ile Gösterimi



Genel olarak bulgular, belirli bir günde borsa ile ilgili tweetlerden elde edilen özniteliklerin özellikle retweetlerin toplam tweetlere oranı, beğeni tweetlerin toplam tweetlere oranı, retweet sayısı, beğeni tweet sayısı, tweet sayısı, takipçi sayısı, arkadaş sayısı, etkileşim skoru gibi etkileşime dayalı olanlarının BIST Pay Endekslerinin volatilitelerini açıklamada önemli katkılarına olduğuna işaret etmektedir. Ayrıca modelin açıklayıcılığında bilişsel boyutta yer alan yükseliş potansiyeli indeksi ile bilişsel, davranışsal ve sosyal boyutları olan nötr tweetlerle ilgili göstergelerin de önemli bir paya sahip olduğu görülmektedir. Bu bulgular, borsa bağlamında olduğu belirlenen tweetlerin hem fikri ve duygusal içeriklerinin hem içerikten bağımsız olarak sayılarının hem de bu tweetlerin kullanıcılar arasında, yani sosyal bir çevrede, sosyal bir etkileşim doğuracak şekilde paylaşılmasının, BIST Pay Endekslerindeki işlem hacimlerini etkilediğine işaret etmektedir.

4.6 BULGULARIN DEĞERLENDİRİLMESİ VE ÖNERİLER

Davranışsal finans, geleneksel finans yaklaşımlarının soyut modellerinden dışlanmış olan insanın psikolojik ve sosyolojik gerçekliğini çözümlemenin odağına yerleştirme çabasının bir sonucudur. Başka bir ifadeyle davranışsal finans sınırlı rasyonaliteye ve davranışsal eğilimlere açık bir insanın finansal kararlarına odaklanmaktadır. Bu perspektife sahip ilk

nesil çalışmalarda rasyonalite standart kabul edilmiş, rasyonaliteden sapmaya neden olan bilişsel ve duygusal eğilimler irrasyonel birer kusur olarak görülmüştür (Statman, 2019). İkinci nesil çalışmalar ise insanın gerçekliği ile daha fazla bağdaşan hipotezler ortaya koyarak kusurları, eğilimleri ve istekleri birbirinden ayırıp insana kusurları hususunda yol gösterici olmuştur.

Davranışsal finans çalışmalarının ilk iki kuşağında, geleneksel finans yaklaşımlarının rasyonalite varsayımı gevşetilmiş, ancak insanların sosyal etkiden korunmuş birer karar vericiler olduğu varsayımı büyük oranda varlığını sürdürmüştür (Kamoune ve Ibenrissoul, 2022). Oysa insan, bilişsel ve davranışsal eğilimlerinin yanı sıra bir toplum içerisinde yaşayan, çevresini etkileyen ve çevresinden etkilenen bir varlıktır. Önceki kuşak çalışmalarda insanın toplumsal yönüne sürü psikoloji bağlamında kısmen kapı aralansa da, bu çalışmalardaki insan modeli büyük oranda, bilinçsizce diğerlerini taklit eden bir nitelikte tezahür etmektedir. Başka bir ifadeyle insanın toplumsal yönü, ilk iki kuşak çalışmalara tam olarak nüfuz edememiştir. Bu nedenle davranışsal finans çalışmalarında ortaya konan psikolojik faktörlerin insan kararlarına etkisine ilave olarak sosyolojik faktörlerin etkisinin de çözümlemenin odağında daha fazla yer alması gerekmektedir. Bu çalışmanın amaçlarından biri bireysel yatırımcıların finansal kararlarında sosyolojik etkiyi analiz etmek olmuştur. Sosyal medya platformları, bu anlamda bireylerin ya da bireysel yatırımcıların sosyal çevreleriyle etkileşimlerini gözlemlemek için önemli araçlar sunmaktadır.

Davranışsal finans yaklaşımlarının ilk iki kuşağında insanın her ne kadar tam bilgiye sahip olmadığı kabul edilse de, eksik bilgiye sahip insan gerçekliği de çözümlemenin merkezine tam olarak yerleşememiştir. Başka bir deyişle insanın eksik bilgiye sahip olduğu gerçekliği davranışsal finans çalışmalarının yarı güçlü formda etkinlik testleriyle ortaya konulmuş, yani eksik bilginin yarı güçlü formda piyasanın mimarı olduğu tespit edilmiş, ancak eksik bilginin insanın doğasına yerleşik olmasının sonuçları yeterince araştırılamamıştır. Bu bağlamda bilginin bireyler arasında eksik ve parçacıklı bir şekilde dağılmış olması (Hayek 1937; 1945), her biri yalnızca kısmî bilgiye sahip olan insanların etkileşimlerine ve bilginin iletim yollarına, yani farklı zihinlerde var olan bilginin insandan insana nasıl aktarıldığına ve bu

mekanizmanın insanların kararlarındaki etkisine odaklanmayı gerekli kılmaktadır. Dolayısıyla bu çalışmadaki amaçlardan bir diğeri insanların eksik bilgiye sahip olmasının, sosyal medya mecralarında bilgi arayışlarına ve elde ettiği bilginin kararlarına etkisini araştırmak olmuştur. Zira insanların birbirleri ile iletişim kurmalarına olanak sağlayan sosyal medya platformları, bu anlamda da bireylerin ya da bireysel yatırımcıların bilgi paylaşımlarını ve arayışlarını gözlemlemek için önemli araçlar sunmaktadır. Örneğin duygu ya da fikir içeren bir tweetin paylaşılması ve daha sonra bu tweetin retweet edilmesi ya da beğenilmesi o içeriğin bireysel kararlardaki etki potansiyelinin bir işareti olarak kabul edilebilir. Nitekim araştırmadaki bulgulardan biri, retweet ve beğeni oranının BIST Pay Endekslerinin değeri, işlem hacmi ve volatilitelerini açıklayabildiğini göstermektedir. Bu bulgu aynı zamanda sosyal etkinin de bir göstergesidir. Zira bir paylaşımın beğenilmesi ya da retweet edilmesi sosyal bir çevrede gerçekleşen bilgi alışverişinin bir tezahürüdür. Ayrıca tweetin retweet edilmesi ya da beğenilmesi, tweetlerin içeriğinde yer alan her türlü bilginin tweeti paylaşan kişinin sosyal çevresinin ötesine taşınmasını sağlar. Bilgi, insanlar arasındaki iletişim kanalları üzerinden akıp giderken geçtiği yerlerde kararlara etki etme potansiyeline sahiptir. Böylece hem eksik bilgiye sahip hem de sosyal bir çevreye yerleşik olan insanın bilgi arayışlarının eylem ver davranışlarına etkisini gözlemleme imkânı doğmuştur.

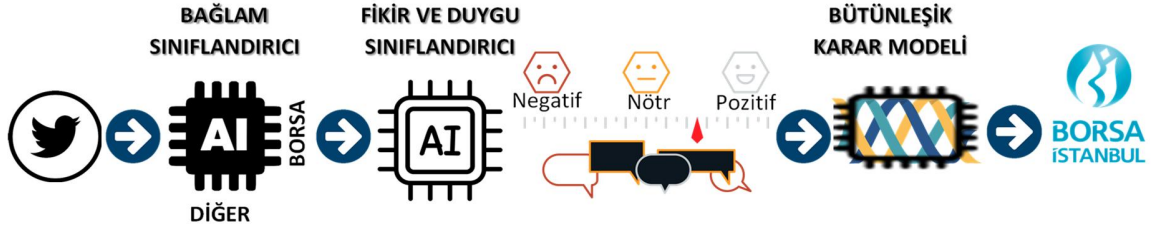
Davranışsal finansın eleştirilen bir diğer yönü insanların bilinçli eylemlerinden ziyade bilişsel kestirmelerine ya da bir takım davranışsal eğilimlerine odaklanılmış olmasıdır. Bu bağlamda bir diğer gerçeklik, etki tepki, dürtüsellik ya da eğilimleri nedeniyle şüursuzca ortaya çıkan davranışlardan ziyade, ekonominin bilinçli yaşayan insanlardan oluştuğu ve bu insanların kendi eylemlerini iradeli bir şekilde icra ettikleridir. Bu bağlamda Shiller'ın (2019/2021) anlatı iktisadi hipotezinde vurguladığı gibi, duygu ve fikir yüklü anlatıların insanların hem iktisadi beklenti ve inançlarının inşasındaki hem de iktisadi kararlarındaki payı göz ardı edilemez bir gerçekliktir. İnsanın yönelimsel eyleminin çözümlemeye dâhil edilmesi davranışsal finansa bilişsel bir boyut kazandırır. Dolayısıyla bu çalışmanın amaçlarından bir diğeri, insanların sosyal medya mecralarında paylaştıkları finans içerikli metinlerdeki anlamın (fikri ve duygusal yönelimin) kararlara etkisini araştırmak olmuştur. Bu anlamda sosyal medya, bireylerin ya da bireysel yatırımcıların fikir ve duygu dünyasına,

geleceğe dair beklenti ve inançlarına açılan bir pencere niteliğindedir. Ancak finans içerikli milyonlarca metnin fikri ve duygusal yönelimlerinin belirlenebilmesi insan gücünü aştığından yapay zekâ desteği ile bağlam, fikir ve duygu sınıflandırıcıları geliştirilmiş ve böylece bilgi arayışı içerisindeki insanın bilinçli kararlarını şekillendirme potansiyeline sahip olan duygu ve bilgi yüklü metinlerin insanların yönelimsel eylemlerine etkisi çözümlemeye dâhil edilmiştir.

Analizde 01.01.2012-28.02.2020 döneminde paylaşılan tweetler toplanmıştır. Tweetlerin borsa ile ilgili olup olmadığını belirlemek için hem anahtar kelime tabanlı hem makine öğrenmesi tabanlı iki farklı bağlam sınıflandırıcı geliştirilmiştir. Makine öğrenmesi tabanlı bağlam sınıflandırıcı derin öğrenme yaklaşımlarından GRU yöntemi ile eğitilmiş ve %98 sınıflandırma başarısı elde edilmiştir. Bağlam sınıflandırıcılar yardımıyla anahtar kelime tabanlı yaklaşıma göre 953.965.658 tweetin, makine öğrenmesi yaklaşımına göre 90.344.468 tweetin borsa ile ilgili olduğu belirlenmiştir. Tweetlerin fikri ve duygusal yöneliminin pozitif, negatif ya da nötr olarak sınıflandırılabilmesi için fikir ve duygu sınıflandırıcısı, ön eğitilmiş BERT yöntemiyle eğitilmiş ve pozitif ve negatif sınıflar için %93, nötr sınıf için %89 sınıflandırma başarısı elde edilmiştir. Borsa ile ilgili olduğu saptanan tweetlerin fikir ve duygu sınıflandırıcısı yardımıyla duygusal yönelimleri belirlenmiştir.

Bir tweet içerdiği fikri ve duygusal anlam bağlamında *bilişsel*, insanların birbirleri arasında paylaşılması ve bir etkileşim doğurması bağlamında *sosyal*, belirli eğilimleri tetikleyebilmesi ya da belirli şeylere tepki niteliğinde doğması bağlamında *davranışsal* boyutlara sahiptir. Çoğu durumda bu üç boyut iç içe geçmiş bir şekilde tezahür ettiğinden bu etkileri birbirinden ayırt edebilmek pek mümkün olamamaktadır. Bu nedenle bu üç boyutu temsil edebilecek nitelikte Twitter'dan elde edilen özniteliklerin BIST Pay Endeksleri ile ilişkisi bütünleşik karar modeli ile test edilmiştir. Analizde farklı avantaj ve dezavantajlara sahip, doğrusal ve doğrusal olmayan dört farklı makine öğrenmesi yöntemi kullanılmıştır. Burada ifade edilen ve çalışma kapsamında kullanılan bireysel yatırımcı fikir ve duygu analizi yaklaşımının bir özeti Şekil 4.21'de yer almaktadır.

Şekil 4.21: Bireysel Yatırımcı Duygu ve Fikir Analizi Yaklaşımı Özeti



Bütünleşik karar modellerinden elde edilen bulgular, sosyal medya verilerinden elde edilen özniteliklerle BIST Pay Endeksleri arasında anlamlı ve güçlü bir ilişki olduğunu ortaya koymaktadır. Başka bir ifadeyle sosyal medyadan elde edilen özniteliklerin, özellikle rassal orman ve XGBoost gibi topluluk öğrenmesine dayalı ve doğrusal olmayan yöntemler ile BIST Pay Endekslerinin açılış değerlerini %90'nın üzerinde açıklayıcılığa sahip olduğunu ortaya koymaktadır.

Çalışmanın bir diğer bulgusu işlem hacimleri ile ilgilidir. Belirli bir günde borsa bağlamında paylaşılan tweetlerin fikri ve duygusal yönelimlerinden elde edilen yükseliş potansiyeli indeksi ile işlem hacmi arasında Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulgularına benzer şekilde pozitif ilişkili olduğu sonucuna ulaşılmıştır. Ancak söz birliği indeksi için literatürle benzer nitelikte bir bulguya ulaşılamamıştır. Bu bulgu tam fikir birliği olmadığı gözönünde bulundurularak değerlendirildiğinde, esasında beklentilere aykırı değildir. Buna ek olarak olumsuz fikirlerin yaygınlaşmasının işlem hacmini düşürebileceği, beklentilere uygun bir şekilde ortaya konulmuştur. Ayrıca yatırımcı ilgisi bağlamında düşünüldüğünde belirli bir günde borsa bağlamında paylaşılan tweetlerin artması Wei ve diğerleri (2016), Alostad ve Davulcu (2017), Garcia-Lopez ve diğerlerinin (2018) de ortaya koyduğu gibi işlem hacimlerini artırmaktadır. Sosyal etkileşim boyutunda ise hem takipçi sayısı hem retweet ve beğeni tweetleri sayısı hem de sosyal etkileşim skoru beklentilere uygun şekilde işlem hacmi ile pozitif ilişkilidir. Bütünleşik tahmin modelinde ise tweetlerden elde edilen bu özniteliklerin işlem hacmi için %70'in üzerinde açıklayıcılığa sahip olduğunu göstermektedir.

Çalışmanın bir diğer bulgusu BIST Pay Endekslerinin volatilitesi ile ilgilidir. Volatilite esasında heterojen inançların ve belirsizliğin bir ölçüsüdür. Yatırımcılar gelecek beklentileri ve fiyatlar konusunda hemfikir olmadıklarında volatiliteler yüksek olacaktır. Bu çerçevede belirli bir günde borsa ile ilgili paylaşılan tweetlerin duygusal yönelimini gösteren hem yükseliş potansiyeli indeksi ile hem de söz birliği indeksi, oldukça düşük, fakat anlamlı ve beklentilere uygun şekilde negatif bir ilişki olduğuna işaret etmektedir. Bu bulgu, Antweiler ve Frank (2004) ile Sprenger ve diğerlerinin (2014) bulgularına benzer şekilde yükseliş potansiyeli azaldığında ya da söz birliği azaldığında volatilitenin artacağı anlamına gelmektedir. Yatırımcı ilgisi bağlamında tweet sayısı ile volatiliteler arasındaki ilişkiye bakıldığında ise, borsa ile ilgili tweetlerin sayısı ile volatilitenin de aynı yönde artış ya da azalış göstereceğine işaret etmektedir. Bu bağlamda tweet sayısı ile volatiliteler arasındaki pozitif yönlü bu bulgu, daha çok tweet paylaşılmasının fikri ve duygusal anlamda farklı düşüncelerin artması anlamına gelebileceğinden, fiyatların daha volatil olması da beklenen bir durumdur. Benzer bulgular Wei ve diğerleri (2016), Alstad ve Davulcu (2017), Garcia-Lopez ve diğerleri (2018) tarafından da raporlanmıştır. Volatilitenin bir diğer boyutu da etkileşim ile ilgilidir. Bu bağlamda volatiliteler ile retweet ve beğeni tweetlerin sayısı, takipçi sayısı ve etkileşim skoru pozitif yönlü bir ilişkiye sahip olduğu bulgusuna ulaşılmıştır. Bu durum etkileşim ölçütlerindeki artış ya da azalışların volatiliteleri aynı yönde etkilediği anlamına gelmektedir. Bütünleşik tahmin modelleri ise tweetlerden elde edilen bu özniteliklerin volatiliteler için %60'ın üzerinde açıklayıcılığa sahip olduğunu göstermektedir.

Son olarak sosyal medya öznitelikleri ile BIST Pay Endekslerinin getirileri arasındaki ilişkisi araştırılmış, ancak getiriler üzerinde (toplam tweet ve toplam retweet sayısı ile negatif ve önemsiz düzeyde zayıf bir ilişki dışında) anlamlı ve güçlü bir etkinin olmadığı bulgusuna ulaşılmıştır. Bu durum sosyal medya öznitelikleri ile volatiliteler ve işlem hacmi arasında olduğu ortaya konan ilişki çerçevesinde düşünüldüğünde, sosyal medya paylaşımlarının daha yüksek işlem hacmi ve volatiliteler ile ilişkili olduğu, daha yüksek işlem hacmi ve volatilitenin ise endeks düzeyinde getirinin sosyal medya öznitelikleri ile açıklanabilirliğini anlamsız kıldığı düşünülebilir. Bu durum aynı zamanda analizin endeks düzeyinde yapılmış olmasından da kaynaklanıyor olabilir. Zira literatürde sosyal medya

öznitelikleri ile getiri arasında ilişki olduğu bulgusu, ekseriyetle olay çalışması yöntemiyle ve hisse senedi düzeyinde yapılan çalışmalarda raporlanmıştır. Dolayısıyla, gelecekteki çalışmalarda analizin ölçeğinin değiştirilerek hisse senedi düzeyinde tekrarlanması halinde sosyal medya özniteliklerinin hisse senedi getirisi ile ilişkisi bağlamında farklı bulgulara ulaşılabilir.

Bulgular sosyal medya aracılığıyla hisse senedi piyasalarının hem öngörülebileceğine hem de kontrol edilebileceğine işaret etmektedir. Bu bağlamda öngörülebilirliğin iki boyutu vardır. Bunlardan birincisi yatırımcıların duygu ve düşüncelerini sosyal medyada paylaştıkları için ne yapacaklarının tahmin edilebilirliğidir. Bu açıdan sosyal medya, yatırımcıların duygu ve düşüncelerinin yansıdığı bir gözlem yeridir. İkinci anlamda öngörülebilirlik, eksik bilgiye sahip yatırımcıların sosyal medya platformlarında yatırım fikirleri arayışları neticesinde kararlarının şekillenmesiyle ilgilidir. Sosyal medya kullanıcılarının, sosyal medyayı bilgi aramak için kullandıkları istatistiklerle birlikte düşünüldüğünde, bu öngörülebilirliğin daha geçeye uygun olduğu düşünülmektedir. Bu çerçevede öngörülebilirlik kontrol edilebilirliğe de kapı aralamaktadır.

Bu bulgular, genel anlamda insanların özel anlamda bireysel yatırımcıların eksik bilgiye sahip olduklarının ve eksik bilgilerini tamamlamak için sosyal medya mecralarında bilgi arayışına yöneldiklerinin bir göstergesidir. Bu bağlamda bireysel yatırımcıların sosyal medyada aradığı bilgi türü *teorik bilgi*, *pratik bilgi* ve *hazır bilgi* olarak üç gruba ayrılabilir.

Teorik bilgi bir konunun uzmanlığına dair bilgi türüdür. Örneğin finans konusunda herhangi bir eğitim almamış bireysel yatırımcıların, kişisel finans yönetimi için temel finans bilgisi konusunda paylaşımlar yapan hesapları takip etmesi teorik bilgi arayışına tekabül etmektedir. Bu çerçevede düşünüldüğünde sosyal medya finansal okuryazarlığın artırılması hususunda önemli katkılar sunma potansiyeline sahiptir.

Bireysel yatırımcıların sosyal medyada aradığı bir diğer bilgi türü pratik bilgidir. Pratik bilgi, yetenek kazanmak ya da uzmanlaşmak şeklinde ifade edilen, bir şeyi yapabilmenin

uygulamalı bilgisidir. Bu bağlamda teorik bilgi arayışı bir şeyin *ne* olduğu ile ilgili iken, pratik bilgi arayışı bir şeyin nasıl olduğu ile ilgilidir. Teorik bilgiyi de kısmen içinde barındıran bu bilgi türünde yatırımcılar, örneğin temel finansal tabloların nasıl analiz edileceği, elde edilen sonuçların nasıl yorumlanacağı, fiyat hareketlerine dayanan teknik göstergelerin ne anlama geldiği ve bütün bunları yatırım kararı verirken nasıl kullanacağının bilgisini aramaktadırlar. Bu çerçevede pratik bilgi yatırımcının elde ettiği verileri kendi muhakemesi sonucunda yorumlamasına dayanmaktadır. Bu bilgi türünde de sosyal medyanın finansal okuryazarlığın artırılmasına katkı sunma potansiyeli yüksektir.

Bireysel yatırımcıların sosyal medyada aradığı bir diğer bilgi türü hazır bilgidir. Hazır bilgi herhangi bir çaba sarf etmeden ulaşılan özet bilgidir. Bu tür bilgi arayışındaki yatırımcılar ya analiz edilip yorumlanmış nitelikli hazır bir bilgi ya da tüyo ve söylenti tarzı finansal temeli olmayan bilgi aramaktadırlar. Bu nedenle hazır bilgi bilişsel emek gerektirmediğinden yatırımcıların hem davranışsal eğilimlerin etkisine hem de manipülasyona en açık olduğu bilgi türü budur. Ayrıca hazır bilgi arayan yatırımcıların karşılaşabileceği en önemli risklerden biri de *algılanan uzmanlık riski*dir. Algılanan uzmanlık riski, sosyal medya mecralarında temel finansal bilgileri paylaşan hesapların uzmanlıklarının bilgi arayıcıları tarafından takipçi sayısı ve popülaritesine göre değerlendirilmesinden kaynaklanmaktadır. Bu durum, esasında uzmanlığı olmayan birinin uzmanmış gibi algılanmasına neden olabilmektedir.

Tweetlerden elde edilen öznitelikler ile BIST Pay Endekslerinin arasında ilişki olduğunu ortaya koyan bulgular, bireysel yatırımcıların yatırım fikirleri için sosyal medyada bilgi ararken yoğun bir şekilde davranışsal eğilimlerinin etkisinde kalmış olabileceklerine de işaret etmektedir. Zira çevrimiçi ortamlarda insanlar aşırı bir bilgi yüküyle karşı karşıyadırlar. Bu nedenle sosyal medyada bilgi arayan yatırımcıların maruz kalma ihtimallerinin en yüksek olduğu eğilimler bilgi kaynaklı davranışsal eğilimlerdir. Zira internetin, sosyal medyanın ve mobil yatırım uygulamalarının yaygın bir şekilde kullanılmaya başlamasından bu yana yatırımcı davranışları büyük ölçüde değişiklik göstermiştir. Bunun temel nedenleri hem bireysel yatırımcıların sayısındaki artış hem bireysel yatırımcıların demografisindeki değişim

hem de çevrimiçi platformlarda yer alan çok miktarda bilgi ve bu bilgiye erişimin çok düşük maliyetle sağlanabilmesidir. Bu durum teorik olarak bireylerin daha bilgili olmalarını ve yeni yatırım fikirlerine erişebilmesini sağlamaktadır. Ancak piyasada yeni yatırımcıların çok olması, doğrudan alım satım yapmaları ve bilgiye maliyetsiz bir şekilde erişmeleri bir takım riskleri de beraberinde getirmektedir. Özellikle bilginin niteliği ile ilgili olarak iki durum ortaya çıkmaktadır. Bunlardan biri bilgi kaynaklı davranışsal eğilimler ile diğeri yanlış bilgilendirme, yani hileli yönlendirme (manipülasyon) ile ilgilidir.

Yatırımcıların sosyal medyadan edindikleri teorik ve pratik bilgi ile analizlerini kendileri yaparak yeterince tecrübe kazanmadan alım satım yapmalarının, işlemlerinin sonucu üzerinde abartılı bir kontrol duygusu vererek *kontrol yanılısaması* ihtimalini güçlendirmektedir. Bir diğere ifadeyle hem *kendine atıf yanılılığı* hem de *bilgi ve kontrol yanılısamlarıyla* artan *aşırı güven*, yatırımcıların işlemlerinde artışa ve performansında düşüşe yol açabilmektedir. Ayrıca hâlihazırdaki sosyal medyada yer alan yatırım fikirleri, yatırımcıların önceki inançlarını *doğrulamalarını* sağlayarak kendilerine *aşırı güvenmelerine* yol açabilmektedir.

Sosyal medyada *doğrulama eğilimi* yankı odalarının hem sebebi hem de sonucu olarak ortaya çıkabilmektedir. Başka bir ifadeyle doğrulama eğilimi, sosyal medyada bireyi *yankı odalarına* hapsederek düşüncelerini hep benzer fikirlerle doğrulamasına, kendi inancına aykırı olan fikirleri tamamen göz ardı etmesine neden olabilmektedir. Doğrulama eğilimine maruz kalan bir yatırımcı sürekli benzer örnekleri göreceğinden hatalı örnekleme eğiliminin etkilerine de açıktır. Aynı zamanda retweetler ve beğeni tweetler de aynı fikrin farklı kişiler üzerinden tekrar tekrar görünmesine yol açacağından genel anlamda bir yankı oluşturma potansiyeline sahiptir. Bu durumda da kullanıcılar farklı kişiler üzerinden hep benzer fikirleri göreceğinden dolaylı ve gayriiradi bir şekilde doğrulama eğilimine maruz kalabilmektedirler. Yankı odaları bazen bireyin kendi tercihindan ziyade sosyal medya platformlarının kullandığı benzer içerikler önermeye dayalı algoritmaların oluşturduğu bilgi kozalarından da kaynaklanabilmektedir. Özellikle hazır bilgi arayışında olan yatırımcıların *hazırda bulunma* eğiliminin etkisinde kalma ihtimalleri de yüksektir. Zira son dönemlerde haberlere konu olan

ya da sosyal medyada çokça görünür olan hisse senetleri yatırımcıların dikkatini daha fazla çekebilmekte ve yakın zamanlarda görünür olan içerikler hafızada daha kalıcı olabilmektedir. Bu yanlılık tetiklendiğinde insanlar sosyal medya akışında gördükleri son bilgilere daha fazla ağırlık vererek önceki bilgiyi göz ardı etme eğilimi sergilerler. Örneğin viral popüler hisse senetleri bu eğilimin en önemli göstergesidir. Ayrıca bu eğilim insanların elde edilmesi daha zor olan bilgileri de göz ardı ederek erişilmesi daha kolay olan bilgilere ağırlık vermesine neden olmaktadır.

Hazırda bulunma eğilimi aynı zamanda *aşinalık hissini* besleyen bir eğilimdir. Zira sosyal medya paylaşımlarında tekrar tekrar görünür olan hisse senetleri bilişsel bir rahatlık ve güven duygusu yaratarak beklenen getirilerini ve risklerini değerlendirirken *aşırı iyimser* olmalarına ya da aşına olmadıkları hisse senetleriyle ilgili olarak *aşırı karamsar* olmalarına yol açmış olabilir. Aşinalık eğilimi aynı zamanda, daha kolay hatırlanmayı sağladığından hafızaya dayalı olarak hazırda bulunanı tercih etme eğilimiyle de ilgilidir. Buna ek olarak aşinalık eğilimi bilgi edinme maliyetini düşürerek belirsizliği azaltıcı bir rol oynamakta ve bilinmeyenin getirebileceği asimetrik bilgi riskini azaltabilmektedir. Bu çerçevede düşünüldüğünde yatırımcının kendi içerisindeki rasyonelliğinin bir tezahürü olmaktadır. Bu durum aynı zamanda olasılıkların bilinen, yani riskli seçenekleri olasılıkları bilinmeyen, yani belirsizlik içeren seçeneklere tercih etme eğilimi olan *belirsizlikten kaçınmanın* bir sonucudur.

Sosyal medyada aynı zamanda *pişmanlıktan kaçınma* eğilimini tetikleyecek içerikler de yer alabilmektedir. Pişmanlıktan kaçınma alınan kararların neticesinde yaşanacak muhtemel bir pişmanlığın önsel tahminin yatırımcı kararlarına etkisiyle ilgilidir. Paradoksal bir biçimde bazen riskten kaçınmaya neden olurken bazen de risk almayı teşvik edebilmektedir. Özellikle viral popüler hisse senetlerinde çoğunluğun bu hisselerle yöneldiği algısı baskın olduğunda pişmanlıktan kaçınma eğilimi, fırsatı kaçırma korkusunu besleyerek yatırımcının risk alması ihtimalini artırmaktadır. Bunun tersi de geçerlidir. Sosyal medyada piyasalar için yaklaşan riske dair korkutucu paylaşımlar yatırımcıları *kayıptan kaçınma* eğilimi nedeniyle tepkisel ve aceleci kararlar vermeye sürükleyebilmektedir.

Sosyal medya aynı zamanda yatırımcıların birbirlerini izledikleri bir alandır. Bu bağlamda sosyal medyada çoğunluğun belirli hisseler ya da belirli pozisyonlara yöneldiği algısı, yatırımcıları sürü psikolojisine sevk edebilmektedir. Sürü psikolojisi ya herkes yapıyor diye meşrulaştırmaya ya da fırsatı kaçırma korkusuyla hemen karar vermeye neden olabilmektedir. Bu duruma aynı zamanda çoğunluk etkisi de denilmektedir. Genellikle inançların ve fikirlerin benimsenme oranının başkaları tarafından benimsedikçe artmasından kaynaklanmaktadır. Sosyal medya aynı zamanda bazı haberlerin geleneksel medyadan çok daha hızlı ve daha geniş kitlelere ulaşmasını sağlamaktadır. Bu durum yatırımcıların beklenmedik haberlere karşı *aşırı tepki vermesine* ve böylece abartılı fiyat hareketlerine yol açabilmektedir.

Sonuç olarak bilgi kaynaklı davranışsal eğilimler yatırımcıyı iki şekilde nüfuz edebilmektedir. Bunlardan ilki yatırımcının kendi davranışsal eğilimlerini tetiklemesi, ikincisi ise başkalarının davranışsal eğilimlerinin bu platformlardaki bilgiye içkin olarak yatırımcı davranışlarına dolaylı yoldan tesir etmesidir. Başka bir ifadeyle başkasının sosyal medya paylaşımlarının içeriğine bulaşmış davranışsal etkinin kişinin düşüncelerine sızarak kararlarını etkilemesidir. Sonuç olarak sosyal medya her iki şekilde de davranışsal eğilimleri tetikleyerek yatırımcı kararlarını etkileme potansiyeline sahiptir. Bu tür riskleri asgari düzeye çekebilmek amacıyla, finansal okuryazarlık çerçevesinde yatırımcılar sosyal medyanın besleyebileceği davranışsal eğilimler hakkında bilgilendirilebilir. Tam bu noktada cevap aranması gereken asıl soru, bireysel yatırımcının davranışsal eğilimleri hakkında bilgi sahibi olmasının, davranışsal eğilimlerini yönetebilmesine imkân sağlayıp sağlayamayacağıdır. Bilmenin korunmak için yeterli olamayacağı ihtimalinden dolayı, bireysel yatırımcıları davranışsal eğilimlerinin olumsuz sonuçlarından koruyacak yapay zekâ destekli uygulamalar geliştirilebilir. Diğer taraftan yatırımcıyı korumak için piyasaların işleyişine ilişkin düzenlemeler de yapılabilir. Örneğin yatırımcıyı *hemen şimdi* baskısıyla almaya ya da satmaya yönlendiren eğilimlerden korumak için brüt takas uygulamasına benzer tedbirler geliştirilebilir. Benzer şekilde gün içi ya da kısa vadeli alım satımlarda işlem maliyetleri artırılarak yatırımcıların davranışsal eğilimlerinin baskısıyla acele karar vermeleri yerine, düşünerek karar vermeleri için bir fırsat sağlanmış olur. Bu fırsat aynı zamanda piyasa

bozucu etkilerin gücünü zayıflatabilir. Ancak bu tür tedbirlerin piyasanın likiditesi bağlamında etkisinin de göz önünde bulundurularak dengeli bir yaklaşım tercih edilmesi daha uygun olur.

Bilginin diğer boyutu hileli yönlendirme ile ilgilidir. Başka bir deyişle sosyal medyanın finansal piyasalar üzerindeki etkisi düşünüldüğünde, bireysel yatırımcıların bilgi arayışını manipüle edebilecek davranışların bu platformlar üzerinden gerçekleşme ihtimali gözden uzak tutulmamalıdır. Bilgi bazlı piyasa dolandırıcılığı olarak da ifade edilen bu durum, fiyatları ya da yatırımcı kararlarını etkilemek amacıyla gerçekle bağdaşmayan bilgi paylaşımlarıyla ilgilidir. Sosyal medya üzerinden maliyetsiz bir şekilde çok fazla kullanıcıya erişebilir olmak, sermaye piyasası dolandırıcıları için cezbedici olabilmektedir. Yatırımcıları, sosyal medya üzerinden yaydıkları yalan, yanlış ve yanıltıcı bilgilerle belirli kararlar almaya sevk edebilen sermaye piyasası dolandırıcıları, aynı zamanda sosyal medya aracılığıyla yatırımcıların davranışsal eğilimlerini istismar ederek de manipülasyon yapabilmektedir. Bu bağlamda öncelikle yatırımcıların çevrimiçi platformlarda piyasa manipülasyonu potansiyelinin farkında olmaları ve sosyal medyadan gelen bilgi veya tavsiyeleri kullanırken dikkatli olmaları son derece önemlidir. Mutlaka araştırma kaynaklarını çeşitlendirmeli, manipülasyona karşı proaktif olmalıdırlar. Zira cezai yaptırımlar zarar ortaya çıktıktan sonra devreye girmektedir. Bu çerçevede sosyal medya aracılığıyla manipülasyon yöntemlerine ilişkin bir rehber hazırlanarak yatırımcıların istifadesine sunulabilir. Ancak yatırımcıları botlar ve yüksek başarımlı hesaplamalar kullanılarak yapılabilecek manipülasyonlardan koruyabilmek için benzer niteliklere sahip, yapay zekâ destekli uygulamalar geliştirilmelidir. Bu uygulamalar yatırımcıların maruz kalabilecekleri manipülatif girişim ihtimalleri karşısında olasılıklar sunarak yatırım kararlarında rehberlik edebilir.

Sosyal medyanın bireysel yatırımcıların kararları ve finansal piyasalar üzerindeki etkisini ortaya koyan bulgular, finansal eğitimde geleneksel yaklaşım biçimlerinin gözden geçirilmesini gerekli kılmaktadır. Başka bir ifadeyle değişen demografik yapıları, bireysel yatırımcıları finansal piyasalara girmeye ve özellikle sosyal medya mecralarından öğrenme fırsatları aramaya istekli olduklarını göstermektedir. Dolayısıyla bireysel yatırımcıların

dijital dönüşümün avantajlarından yararlanırken hem manipülasyondan korunmak hem de bilinçli yatırım kararları verebilmek için gerekli olan bilgiyle nasıl donatılabileceğinin büyük ölçüde yeniden düşünülmesi gerekmektedir. Zira bireysel yatırımcılar daha fazla bilgi sahibi olduklarında daha fazla yatırım yapacaklar, böylece yatırımcılar finansal piyasalardaki fırsatlardan yararlanırken piyasalar da yatırımcıların sermaye birikimlerinden faydalanabilecektir.

Bu bağlamda finansal influencerlar etkin bir rol üstlenmektedir. Diğer bir deyişle finansal influencerlar sosyal medya üzerinden bireysel yatırımcıların bilgiye erişimine aracılık etmektedirler. Genellikle popülerliklerini en üst düzeye çıkarma amacı güderler. Finansal okuryazarlığı, bireysel yatırımcı katılımını ve piyasaya erişimi yaygınlaştırılmasına katkı sunarlar.

Finansal influencerlar aynı zamanda çok sayıda takipçinin işlemlerini tahmin edilebilir şekillerde etkileme potansiyeline de sahiptirler. Bu nedenle bazı finansal influencerlar takipçilerini ve popüleritelerini kaybetme pahasına yanlış ve yanıltıcı beyanlarda bulunarak manipülatif kazanç elde etmenin cazibesine kapılabilirler. Pek çok ülkede uygun bir lisansa sahip olmayan herhangi bir finansal influencerın katı yaptırımlarla karşı karşıya kalabileceği katı yönergeler bulunmaktadır. Ancak düzenleyici kurumların, ihtiyaç duyulan anda harekete geçebilmek için teknoloji odaklı bir yaklaşım geliştirmeleri gerekmektedir. Finansal influencerlar için bazı durumlarda ise daha farklı teşvik mekanizmaları da devreye girebilmektedir. Diğer pek çok influencerın reklam alması gibi, finansal influencerlar da arka planda belirli firmalar ile reklam anlaşması yapabilmektedirler. Bu tür durumlardan kaynaklanabilecek riskleri minimize edebilmek için finansal influencerların bu kapsamdaki gelir kaynaklarının şeffaf bir şekilde açıklaması istenebilir.

Finansal influencerlar ile ilgili bir diğer husus ise algılanan uzmanlık riskidir. Zira finansal influencerların hepsi, geleneksel anlamda bir uzmanlığa sahip olmayabilir. Hatta bazı kullanıcılar için finansal influencerların takipçi sayısının çokluğu ve popüleritesi geleneksel anlamda bir uzmanmış gibi algılamasına neden olabilir. Bu çerçevede algılanan uzmanlık

düzeyini minimize edecek ya da finansal influencerları asgari uzmanlık düzeyine ulaştıracak düzenlemeler yapılabilir. Düzenleyici kurumlar finansal influencerları uzmanlık sertifikasına sahip olmaları hususunda teşvik edilebilir.

Finansal influencerlar ile ilgili bir başka husus sosyal medya içeriklerinin kişiselleştirilmiş finansal tavsiyelerin yerine geçmediğinin takipçilerine izah edilmesi gerekliliğidir. Başka bir ifadeyle bireysel yatırımcı kişiselleştirilmiş rehberliğe ihtiyaç duyduğunda çevrimiçi finansal fikirlerin genellikle yetersiz kalacağı, farklı teşvik mekanizmaları ile güdülenen finansal bir influencera güvenerek işlem yapmanın maliyetlerinin olabileceği anlatılmalıdır. Zira yatırım farklı risk iştahı, farklı yatırım ufku ve farklı beklentilere göre bireyselleştirilmiş bir eylemdir. Dolayısıyla özellikle yeni başlayan yatırımcılara, giriş seviyesi yatırım dili, piyasa dalgalanmaları ve risk ile getirinin kendi kişisel senaryolarıyla nasıl ilişkili olduğu izah edilmelidir.

Çalışmada sosyal medya kullanımına ilişkin sunulan istatistikler ile sosyal medyanın hisse senedi piyasalarına etkisine ilişkin bulgular birlikte değerlendirildiğinde, hem bireysel yatırımcı sayısındaki artışın hem de bireysel yatırımcı demografisindeki değişimin yakından izlenmesinin ve yeni nesil yatırımcılar için uygun altyapıların ve düzenlemelerin hayata geçirilmesinin gerekliliği ortaya çıkmaktadır. Bu bağlamda örneğin halka açık şirketlerin yatırımcı ilişkileri biriminin (sadece statik bir web sayfası olmanın ötesine taşınarak) belirli sosyal medya platformlarında interaktif bir şekilde yer almaları teşvik edilebilir. Bu durum yatırımcıların kaynağından hızlı ve interaktif bilgiye ulaşılabilirliklerine katkı sağlayabilir.

Bu çalışmanın sonuçlarından bir diğeri de algoritmik finansal işlemlerle ilgilidir. Algoritmik finansal işlemler önceden programlanmış, fiyat, zaman, hacim bilgisi ve çeşitli teknik göstergeleri kullanarak robotların al-sat-tut kararı verip uyguladıkları finansal işlemlerdir. Bu çerçevede özellikle finansla ilgili sayısal ve yapılandırılmış veriler hızla hesaplanmakta ve elde edilen sonuçlar bir stratejinin parçası olarak karar süreçlerine dâhil edilmektedir. Sayısal verilerin yanı sıra, hem insanların bu verileri nasıl yorumlandığı ve bunlar hakkındaki düşüncelerinin hem de haber kaynaklarına ulaşan yeni bir haberin hızla anlamlandırılıp

buradan elde edilen bilgilerin karar süreçlerine dâhil edilmesi oldukça önemlidir. Bu çerçevede özellikle sosyal medya bilginin yatırımcılara iletilme şeklini etkilediğinden yatırımcıların fikir ve duygularına ilişkin zengin bir içeriğe sahiptir. Sosyal medyada paylaşılan finansal fikir ve duygu içerikli metinlerin, piyasalar üzerinde önemli bir etkiye sahip olduğunu ortaya koyan bulgular, sosyal medyadan gerçek zamanlı olarak elde edilen verilerin anlamlandırılarak algoritmik finansal işlem stratejilerine dâhil edilmesinin göreceli bir avantaj sağlayabileceğine işaret etmektedir. Diğer taraftan sosyal medya platformlarının dezenformasyona açık yapısı, sosyal medya bazlı göstergelerin teyit edilmeden karar süreçlerine dâhil edilmesi durumunda, zararlı sonuçlanabilecek hatalı kararlar alınmasına yol açabilir. Bu nedenle algoritmalar her iki durumun avantaj ve dezavantajlarını göz önünde bulunduracak şekilde tasarlanmalıdır.

Algoritmik işlemler için oluşturulan ve belirli bir piyasa ya da hisse senedi ile ilgili paylaşımların fikri ve duygusal yönelimlerdeki değişimleri gerçek zamanlı olarak ölçüp yatırım kararlarında kullanılmasını sağlayan bu gösterge, aynı zamanda yatırımcıların doğrudan istifadesine de sunulabilir. Böylece yatırımcılar belirli piyasalar ya da hisse senetlerine olan ilgiyi takip etmek, izlemek ve geleneksel kaynaklar aracılığıyla bulamayacakları içgörüler sağlamak için bu araçtan faydalanabilirler. Bu çerçevede yatırımcılar fikir ve duygu göstergesini piyasa dalgalanmaları ve davranış değişikliklerini izlemek ya da bir hisse senedinin fiyatının duygulardan etkilenip etkilenmediğini ayırt etmek için kullanabilirler. Bu çerçevede fikir ve duygu göstergesi daha çok kısa vadeli işlemlerle ilgilidir. Fakat aynı zamanda bir hisse senedinin temel değerini analiz ederken alternatif bir bakış açısı sunabilir. Kurumsal yatırımcılar böyle bir göstergeyi bireysel yatırımcıların yönelimlerini izlemek için kullanırken, aracı kurumlar ise bu göstergeyi bireysel yatırımcılara bir hizmet olarak sunmaktadırlar. Bu bağlamda tüm kullanıcıların izleyebileceği bir gösterge de oluşturulabilir. Nitekim Dow Jones Endeksinin S&P 500 için oluşturduğu Twitter Sentiment Endeksi böyle bir hizmet vermektedir. Benzer şekilde tüm yatırımcıların izleyebileceği endeks düzeyinde bir sentiment göstergesi Borsa İstanbul için de geliştirilebilir.

Sosyal medya aynı zamanda finansal krizler ya da finansal panik salgını başlamadan ve kontrolden çıkmadan önce devreye girebilecek bir erken uyarı sistemi kurulmasına yardımcı olabilir. Bu bağlamda sosyal medyadan iki şekilde yararlanılabilir. Birincisi sosyal medyadan büyük ölçekli paylaşım ve etkileşim verileri toplanarak gerçek hayata uyarlanabilecek modeller geliştirmek için kullanılabilir. Aynı zamanda bu modeller mevcut öngörü modellerine dâhil edilerek öngörü başarısı artırılabilir. İkincisi bu modeller sosyal medya etkileşimleri üzerinden gerçek zamanlı çalışan bir erken uyarı sistemi olarak da kurgulanabilir. Ayrıca bu erken uyarı sistemleri yapay zekâ desteği ile gerektiğinde bir finansal panik salgının önüne geçebilecek erken müdahale sistemi şeklinde de tasarlanabilir.

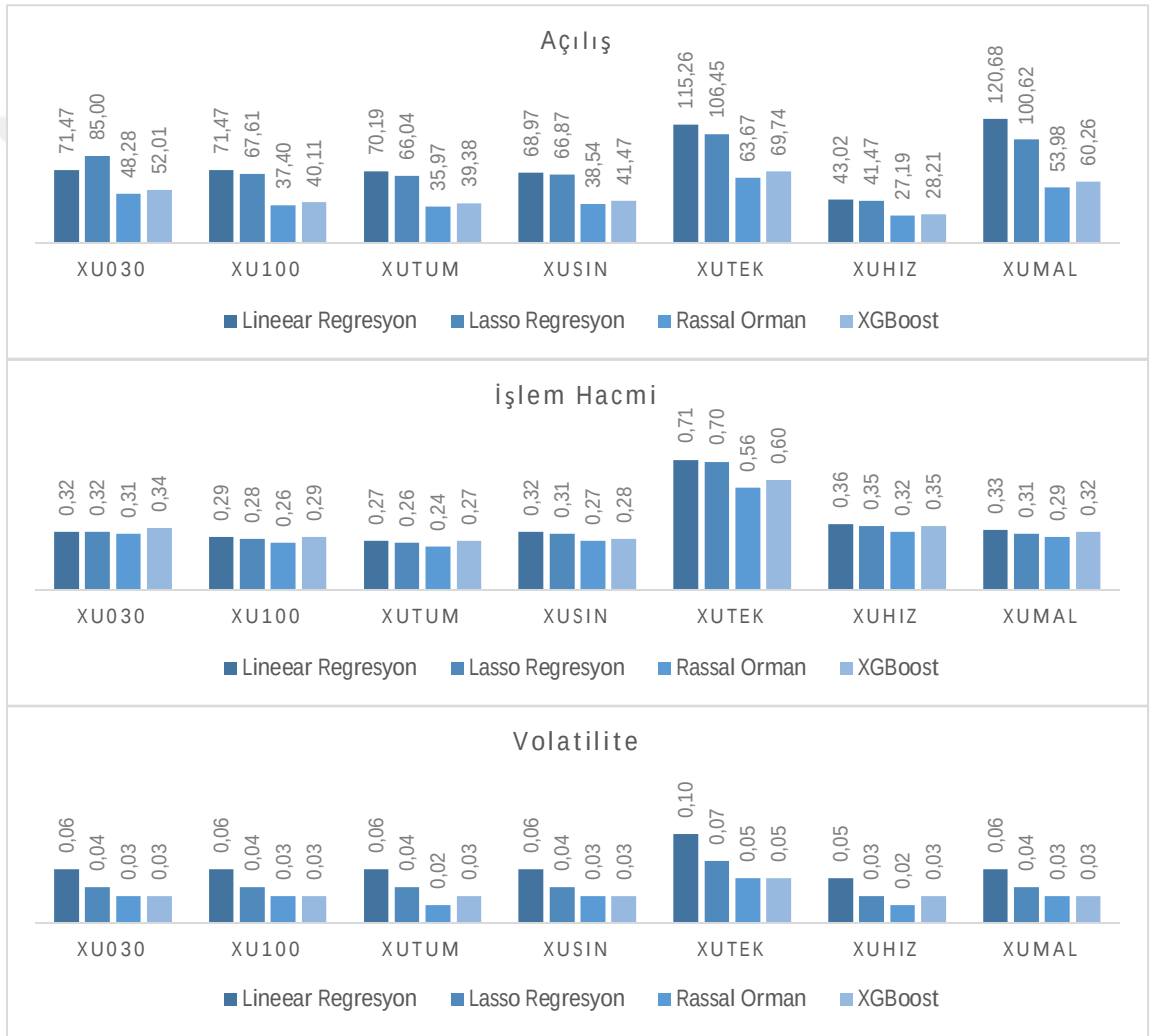
Bu araştırma kapsamında elde edilen bulgular, insanların finansal kararlarında bilişsel, davranışsal ve sosyal etkinin varlığını ortaya koymaktadır. Bu durum psikolojik faktörlere ek olarak sosyolojik faktörlerin yatırımcı kararlarına etkisinin mikro ve mezo ölçekte daha derinlemesine araştırılmasının finans teorisine yeni bakış açıları sağlayabileceğine işaret etmektedir. Özellikle dijitalleşme, blokzinciri ve Endüstri 4.0 ile gündeme gelen köklü dönüşüm, araçların ortadan kalktığı merkeziyetsiz finans (DeFi) bağlamında düşünüldüğüne, insanların eylem ve davranışlarına etki eden sosyolojik faktörlerin daha iyi anlaşılmasının, potansiyel yeni finansal kurumların inşasında önemini gözler önüne sermektedir. Dolayısıyla davranışsal finans insanların eylem ve davranışlarında sosyolojik faktörlerin etkisine ilişkin yeni araştırmalarla zenginleştirilmelidir.

Son olarak bu çalışmada amaçlardan biri de farklı tekniklere, avantaj ve dezavantajlara sahip makine öğrenmesi yöntemlerinin performansını karşılaştırmaktır. Bu bağlamda RMSE skoru farklı regresyon modelleri arasındaki performansı karşılaştırmak için kullanılmaktadır. Grafik 4.23 açılış, işlem hacmi ve volatilitenin tahmininde kullanılan regresyon modellerinin RMSE skorlarını karşılaştırmaktadır⁵⁸. Grafikten de görüleceği üzere bütün endeksler için en düşük RMSE skoruna sahip olan yöntem rassal orman yöntemidir. XGBoost yöntemi de

⁵⁸ Tweetlerden elde edilen öznetelikler ile getiri arasında istatistiksel olarak anlamlı bir ilişki bulunamadığı için, getiri ile ilgili regresyon modellerinin RMSE skorları, makine öğrenmesi yöntemlerinin performanslarının karşılaştırılmasına dahil edilmemiştir.

aslında rassal orman yöntemlerinden elde edilen sonuca oldukça yakın değerlere sahip olduğu görülmektedir. Bu durum her iki yöntemin de hem topluluk öğrenmesi tekniğine dayalı olmasından hem de nonlineer bir teknik olmasından kaynaklanıyor olabileceğine işaret etmektedir.

Grafik 4.44: Makine Öğrenmesi Yöntemlerinin Performanslarının Karşılaştırılması



SONUÇ

Bu çalışmada Borsa İstanbul’da işlem gören şirketler bağlamında belirli bir günde paylaşılan tweetlerin hem fikri ve duygusal olarak yönelimleri ile ilgili özniteliklerin hem de bu tweetleri paylaşan kullanıcıların etkileşimlerini temsil eden özniteliklerin BIST Pay Endeksleri ile ilişkisi araştırılmıştır. Araştırmanın birinci aşamasında borsa ile ilgili tweetleri belirlemek için bağlam sınıflandırıcı, tweetlerin yönelimlerini belirlemek için fikir ve duygu sınıflandırıcı geliştirilmiştir. Analizin ikinci aşamasında, farklı avantaj ve dezavantajlara sahip makine öğrenmesi yöntemleri kullanılarak, tweetlerden elde edilen öznitelikler ile Borsa İstanbul pay piyasası endekslerinin açılış değerlerinin, işlem hacimlerinin ve volatilitelerinin açıklanabileceği bulgusuna ulaşılmıştır. Tweetler, içerdiği fikri ve duygusal yönelimler açısından, Twitter’da yatırım fikri arayan yatırımcıların beklenti ve inançlarının ve dolayısıyla iradeli eylemlerinin inşasında pay sahibidirler. Nitekim fikri ve duygusal yönelimlerin borsa değişkenlerini açıklama kabiliyeti bu duruma işaret etmektedir. Benzer şekilde tweetler, davranışsal eğilimlerden arındırılmış değildir. Başka bir ifadeyle tweetler yatırımcıların bir takım davranışsal eğilimlerini tetikleyebilirler. Aynı zamanda tweetler sosyal bir ortamda insanlar arasında paylaşılmakta, içerdiği bilgi ve duyguyu insandan insana aktarmaktadırlar. Bu çerçevede bu özellikleri kısmen temsil ettiği düşüncesiyle oluşturulan özniteliklerin borsa değişkenlerini açıklıyor olması, sosyal medyanın yatırımcılar üzerinde bilişsel, sosyal ve davranışsal etkileri olduğuna işaret etmektedir.

Teorik bulguların dışında bu çalışmada kullanılan yöntemler ve yaklaşımlarla ilgili bir değerlendirme de yapılabilir. Zira bu çalışmada farklı avantaj ve dezavantajlara sahip makine öğrenmesi yöntemlerinin performansları da kıyaslanmıştır. Örneğin bağlam sınıflandırıcısının eğitiminde kullanılan derin öğrenme yaklaşımlarından GRU mimarisi ve LSTM mimarisi kıyaslandığında başarı performansı açısından önemli bir farklılık olmasa da işlem zamanını %40’a yakın azaltan GRU mimarisi bariz bir avantaja sahiptir. Özellikle büyük veri çalışmalarında bu zamansal avantajından dolayı GRU mimarisi tercih edilebilir. Bir diğer kıyaslama bütünleşik karar modelinin eğitiminde kullanılan lineer regresyon, lasso regresyon, rassal orman ve XGBoost yöntemleri için yapılmıştır. Bu yöntemlerin

performansları karşılaştırıldığında topluluk öğrenmesine dayalı rassal orman ya da XGBoost gibi doğrusal olmayan yöntemlerin daha başarılı sonuçlar elde ettiği görülmüştür. Özellikle XGBoost yöntemi paralel işlemeye de imkân verdiği için büyük veri analizlerinde tercih edilebilir. Ancak bu alanın, mevcutlardan çok daha başarılı bir performans sergileyebilecek yeni yöntemlerin gelişimine çok açık olduğu ve gelişmelerin yakından izlenmesinin gerekliliği de gözden uzak tutulmamalıdır.

Bu çalışmada ayrıca anahtar kelime tabanlı sınıflandırma yaklaşımı ile makine öğrenmesi tabanlı sınıflandırma yaklaşımlarının metnin bağlamının belirlenmesinde benzer performans sergiledikleri ortaya konulmuştur. Dikkat edilirse makine öğrenmesi tabanlı yaklaşımda borsa ile ilgili olarak daha az veri sınıflandırılmış, ancak tahmin performansı açısından kelime tabanlı yaklaşımla kıyaslandığında oldukça yakın sonuçlar elde edilmiştir. Bu durumda verilerin toplanması aşamasında kelime tabanlı yaklaşımlar büyük bir kolaylık sağlamaktadır. Ancak diğerine kıyasla çok fazla veri topladığı için bu verileri depolama ve işleme maliyeti açısından dezavantaja sahiptir. Makine öğrenmesi tabanlı yaklaşımda ise önce bütün veriyi görmek ve bunları sınıflandırmak maliyetlidir. Ancak bu sınıflandırmadan sonra daha küçük bir veri kümesi üzerinde çalışılacağı için daha avantajlı olmaktadır. Bu durumda birini diğerine tercih etmek yerine hibrit bir sistem kurularak bu iki yaklaşımın avantajlı yönleri birleştirilebilir. Örneğin anahtar kelimelerle veriler toplandıktan sonra makine öğrenmesi tabanlı sınıflandırıcı ile ikinci bir filtreleme yapan bir akış kurgulanabilir.

Bu çalışmada karşılaşılan kısıtlardan biri Türkçe dili için geliştirilen duygu analiz araçlarının sınırlı olmasıyla ilgiliydi. Bu bağlamda Türkçe için finans alanına özgü Harvard IV-4 ya da Loughran'a benzer bir sözlük geliştirilebileceği gibi daha büyük bir finans derlemi oluşturularak hem bağlam sınıflandırıcı için hem de fikir ve duygu sınıflandırıcı için (BERT yönteminde kullanılan) ön eğitilmiş modellerin geliştirilmesi sağlanabilir.

Bu çalışma sadece borsaya odaklanmıştır. Gelecek çalışmalarda fikir ve duygu analizi farklı piyasalar ve finansal enstrümanlar için de yapılabilir. Başka bir ifadeyle gelecekte çalışmalarda

fikir ve duygu analizi ile döviz kurları, kripto para piyasaları, değerli metaller ve benzeri finansal enstrümanlar ile ilişkisi araştırılabilir.

Analiz döneminde finansal bilgi arayışlarında öncelikle başvuru alan sosyal medya platformlarından biri Twitter idi. Ancak zaman içerisinde hem sosyal medya platformları hem de bireysel yatırımcıların finansal bilgi arayışlarında başvurdukları platformlar çeşitlenmiştir. Bu çerçevede gelecek çalışmalar Whatsapp, Telegram, TikTok, Instagram gibi farklı platformlar için gerçekleştirilebilir.

Bu çalışma tweetler ile borsa arasındaki ilişkiyi makro ölçekte analiz etmiştir. Sosyal medyanın yatırımcı kararlarına etkisine daha yakından bakmak için analizin, borsa tarafında hisse senedi, sosyal medya tarafında kullanıcı düzeyinde yapılması daha derin çıkarımlar elde edilmesini sağlayabilir. Bu çerçevede sosyal ağ analizi, fikir ve duygu analizi ile birleştirilerek mikro ve mezo ölçekte çözümleme aracı olarak kullanılabilir.

YARARLANILAN KAYNAKLAR

- Ackert, L. F. (2014). Traditional and behavioral finance. In H. K. Baker, & V. Ricciardi (Eds.), *Investor behavior: The psychology of financial planning and investing* (pp. 25-41). John Wiley & Sons.
- AFINN. (2011). <http://www2.imm.dtu.dk/pubdb/pubs/6010-full.html>
- Akerlof, G. A., & Shiller, R. J. (2010). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.
- Alostad, H., & Davulcu, H. (2017, January). Directional prediction of stock prices using breaking news on Twitter. In *Web Intelligence* (Vol. 15, No. 1, pp. 1-17). IOS Press. <https://doi.org/10.3233/web-170349>
- Alpar, R. (2022). *Spor, sađlık ve eđitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlilik-güvenilirlik* (7. baskı). Detay Yayıncılık.
- Altman, M. (2014). Behavioral economics, thinking processes, decision making, and investment behavior. In H. K. Baker, & V. Ricciardi (Eds.), *Investor behavior: The psychology of financial planning and investing* (pp. 43-61). John Wiley & Sons.
- Altug, S., Yesiltas, S., Şen, A., & Arslan, B. (2021). A twitter-based economic policy uncertainty index for Turkey. *Available at SSRN 3954696*.
- Aman, H., Kasuga, N., & Moriyasu, H. (2018). Mass media effects on trading activities: television broadcasting evidence from Japan. *Applied Economics*, 50(42), 4522-4539.
- Amazon SageMaker Documentation. (n.d.). XGBoost. Amazon SageMaker: Developer Guide. In [aws.amazon.com. https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost-HowItWorks.html](https://docs.aws.amazon.com/sagemaker/latest/dg/xgboost-HowItWorks.html) adresinden 23 Ocak 2023 tarihinde alınmıştır.
- Amidi, A., & Amidi, S. (n.d.-a). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-deep-learning-tips-and-tricks>
- Amidi, A., & Amidi, S. (n.d.-b). Deep learning tips and tricks cheatsheet. <https://stanford.edu/~shervine/l/tr/teaching/cs-229/cheatsheet-machine-learning-tips-and-tricks>
- Andrei, D., & Hasler, M. (2015). Investor attention and stock market volatility. *The review of financial studies*, 28(1), 33-72.
- Antweiler, W., & Frank, M. Z. (2004). Is all that talk just noise? The information content of internet stock message boards. *The Journal of finance*, 59(3), 1259-1294.

- Apaydın, E. (2012). *Yapay öğrenme* (2. baskı). Boğaziçi Üniversitesi Yayınevi.
- Arkansas Securities Department. (2022, August 04) Financial Advice via Social Media: The Rise of the “Finfluencer”. <https://securities.arkansas.gov/financial-advice-via-social-media-the-rise-of-the-finfluencer/> adresinden 30 Nisan 2023 tarihinde alınmıştır.
- Arora, K. (2022, November 4). The Rise Of The Retail Investor. <https://www.forbes.com/sites/forbesagencycouncil/2022/11/04/the-rise-of-the-retail-investor/?sh=2a29cd231755>
- Arrow, K. J. (1970). *Social choice and individual values* (2nd ed.). Yale University Press.
- Arrow, K. J. (1986). Rationality of self and others in an economic system. *Journal of business*, 59(4), S385-S399. <https://www.jstor.org/stable/2352770>
- Arslan, S. (n.d.). LSTM ile Türkçe duygu analizi. <https://github.com/serkanars/turkishsentimentanalysis> adresinden 01 Eylül 2020 tarihinde alınmıştır.
- ASIC. (2022, March). Discussing financial products and services online: Information Sheet 269 (INFO 269). <https://asic.gov.au/regulatory-resources/financial-services/giving-financial-product-advice/discussing-financial-products-and-services-online/>
- Atefeh, F., & Khreich, W. (2015). A survey of techniques for event detection in twitter. *Computational Intelligence*, 31(1), 132-164.
- Aydoğan, M., & Karcı, A. (2019). Kelime temsil yöntemleri ile kelime benzerliklerinin incelenmesi. *Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 34(2), 181-196. <https://doi.org/10.21605/cukurovaummfd.609119>
- Baddeley, M. (2012). *Behavioural Economics and Finance*. Routledge.
- Baker H. K., & Nofsinger, J. R. (2010). Behavioral finance: An overview. In H. K. Baker, & J. R. Nofsinger (Eds.), *Behavioral finance: Investors, corporations, and markets* (pp. 3-21). John Wiley & Sons.
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A., & Nielsen, H. (2000). Assessing the accuracy of prediction algorithms for classification: an overview. *Bioinformatics*, 16(5), 412-424. <https://doi.org/10.1093/bioinformatics/16.5.412>
- Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, 107(3), 797-817.
- Barber, B. M., & Odean, T. (2001). The internet and the investor. *Journal of Economic Perspectives*, 15(1), 41-54.

- Barber, B. M., & Odean, T. (2002). Online investors: do the slow die first?. *The Review of financial studies*, 15(2), 455-488.
- Barber, B. M., & Odean, T. (2008). All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The review of financial studies*, 21(2), 785-818.
- Bartov, E., Faurel, L., & Mohanram, P. S. (2018). Can Twitter help predict firm-level earnings and stock returns?. *The Accounting Review*, 93(3), 25-57. <https://doi.org/10.2308/accr-51865>
- Basar, S. (2021, May 24.). *As the algo wheel turns*. <https://www.marketsmedia.com/as-the-algo-wheel-turns>
- Basık, F. O. (2011). Ansiklopedik muhasebe ve finans terimleri sözlüğü (1st ed., p. 528), Türkiye İş Bankası Kültür Yayınları.
- Batra, R., & Daudpota, S. M. (2018, March). Integrating StockTwits with sentiment analysis for better prediction of stock price movement. In *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICOMET.2018.8346382>
- Bayrak, O. K. (2012). Davranışsal Finans. *Sermaye Piyasasında Gündem*, 120, 6-17.
- Benou, G., & Richie, N. (2003). The reversal of large stock price declines: The case of large firms. *Journal of Economics and Finance*, 27(1), 19-38.
- Bernardo, I., Henriques, R., & Lobo, V. (2018). Social market: Stock market and twitter correlation. In *Intelligent Decision Technologies 2017: Proceedings of the 9th KES International Conference on Intelligent Decision Technologies (KES-IDT 2017)–Part II 9* (pp. 341-356). Springer International Publishing. https://doi.org/10.1007/978-3-319-59424-8_32
- Bernoulli, D. (1954). Exposition of a new theory on the measurement of risk. *Econometrica: Journal of the Econometric Society*, 22(1), 23-36. <https://doi.org/10.2307/1909829>
- Bharathi, S., Geetha, A., & Sathiynarayanan, R. (2017). Sentiment analysis of twitter and RSS news feeds and its impact on stock market prediction. *International Journal of Intelligent Engineering and Systems*, 10(6), 68-77. <https://doi.org/10.22266/ijies2017.1231.08>
- Bhatia, A., Hagras, H., & Lepley, J. J. (2018, September). Machine Learning Approach to Extracting Emotions Information from Open Source Data for Relative Forecasting of Stock Prices. In *2018 10th Computer Science and Electronic Engineering (CEECE)* (pp. 142-147). IEEE. <https://doi.org/10.1109/CEECE.2018.8674180>

- Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. *IMF Staff papers*, 47(3), 279-310.
- BING. (n.d.). <https://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html>
- Bishop, M. (2019). Healthcare social media for consumer informatics. M. Edmunds, C. Hass, & E. Holve (Eds.) In *Consumer informatics and digital health* (pp. 61-86). Springer. https://doi.org/10.1007/978-3-319-96906-0_4
- Black, F. (1986). Noise. *The journal of finance*, 41(3), 528-543.
- Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy*, 81(3), 637-654. <https://www.jstor.org/stable/1831029>
- Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests.
- Blackbird.AI. (2022, December 23). Social media-fueled investment fraud: Exploring the dark side of public perception manipulation. <https://www.blackbird.ai/blog/2022/12/23/social-media-fueled-investment-fraud-exploring-the-dark-side-of-public-perception-manipulation>
- Bogan, V. (2008). Stock market participation and the internet. *Journal of Financial and Quantitative Analysis*, 43(1), 191-211.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the association for computational linguistics*, 5, 135-146. https://doi.org/10.1162/tacl_a_00051
- Bollen, J., Mao, H., & Zengb, X. (2011). Twitter mood as a stock market predictor. *Journal of Computational Science*, 2(1), 1-8. <https://doi.ieeecomputersociety.org/10.1109/MC.2011.323>
- Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, 92(5), 1170-1182. <https://doi.org/10.1086/228631>
- Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social networks*, 29(4), 555-564. <https://doi.org/10.1016/j.socnet.2007.04.002>
- Bouktif, S., Fiaz, A., & Awad, M. (2020). Augmented textual features-based stock market prediction. *IEEE Access*, 8, 40269-40282. <https://doi.org/10.1109/ACCESS.2020.2976725>
- Boyd, D. M., & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of computer-mediated Communication*, 13(1), 210-230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>

- Bozma, G., & Sinan, K. U. L. (2020). Twitter ile Hisse Senetleri Oynaklığı Tahmin Edilebilir mi?. *Sosyoekonomi*, 28(45), 315-326.
- Bremer, M., & Sweeney, R. J. (1991). The reversal of large stock-price decreases. *The Journal of Finance*, 46(2), 747-754
- Brown, K. C., & Harlow, W. V. (1988). Market overreaction: Magnitude and intensity. *Journal of Portfolio Management*, 14(2), 6.
- Brown, R. (2011). Incorporating news into algorithmic trading strategies: increasing the signal-to-noise ratio. In G. Mitra, & L. Mitra (Eds.), *The handbook of news analytics in finance* (pp. 305-310). John Wiley & Sons.
- Brunswick. (2023). 2023 Digital Investor Survey: The Age of Information Without Limitation. From <https://www.brunswickgroup.com/perspectives/digital-investor-survey/2023/>
- Buğra A. (2011). *İktisatçılar ve insanlar* (8. baskı). İletişim Yayınları.
- Bujari, A., Furini, M., & Laina, N. (2017, January). On using cashtags to predict companies stock trends. In *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)* (pp. 25-28). IEEE. <https://doi.org/10.1109/CCNC.2017.7983075>
- Çağlayan Akay, E. (2020), *Ekonometride büyük veri ve makine öğrenmesi: Temel kavramlar*. Der yayınları.
- Canan, S. (2015). *Değişen beynim*. Tuti Kitap Nefes Yayıncılık.
- Cao, L. (2017). Data science: a comprehensive overview. *ACM Computing Surveys (CSUR)*, 50(3), 1-42. <https://doi.org/10.1145/3076253>
- Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic journal of communication*, 23(1), 46-65. <https://doi.org/10.1080/15456870.2015.972282>
- Cazzoli, L., Sharma, R., Treccani, M., & Lillo, F. (2016, September). A large scale study to understand the relation between twitter and financial market. In *2016 third European network intelligence conference (ENIC)* (pp. 98-105). IEEE. <https://doi.org/10.1109/ENIC.2016.022>
- Cevizci, A. (2014). *Paradigma felsefe sözlüğü*. Paradigma Yayınları.
- Chaubard, F., Fang, M., Genthial, G., Mundra, R., & Socher, R. (2019). *CS224n: Natural Language Processing with DeepLearning, Lecture Notes: Part I: Word Vectors I: Introduction, SVD and Word2Vec*. CS224N@Stanford. <http://web.stanford.edu/class/cs224n/readings/cs224n-2019-notes01-wordvecs1.pdf>

- Chaubard, F., Mundra, R., & Socher, R. (2016). *CS 224D: Deep learning for natural language processing*, Lecture Notes: Part I. CS224N@Stanford.
http://cs224d.stanford.edu/lecture_notes/notes1.pdf
- Chen, J. (2022, December 30). "Headline Risk". Investopedia.
<https://www.investopedia.com/terms/h/headline-risk.asp> adresinden 30 Aralık 2022 tarihinde alınmıştır.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
- Cherrington, M., Thabtah, F., Lu, J., & Xu, Q. (2019, April). Feature selection: filter methods performance challenges. In *2019 International Conference on Computer and Information Sciences (ICCIS)* (pp. 1-4). IEEE.
- Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*.
<https://doi.org/10.48550/arXiv.1406.1078>
- Chollet, F., *Python ile derin öğrenme* (1. baskı). Buzdağı Yayınevi, 2019.
- Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, 357(4), 370-379.
<https://doi.org/10.1056/NEJMsa066082>
- Christakis, N. A., & Fowler, J. H. (2012). *Sosyal ağların şaşırtıcı gücü ve yaşantımızı biçimlendiren etkisi* (D. Yüksel, Çev.). Varlık Yayınları. (Orijinal eserin basım tarihi 2011).
- Christakis, N. A., & Fowler, J. H. (2013). Social contagion theory: examining dynamic social networks and human behavior. *Statistics in medicine*, 32(4), 556-577.
<https://doi.org/10.1002/sim.5408>
- Chun, J., Ahn, J., Kim, Y., & Lee, S. (2021). Using deep learning to develop a stock price prediction model based on individual investor emotions. *Journal of Behavioral Finance*, 22(4), 480-489. <https://doi.org/10.1080/15427560.2020.1821686>
- Clark, W. A. (1991). Residential preferences and neighborhood racial segregation: A test of the Schelling segregation model. *Demography*, 28(1), 1-19.
<https://doi.org/10.2307/2061333>
- Coleman, J. S. (1994). *Foundations of social theory*. Harvard university press.
- Colman, A. A. (2006). *A Dictionary of Psychology*, (3.rd). Oxford University Press.

- Conlisk, J. (1996). Why bounded rationality?. *Journal of economic literature*, 34(2), 669-700. <http://www.jstor.org/stable/2729218>
- Cook, K. S., Emerson, R. M., Gillmore, M. R., & Yamagishi, T. (1983). The distribution of power in exchange networks: Theory and experimental results. *American journal of sociology*, 89(2), 275-305. <https://doi.org/10.1086/227866>
- Cookson, J. A., Engelberg, J. E., & Mullins, W. (2023). Echo chambers. *The Review of Financial Studies*, 36(2), 450-500.
- Cooper, M. J., Dimitrov, O., & Rau, P. R. (2001). A rose. com by any other name. *The journal of Finance*, 56(6), 2371-2388.
- Cootner, P. H. (1964). *The random character of stock market prices*. MIT Press.
- Cormode, G., & Krishnamurthy, B. (2008). Key differences between Web 1.0 and Web 2.0. *First Monday*. <https://doi.org/10.5210/fm.v13i6.2125>
- Coviello, L., Sohn, Y., Kramer, A. D., Marlow, C., Franceschetti, M., Christakis, N. A., & Fowler, J. H. (2014). Detecting emotional contagion in massive social networks. *PloS one*, 9(3), e90315.
- Cui, X., Lam, D. & Verma, A. (2016). Embedded Value In Bloomberg News & Social Sentiment Data [White paper]. <https://www.bloomberg.com/professional/sentiment-analysis-white-papers/>
- Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı ve Sanayi ve Teknoloji Bakanlığı. (2021). *Ulusal Yapay Zekâ Stratejisi 2021-2025*. <https://cbddo.gov.tr/SharedFolderServer/Genel/File/TR-UlusalYZStratejisi2021-2025.pdf>
- Cunningham, P., Kathirgamanathan, B., & Delany, S. J. (2021). Feature selection tutorial with python examples. *arXiv preprint arXiv:2106.06437*.
- Cybulska, E. (2015). Freud's burden of debt to Nietzsche and Schopenhauer. *Indo-Pacific Journal of Phenomenology*, 15(2), 1-15. <https://doi.org/10.1080/20797222.2015.1101836>
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Avon Books.
- Damasio, A. (2010). *Self comes to mind: Constructing the conscious brain*. Vintage.
- Damasio, A. (2018). *Spinoza'yı Ararken: Haz, Acı ve Hisseden Beyin*. (E. Kumral ve İ. Çetiner, Çev.), ODTÜ Yayıncılık. (Orijinal eserin basım tarihi 2018).

- Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market under- and Overreactions. *The Journal of Finance*, 53(6), 1839–1885. <http://www.jstor.org/stable/117455>
- Daniel, M., Neves, R. F., & Horta, N. (2017). Company event popularity for financial markets using Twitter and sentiment analysis. *Expert Systems with Applications*, 71, 111-124. <https://doi.org/10.1016/j.eswa.2016.11.022>
- Das, S. R. (2011). News analytics: Framework, techniques and metrics. In G. Mitra, & L. Mitra (Eds.), *The handbook of news analytics in finance* (pp. 43-71). John Wiley & Sons. <https://doi.org/10.1002/9781118467411.ch2>
- Das, S. R., & Chen, M. Y. (2007). Yahoo! for Amazon: Sentiment extraction from small talk on the web. *Management science*, 53(9), 1375-1388. <https://doi.org/10.1287/mnsc.1070.0704>
- Data science central. 3V of big data. <https://www.datasciencecentral.com> adresinden 08 Mayıs 2020 tarihinde alınmıştır.
- De Jong, P., Elfayoumy, S., & Schnusenberg, O. (2017). From returns to tweets and back: An investigation of the stocks in the dow jones industrial average. *Journal of Behavioral Finance*, 18(1), 54-64. <https://doi.org/10.1080/15427560.2017.1276066>
- De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy*, 98(4), 703–738. <http://www.jstor.org/stable/2937765>
- de Regt, A., Cheng, Z., & Fawaz, R. (2023, March). Young People Under ‘Finfluencer’: The Rise of Financial Influencers on Instagram: An Abstract. In *Optimistic Marketing in Challenging Times: Conference, May 25-27, Monterey, CA, USA* (pp. 271-272). Cham: Springer Nature Switzerland.
- Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science*, 41(6), 391-407.
- Dehkharghani, R., Saygin, Y., Yanikoglu, B., & Oflazer, K. (2016). SentiTurkNet: a Turkish polarity lexicon for sentiment analysis. *Language Resources and Evaluation*, 50, 667-685. <https://doi.org/10.1007/s10579-015-9307-6>
- Deloitte. (2022, June 28). Regulatory posture on social media advertising and finfluencers. <https://www2.deloitte.com/vn/en/pages/financial-services/articles/regulatory-posture-social-media-advertising-finfluencers.html>
- Demir, Ö. (2013). *Akıl ve çıkar: Davranışsal iktisat açısından rasyonel olmanın rasyonelliği*. Sentez Yayınları.

- Demir, Ö., & Acar, M. (2006). *Sosyal Bilimler Sözlüğü* (6. baskı). Adres Yayınları.
<http://www.sosyalbilimlervakfi.org/tr/sozluk/>
- Dennis, M. A. & Kahn, R. (2023, May 4). *internet*. In *Encyclopedia Britannica*.
<https://www.britannica.com/technology/Internet>
- Devlin, J. (2022). *Contextual Word Representations with BERT and Other Pre-trained Language Models* [PowerPoint slides]. CS224N@Stanford.
http://web.stanford.edu/class/cs224n/slides/Jacob_Devlin_BERT.pdf
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. <https://doi.org/10.48550/arXiv.1810.04805>
- Dewally, M. (2003). Internet investment advice: Investing with a rock of salt. *Financial Analysts Journal*, 59(4), 65-77.
- Digital 2023. (2023). Global Overview Report. <https://datareportal.com/reports/digital-2023-global-overview-report>
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs, *Numerische Mathematik*, 1(1), 269-271. <https://doi.org/10.1145/3544585.3544600>
- dmlcXGBoost. (n.d.). <https://xgboost.readthedocs.io/en/latest/tutorials/model.html>
adresinden 23 Ocak 2023 tarihinde alınmıştır.
- Dollarhide, M. (August 31, 2021). Social media: Definition, effects, and list of top apps. v
<https://www.investopedia.com/terms/s/social-media.asp>
- Domeniconi, G., Moro, G., Pagliarani, A., & Pasolini, R. (2017, November). Learning to Predict the Stock Market Dow Jones Index Detecting and Mining Relevant Tweets. In *Proceedings of the 9th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 1, (pp. 165-172).
<https://doi.org/10.5220/0006488201650172>
- Domingos, P., & Richardson, M. (2001, August). Mining the network value of customers. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 57-66). <https://doi.org/10.1145/502512.502525>
- Domo. (2021). *Data Never Sleeps 9.0* [Infographic]. Domo.
<https://www.domo.com/learn/infographic/data-never-sleeps-9>
- Dougal, C., Engelberg, J., Garcia, D., & Parsons, C. A. (2012). Journalists and the stock market. *The Review of Financial Studies*, 25(3), 639-679.

- Dowling, M., & Lucey, B. (2010). Behavioral finance: An overview. In H. K. Baker, & J. R. Nofsinger (Eds.), *Behavioral finance: Investors, corporations, and markets* (pp. 313-330). John Wiley & Sons.
- Downey, L. (2022, May 09). "Noise Trader Risk". Investopedia. <https://www.investopedia.com/terms/n/noisetraderisk.asp>. adresinden 09 Mayıs 2022 tarihinde alınmıştır.
- Duz Tan, S., & Tas, O. (2021). Social media sentiment in international stock returns and trading activity. *Journal of Behavioral Finance*, 22(2), 221-234. <https://doi.org/10.1080/15427560.2020.1772261>
- Eagleman, D. (2011). *Incognito: Beynin gizli hayatı* (19. baskı) Bkz Yayıncılık.
- Eagleman, D. (2016). *Beyin: Senin hikayen* (6. baskı), Bkz Yayıncılık.
- Edgeworth, F. Y. (1881). *Mathematical psychics: An essay on the application of mathematics to the moral sciences* (Vol. 10). CK Paul.
- Eliacik, A. B., & Erdogan, N. (2018). Influential user weighted sentiment analysis on topic based microblogging community. *Expert Systems with Applications*, 92, 403-418. <https://doi.org/10.1016/j.eswa.2017.10.006>
- Eltantawy, N., & Wiest, J. B. (2011). The Arab spring| Social media in the Egyptian revolution: reconsidering resource mobilization theory. *International journal of communication*, 5, 18.
- Erdem, S. E. (2010). İnsan Şahsiyetinin Gelişimine Holistik Bakış: Felsefi Perspektiften Davranışçı ve Bilişselci Ekolün Eleştirisi. *İnönü Üniversitesi İlahiyat Fakültesi Dergisi*, 1(1), 179-194.
- ESMA. (2021, October 28). ESMA's Statement on Investment Recommendations on Social Media. https://www.esma.europa.eu/sites/default/files/library/esma70-154-2780_esmas_statement_on_investment_recommendations_on_social_media.pdf
- Fama, E. F. (1965a). The behavior of stock-market prices. *The journal of Business*, 38(1), 34-105.
- Fama, E. F. (1965b). Random Walks in Stock Market Prices. *Financial Analysts Journal*, 21(5), 55-59. <http://www.jstor.org/stable/4469865>
- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2), 383-417. <https://doi.org/10.2307/2325486>
- Fama, E. F., & French, K. R. (2004). The Capital Asset Pricing Model: Theory and Evidence. *The Journal of Economic Perspectives*, 18(3), 25-46. <http://www.jstor.org/stable/3216805>

- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1), 1-22. <https://doi.org/10.1016/j.jfineco.2014.10.010>
- Faust, K. (1997). Centrality in affiliation networks. *Social networks*, 19(2), 157-191. [https://doi.org/10.1016/S0378-8733\(96\)00300-0](https://doi.org/10.1016/S0378-8733(96)00300-0)
- Feld, S. L. (1981). The focused organization of social ties. *American journal of sociology*, 86(5), 1015-1035. <https://doi.org/10.1086/227352>
- Feld, S. L. (1982). Social structural determinants of similarity among associates. *American sociological review*, 47(6), 797-801. <https://doi.org/10.2307/2095216>
- Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4), 82-89. <https://doi.org/10.1145/2436256.2436274>
- Fellner-Röhling, G. (2004). *Illusion of control as a source of poor diversification: An experimental approach*. Max Planck Institute of Economics, Strategic Interaction Group.
- Fenton-O'Creevy, M., Nicholson, N., Soane, E., & Willman, P. (2003). Trading on illusions: Unrealistic perceptions of control and trading performance. *Journal of occupational and organizational psychology*, 76(1), 53-68.
- Fernandez, R. M., & Gould, R. V. (1994). A dilemma of state power: Brokerage and influence in the national health policy domain. *American journal of Sociology*, 99(6), 1455-1491. <https://doi.org/10.1086/230451>
- Ferrara, E., & Yang, Z. (2015). Measuring emotional contagion in social media. *PloS one*, 10(11), e0142390.
- Festinger, L. (1954). A theory of social comparison processes. *Human relations*, 7(2), 117-140. <https://doi.org/10.1177/001872675400700202>
- Festinger, L., Back K. W., & Schachter, S. (1950). *Social pressures in informal groups: A study of human factors in housing*. Stanford University Press.
- Firth, M. (1977). The Capital Asset Pricing Model. In: *The Valuation of Shares and the Efficient-Markets Theory*. Studies in Finance and Accounting. Palgrave. https://doi.org/10.1007/978-1-349-15819-5_5
- FMA. (2021, June 28). A guide to talking about money online. <https://www.fma.govt.nz/library/articles/guide-to-talking-about-money-online/>
- Freud, S. (1989). The ego and the id (1923). *TACD Journal*, 17(1), 5-22. <https://doi.org/10.1080/1046171X.1989.12034344>

- Friedman, M. 1953. The methodology of positive economics. In M. Friedman (Ed.), *Essays in Positive Economics* (pp. 3–43). University of Chicago Press.
- Frizzo-Barker, J., Chow-White, P. A., Mozafari, M., & Ha, D. (2016). An empirical study of the rise of big data in business scholarship. *International Journal of Information Management*, 36(3), 403-413. <https://doi.org/10.1016/j.ijinfomgt.2016.01.006>
- Fung, G. P. C., Yu, J. X., & Lu, H. (2005). The Predicting Power of Textual Information on Financial Markets. *IEEE Intelligent Informatics Bulletin*, 5(1), 1-10.
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *The Journal of Finance*, 77(1), 5-47. <https://doi.org/10.1111/jofi.13090>
- Gadarowski, C. (2002). Financial press coverage and expected stock returns. *EFMA 2002 London Meetings*.
- Garcia-Lopez, F. J., Batyrshin, I., & Gelbukh, A. (2018). Analysis of relationships between tweets and stock market trends. *Journal of Intelligent & Fuzzy Systems*, 34(5), 3337-3347. <https://doi.org/10.3233/JIFS-169515>
- Gartner. (2012). *Big data*. <https://www.gartner.com/en/information-technology/glossary/big-data> adresinden 4 Mart 2022 tarihinde alınmıştır.
- Géron, A. (2021). *Scikit-Learn, Keras ve TensorFlow ile Uygulamalı Makine Öğrenmesi* (1. baskı). (B. Aksoy, & Ö. Kaya, Çev.). Buzdağı Yayınevi. (Orijinal eserin basım tarihi 2019).
- Gidofalvi, G., & Elkan, C. (2001). Using news articles to predict stock price movements. *Department of computer science and engineering, University of California*.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. *European review of social psychology*, 2(1), 83-115. <https://doi.org/10.1080/14792779143000033>
- Gigerenzer, G. (1996). On narrow norms and vague heuristics: A reply to Kahneman and Tversky. *Psychological Review*, 103(3), 592-596. <https://psycnet.apa.org/doi/10.1037/0033-295X.103.3.592>
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Penguin.
- Gigerenzer, G. (2008). *Rationality for mortals: How people cope with uncertainty*. Oxford University Press.
- Gigerenzer, G. (2008). Why heuristics work. *Perspectives on psychological science*, 3(1), 20-29. <https://doi.org/10.1111/j.1745-6916.2008.00058.x>

- Gigerenzer, G. (2014). *Risk savvy: How to make good decisions*. Penguin Books.
- Gigerenzer, G. (2018). The bias bias in behavioral economics. *Review of Behavioral Economics*, 5(3-4), 303-336. <https://doi.org/10.1561/105.00000092>
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual review of psychology*, 62(1), 451-482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological review*, 103(4), 650-669. <https://psycnet.apa.org/doi/10.1037/0033-295X.103.4.650>
- Glassman, D. A., & Riddick, L. A. (2001). What causes home asset bias and how should it be measured?. *Journal of Empirical Finance*, 8(1), 35-54.
- Goldberg, Y. (2017). *Neural network methods for natural language processing*. Morgan & Claypool.
- Golder, S. A., & Macy, M. W. (2014). Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology*, 40(1), 129-152. <https://doi.org/10.1146/annurev-soc-071913-043145>
- Gomez-Carrasco, P., & Michelon, G. (2017). The power of stakeholders' voice: The effects of social media activism on stock markets. *Business Strategy and the Environment*, 26(6), 855-872. <https://doi.org/10.1002/bse.1973>
- Goodfellow, I., & Bengio, Y., & Courville, A. (2018) *Derin öğrenme*. (F. Yarman Vural, R. G. Cinbiş ve S. Kalkan, Çev.). Buzdağı Yayınevi. (Orijinal eserin basım tarihi 2016).
- Gorodkin, J. (2004). Comparing two K-category assignments by a K-category correlation coefficient. *Computational biology and chemistry*, 28(5-6), 367-374. <https://doi.org/10.1016/j.compbiolchem.2004.09.006>
- Gould, R. V., & Fernandez, R. M. (1989). Structures of mediation: A formal approach to brokerage in transaction networks. *Sociological methodology*, 9, 89-126. <https://doi.org/10.2307/270949>
- Griffin, J. M., Hirschey, N. H., & Kelly, P. J. (2011). How Important Is the Financial Media in Global Markets? *The Review of Financial Studies*, 24(12), 3941–3992. <http://www.jstor.org/stable/41302005>
- Groß-Klußmann, A., König, S., & Ebner, M. (2019). Buzzwords build momentum: Global financial Twitter sentiment and the aggregate stock market. *Expert Systems with Applications*, 136, 171-186. <https://doi.org/10.1016/j.eswa.2019.06.027>

- Groth, S. S., & Muntermann, J. (2011). An intraday market risk management approach based on textual analysis. *Decision Support Systems*, 50(4), 680-691. <https://doi.org/10.1016/j.dss.2010.08.019>
- Guan, S. S. (2022). The Rise of the Finfluencer. *New York University Journal of Law and Business*, *Forthcoming*, Santa Clara Univ. Legal Studies Research Paper *Forthcoming*. Paper No. 4400042.
- Guijarro, F., Moya-Clemente, I., & Saleemi, J. (2019). Liquidity risk and investors' mood: Linking the financial market liquidity to sentiment analysis through Twitter in the S&P500 index. *Sustainability*, 11(24), 7048. <https://doi.org/10.3390/su11247048>
- Guild, S. & Johnson, K. (2022, October 27) . A fresh look at how to empower retail investors. <https://www.weforum.org/agenda/2022/10/a-fresh-look-at-how-to-empower-retail-investors/>
- Güran, A., & Ates, E. (2021). Pearson correlation and Granger causality analysis of Twitter sentiments and the daily changes in Bist30 index returns. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 36(3), 1687-1701. <https://doi.org/10.17341/gazimmfd.660018>
- Guttmann, A. (2023, Jan 9). Time spent with digital vs. traditional media in the U.S. 2011-2023. <https://www.statista.com/statistics/565628/time-spent-digital-traditional-media-usa/>
- Hansen, D., Shneiderman, B., & Smith, M. A. (2010) *Analyzing social media networks with nodexl: Insights from a connected world*, Morgan Kaufmann.
- Hariri, R. H., Fredericks, E. M., & Bowers, K. M. (2019). Uncertainty in big data analytics: survey, opportunities, and challenges. *Journal of Big Data*, 6(1), 1-16. <https://doi.org/10.1186/s40537-019-0206-3>
- Harvard IV-4. (2018) Descriptions of Inquirer Categories and Use of Inquirer Dictionaries, <https://inquirer.sites.fas.harvard.edu/homecat.htm> adresinden 11 Temmuz 2018 tarihinde alınmıştır.
- Hassanein, A., Mostafa, M. M., Benameur, K. B., & Al-Khasawneh, J. A. (2021). How do big markets react to investors' sentiments on firm tweets?. *Journal of Sustainable Finance & Investment*, 1-23. <https://doi.org/10.1080/20430795.2021.1949198>
- Hastie, R., & Dawes, R. M. (2009), *Rational choice in an uncertain world: The psychology of judgment and decision making*. Sage Publications.
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. *Current directions in psychological science*, 2(3), 96-100. <https://doi.org/10.1111/1467-8721.ep10770953>

- Hayek, F. A. (1937). Economics and knowledge. *Economica*, 4(13), 33-54. <https://doi.org/10.2307/2548786>
- Hayek, F. A. (1945). The use of knowledge in society. *American Economic Review*, 35(4), 519-530. <https://www.jstor.org/stable/1809376>
- Hayek, F. A. (1955). *The counter-revolution of science: Studies on the abuse of reason*. The Free Press of Glencoe.
- Hayek, F. A. (2012). *Law, legislation and liberty: a new statement of the liberal principles of justice and political economy*. Routledge.
- Haythornthwaite, C. (2005). Social networks and Internet connectivity effects. *Information, Community & Society*, 8(2), 125-147. <https://doi.org/10.1080/13691180500146185>
- He, W., Guo, L., Shen, J., & Akula, V. (2016). Social media-based forecasting: A case study of tweets and stock prices in the financial services industry. *Journal of Organizational and End User Computing (JOEUC)*, 28(2), 74-91. <https://doi.org/10.4018/JOEUC.2016040105>
- Healy, A. D., & Lo, A. W. (2011). Managing real-time risks and returns: the thomson reuters newsscope event indices. In G. Mitra, & L. Mitra (Eds.), *The handbook of news analytics in finance* (pp. 73-108). John Wiley & Sons.
- Hegde, M. S., Krishna, G., & Srinath, R. (2018, September). An ensemble stock predictor and recommender system. In *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1981-1985). IEEE. <https://doi.org/10.1109/ICACCI.2018.8554424>
- Hendler, J. (2009). Web 3.0 emerging. *Computer*, 42(1), 111-113. <https://doi.org/10.1109/MC.2009.30>
- Hicks, J. R., & Allen, R. G. (1934). A reconsideration of the theory of value. Part I. *Economica*, 1(1), 52-76. <https://doi.org/10.2307/2548574>
- Hira, İ. (2000). Sosyal bilimler: yasa koyucu tasarımdan yorumcu tasarıma. *Bilgi Sosyal Bilimler Dergisi*, 2(2), 81-97.
- Hirshleifer, D., & Hong Teoh, S. (2003). Herd behaviour and cascading in capital markets: A review and synthesis. *European Financial Management*, 9(1), 25-66.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

- Houghton, S. M., Simon, M., Aquino, K., & Goldberg, C. B. (2000). No safety in numbers: Persistence of biases and their effects on team risk perception and team decision making. *Group & Organization Management*, 25(4), 325-353.
- Huang, B., Yu, G., & Karimi, H. R. (2014). The finding and dynamic detection of opinion leaders in social network. *Mathematical problems in engineering*, 2014. <https://doi.org/10.1155/2014/328407>
- Huang, H., & Zhao, T. (2021, April). Stock market prediction by daily news via natural language processing and machine learning. In *2021 International Conference on Computer, Blockchain and Financial Development (CBFD)* (pp. 190-196). IEEE. <https://doi.org/10.1109/CBFD52659.2021.00044>
- IBM Cloud Education. (n.d.). Structured vs unstructured data. *IBM Cloud Education*. <https://www.ibm.com/cloud/blog/structured-vs-unstructured-data> adresinden 16 Mart 2020 tarihinde alınmıştır.
- İlhan, B., & Bayir, M. (2021). BİST Sınai ve BİST Mali Endeksi ile CDS, Faiz, Döviz Kuru, Toplam Krediler ve COVID-19 Arasındaki Dinamik İlişki. *Third Sector Social Economic Review*, 56(4), 3090-3110.
- İnan, İ. (2012). *Dil Felsefesi*. Anadolu Üniversitesi Yayınları.
- Investopedia. (n.d.). <https://www.investopedia.com/>
- Jevons, W. S. (1871). *The theory of political economy*. Macmillan and Company.
- Jurman, G., Riccadonna, S., & Furlanello, C. (2012). A comparison of MCC and CEN error measures in multi-class prediction.
- Kağıtçıbaşı, Ç., & Cemalcılar, Z. (2014). *Dünden Bugüne İnsan ve İnsanlar: Sosyal Psikolojiye Giriş*. (14. Baskı). Evrim Yayınevi
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American economic review*, 93(5), 1449-1475. <https://doi.org/10.1257/00028280322655392>
- Kahneman, D. (2011). *Thinking, fast and slow*. Penguin Books.
- Kahneman, D. (2015). *Hızlı ve yavaş düşünme* (2. baskı). (F. Deniztekin ve O. Ç. Deniztekin, Çev.). Varlık Yayınları. (Orijinal eserin basım tarihi 2011).
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49-81). Cambridge university press. <https://psycnet.apa.org/doi/10.1017/CBO9780511808098.004>

- Kahneman, D., & Frederick, S. (2004). Attribute substitution in intuitive judgment. *Models of a man: Essays in memory of Herbert A. Simon*, 411-432.
- Kahneman, D., & Frederick, S. (2005). A model of heuristic judgment. In K. J. Holyoak, & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 267-294). Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive psychology*, 3(3), 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237.
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103(3), 582–591. <https://doi.org/10.1037/0033-295X.103.3.582>
- Kalev, P. S., & Duong, H. N. (2011). Firm-specific news arrival and the volatility of intraday stock index and futures returns. In G. Mitra, & L. Mitra (Eds.), *The Handbook of News Analytics in Finance* (pp. 299-288). John Wiley & Sons. <https://doi.org/10.1002/9781118467411.ch12>
- Kaminski, J. (2014). Nowcasting the bitcoin market with twitter signals. *arXiv preprint arXiv:1406.7577*. <https://doi.org/10.48550/arXiv.1406.7577>
- Kamoune, A., & Ibenrissoul, N. (2022). Traditional versus Behavioral Finance Theory. *International Journal of Accounting, Finance, Auditing, Management and Economics*, 3(2-1), 282-294.
- Kandel, E. R. (2011). *Belleğin peşinde: Yeni bir zihin biliminin doğuşu* (M. Doğan, Çev.). Boğaziçi Üniversitesi Yayınevi. (Orijinal eserin basım tarihi 2007).
- Kapoor, K. K., Tamilmani, K., Rana, N. P., Patil, P., Dwivedi, Y. K., & Nerur, S. (2018). Advances in social media research: Past, present and future. *Information Systems Frontiers*, 20(3), 531-558. <https://doi.org/10.1007/s10796-017-9810-y>
- Karan, M. B. (2022). *Yatırım Analizi ve Portföy Yönetimi* (8. Baskı). Gazi Kitabevi.
- Kearney, M. W. (2020). *rtweet: collecting twitter data*. R package version 0.7.0. <https://cran.r-project.org/package=rtweet> adresinden 07 Nisan 2020 tarihinde alınmıştır.
- Keynes, J. M. (1935). *The general theory of employment interest and money*. Harcourt Brace & Co.

- Keynes, J. M. (1937). The general theory of employment. *The quarterly journal of economics*, 51(2), 209-223. <https://doi.org/10.2307/1882087>
- Khalil, S., & Fakir, M. (2017). RCrawler: An R package for parallel web crawling and scraping. *SoftwareX*, 6, 98-106. <https://doi.org/10.1016/j.softx.2017.04.004>
- Kindleberger, C. P., Aliber, R., & Solow, R. (2011). *Manias, Panics and Crashes: A History of Financial Crises*. Palgrave Macmillan.
- Kırmızıaltın, E. (2017). *İktisat ve Rasyonalite*. Heretik Yayınları.
- Kirzner, I. M. (1992). *The meaning of the market process: Essays in the development of modern Austrian economics*. Routledge.
- Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. *Big Data & Society*, 3(1), 1-10. <https://doi.org/10.1177/2053951716631130>
- Kıyılar, M. & Akkaya, M. (2020). *Davranışsal finans* (2. Baskı). Literatür Yayınları.
- Kliger, D., & Kudryavtsev, A. (2010). The availability heuristic and investors' reaction to company-specific events. *The Journal of Behavioral Finance*, 11(1), 50-65.
- Knight, F. H. (1940). "What is truth" in economics?. *Journal of Political Economy*, 48(1), 1-32. <https://doi.org/10.1086/255500>
- Knight, F. H. (2016). *Risk, Uncertainty and Profit*. Forgotten Books. (Original eserin basım tarihi 1921).
- Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24), 8788-8790.
- Kuhn, M., & Johnson, K. (2020). *Feature Engineering and Selection: A Practical Approach for Predictive Models*. CRC Press
- Kuhn, S. (2019). Prisoner's Dilemma, *The Stanford Encyclopedia of Philosophy* (Winter 2019 ed.).(E. N. Zalta (Ed.)). <https://plato.stanford.edu/archives/win2019/entries/prisoner-dilemma>
- Kulkarni, S. V., & Dhage, S. N. (2019). Advanced credit score calculation using social media and machine learning. *Journal of Intelligent & Fuzzy Systems*, 36(3), 2373-2380. . <https://doi.org/10.3233/JIFS-169948>

- Kwak, H., Lee, C., Park, H., & Moon, S. (2010, April). What is Twitter, a social network or a news media?. In *Proceedings of the 19th international conference on World wide web* (pp. 591-600). <https://doi.org/10.1145/1772690.1772751>
- L'heureux, A., Grolinger, K., Elyamany, H. F., & Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *Ieee Access*, 5, 7776-7797. <https://doi.org/10.1109/ACCESS.2017.2696365>
- Langager, C. (2022). Industry vs. sector: What's the difference?. . In *investopedia.com* <https://www.investopedia.com/ask/answers/05/industrysector.asp> adresinden 22 Aralık 2022 tarihinde alınmıştır.
- Langer, E. J. (1975). The illusion of control. *Journal of personality and social psychology*, 32(2), 311.
- Lavoie, D. (Ed.). (1994). *Expectations and the Meaning of Institutions: Essays in Economics by Ludwig M. Lachmann* (1st ed.). Routledge. <https://doi.org/10.4324/9780203976555>
- Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., & Van Alstyne, M. (2009). Social science. Computational social science. *Science*, 323(5915), 721-723. <https://doi.org/10.1126/science.1167742>
- Leenders, R. T. A. (1997). Longitudinal behavior of network structure and actor attributes: modeling interdependence of contagion and selection. In P. Doreian, & F. N. Stokman (Eds.), *Evolution of social networks* (pp. 165-184). Routledge.
- Leinweber, D., & Sisk, J. (2011). Relating news analytics to stock returns. In G. Mitra, & L. Mitra (Eds.), *The Handbook of News Analytics in Finance* (pp. 149-172). John Wiley & Sons. <https://doi.org/10.1002/9781118467411.ch6>
- Lewis, K., Gonzalez, M., & Kaufman, J. (2012). Social selection and peer influence in an online social network. *Proceedings of the National Academy of Sciences*, 109(1), 68-72. <https://doi.org/10.1073/pnas.1109739109>
- Li, B., Chan, K. C., & Ou, C. (2014, November). Public sentiment analysis in Twitter data for prediction of a company's stock price movements. In *2014 IEEE 11th International Conference on e-Business Engineering* (pp. 232-239). IEEE. <https://doi.org/10.1109/ICEBE.2014.47>
- Li, B., Chan, K. C., Ou, C., & Ruifeng, S. (2017). Discovering public sentiment in social media for predicting stock movement of publicly listed companies. *Information Systems*, 69, 81-92. <https://doi.org/10.1016/j.is.2016.10.001>

- Li, F-F., Krishna, R., & Xu, D. (2020). *Lecture 10: Recurrent neural networks*. [PowerPoint slides]. CS231n: Convolutional Neural Networks for Visual Recognition. CS231n @Stanford. http://cs231n.stanford.edu/slides/2020/lecture_10.pdf
- Li, K. (2004). Confidence in the familiar: An international perspective. *Journal of Financial and Quantitative Analysis*, 39(1), 47-68.
- Li, M., Yang, C., Zhang, J., Puthal, D., Luo, Y., & Li, J. (2018a, January). Stock market analysis using social networks. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-10). <https://doi.org/10.1145/3167918.3167967>
- Li, Q., Zhou, B., & Liu, Q. (2016, July). Can twitter posts predict stock behavior?: A study of stock market with twitter social emotion. In *2016 IEEE International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)* (pp. 359-364). IEEE. <https://doi.org/10.1109/ICCCBDA.2016.7529584>
- Li, T., van Dalen, J., & van Rees, P. J. (2018b). More than just noise? Examining the information content of stock microblogs on financial markets. *Journal of Information Technology*, 33(1), 50-69. <https://doi.org/10.1057/s41265-016-0034-2>
- Li, Y., Ma, S., Zhang, Y., & Huang, R. (2013). An improved mix framework for opinion leader identification in online learning communities. *Knowledge-Based Systems*, 43, 43-51. <https://doi.org/10.1016/j.knosys.2013.01.005>
- Liew, J., & Budavari, T. (2017). The “sixth” factor—a social media factor derived directly from tweet sentiments. *The Journal of Portfolio Management*, 43(3), 102-111. <https://doi.org/10.3905/jpm.2017.43.3.102>
- Lin, J. T., Bumcrot, C., Mottola, G., Valdes, O., and Walsh, G. (2022). *The Changing Landscape of Investors in the United States: A Report of the National Financial Capability Study*. FINRA Investor Education Foundation. <https://www.finrafoundation.org/sites/finrafoundation/files/NFCS-Investor-Report-Changing-Landscape.pdf>
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1), 13–37. <https://doi.org/10.2307/1924119>
- Lisa, M., & Bot, H. (2017). My Research Software (Version 2.0.4) [Computer software]. <https://doi.org/10.5281/zenodo.1234>
- Liu, H. (2011). Feature Selection. In: Sammut, C., Webb, G.I. (Eds) *Encyclopedia of machine learning*. Springer. https://doi.org/10.1007/978-0-387-30164-8_306

- Liu, P., Smith, S. D., & Syed, A. A. (1990). Stock price reactions to the Wall Street Journal's securities recommendations. *Journal of financial and Quantitative Analysis*, 25(3), 399-410.
- Liu, X. (2020). Analyzing the impact of user-generated content on B2B Firms' stock performance: Big data analysis with machine learning methods. *Industrial marketing management*, 86, 30-39. <https://doi.org/10.1016/j.indmarman.2019.02.021>
- Liu, Y., Yang, M., Wang, Y., Li, Y., Xiong, T., & Li, A. (2022). Applying machine learning algorithms to predict default probability in the online credit market: Evidence from China. *International Review of Financial Analysis*, 79, 101971. <https://doi.org/10.1016/j.irfa.2021.101971>
- Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368), 805-824.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of finance*, 66(1), 35-65. <https://doi.org/10.1111/j.1540-6261.2010.01625.x>
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30.
- Maas, A., & Tandon, S. (2013, September 26). Stanford Unsupervised Feature Learning and Deep Learning Tutorial. <http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNeuralNetworks/> adresinden 23 Mayıs 2020 tarihinde alınmıştır.
- MacFeely, S., Me, A., Fu, H., & Schweinfest, S. (2020, Nov 13). *We urgently need a Global Data Convention. Here's why*. <https://www.weforum.org/agenda/2020/11/global-data-convention> adresinden 13 Mart 2022 tarihinde alınmıştır.
- Mackay, C. (2012). *Extraordinary popular delusions and the madness of crowds*. Simon and Schuster.
- Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual review of sociology*, 28, 143-166. <https://doi.org/10.1146/annurev.soc.28.110601.141117>
- Malkiel, B. G. (1973). *A Random Walk Down Wall Street: Including A Life-Cycle Guide To Personal Investing*. W.W. Norton & Company, Inc.
- Manning, C. (2020a). *Lecture 1: Introduction and Word Vectors*. [PowerPoint slides]. Natural Language Processing with Deep Learning CS224N/Ling284. CS224N@Stanford. <http://web.stanford.edu/class/cs224n/slides/cs224n-2020-lecture01-wordvecs1.pdf>

- Manning, C. (2020b). *Lecture 2: Word vectors, word senses, and classifier review*. [PowerPoint slides]. Natural Language Processing with Deep Learning CS224N/Ling284. CS224N@Stanford. <http://web.stanford.edu/class/cs224n/slides/cs224n-2020-lecture02-wordvecs2.pdf>
- Mao, Y., Wei, W., Wang, B., & Liu, B. (2012, August). Correlating S&P 500 stocks with Twitter data. In *Proceedings of the first ACM international workshop on hot topics on interdisciplinary social networks research* (pp. 69-72). <https://doi.org/10.1145/2392622.2392634>
- Marchionatti, R., & Gambino, E. (1997). Pareto and political economy as a science: methodological revolution and analytical advances in economic theory in the 1890s. *Journal of Political Economy*, 105(6), 1322-1348. <https://doi.org/10.1086/516395>
- Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1), 77-91. <https://doi.org/10.2307/2975974>
- Markowitz, H. (1959). *Portfolio selection: efficient diversification of investments*, Cowies Foundation for Research in Economics at Yale University Monograph 16, John Wiley & Sons. <https://www.jstor.org/stable/j.ctt1bh4c8h>
- Marwick, B. (n.d.). *Convert pdfs to text files or csv files (dfr format) with r*. In *github.com*. <https://gist.github.com/benmarwick/11333467> adresinden 25 Aralık 2019 tarihinde alınmıştır.
- Matta, M., Lunesu, I., & Marchesi, M. (2015, June). Bitcoin Spread Prediction Using Social and Web Search Media. In *UMAP workshops* (pp. 1-10).
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. Houghton Mifflin Harcourt.
- McCabe, C.(2022, January 16). Day Traders as ‘Dumb Money’? The Pros Are Now Paying Attention. Wall Street Journal. <https://www.wsj.com/articles/fund-managers-pay-attention-to-retail-day-traders-11642132135>
- McCann, B., Bradbury, J., Xiong, C., & Socher, R. (2017). Learned in translation: Contextualized word vectors. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in neural information processing systems 30* (pp. 6294-6305). Curran Associates.
- McCormick, C., & Ryan, N. (2019). BERT fine-tuning tutorial with pytorch. <https://mccormickml.com/2019/07/22/BERT-fine-tuning/#sentence-length--attention-mask>
- MDZ Digital Library. (n.d.). Turkish BERT model. <https://huggingface.co/dbmdz/bert-base-turkish-uncased>

- Merriam-Webster. (n.d.). Social media. In *Merriam-Webster.com dictionary*. <https://www.merriam-webster.com/dictionary/social%20media> adresinden 13 Nisan 2022 tarihinde alınmıştır.
- Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2), 125-144. [https://doi.org/10.1016/0304-405X\(76\)90022-2](https://doi.org/10.1016/0304-405X(76)90022-2)
- Metz, C. (2016, January 25). *The rise of the artificially intelligent hedge fund*. <https://www.wired.com/2016/01/the-rise-of-the-artificially-intelligent-hedge-fund>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*. <https://doi.org/10.48550/arXiv.1301.3781>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In C.J. Burges, L. Bottou, M. Welling, Z. Ghahramani, & K.Q. Weinberger (Eds.), *Advances in neural information processing systems* 26 (pp. 3111-31190). Curran Associates.
- Milgram, S. (1967). The small world problem. *Psychology today*, 2(1), 60-67.
- Mill, J. S. (2008). On the definition and method of political economy. In D. M. Hausman (Ed.), *The philosophy of economics: An anthology* (pp. 41-58.). Cambridge University Press.
- Miller, M. H., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *The journal of business*, 34(4), 411-433. <https://www.jstor.org/stable/2351143>
- Ming, F., Wong, F., Liu, Z., & Chiang, M. (2014, December). Stock market prediction from WSJ: text mining via sparse matrix factorization. In *2014 IEEE International Conference on Data Mining* (pp. 430-439). IEEE. <https://doi.org/10.1109/ICDM.2014.116>
- Mirowski, P., & Cook, P. (1990). Walras’“economics and mechanics”: Translation, commentary, context. In *Economics as Discourse* (pp. 189-224). Springer, Dordrecht.
- Mises, L. (2008). *İnsan eylemi: Ekonomi üzerine bir inceleme*. Liberte Yayınları.
- Mishra, S., & Datta-Gupta, A. (2017). *Applied statistical modeling and data analytics: A practical guide for the petroleum geosciences*. Elsevier.
- Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill.

- Mitra, L., & Mitra, G. (2011). Applications of news analytics in finance: A review. In G. Mitra, & L. Mitra (Eds.), *The Handbook of News Analytics in Finance* (pp. 1-39). John Wiley & Sons. <https://doi.org/10.1002/9781118467411.ch1>
- Mitra, L., Mitra, G., & Dibartolomeo, D. (2009). Equity portfolio risk estimation using market information and sentiment. *Quantitative Finance*, 9(8), 887-895. <https://doi.org/10.1080/14697680903448361>
- Mittermayer, M. A. (2004, January). Forecasting intraday stock price trends with text mining techniques. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the* (pp. 10-pp). IEEE. <https://doi.org/10.1109/HICSS.2004.1265201>
- Moat, H. S., Preis, T., Olivola, C. Y., Liu, C., & Chater, N. (2014). Using big data to predict collective behavior in the real world. *Behavioral and Brain Sciences*, 37(1), 92-93. <https://doi.org/10.1017/S0140525X13001817>
- Modigliani, F., & Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American economic review*, 48(3), 261-297. <https://www.jstor.org/stable/1809766>
- Mohammad, S. M., & Turney, P. (2011). *NRC Word-Emotion Association Lexicon*. NRC: <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>
- MongoDB. (n.d.). *Unstructured data*. <https://www.mongodb.com/unstructured-data> adresinden 16 Mart 2020 tarihinde alınmıştır.
- Montier, J. (2007). *Behavioural investing: a practitioner's guide to applying behavioural finance*. John Wiley & Sons.
- Moscatelli, M., Parlapiano, F., Narizzano, S., & Viggiano, G. (2020). Corporate default forecasting with machine learning. *Expert Systems with Applications*, 161, 113567. <https://doi.org/10.1016/j.eswa.2020.113567>
- Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica*, 34(4), 768–783. <https://doi.org/10.2307/1910098>
- Munzert, S., Rubba, C., Meißner, P., & Nyhuis, D. (2014). *Automated data collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons.
- Naik, U., & Shivalingaiah, D. (2008). Comparative study of web 1.0, web 2.0 and web 3.0. (pp. 499-507). <http://ir.inflibnet.ac.in/ir40/bitstream/1944/1285/1/54.pdf>
- Narula, S. K. (2022, January 25). What Generation Z Wants From Financial Technology. *Wall Street Journal*. <https://www.wsj.com/articles/generation-z-financial-technology-11642714326>

- Nasaa. (2022, August) Informed Investor Advisory: Finfluencers. Are you an informed investor? Financial Advice via Social Media – the Rise of the “Finfluencer”. <https://www.nasaa.org/64940/informed-investor-advisory-finfluencers/> adresinden 30 Nisan 2023 tarihinde alınmıştır.
- Nasseri, A., Tucker, A., & De Cesare, S. (2015). Quantifying StockTwits semantic terms' trading behavior in financial markets: An effective application of decision tree algorithms. *Expert systems with applications*, 42(23), 9192-9210. <https://doi.org/10.1016/j.eswa.2015.08.008>
- Netzer, O., Lemaire, A., & Herzenstein, M. (2019). When words sweat: Identifying signals for loan default in the text of loan applications. *Journal of Marketing Research*, 56(6), 960-980. <https://doi.org/10.1177/0022243719852959>
- Neumann, J. V., & Morgenstern, O. (1953). *Theory of games and economic behavior*. Princeton University Press.
- Neural networks. (n.d.). CS231n convolutional neural networks for visual recognition. CS231n @Stanford. <https://cs231n.github.io/neural-networks-1>
- Nicholson, W., & Snyder C. M. (2011). *Microeconomic theory: Basic principles and extensions* (11th ed.). South-Western College Pub.
- Nielsen, F. Å. (2011, May). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages* 718 in *CEUR Workshop Proceedings* 93-98. <https://arxiv.org/abs/1103.2903>. <https://doi.org/10.48550/arXiv.1103.2903>
- Nietzsche, F. (2008). *The gay science*. Cambridge University Press.
- Nilsson, N. J. (2011). *Yapay zekâ: Geçmişi ve geleceği* (M. Doğan, Çev.). Boğaziçi Üniversitesi Yayınevi.
- Nisar, T. M., & Yeung, M. (2018). Twitter as a tool for forecasting stock market movements: A short-window event study. *The journal of finance and data science*, 4(2), 101-119. <https://doi.org/10.1016/j.jfds.2017.11.002>
- Niu, B., Ren, J., & Li, X. (2019). Credit scoring using machine learning by combing social network information: Evidence from peer-to-peer lending. *Information*, 10(12), 397. <https://doi.org/10.3390/info10120397>
- Nofsinger, J. R. (2005). *The Psychology of Investing* (2nd ed.). Pearson Prentice Hall.
- Nolen-Hoeksema, S. (2009). *Atkinson & Hilgard's introduction to psychology* (15th ed.). Cengage Learning.

- Novick, B., Mayston, D., Marcus, S., Barry, R., Fox, G., Betts, B., Pasqual, S., & Eisenmann, K. (2019). *Artificial intelligence and machine learning in asset management* [White paper]. Blackrock. <https://www.blackrock.com/corporate/literature/whitepaper/viewpoint-artificial-intelligence-machine-learning-asset-management-october-2019.pdf> adresinden 25 Nisan 2020 tarihinde alınmıştır.
- O'Brien, D. P. (2004). *The classical economists revisited*. Princeton University Press.
- Oğuz, F. (1999). Hayek'in Pratik Bilgi Anlayışı Üzerine Kısa Bir Yorum. *Liberal Düşünce*, 4(16), 92-102.
- Oğuz, F. (2001). Piyasa Süreci Teorisi: Tarihsel Gelişim, *Liberal Düşünce*, 6(21), 73-85.
- Olah, C. (2015, August 27). Understanding LSTM Networks. *Colah's Blog*. <https://colah.github.io/posts/2015-08-Understanding-LSTMs>
- Onay, C., & Öztürk, E. (2018). A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance*, 26(3), 382-405. <https://doi.org/10.1108/JFRC-06-2017-0054>
- Ooms, J. (October 16, 2022). *Text extraction, rendering and converting of pdf documents*. (Version 3.3.2), [Package for R]. Cran.r-project. <https://cran.r-project.org/web/packages/pdftools/index.html>
- Organisation for Economic Co-operation and Development. (2019). *Scoping the OECD AI principles: Deliberations of the Expert Group on Artificial Intelligence at the OECD (AIGO)*. OECD Digital Economy Papers, 291, OECD Publishing. <https://doi.org/10.1787/d62f618a-en>.
- Organisation for Economic Co-operation and Development. (2021a). *Artificial Intelligence, Machine Learning and Big Data in Finance: Opportunities, Challenges, and Implications for Policy Makers*. OECD Publishing. <https://www.oecd.org/finance/artificial-intelligence-machine-learning-big-data-in-finance.htm>
- Organisation for Economic Co-operation and Development. (2021b). *Scoping The OECD AI Principles Deliberations of The Expert Group On Artificial Intelligence at The OECD (AIGO)*. OECD Publishing. <https://www.oecd.org/finance/artificial-intelligence-machine-learning-big-data-in-finance.htm>
- Othan, D., & Kilimci, Z. H. (2021, August). Stock Market Prediction with New Generation Deep Contextualized Word Representations and Deep Learning Models using User Sentiments. In *2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-6). IEEE. <https://doi.org/10.1109/INISTA52262.2021.9548419>

- Oussalah, M., & Zaidi, A. (2018, July). Forecasting weekly crude oil using twitter sentiment of us foreign policy and oil companies data. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)* (pp. 201-208). IEEE. <https://doi.org/10.1109/IRI.2018.00037>
- Özcan, K. (2011). Kurumsal söylemin rasyonel temelleri: Yeni kurumsal kuram bağlamında rasyonelite tartışması. *Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (13), 297-326.
- Özel, H. (2016). Sosyal Bilimin “İnsansızlaşması”. *İktisat ve Toplum*, 69, 4-10.
- Pagolu, V. S., Reddy, K. N., Panda, G., & Majhi, B. (2016, October). Sentiment analysis of Twitter data for predicting stock market movements. In *2016 international conference on signal processing, communication, power and embedded system (SCOPES)* (pp. 1345-1350). IEEE. <https://doi.org/10.1109/SCOPES.2016.7955659>
- Parau, P., Lemnaru, C., & Potolea, R. (2015, November). Assessing vertex relevance based on community detection. In *2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K)* (Vol. 1, pp. 46-56). IEEE.
- Parau, P., Lemnaru, C., Dinsoreanu, M., & Potolea, R. (2017). Opinion leader detection. In F. A. Pozzi, E. Fersini, E. Messina, & B. Liu (Eds.), *Sentiment analysis in social networks* (pp. 157-170). Morgan Kaufmann. <https://doi.org/10.1016/B978-0-12-804412-4.00010-3>
- Park, J., Konana, P., Gu, B., Kumar, A., & Raghunathan, R. (2013). Information valuation and confirmation bias in virtual communities: Evidence from stock message boards. *Information Systems Research*, 24(4), 1050-1067.
- Parkinson, M. (1980). The Extreme Value Method for Estimating the Variance of the Rate of Return. *The Journal of Business*, 53(1), 61–65. <http://www.jstor.org/stable/2352357>
- Pasupulety, U., Anees, A. A., Anmol, S., & Mohan, B. R. (2019, June). Predicting stock prices using ensemble learning and sentiment analysis. In *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)* (pp. 215-222). IEEE. <https://doi.org/10.1109/AIKE.2019.00045>
- Pedersen, L. H. (2022). Game on: Social networks and markets. *Journal of Financial Economics*, 146(3), 1097-1119.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830. https://scikit-learn.org/stable/modules/cross_validation.html

- Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).
- Peramunetilleke, D., & Wong, R. K. (2002). Currency exchange rate forecasting from news headlines. *Australian Computer Science Communications*, 24(2), 131-139.
- Peress, J. (2014). The media and the diffusion of information in financial markets: Evidence from newspaper strikes. *the Journal of Finance*, 69(5), 2007-2043.
- Peters, M. E., Ammar, W., Bhagavatula, C., & Power, R. (2017). Semi-supervised sequence tagging with bidirectional language models. *arXiv preprint arXiv:1705.00108*. <https://doi.org/10.48550/arXiv.1705.00108>
- Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., & Zettlemoyer, L. (2018). Deep Contextualized Word Representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, (pp 2227–2237). *arXiv:1802.05365*. <https://doi.org/10.48550/arXiv.1802.05365>
- Petrasic, K., Saul, B., Greig, J., & Lamberth, K. (2017, January 20). Algorithms and bias: What lenders need to know. <https://www.whitecase.com/publications/insight/algorithms-and-bias-what-lenders-need-know> adresinden 27 Nisan 2022 tarihinde alınmıştır.
- Picasso, A., Merello, S., Ma, Y., Oneto, L., & Cambria, E. (2019). Technical analysis and sentiment embeddings for market trend prediction. *Expert Systems with Applications*, 135, 60-70. <https://doi.org/10.1016/j.eswa.2019.06.014>
- Pindyck, R. S., & Rubinfeld, D. L. (2013). *Microeconomics* (8th ed.). Pearson.
- Plato. (2020). *Phaedrus* (B. Jowett, Trans.). Independently published. (Original eserin basım tarihi milattan önce 360). <http://classics.mit.edu/Plato/phaedrus.html>
- Plous, S. (1993). *The psychology of judgment and decision making*. McGraw-Hill Book Company.
- Pompian, M. (2006). *Behavioral finance and wealth management: building optimal portfolios that account for investor biases*. Wiley.
- Pompian, M. M. (2012). *Behavioral finance and investor types: Managing behavior to make better investment decisions* (1st ed.). John Wiley & Sons.
- Portes, J. (2020, March). Book Reviews: Narrative Economics. <https://www.imf.org/en/Publications/fandd/issues/2020/03/book-review-narrative-economics-portes>

- Pouget, S., & Villeneuve, S. (2009). Price formation with confirmation bias. *Working Paper. Toulouse School of Economics, University of Toulouse*. <http://www.creedexperiment.nl/enable2008/pouget.pdf>
- Raafat, R. M., Chater, N., & Frith, C. (2009). Herding in humans. *Trends in cognitive sciences*, 13(10), 420-428. <https://doi.org/10.1016/j.tics.2009.08.002>
- Rakowski, D., Shirley, S. E., & Stark, J. R. (2021). Twitter activity, investor attention, and the diffusion of information. *Financial Management*, 50(1), 3-46.
- Ranco, G., Aleksovski, D., Caldarelli, G., Grčar, M., & Mozetič, I. (2015). The effects of Twitter sentiment on stock price returns. *PloS one*, 10(9), e0138441. <https://doi.org/10.1371/journal.pone.0138441>
- Rane, H., & Salem, S. (2012). Social media, social movements and the diffusion of ideas in the Arab uprisings. *Journal of international communication*, 18(1), 97-111.
- Rao, T., & Srivastava, S. (2012, November). Tweetsmart: Hedging in markets through twitter. In *2012 Third international conference on emerging applications of information technology* (pp. 193-196). IEEE. <https://doi.org/10.1109/EAIT.2012.6407894>
- Rao, T., & Srivastava, S. (2013, May). Modeling movements in oil, gold, forex and market indices using search volume index and twitter sentiments. In *Proceedings of the 5th annual ACM Web science conference* (pp. 336-345). <https://doi.org/10.1145/2464464.2464521>
- Rautiainen, A., & Jokinen, J. (2022). The value-relevance of social media activity of Finnish listed companies. *International Journal of Accounting and Information Management*, 30(2), 301-323. <https://doi.org/10.1108/IJAIM-04-2021-0076>
- Rawell, R. (2023, April 13). Impact Of Social Media On Investments. <https://www.fisd.com/impact-of-social-media-on-investments/>
- Ricciardi, V., & Simon, H. K. (2000). What is behavioral finance?. *Business, Education & Technology Journal*, 2(2), 1-9.
- Richardson, M., & Domingos, P. (2002, July). Mining knowledge-sharing sites for viral marketing. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 61-70). <https://doi.org/10.1145/775047.775057>
- Robins, G. (2015). *Doing social network research: network-based research design for social scientist*. Sage Publications.

- Robins, G., Elliott, P., & Pattison, P. (2001). Network models for social selection processes. *Social networks*, 23(1), 1-30. [https://doi.org/10.1016/S0378-8733\(01\)00029-6](https://doi.org/10.1016/S0378-8733(01)00029-6)
- Rosario, B. (2000). Latent semantic indexing: An overview. Techn. rep. INFOSYS, 240, 1-16.
- Roslan, M. A. A., & Rahiman, M. H. F. (2018, August). Stock prediction using sentiment analysis in twitter for day trader. In *2018 9th IEEE Control and System Graduate Research Colloquium (ICSGRC)* (pp. 177-182). IEEE. <https://doi.org/10.1109/ICSGRC.2018.8657614>
- Roy, N. (2023, February 16). Other ML jargons: Sparse and dense representations of texts for machine learning. Medium. <https://towardsdatascience.com/other-ml-jargons-sparse-and-dense-representations-of-texts-for-machine-learning-21fcd7a01410>
- Ruan, Y., Alfantoukh, L., & Duresi, A. (2015, March). Exploring stock market using twitter trust network. In *2015 IEEE 29th International Conference on Advanced Information Networking and Applications* (pp. 428-433). IEEE. <https://doi.org/10.1109/AINA.2015.217>
- Ruben, E., & Dumludağ, D. (2015). Davranışsal İktisadın Gelişimi. *İktisat ve Toplum*, 58, 4-9.
- Rubin, Z. (1973). *Liking and loving: An invitation to social psychology*. Holt, Rinehart & Winston.
- Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information processing & management*, 24(5), 513-523.
- Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61-71. <https://doi.org/10.2307/2548836>
- Sarfati, M. (2011). Spinoza-Smith ve İktisat Teorisine Bir Eleştiri. L. Cantek (Ed.), *İktisatta Yeni Yaklaşımlar* içinde (s. 105-134). İletişim Yayınları.
- Sarıyer, M., Akıl, A., Bulgurcu, F. N., Öge, F. E., & Ganiz, M. C. (2022, August). Individual Stock Price Prediction by Using KAP and Twitter Sentiments with Machine Learning for BIST30. In *2022 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-6). IEEE. <https://doi.org/10.1109/INISTA55318.2022.9894172>
- Savage, L. J. (1954). *The foundations of statistics*, John Wiley & Sons Inc.
- Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *The American economic review*, 465-479.

- Schnaubelt, M., Fischer, T. G., & Krauss, C. (2020). Separating the signal from the noise—financial machine learning for twitter. *Journal of Economic Dynamics and Control*, 114, 103895. <https://doi.org/10.1016/j.jedc.2020.103895>
- Schoen, H., Gayo-Avello, D., Takis Metaxas, P., Mustafaraj, E., Strohmaier, M., & Gloor, P. (2013). The power of prediction with social media. *Internet research*, 23(5), 528-543.
- Schoenewolf, G. (1990). Emotional contagion: Behavioral induction in individuals and groups. *Modern Psychoanalysis*, 15(1), 49-61.
- Schopenhauer, A. (1966). *World as will and representation* (Vol. 1-2). Dover Publications.
- Schumaker, R. P., & Chen, H. (2006). Textual Analysis of Stock Market Prediction Using Financial News Articles. In *Proceedings of the Twelfth Americas Conference on Information Systems* (pp. 1432-1440).
- Schumaker, R. P., & Chen, H. (2009). A quantitative stock prediction system based on financial news. *Information Processing & Management*, 45(5), 571-583. <https://doi.org/10.1016/j.ipm.2009.05.001>
- Schumaker, R. P., Zhang, Y., Huang, C. N., & Chen, H. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems*, 53(3), 458-464. <https://doi.org/10.1016/j.dss.2012.03.001>
- Scott, J. (2013). *Social network analysis* (3rd ed.), Sage Publications.
- Scweet. (n.d.). *A simple and unlimited Twitter scraper*. In *github.com*. <https://github.com/Altimis/Scweet> adresinden 03 Şubat 2022 tarihinde alınmıştır.
- SEC. (2010, August 30). Analyzing Analyst Recommendations. From <https://www.sec.gov/about/reports-publications/investor-publications/analyzing-analyst-recommendations>
- Şeker, S. E. (2015a). Doğal Dil İşleme (Natural Language Processing). *YBS Ansiklopedi*, 2(4), 14-31.
- Şeker, S. E. (2015b). Sosyal ağlarda veri madenciliği (data mining on social networks). *YBS Ansiklopedi*, 2(2), 30-39.
- Sen, A. (1988). *On Ethics and Economics*. Blackwell Publishing.
- Sen, A. (1993). Internal Consistency of Choice. *Econometrica: Journal of the Econometric Society*, 61(3), 495-521. <https://doi.org/10.2307/2951715>

- Sezer, T. (2015, January 9). *Internet Slang – İnternet Argosu – Turkish Twitter Corpus* – <http://tanersezer.com/?p=239> adresinden 01 Eylül 2020 tarihinde alınmıştır.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, 19(3), 425-442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- Shi, L., Teng, Z., Wang, L., Zhang, Y., & Binder, A. (2018). DeepClue: visual interpretation of text-based deep stock prediction. *IEEE Transactions on Knowledge and Data Engineering*, 31(6), 1094-1108. <https://doi.org/10.1109/TKDE.2018.2854193>
- Shi, Y. (2022). Feature selection. In: *Advances in big data analytics*. Springer. https://doi.org/10.1007/978-981-16-3607-3_4
- Shiller, R. J. (1992). *Market volatility*. MIT press.
- Shiller, R. J. (1999). Human behavior and the efficiency of the financial system. *Handbook of macroeconomics*, 1, 1305-1340.
- Shiller, R. J. (2006). *Irrational Exuberance*. Crown Business.
- Shiller, R. J. (2021), Anlatı iktisadı: Hikâyeler nasıl yayılır ve büyük iktisadi olayları yönlendirir? (G. Kafa, Çev.). Albaraka Yayınları. (Orijinal eserin basım tarihi 2019).
- Shynkevich, Y., McGinnity, T. M., Coleman, S. A., & Belatreche, A. (2016). Forecasting movements of health-care stock prices based on different categories of news articles using multiple kernel learning. *Decision Support Systems*, 85, 74-83. <https://doi.org/10.1016/j.dss.2016.03.001>
- Si, J., Mukherjee, A., Liu, B., Pan, S. J., Li, Q., & Li, H. (2014, October). Exploiting social relations and sentiment for stock prediction. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1139-1145).
- Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior & Organization*, 107, 730-743. <https://doi.org/10.1016/j.jebo.2014.06.004>
- Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2017). Divergence of sentiment and stock market trading. *Journal of Banking & Finance*, 78, 130-141. <https://doi.org/10.1016/j.jbankfin.2017.02.005>
- Silva, J. M., Mahfujur Rahman, A. S. M., & El Saddik, A. (2008, October). Web 3.0: a vision for bridging the gap between real and virtual. In *Proceedings of the 1st ACM international workshop on Communicability design and evaluation in cultural and ecological multimedia system* (pp. 9-14). <https://doi.org/10.1145/1462039.1462042>

- Simon, H. A. (1955). A behavioral model of rational choice. *The quarterly journal of economics*, 69(1), 99-118. <https://doi.org/10.2307/1884852>
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63(2), 129-138. <https://psycnet.apa.org/doi/10.1037/h0042769>
- Simon, H. A. (1984). Models of bounded rationality: economic analysis and public policy. Vol. 1. MIT Press.
- Simon, H. A. (1990). Invariants of human behavior. *Annual review of psychology*, 41(1), 1-20. <https://psycnet.apa.org/doi/10.1146/annurev.ps.41.020190.000245>
- Simon, H. A. (2000). Bounded rationality in social science: Today and tomorrow. *Mind & Society*, 1(1), 25-39. <https://doi.org/10.1007/BF02512227>
- Simonson, I. (1992). The influence of anticipating regret and responsibility on purchase decisions. *Journal of consumer research*, 19(1), 105-118.
- Slovic, P., Finucane, M. L., Peters, E., & MacGregor, D. G. (2007). The affect heuristic. *European journal of operational research*, 177(3), 1333-1352. <https://doi.org/10.1016/j.ejor.2005.04.006>
- Smailović, J., Grčar, M., Lavrač, N., & Žnidaršič, M. (2014). Stream-based active learning for sentiment analysis in the financial domain. *Information sciences*, 285, 181-203. <https://doi.org/10.1016/j.ins.2014.04.034>
- Smith, A. (2016). *Milletlerin Zenginliği* (11. baskı). (H. Derin, Çev.). Türkiye İş Bankası Kültür Yayınları. (Orijinal eserin basım tarihi 1776).
- Smith, A. (2018). *Ahlaki duygular kuramı*. (1. baskı). (D. Kızılay, Çev.). Pinhan Yayıncılık. (Orijinal eserin basım tarihi 1759).
- Som, R. K. (1995). *Practical sampling techniques*. CRC press.
- Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, 28(1), 11-21.
- Sprenger, T. O., Tumasjan, A., Sandner, P. G., & Welpe, I. M. (2014). Tweets and trades: The information content of stock microblogs. *European Financial Management*, 20(5), 926-957. <https://doi.org/10.1111/j.1468-036X.2013.12007.x>
- Statman, M. (1999). Behavioral finance: Past battles and future engagements. *Financial analysts journal*, 55(6), 18-27. <https://doi.org/10.2469/faj.v55.n6.2311>
- Statman, M. (2008). What is behavioral finance. *Handbook of finance*, 2(9), 79-84.

- Statman, M. (2019). *Behavioral finance: The second generation*. CFA Institute Research Foundation.
- Stepanova, E. (2011). The role of information communication technologies in the “Arab Spring”. *Ponars Eurasia*, 15(1), 1-6.
- Strauß, N., Vliegenthart, R., & Verhoeven, P. (2018). Intraday news trading: The reciprocal relationships between the stock market and economic news. *Communication Research*, 45(7), 1054-1077. <https://doi.org/10.1177/0093650217705528>
- Sugden, R. (1989). Spontaneous order. *Journal of Economic perspectives*, 3(4), 85-97. <https://doi.org/10.1257/jep.3.4.85>
- Sugden, R. (1991). Rational choice: a survey of contributions from economics and philosophy. *The economic journal*, 101(407), 751-785. <https://doi.org/10.2307/2233854>
- Sul, H. K., Dennis, A. R., & Yuan, L. (2017). Trading on twitter: Using social media sentiment to predict stock returns. *Decision Sciences*, 48(3), 454-488.
- Sulphrey, M. M. (2014). *Behavioural finance*. PHI Learning Pvt. Ltd..
- Sun, A., Lachanski, M., & Fabozzi, F. J. (2016). Trade the tweet: Social media text mining and sparse matrix factorization for stock market prediction. *International Review of Financial Analysis*, 48, 272-281. <https://doi.org/10.1016/j.irfa.2016.10.009>
- Tabari, N., Seyeditabari, A., Peddi, T., Hadzikadic, M., & Zadrozny, W. (2019). A comparison of neural network methods for accurate sentiment analysis of stock market tweets. In *ECML PKDD 2018 Workshops: MIDAS 2018 and PAP 2018, Dublin, Ireland, September 10-14, 2018, Proceedings 3* (pp. 51-65). Springer International Publishing. https://doi.org/10.1007/978-3-030-13463-1_4
- Tardy, C. M. (2023). How epidemiologists exploit the emerging genres of twitter for public engagement. *English for Specific Purposes*, 70, 4-16.
- TDK. (2011). Manipülasyon (Yönlendirim). In *İktisat Terimleri Sözlüğü*. Türk Dil Kurumu Yayınları. (p. 471)
- TDK. (2009). Türk Dil Kurumu Yayınları. <https://sozluk.gov.tr>
- Temel, E. (2012). *Yapay Sinir Ağları* (3. baskı). Papatya Yayıncılık.
- Teplova, T., Sokolova, T., Tomtosov, A., Buchko, D., & Nikulin, D. (2022). The sentiment of private investors in explaining the differences in the trade characteristics of the Russian market stocks. *Journal of the New Economic Association*, 53(1), 53-84.

- Termbank. (n.d.). <http://www.termbank.net/psychology/5956.html> adresinden 07 Ağustos 2015 tarihinde alınmıştır.
- Tetlock, P. C. (2007). Giving Content to Investor Sentiment: The Role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139–1168. <http://www.jstor.org/stable/4622297>
- Thaler, R. (1985). Mental accounting and consumer choice. *Marketing science*, 4(3), 199-214. <https://doi.org/10.1287/mksc.4.3.199>
- Tiwari, S., Bharadwaj, A., & Gupta, S. (2017, December). Stock price prediction using data analytics. In *2017 International Conference on Advances in Computing, Communication and Control (ICAC3)* (pp. 1-5). IEEE. <https://doi.org/10.1109/ICAC3.2017.8318783>
- Torrey, L., & Shavlik, J. (2010). Transfer learning. In E. S. Olivas (Ed.), *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (pp. 242-264). IGI global. <https://doi.org/10.4018/978-1-60566-766-9.ch011>
- Transfer Learning. (n.d.). CS231n convolutional neural networks for visual recognition. CS231n @Stanford. <https://cs231n.github.io/transfer-learning>
- Travers, J., & Milgram, S. (1977). An experimental study of the small world problem. In S. Leinhardt (Ed.), *Social networks: A developing paradigm* (pp. 179-197). Academic Press. <https://doi.org/10.1016/C2013-0-11063-X>
- Treynor, Jack L. (1962). *Toward a Theory of Market Value of Risky Assets*. Unpublished manuscript. A final version was published in 1999, in *Asset Pricing and Portfolio Performance: Models, Strategy and Performance Metrics*. Robert A. Korajczyk (editor) London: Risk Books, pp. 15–22.
- Tsai, C. I., Klayman, J., & Hastie, R. (2008). Effects of amount of information on judgment accuracy and confidence. *Organizational Behavior and Human Decision Processes*, 107(2), 97-105.
- Tumarkin, R., & Whitelaw, R. F. (2001). News or noise? Internet postings and stock prices. *Financial Analysts Journal*, 57(3), 41-51.
- Türk Dilleri için Dil Kaynakları ve Araçları. (n.d.). <http://ddi.itu.edu.tr/araclarkaynaklar>
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological bulletin*, 76(2), 105.
- Tversky, A., & Kahneman, D., (1974), Judgment under Uncertainty: Heuristics and Biases, *Science*, 185(4175), 1124-1131.

- Twint. (n.d.). *TWINT project: Open-source Twitter intelligence*. In *github.com*. <https://github.com/twintproject/twint> adresinden 07 Nisan 2019 tarihinde alınmıştır.
- Twitter Developer Platform. (n.d.). <https://developer.twitter.com>
- U.S. Securities and Exchange Commission. (2015, November 5). Updated Investor Alert: Social Media and Investing -- Stock Rumors. <https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-alerts/updated-1>
- U.S. Securities and Exchange Commission. (2021, January 29). Investor Alert: Thinking About Investing in the Latest Hot Stock? Understand the Significant Risks of Short-Term Trading Based on Social Media. <https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-alerts/investor-62>
- U.S. Securities and Exchange Commission. (2022, August 29). Social Media and Investment Fraud – Investor Alert. <https://www.investor.gov/introduction-investing/general-resources/news-alerts/alerts-bulletins/investor-alerts/social-media>
- Urolagin, S. (2017, September). Text mining of tweet for sentiment classification and association with stock prices. In *2017 International Conference on Computer and Applications (ICCA)* (pp. 384-388). IEEE. <https://doi.org/10.1109/COMAPP.2017.8079788>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems* 30 (pp. 5998-6008). Curran Associates.
- Volle, A. (2023, February 10). *mass communication*. *Encyclopedia Britannica*. <https://www.britannica.com/topic/mass-communication>
- Walras, L. (1954). *Elements of pure economics or the theory of social wealth*. Allen & Unwin.
- Wang, L., Lou, T., Tang, J., & Hopcroft, J. E. (2011, December). Detecting community kernels in large social networks. In *2011 IEEE 11th International Conference on Data Mining* (pp. 784-793). IEEE. <https://doi.org/10.1109/ICDM.2011.48>
- Wang, S., Tang, J., Liu, H. (2016). Feature Selection. In: Sammut, C., Webb, G. (Eds) *Encyclopedia of machine learning and data mining*. Springer. https://doi.org/10.1007/978-1-4899-7502-7_101-1
- Wärneryd, K.E. (2008). The Economic Psychology of the Stock Market. In A. Lewis (ed.), *The Cambridge Handbook of Psychology and Economic Behaviour* (pp. 39-63). Cambridge University Press.

- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological review*, 20(2), 158-177. <https://psycnet.apa.org/doi/10.1037/h0074428>
- Wei, W., Mao, Y., & Wang, B. (2016). Twitter volume spikes and stock options pricing. *Computer Communications*, 73, 271-281. <https://doi.org/10.1016/j.comcom.2015.06.018>
- Werner F. M. De Bondt, & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, 40(3), 793–805. <https://doi.org/10.2307/2327804>
- White, D. R., & Reitz, K. P. (1983). Graph and semigroup homomorphisms on networks of relations. *Social Networks*, 5(2), 193-234. [https://doi.org/10.1016/0378-8733\(83\)90025-4](https://doi.org/10.1016/0378-8733(83)90025-4)
- Wolfsfeld, G., Segev, E., & Sheafer, T. (2013). Social media and the Arab Spring: Politics comes first. *The International Journal of Press/Politics*, 18(2), 115-137.
- World Bank. (2014). *World development report 2015: Mind, society, and behavior*. The World Bank. <https://doi.org/10.1596/978-1-4648-0342-0>
- World Economic Forum. (2019, April 17). *How much data is generated each day?* (J. Desjardins, Ed.). <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f>
- World Economic Forum. (2019, July 24). *Why Big Data Keeps Getting Bigger* (J. Desjardins, Ed.). <https://www.weforum.org/agenda/2019/07/why-big-data-keeps-getting-bigger>
- World Health Organization. (2018, May 24). The top 10 causes of death. <https://www.who.int/news-room/factsheets/detail/the-top-10-causes-of-death>
- Wortman, C. B. (1975). Some determinants of perceived control. *Journal of Personality and Social Psychology*, 31(2), 282.
- Wu, C., Buyya, R., & Ramamohanarao, K. (2016). Big data analytics= machine learning+ cloud computing. In R. Buyya, R. N. Calheiros, & A. V. Dastjerdi (Eds.), *Big Data: Principles and Paradigms* (pp. 3-38). Morgan Kaufmann. <https://doi.org/10.48550/arXiv.1601.03115>
- Wu, D. (2019). Does social media get your attention?. *Journal of Behavioral Finance*, 20(2), 213-226. <https://doi.org/10.1080/15427560.2018.1505729>
- Wuthrich, B., Cho, V., Leung, S., Permuntilleke, D., Sankaran, K., & Zhang, J. (1998, October). Daily stock market forecast from textual web data. In *SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 98CH36218)* (Vol. 3, pp. 2720-2725). IEEE. <https://doi.org/10.1109/ICSMC.1998.725072>

- XGBoost. (n.d.). In *nvidia.com*. <https://www.nvidia.com/en-us/glossary/data-science/xgboost/> adresinden 23 Ocak 2023 tarihinde alınmıştır.
- XGBoost. (n.d.). XGBoost Documentation. <https://xgboost.readthedocs.io/en/stable/> adresinden 18 Mayıs 2023 tarihinde alınmıştır.
- Xu, Y., & Cohen, S. B. (2018, July). Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1970-1979). <http://dx.doi.org/10.18653/v1/P18-1183>
- Xun, J., & Guo, B. (2017). Twitter as customer's eWOM: an empirical study on their impact on firm financial performance. *Internet Research*, 27(5), 1014-1038. <https://doi.org/10.1108/IntR-07-2016-0223>
- Yang, S. Y., Mo, S. Y. K., & Liu, A. (2015). Twitter financial community sentiment and its predictive relationship to stock market movement. *Quantitative Finance*, 15(10), 1637-1656. <https://doi.org/10.1080/14697688.2015.1071078>
- Yay, T. (2004). Avusturya İktisat Okulu'nun Tarihsel Gelişimi ve Metodolojisi. *Piyasa*, 11, 1-29.
- Yılmaz, F. (2004). Avusturya İktisadı ve Sübjektivizm. *Piyasa*, 3(11), 101-119.
- Yılmaz, F. (2009). *Rasyonelite: İktisat Özelinde Bir Tartışma*. Paradigma Yayıncılık.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of personality and social psychology*, 9(2p2), 1.
- Zarowin, P. (1990). Size, Seasonality, and Stock Market Overreaction. *The Journal of Financial and Quantitative Analysis*, 25(1), 113-125. <https://doi.org/10.2307/2330891>
- Zarowin, P. (1990). Size, Seasonality, and Stock Market Overreaction. *Journal of Financial and Quantitative Analysis*, 25(1), 113-125. doi:10.2307/2330891
- Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of behavioral decision making*, 12(2), 93-106.
- Zhang, A., Zachary C. L., Li, M., & Smola, A. J. (2021). Dive into Deep Learning. https://d2l.ai/chapter_natural-language-processing-applications/index.html
- Zhang, X., Fuehres, H., & Gloor, P. A. (2011). Predicting stock market indicators through twitter "I hope it is not as bad as I fear". *Procedia-Social and Behavioral Sciences*, 26, 55-62. <https://doi.org/10.1016/j.sbspro.2011.10.562>

- Zhang, X., Fuehres, H., & Gloor, P. A. (2012). Predicting asset value through twitter buzz. In In Altmann J., Baumöl U., & Krämer B. (Eds.), *Advances in collective intelligence 2011* (Vol. 13, pp. 23-34). Springer. https://doi.org/10.1007/978-3-642-25321-8_3
- Zhang, X., Shi, J., Wang, D., & Fang, B. (2018). Exploiting investors social network for stock prediction in China's market. *Journal of computational science*, 28, 294-303. <https://doi.org/10.1016/j.jocs.2017.10.013>
- Zhao, Q., Erdogdu, M. A., He, H. Y., Rajaraman, A., & Leskovec, J. (2015, August). Seismic: A self-exciting point process model for predicting tweet popularity. In *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1513-1522). <https://doi.org/10.1145/2783258.2783401>
- Zhao, S., Tong, Y., Meng, X., Yang, X., & Tan, S. (2016, September). Correlating tweets with the daily return reversal of the US stock market. In *2016 IEEE International Conference on Knowledge Engineering and Applications (ICKEA)* (pp. 83-87). IEEE. <https://doi.org/10.1109/ICKEA.2016.7802997>
- Zheng, A., & Casari, A. (2018). Feature engineering for machine learning: Principles and techniques for data scientists. O'Reilly Media.

ÖZGEÇMİŞ

Adı-Soyadı	Yunus Emre		AKDOĞAN
Doğum Yeri ve Yılı			
Bildiği Yabancı Diller	İngilizce		
Eğitim Durumu	Başlama - Bitirme		Kurum Adı
Lise	1995	1998	Sarayönü Lisesi
Lisans	2000	2004	Selçuk Üniversitesi İİBF İşletme Bölümü
Yüksek Lisans	2006	2009	Selçuk Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı
Doktora	2014	2023	Bursa Uludağ Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim Dalı
Çalıştığı Kurum (lar)	Başlama - Ayrılma		Çalışılan Kurumun Adı
1.	2008		Yozgat Bozok Üniversitesi
Üye Olduğu Bilimsel ve Meslekî Kuruluşlar			
Katıldığı Proje ve Toplantılar			
Yayınlar:	https://akademik.yok.gov.tr/AkademikArama/ adresinden erişilebilir.		
Diğer:			
İletişim (e-posta):			
		Tarih: İmza: Adı-Soyadı:	