

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ MÜHENDİSLİĞİ
ANABİLİM DALI

YÜKSEK LİSANS TEZİ

ALTIN GERİ DÖNÜŞ TAHMİNLEMESİ: ÖNEMLİ OLAYLAR
TEMELLİ DEĞERLENDİRME

AHMET YAVUZ

KOCAELİ 2023

KOCAELİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLİŞİM SİSTEMLERİ MÜHENDİSLİĞİ
ANABİLİM DALI

YÜKSEK LİSANS TEZİ

ALTIN GERİ DÖNÜŞ TAHMİNLEMESİ: ÖNEMLİ OLAYLAR
TEMELLİ DEĞERLENDİRME

AHMET YAVUZ

Doç. Dr. Süleyman EKEN
Danışman, Kocaeli Üniversitesi

.....

Dr. Öğr. Üyesi Kaplan KAPLAN
Jüri Üyesi, Kocaeli Üniversitesi

.....

Dr. Öğr. Üyesi Ali HAMİTOĞLU
Jüri Üyesi, İstinye Üniversitesi

.....

Tezin Savunulduğu Tarih: 20.10.2023

ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ

Kocaeli Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez/proje çalışmada,

- Bu tezin/projenin bana ait, özgün bir çalışma olduğunu,
- Çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı,
- Bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi,
- Bu çalışmanın Kocaeli Üniversitesi'nin abone olduğu intihal yazılım programı kullanılarak Fen Bilimleri Enstitüsü'nün belirlemiş olduğu ölçütlere uygun olduğunu,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Tezin/Projenin herhangi bir bölümünü bu üniversite veya başka bir üniversitede başka bir tez/proje çalışması olarak sunmadığımı,

beyan ederim.

Bu tez/proje çalışmasının herhangi bir aşaması hiçbir kurum/kuruluş tarafından maddi/alt yapı desteği ile desteklenmemiştir.

Bu tez/proje çalışması kapsamında üretilen veri ve bilgiler tarafından no'lu proje kapsamında maddi/alt yapı desteği alınarak gerçekleştirilmiştir.

Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

(İmza)

Ahmet YAVUZ

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI

Fen Bilimleri Enstitüsü tarafından onaylanan lisansüstü tezimin/projemin tamamını veya herhangi bir kısmını, basılı ve elektronik formatta arşivleme ve aşağıda belirtilen koşullarla kullanıma açma izninin Kocaeli Üniversitesi'ne verdiğimi beyan ederim. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin/projemin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanımı bana ait olacaktır.

Tezim/projenin kendi özgün çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin/projenin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim kurulu tarafından yayınlanan "**Lisanüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**" kapsamında tezim aşağıda belirtilen koşullar haricinde YÖK Ulusal Tez Merkezi/ Kocaeli Üniversitesi Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü yönetim kurulu kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 2 yıl ertelenmiştir.
- Enstitü yönetim kurulu gerekçeli kararı ile tezimin/projemin erişime açılması mezuniyet tarihinden itibaren 6 ay ertelenmiştir.
- Tezim/projem ile ilgili gizlilik kararı verilmemiştir.

(İmza)

Ahmet YAVUZ

ÖNSÖZ VE TEŞEKKÜR

Yüksek lisans eğitimim boyunca bilgi ve tecrübesini benden esirgemeyen akademik danışmanım Doç. Dr. Süleyman EKEN hocama teşekkür ederim.

Ekim – 2023

Ahmet YAVUZ



İÇİNDEKİLER

ETİK BEYAN VE ARAŞTIRMA FONU DESTEĞİ.....	i
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI	ii
ÖNSÖZ VE TEŞEKKÜR.....	iii
İÇİNDEKİLER.....	iv
ŞEKİLLER DİZİNİ	vi
TABLolar DİZİNİ.....	vii
SİMGELER VE KISALTMALAR DİZİNİ	viii
1. GİRİŞ.....	1
2. GENEL BİLGİLER	5
2.1. Zaman Serileri	5
2.1.1. Zaman Serisi Bileşenleri.....	5
2.1.1.1. Mevsimsel Bileşen	6
2.1.1.2. Trend	7
2.1.1.3. Döngüsel Bileşen	7
2.1.1.4. Rasgele (Düzensiz) Bileşen.....	8
2.2. Piyasayı Etkileyen Kritik Dönemler	8
2.2.1. Gezi Parkı Eylemleri	9
2.2.2. 15 Temmuz Darbe Girişimi.....	9
2.2.3. Papaz Andrew Brunson Olayı	10
2.2.4. Covid-19 Pandemisi	10
2.2.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi	11
3. İLGİLİ ÇALIŞMALAR.....	12
3.1. Zaman Serisi Tahmini.....	12
3.2. Altın Geri Dönüş Tahmini	13
3.3. Değer Tahminlemede Sosyal Medya Kullanımı.....	14
4. FİNANSAL VERİLER ÜZERİNDE ALTIN GERİ DÖNÜŞ TAHMİNİ	17
4.1. Veri Seti Hazırlama	17
4.2. Çalışma Ortamı Hazırlıkları.....	17
4.3. Keşifsel Veri Analizi	17
4.4. Zaman Serisi Tahmini Performans Sonuçları ve Tartışma	18
4.4.1. Süreklilik Modeli ile Tahmin	18
4.4.2. ARIMA Modeli ile Tahmin	18
4.4.3. Hareketli Ortalama Modeli ile Tahmin	18
4.4.4. Tek Üstel Düzeltme Modeli ile Tahmin	21
4.4.5. Çift Üstel Düzeltme Modeli ile Tahmin	21
4.4.6. Zaman Serisi Tahmini Modelleri Karşılaştırma	22
4.5. Danışmanlı Öğrenme Performans Sonuçları ve Tartışma	22
4.5.1. Özellik Mühendisliği İşlemleri.....	22
4.5.2. Doğrusal Regresyon ile Tahmin	23
4.5.3. Rassal Orman Regresyonu ile Tahmin	23
4.5.4. Karar Ağacı Regresyonu ile Tahmin	24
4.5.5. Danışmanlı Öğrenme Modelleri Karşılaştırma.....	24
4.6. Kritik Olaylar ile İlgili Değerlendirme	25

5. SOSYAL MEDYA VERİSİ ÜZERİNDE ALTIN GERİ DÖNÜŞ KONU MODELLEMESİ	28
5.1. Veri Toplama	28
5.2. Veri Ön İşleme.....	29
5.3. Konu Modelleme	29
5.4. Duygu Analizi.....	30
5.4.1. Destek Vektör Makineleri	30
5.4.2. Çok Katmanlı Algılayıcı.....	31
5.4.3. Evrişimli Sinir Ağları	32
5.4.4. Uzun Kısa Süreli Bellek	32
5.4.5. Transformatörlerden Çift Yönlü Enkoder Gösterimleri	33
5.4.6. M-BERT	33
5.4.7. RoBERTa	34
5.4.8. Distilled BERT	34
5.4.9. OpenAI GPT-2	34
5.5. Konu Bazlı Kritik Olaylar ile İlgili Değerlendirme.....	34
6. SONUÇLAR VE ÖNERİLER.....	41
6.1. Çalışmanın Kısıtları	41
6.2. Gelecek Çalışmalar	41
KAYNAKLAR.....	43
KİŞİSEL YAYIN VE ESERLER.....	45
ÖZGEÇMİŞ.....	46

ŞEKİLLER DİZİNİ

Şekil 2.1. Zaman Serisi Çizelgesi.....	6
Şekil 2.2. Mevsimsel Bileşen Grafik.....	6
Şekil 2.3. Olası trent grafikleri.	7
Şekil 2.4. Döngüsel Dalgalanmaların Gösterimi.....	8
Şekil 2.5. Döngüsel Bileşen Grafik Gösterimi.....	8
Şekil 4.1. Gezi Parkı Eylemleri Dönemi Altın Fiyatı (₺).....	25
Şekil 4.2. 15 Temmuz Darbe Girişimi Dönemi Altın Fiyatı (₺).....	26
Şekil 4.3. Papaz Andrew Brunson Olayı Dönemi Altın Fiyatı.....	26
Şekil 4.4. Covid-19 Pandemisi Dönemi Altın Fiyatı.....	27
Şekil 4.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi.....	27
Şekil 5.1. Gezi Parkı Olayları Konu Modelleme.....	35
Şekil 5.2. 15 Temmuz Darbe Girişimi Konu Modelleme.....	36
Şekil 5.3. Papaz Andrew Brunson Davası Konu Modelleme.....	37
Şekil 5.4. Covid-19 Pandemisi Konu Modelleme.....	38
Şekil 5.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi Konu Modelleme.....	39

TABLULAR DİZİNİ

Tablo 2.1. Zaman Serisi Örneği	5
Tablo 2.2. Piyasayı Etkileyen Kritik Dönemler	8
Tablo 4.1. Zaman Serisi Tahmini Modelleri Karşılaştırma.....	22
Tablo 4.2. Danışmanlı Öğrenme Modelleri Karşılaştırma	24
Tablo 4.3. Piyasayı Etkileyen Kritik Dönemler	25
Tablo 5.1. Kritik Dönemlere ait Tweet Sayıları ve Veri Boyutu	28
Tablo 5.2. Gezi Parkı Olayları Duygu Analizi Sonuçları.....	35
Tablo 5.3. 15 Temmuz Darbe Girişimi Olayları Duygu Analizi.....	37
Tablo 5.4. Papaz Andrew Brunson Davası Duygu Analizi	38
Tablo 5.5. Covid-19 Pandemisi Duygu Analizi Sonuçları	39
Tablo 5.6. Türkiye Hükümeti Doğrudan Döviz Müdahalesi.....	40



SİMGELER VE KISALTMALAR DİZİNİ

Kısaltmalar

BERT	: Bidirectional Encoder Representations from Transformers (Transformatörlerden Çift Yönlü Enkoder Gösterimleri)
CNN	: Convolutional Neural Network (Evrışimli sinir ağları)
COVID-19	: Coronavirus disease 2019 (Koronavirüs hastalığı 2019)
CP	: COVID passports (Covid Aşı Kartı)
DHPs	: Digital Health Passes (Dijital Sağlık Kartları)
EU	: European Union (Avrupa Birliği)
HIT	: Health Information Technology (Sağlık Bilgi Teknolojisi)
IBM	: International Business Machines (Uluslararası İş Makineleri)
ID	: Identification (Tanılama)
IIS	: Immunization Information System (Bağışıklama Bilgi Sistemi)
IP	: Immunity passports (Bağışıklık Kartı)
LMIC	: Low and Middle-income Countries (Düşük ve Orta Gelirli Ülkeler)
LSTM	: Long short-term memory (Uzun kısa süreli hafıza)
MES	: Multidimensional Ethical Scale (Çok Boyutlu Etik Ölçek)
MLP	: Multi-layer Perceptron (Çok Katmanlı Algılayıcı)
MÖ	: Milattan Önce
NLP	: Natural Language Processing (Doğal Dil İşleme)
NMF	: Nonnegative Matrix Factorization (Negatif Olmayan Matris Çarpanlara Ayrıştırma)
RoBERTa	: A Robustly Optimized Bert (Sağlam Optimize Edilmiş Bert)
S&P	: Standard & Poor (Standart ve Zayıf)
STM	: Structural Topic Modeling (Yapısal Konu Modelleme)
SVM	: Support Vector Machine (Destek Vektör Makinesi)
TCMB	: Türkiye Cumhuriyeti Merkez Bankası

ALTIN GERİ DÖNÜŞ TAHMİNLEMESİ: ÖNEMLİ OLAYLAR TEMELLİ DEĞERLENDİRME

ÖZET

Ekonomik krizler, enflasyon, jeopolitik gerilimler, faiz oranları gibi önemli olaylar altının fiyatını ve getirisini önemli ölçüde etkileyebilmektedir. Bu çalışmada Türkiye'de meydana gelmiş beş önemli olayda altın getirisi tahminine odaklanılmıştır. Biri metin tabanlı, diğeri finansal veri olmak üzere iki veri seti üzerinde çalışma yapılmıştır. Finansal veriler üzerinde birçok algoritma test edilmiş ve çoğu metrikte Extra Trees Regressor'un en iyi sonuçları verdiği tespit edilmiştir. Metin tabanlı veri üzerinde duygu analizi ve konu modellemesi uygulanmıştır. Veriler üzerinde iki farklı modda (sayısal ve metin) çalışmak farklı bakış açıları sunmuştur. Altın getirilerini tahmin etmek için yalnızca duyarlılık analizinin kullanılması tavsiye edilmez. Altın getirilerine ilişkin daha kesin ve güvenilir bir tahmin üretmek için faiz oranları, enflasyon, jeopolitik kaygılar, arz ve talep gibi ek temel ve teknik unsurların da dikkate alınması gereklidir.

Anahtar Kelimeler: Altın Geri Dönüş Tahmini, Duygu Analizi, Makine Öğrenmesi, Metin Analizi.

GOLD RETURNS PREDICTION: ASSESMENT BASED ON MAJOR EVENTS

ABSTRACT

Major events such as economic crises, inflation, geopolitical tensions, and interest rates can have a significant impact on the price and returns of gold. In this work, we focus on gold return prediction in five major events that occurred in Turkey. We work on two data, one of which is text-based and the other is financial data. In the financial part, many algorithms are tested and it is found that Extra Trees Regressor gives the best results in most metrics. In text-based part, we implement sentiment analysis and topic modelling. Working on data with two different modes (numeric and text) offers different perspectives. The use of sentiment analysis alone to forecast gold returns is not advised. To produce a more precise and trustworthy estimate of gold returns, additional fundamental and technical elements including interest rates, inflation, geopolitical concerns, and supply and demand should also be taken into account.

Keywords: Gold Return Prediction, Sentiment Analysis, Machine Learning, Text Analysis.

1. GİRİŞ

Altının geçmişi MÖ 5000’li yıllara kadar uzanır. Günümüze kadar değerini korumuş olmasında önemli etkenler vardır. Altının ziynet ve süslenme aracı olarak kullanılmasının yanında bir değişim aracı olarak da kullanılmaktadır. Bunların yanı sıra doğada az bulunduğu için üretiminin sınırlı düzeyde olması, kolay işlenebilir olması, ileri seviyede hem ısı hem de elektriği iletme özelliğinin bulunması, birçok kimyasala karşı dayanıklı olması endüstriyel alanda da aranan bir maden olmuştur.

Dünyada altın için kullanılan ağırlık ölçü birimi Ons’tur. Standart olarak 1 Ons (31,1038079) Gram’dır. Altının saflık derecesini ölçen birim ise ayar (kırat) ve milyemdir. Altının 24 ayar olması 1000 milyem saflığı ifade etmektedir. Ticarete kullanılan külçe altının saflığı 995 milyem ve üstüdür (Baş, 2010).

Altın bireysel bir yatırım aracı olduğu gibi, merkez bankalarının rezervlerinde de bulundurulur. Dünyanın önemli finans piyasaları tarafından, altının arz ve talep durumu da dikkate alınarak, altın fiyatları günlük olarak belirlenmektedir. Altın her çağda değerini koruyabilen bir yatırım aracı olması itibari ile güvenli liman olarak ifade edilmektedir. Altının arz ve talep durumunu dolayısı ile de altın fiyatlarını etkileyen faktörler hem araştırmacılar hem de yatırımcılar için önemli bir araştırma konusu haline gelmiştir. Alternatif yatırım araçlarının, enflasyonun, uluslararası önem düzeyi yüksek para birimlerinin (başta ABD Doları ve Euro), faiz oranlarının, petrol fiyatlarının altın fiyatlarına etkisi üzerine çok sayıda araştırmalar yapılmıştır.

Wang ve Chueh 2013 yılında, 2 Ocak 1989 - 20 Aralık 2007 dönemindeki günlük verileri kullanarak, altın fiyatının, petrol fiyatının, ABD Dolar’ının ve faiz oranlarının birbiri üzerinde nasıl bir etkisi olduğunu araştırmak için bir çalışma yapmışlardır. Bu çalışmalarında altın fiyatının ve petrol fiyatının birbiri üzerinde pozitif etki oluşturduğunu, ABD Doları ve faiz oranlarının ise altın fiyatları üzerinde negatif etki oluşturduğunu gözlemlemişlerdir.

Ciner, Gurdgiev ve Lucey 2013 yılında, ABD ve İngiltere’ye Ocak 1990 – Haziran 2010 dönemine ait günlük verileri kullanarak altın fiyatı, petrol fiyatı, döviz kuru, tahvil ve hisse senedi arasında nasıl bir etkileşim olduğunu araştırmışlardır. Bu araştırmada altının

döviz kuruna karşı değer kaybetmediğini ve güvenli bir yatırım aracı olduğunu tespit etmişlerdir.

Altının fiyatı üzerinde etkisi olabilecek tüm faktörlerin, geçmiş verilere dayalı olarak gelecekteki altın fiyatlarını tahmin etme tekniği “altın getiri tahmini” olarak ifade edilmektedir. Gelecekteki altın fiyatlarını tahmin etmek için teknik analiz, temel analiz ve ekonometrik modeller gibi bazı teknikler kullanılmaktadır.

Altın getiri tahmini özellikle yatırımcılar için önemli bir konudur. Bunun nedeni ise alınacak yatırım kararlarını etkilemesi, risk yönetimi ve ekonominin geleceği hakkında fikir vermesi gösterilebilir. Bir yatırımcı, fiyatın gelecekte yükseleceğini düşünüyorsa altın almaya, fiyatın düşeceğini düşünüyorsa elindeki altınları satmaya karar verebilir. Yatırımcılar açısından güvenli liman olarak kabul edilen altın, yatırım portföylerindeki riski kontrol etmede yardımcı olur. Yatırımcılar altın sayesinde genel risk durumlarını yönetmektedirler. Bazı yatırımcılar altını, enflasyona veya diğer ekonomik risklere karşı riskten korunma planının bir parçası olarak kullanır. Altın fiyatlarındaki büyük hareketlilikler ekonomide de büyük hareketliliğin yaşanabileceği anlamına gelmektedir. Altın fiyatlarındaki tahminler, genel ekonomi durumu hakkında ekonomistlere fikir vermektedir. Altın getiri tahmini genel olarak yatırımcıların daha doğru kararlar vermelerine yardımcı olabileceği gibi, ekonomistlerin ve kara vericilerin genel ekonomi durumunun gelecekteki eğilimleri hakkında fikir sahibi olmalarında önemli bir araçtır. Altın getiri tahminlerinin gelecekteki öngörülemeyen faktörlerden dolayı kesinlik içermediğinin altını çizmek gerekmektedir.

Bu çalışmada, 2013-2021 yılları arasında, Türkiye’de meydana gelen beş önemli olayın yaşandığı dönemlere ait, biri metin tabanlı, diğeri finansal veriler olmak üzere iki veri seti üzerinde çalışma yapılmıştır.

Metin tabanlı verilerin işlendiği bölümde, aşağıdaki önemli olayların yaşandığı dönemlere ait veriler toplanmıştır.

- Gezi Parkı Eylemleri
- 15 Temmuz Darbe Girişimi
- Papaz Andrew Brunson Olayı

- Covid-19 Pandemisi
- Türkiye Hükümeti Doğrudan Döviz Müdahalesi

Yukarıda belirtilen olaylar hem Türkiye gündemini hem de dünya gündemini meşgul etmiş konulardır. Bu dönemlere ait veriler sosyal medya platformu Twitter üzerinden toplanmıştır. BeautifulSoup4 ve Twitter API5 kullanarak Twitter üzerinden toplanan veriler ile metin tabanlı veri seti oluşturulmuştur. Twitter üzerinden veri çekmek için hashtag özelliğinden faydalanılmıştır. Bu veri seti oluşturulurken kullanılan hashtag'ler (etiketler): gold, ons, gold/sell, selling, coins, bitcoin, exchange, buy, buying, brent, euro, dollar, silver, pay, paying, yellow, forrest, expensive, cheap, coin, metal, collect, CreditCards, goldjewelry, goldexchange, electronicgoldreceipts, goldinvestment, buyinggold, goldprice, Diamond, xauusd, xaueur, xautr, bestwaybuygold, bestwaysellgold, goldpriceusd, carat, goldpricehourly. Toplanan bu ham veriler özellik çıkarımı için uygun değildir. Bu verilerin doğru şekilde analiz edilebilmesi için bazı ön işlemlerden geçmesi gerekmektedir. İlk olarak sayı ya da text verisi olmayan ifadeler tweet'lerden silinmiştir. Ardından tüm tweet'ler küçük harflere dönüştürülmüş, URL'ler, emojiler, özel karakterler, diğer kullanıcılara yapılan referanslar tweet'lerden silinmiştir. Daha sonra metinler sırasıyla spaCy ile lemmatize edilir ve Gensim ile temizlenmiştir. Bu veriler üzerinde duygu analizi ve konu modellemesi uygulanmıştır.

Bu çalışma kapsamında birçok enstrümanın son on yılın kapanış fiyatı verilerine ihtiyaç duyulmuştur. Bu verilerin alınabileceği ücretli (Reuters, Bloomberg) ve ücretsiz (IEX, Quandl, Yahoo Finance, Google Finance) kaynaklar mevcuttur. Bu çalışmada "yahoofinancials" paketi tercih edilmiştir. Çalışmaya dahil edilecek enstrümanlara ait verileri çekmek için önce hangi zaman aralığındaki verileri kullanmak istediğimiz belirlenmiştir. Ocak 2013 – Aralık 2023 tarih aralığına ait veriler kullanılmıştır. Finansal verilerin kullanıldığı bölümde çok sayıda algoritma test edilmiştir.

Bu araştırmada şu sorulara cevap aranmıştır;

- Finansal veri setinde elde edilmiş olan Altın, Gümüş, Ham Petrol, S&P500, Russel 2000 Endeksi, 10 Yıllık ABD T-Note vadeli işlemleri, 2 Yıllık ABD T-Note vadeli işlemleri, Platin, Bakır, Dolar Endeksi, Volatilite Endeksi, MSCI EM ETF, Euro,

Euronext100, Nasdaq enstrümanlarına ait günlük veriler ile altın getirisi için tahmin yapılabilir mi?

- Ulusal kritik gelişmeler, uluslararası kritik gelişmeler, ekonomik krizler altın fiyatları üzerinde etki oluşturur mu? Bu gelişmeler ile ilgili toplanacak veriler kullanılarak altın için geri dönüş tahmini yapılabilir mi?
- Özel olarak belirlenmiş 5 önemli olayın (Bölüm 2.2) yaşandığı dönemde toplanan finansal ve sosyal medya verileri üzerinde zaman serisi analizleri ve duygu analizi yöntemleri ile altın getiri tahmini yapılabilir mi?

Birinci bölümde altının tarihçesinden kısaca bahsedildikten sonra altının tarih boyunca neden bu kadar kıymetli olduğu ifade edilmeye çalışılmıştır. Tarih boyunca değerli olan altın günümüzde de yatırımcılar için önemli bir yatırım aracıdır. Bu nedenle geleceği konusunda her zaman merak olan altın ile ilgili, gelecek tahminleri yapılabilmesi noktasında yaptığımız çalışmada, kullandığımız veri setlerinden ve bu veri setlerinin toplandığı dönemlerden, çalışmanın amacından genel olarak bahsedilmiştir. Bundan sonraki bölümlerde veri setleri ve veri setlerinin toplandığı dönemler, bu veri setleri üzerinde uygulanan teknik yöntemler detaylı olarak açıklanmıştır. Son olarak bu yöntemlerden elde edilen sonuçlara yer verilmiştir.

2. GENEL BİLGİLER

Bu bölümde zaman serileri ve metin tabanlı veri setinin elde edildiği dönemler hakkında genel bir bilgilendirme yapılmıştır.

2.1. Zaman Serileri

Zaman serileri, zamana göre, kronolojik olarak sıralanan veri setleridir. Veri setleri saatlik, günlük, haftalık, aylık, yıllık ya da daha uzun dönemleri kapsayan, periyodik bir döngü ile sıralanırlar. Sayısal olarak ifade edilebilen olay ve işlemlerin, belirli bir zaman dilimi ile ilişkilendirildiği veri setleri de zaman serileri olarak ifade edilebilir. Bu anlamda Zaman serilerinin belirli bir tarih aralığı ile sınırlandırılması, zaman serilerinin tarih üzerinde bir örneklem olduğunu gösterir. Bir kişinin günlük atmış olduğu adım sayısının günler ile işaretlenmesi bir zaman serisi örneği oluşturur. Günlük adım sayısı, zaman serisi örneğine ait zaman ve gözlem bilgileri Tablo 2.1'deki gibi gösterilebilir. Tablo 2.1'deki zaman serisi örneğine ait zaman serisi çizelgesini Şekil 2.2'deki gibi gösterebiliriz.

Tablo 2.1. Zaman Serisi Örneği

Günler (Zaman)	Adım Sayısı (Gözlem)
01-12-2023	1299
02-12-2023	1842
03-12-2023	1918
04-12-2023	1909
05-12-2023	2053
....	...

2.1.1. Zaman Serisi Bileşenleri

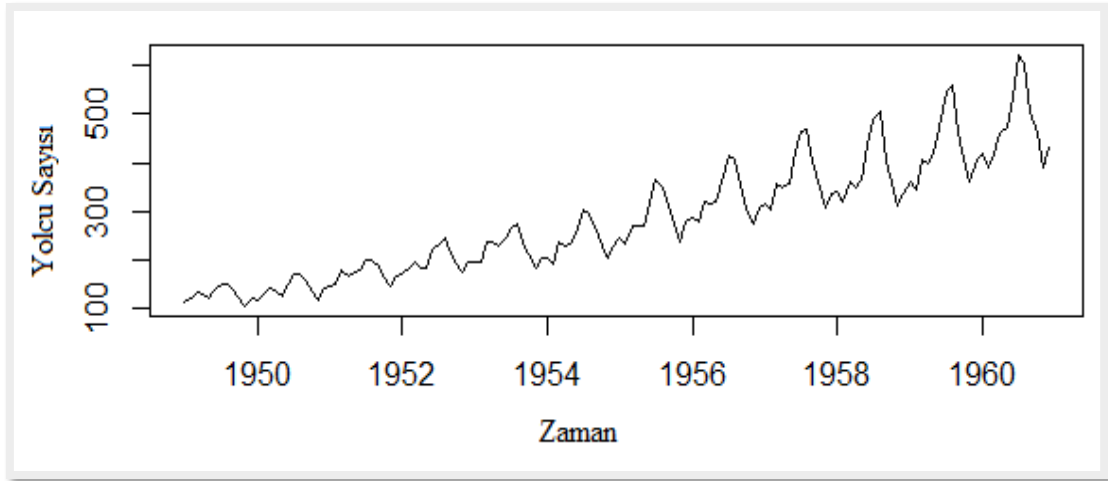
Zaman serileri, birer varsayım olarak kabul edilen 4 temel bileşenden oluşur. Analiz tekniklerinde bu bileşenlerden yararlanır. Zaman serisi bileşenleri, analizler esansında oluşan farklı durumların açıklanmasında büyük yararlar sağlamaktadır.

2.1.1.1. Mevsimsel Bileşen

Mevsimsel bileşen, sabit uzunluktaki periyotlardan oluşan ve sistemli bir döngüsellik içeren zaman serisi örneğidir. Genellikle aylık veriler ya da çeyreklik dönem verileri ile elde edilen zaman serileri için bu bileşen söz konusu olabilir. Yaz mevsiminde dondurma satışlarının artması bu bileşene örnek gösterilebilir. Her yıl yaz mevsiminde dondurma satışları artış gösterir ve diğer mevsimlerde bu satış hacmi düşerek normal satış seyrine geri döner. Şekil 2.3'teki grafik, mevsimsel bileşene ait bir grafik örneğini içermektedir.



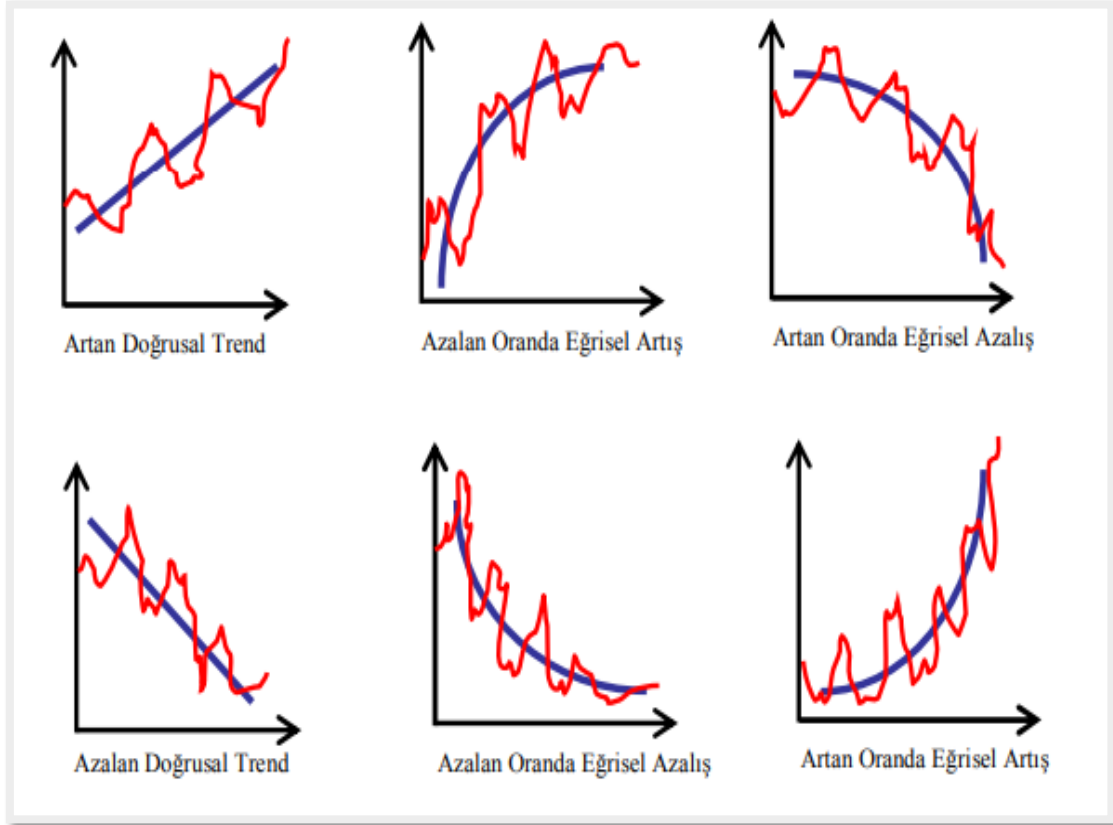
Şekil 2.1. Zaman Serisi Çizelgesi



Şekil 2.2. Mevsimsel Bileşen Grafik

2.1.1.2. Trend

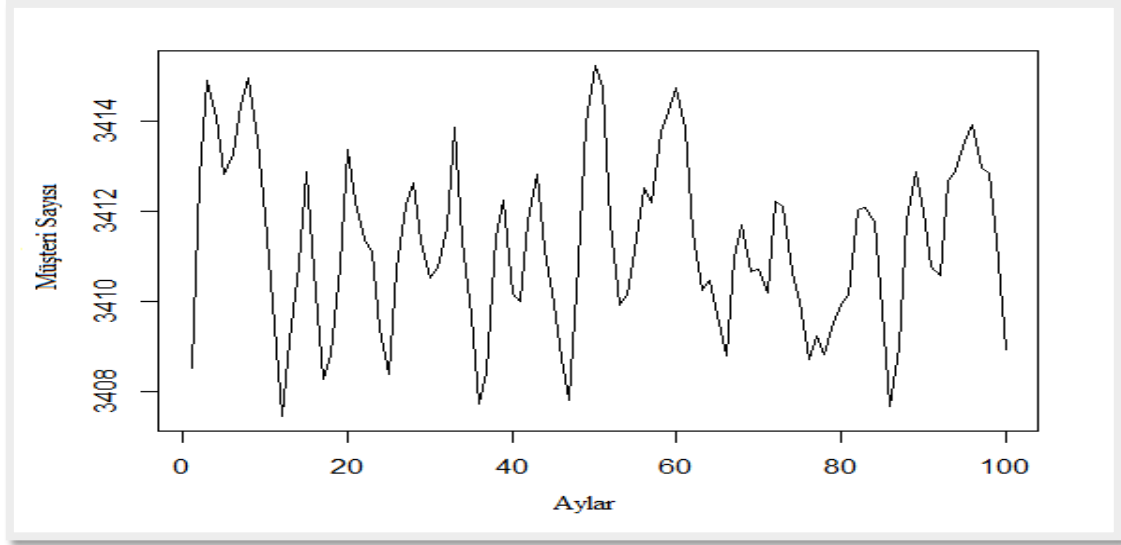
Trend, bir zaman serisinin uzun dönemli artış ya da uzun dönemli azalış eğilimi göstermesidir. Trend, her zaman aynı yönde ve aynı şiddette devam etmez, doğrusal ya da eğrisel eğilim içerisinde olabilir. Şekil 2.4'te eğrisel ve doğrusal trend eğilimlerine örnek grafikler verilmiştir.



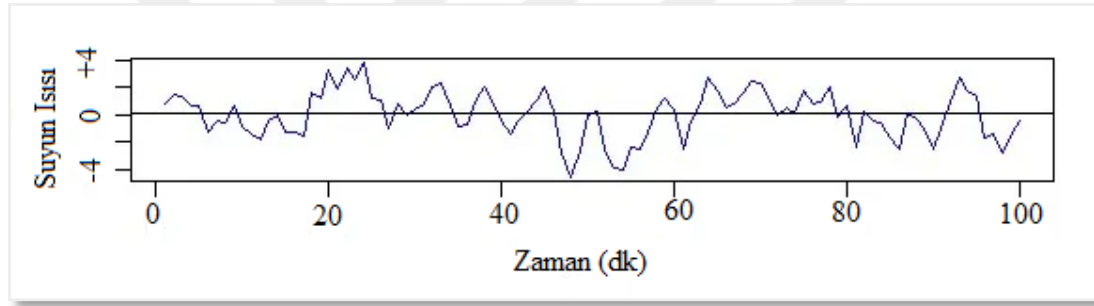
Şekil 2.3. Olası trend grafikleri (Mann, 1995).

2.1.1.3. Döngüsel Bileşen

Bir trend etrafında, bir süre artış eğilimi gösterdikten sonra azalış (ya da tam tersi, önce azalış sonra artış) eğilimi gösteren zaman serileridir. Belirli dönemlerde meydana geldiği için mevsimsel bileşen ile benzerlik gösterir. Ancak sabit olmayan periyotlar ile meydana gelmesi, bu bileşeni mevsimsel bileşenden ayırır. Şekil 2.5'te döngüsel bileşene örnek bir grafiksel gösterim verilmiştir



Şekil 2.4. Döngüsel Dalgaların Gösterimi



Şekil 2.5. Döngüsel Bileşen Grafik Gösterimi

2.1.1.4. Rasgele (Düzensiz) Bileşen

Rasgele bileşen, etkisi sürekli olmayan, varlığı önceden öngörülemeyen, belirli faktörlerden bağımsız olarak ortaya çıkan olayların oluşturduğu zaman serisi örneğidir. Savaşlar, grevler, doğal afetler örnek gösterilebilir. Şekil 2.6’da düzensiz bileşene örnek gösterim verilmiştir

2.2. Piyasayı Etkileyen Kritik Dönemler

Metin tabanlı veriler üzerinde yapılacak çalışmalar için sosyal medya platformlarından veri toplanması kararlaştırılmıştır. Bu verilerin toplanacağı dönemler ise ülke ve dünya gündemini ciddi anlamda meşgul eden olaylar arasından seçilmiştir. Veri setlerinin toplandığı olaylar ve tarih aralıkları Tablo 2.2’de verilmiştir.

Tablo 2.2. Piyasayı Etkileyen Kritik Dönemler

Kritik Olay	Tarih Aralığı
Gezi Parkı Eylemleri	28 Mayıs – 20 Ağustos, 2013
15 Temmuz Darbe Girişimi	20 Mart – 20 Ağustos, 2016
Papaz Andrew Brunson Olayı	1 Ağustos – 10 Kasım, 2018
Covid-19 Pandemisi	1 Haziran 2019 – 30 Haziran 2020
Türkiye Hükümeti Doğrudan Döviz Müdahalesi	1 Aralık – 31 Aralık, 2021

2.2.1. Gezi Parkı Eylemleri

Gezi Parkı Eylemleri, 2013 yılında İstanbul'un Taksim Meydanı'nda başlayıp sonrasında Türkiye'nin diğer şehirlerine yayılan gösteriler olarak bilinmektedir. İstanbul'un Taksim Meydanı'nda bulunan, küçük bir park olan Gezi Parkı'nı geliştirme planları kapsamında, iş makinalarının parka girmesi ile bazı aktivistler parka gidip çalışmalarını durdurmaya çalışmıştır. Bu aktivistlere karşılık polis müdahalede bulunmuş, mevcut hükümet çalışmalarının devam etmesinde ısrarcı açıklamalar yapmıştır. Bunun üzerine bu eylemler İstanbul dışındaki diğer şehirlere de sıçramıştır ve aylarca ülke gündemini meşgul etmiştir. Olaylarda 7 sivil ve 2 polis memuru hayatını kaybetmiş çok sayıda kişi yaralanmıştır (Özel ve Deniz, 2015). Kamu binaları, kamu araçları ve bazı kamu alanları zarar görmüştür (Babacan, 2014). Bu olaylar üzerine turizm maksatlı çok sayıda rezervasyonun iptal edildiği açıklanmış, İstanbul borsasında sert düşüşler görülmüştür.

2.2.2. 15 Temmuz Darbe Girişimi

Türkiye'de, 15 Temmuz 2016 yılında başarısız olan ancak kanlı sonuçlanan bir darbe girişimi yaşanmıştır. 15 Temmuz 2016'da akşam saatlerinde başlayan darbe girişimi, zamanın mevcut hükümetini devirmek isteyen, TSK'da görevli bir grup askerin öncülüğünde gerçekleşmiştir. Darbe girişimini durdurmak amacıyla halk sokaklara çıkmış, darbeciler halka gerçek mermiler ile ateş etmiştir. Hükümet, güvenlik kuvvetlerini seferber ederek, kritik olarak kabul edilen ve darbe girişiminde bulunanların kontrol altına aldığı yerlerin kontrolünü darbecilerin elinden almıştır. 16 Temmuz sabah saatlerinde mevcut hükümet darbe girişimini bastırarak ve ülke genelinde kontrolü ele almıştır. Darbe girişiminde halkın da sokaklara çıkmış olması ve darbe girişiminde

bulunanların halka ateş açmış olması nedeni ile çok sayıda kişi hayatını kaybetmiş, çok sayıda kişi yaralanmıştır. Yaklaşık 2000 kişinin yaralandığı bu darbe girişiminde 250 kişi hayatını kaybetmiştir. Bu darbe girişiminden sonra istihdam verilerinde azalma olmuş, enflasyon oranlarında artış görülmüş, büyüme yavaşlamıştır (Balci, 2018).

2.2.3. Papaz Andrew Brunson Olayı

ABD’li bir papaz olan Andrew Brunson, 2016 yılında Türkiye’de gözaltına alınmıştır. Papaz Andrew Brunson’ın, Türkiye’de hükümeti devirmek isteyen terör örgütleri ile ilişki içinde olması ve casusluk yapması, gözaltı nedeni olarak gösterilmiştir. Papaz Andrew Brunson bu suçlamaları şiddetle reddetmiştir. Bu olay üzerine Türkiye – ABD ilişkileri gerilmiştir ve iki ülke karşılık yaptırımlar uygulamıştır. Dünya kamuoyu bu olayı ilgi ile takip etmiştir. Bu süreçte Türkiye’de, özellikle döviz kurunda yüksek oranda artış görülmüştür. Papaz Andrew Brunson 2 yıl, Türkiye’de bir hapisanede kalmıştır. Ekim 2018’de ABD’ye dönmesine izin verilmiştir. Papaz Andrew Brunson’ın ABD’ye dönmesine izin verildikten sonra dolar kuru ve altın düşüş trendine girmiştir (Akgül, 2019).

2.2.4. Covid-19 Pandemisi

COVID-19 pandemisi, koronavirüs pandemisi olarak da bilinmektedir. Resmi kayıtlara göre ilk defa, 17 Kasım 2019 tarihinde Çin’in Hubei bölgesinin başkenti olan Vuhan’da bir hastada görülmüş ve kısa süre içerisinde bölgedeki birçok kişiye bulaşmıştır. Hakkında tam bir bilgi bulunmayan bu virüs SARS-CoV-2 olarak adlandırılmıştır (Benlagha, 2019). Ortaya çıktığı zamanda, hastalara uygulanan tedavilere tam olarak cevap alınamamıştır. Hava yolu ile bulaşan bu virüsün bulaşma oranı 2020 Ocak ortasında çok ciddi bir artış göstermiştir. Zamanla Avrupa, Kuzey Amerika ve Asya-Pasifik ülkeleri Covid-19 salgınına ait ilk hastalarını rapor etmeye başlamıştır. Bu gelişmeler üzerine, 11 Mart 2020 tarihinde Dünya Sağlık Örgütü Covid-19’u küresel salgın olarak ilan etmiştir. Salgının tam olarak tedavi edilemiyor olması, hızlı bir şekilde yayılıyor olması gibi nedenlerden dolayı karantinalar uygulanmış, belli dönemlerde, belirli sürelerde sokağa çıkma yasakları getirilmiştir. Virüse yakalanan kişilerin sayısı arttıkça, ülkelerin sağlık sistemleri ciddi bir sınavdan geçmiştir. Zorunlu maske kullanımı uygulamaya geçirilmiştir.

Tüm bu yaşananlar üzerine hızlı bir şekilde aşı geliştirme çalışmaları başlamış, 23 Ağustos 2021 tarihinde ilk Covid-19 aşısı için hastalara uygulanabilirlik izni alınmıştır. Hastalar üzerinde oluşturabilecekleri etkiler tartışmalı olsa da zamanla Covid-19 aşılarının sayısı artmıştır. Resmi kayıtlara (8 Nisan 2023 itibari ile açıklanan rakamlar) göre yaklaşık 700 milyon vaka görülmüş, 7 milyon civarında insan hayatını kaybetmiş, 20 milyon civarında aktif vaka olduğu açıklanmıştır (URL-2).

Bu salgına karşı ülkelerin belli dönemlerde, özellikle vaka sayılarında artış olan dönemlerde, almış olduğu çeşitli tedbirler bazı sektörlerde üretimi aksatmıştır. Bu dönemde eve tam kapanma tedbirlerinin alınması, özellikle küçük esnaf diye adlandırdığımız, düşük ciro lu esnafı yüksek oranda etkilenmiştir, çok sayıda küçük işletmenin kapatma kararı aldığı görülmüştür. Olumsuz haber ve söylemlerin yatırımcıları, güvenli yatırım aracı olarak görülen altına yönlendirmiştir.

2.2.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi

Türkiye Cumhuriyeti Merkez Bankası'nın Türk Lirası'ndaki istikrarı sağlamak amacı ile açık piyasalardan doğrudan döviz alımı şeklinde gerçekleşen piyasa müdahalesidir. TCMB tarafından 17 Aralık 2021'de gerçekleştirilen beşinci doğrudan müdahalede altın fiyatlarında çok sert bir düşüş görülmüştür.

3. İLGİLİ ÇALIŞMALAR

Bu bölümde zaman serisi tahmini, Altın Geri Dönüş Tahmini ve Değer Tahminlemede Sosyal Medya Kullanımı ile ilgili yapılmış literatür taraması hakkında bilgi verilmiştir.

3.1. Zaman Serisi Tahmini

Zaman serisi, bir takım gözlem sonuçlarının zamana göre sıralanması ile oluşan veriye denir. Zaman serisi verileri, kurum ve kuruluşlar için zamana bağlı bir takım faktörlere ait detayların izlenmesi ve anlamlandırılmaya çalışılması noktasında önemli çok önemli rol oynar.

Zaman serisi tahmini, belirli olaylar için geçmiş dönük elde edilmiş zaman serisi verilerini kullanarak, bu olayların gelecekte nasıl bir eğilim göstereceklerini tahmin etmek için kullanılan bir tekniktir. Zaman serisi tahmini, jeoloji, ekonomi, meteoroloji gibi birçok alanda kullanılan bir tekniktir. Bu tahminler yapılırken üzerinde çalışma yapılan olayların geçmişte gösterdikleri davranış ve eğilimlerin analiz edilmesi, gelecekte de benzer eğilim ve davranışların gösterileceği varsayımıyla hareket edilmektedir.

Zaman serisi tahmininin kullanıldığı bazı alanlar:

- Deprem tahmini
- İstatistik
- Hava durumu tahmini
- Kontrol mühendisliği
- Sinyal işleme
- Ekonometri
- Astronomi
- Desen tanıma
- Matematiksel finans

Zaman serisi tahmini yapılırken kullanılan bazı bilgisayar teknolojileri:

- Yapay sinir ağları
- Makine öğrenmesi
- Bulanık mantık
- Gauss süreçleri

- Gizli Markov modelleri

Zaman serileri kullanılarak bir çok alanda gelecek tahminleri yapılmaya çalışılmıştır. Nikolopoulos ve diğ. (2016) yapmış oldukları bir çalışmada ilaç tedarik zincirinde talep tahmininde bulunmaya çalışmışlardır. ARIMA, Tek Üstel Düzeltme, Holt Üstel Düzeltme, Difüzyon ve Regresyon modelleri ile yapmış oldukları analiz çalışmasında kısa dönemli tahminlerde ARIMA model ve Holt Üstel Düzeltmenin en iyi sonuçları verdiği tespit edilmiştir. A.G. Malliaris ve Mary Malliaris (2015) “Altın getirilerini ne yönlendirir? Bir karar ağacı analizi” isimli çalışmalarında Karar Ağacı yöntemini kullanarak altın geri dönüşlerinin farklı ülkelerde farklı faktörlere bağlı olduğunu tespit etmişlerdir. Mahmutoğlu ve Öztürk (2015) tarafından yapılan bir çalışmada 2015-2023 yılları arasında, Türkiye’deki brüt elektrik tüketimi tahmin edilmeye çalışılmıştır. 1970-2011 yılları arasındaki elektrik tüketimine ait veriler kullanılmış, ARIMA tahmin modeli uygulanmıştır. Benzer bir çalışmayı Hussain ve diğ. (2016) Pakistan için yapmışlardır. Pakistan’da yaşanan elektrik kesintilerinin etkilerini öngörmek üzere 2012-2020 yılları arasındaki elektrik enerji talebini tahmin edilmeye çalışılmıştır. Bu çalışma kapsamında 1980-2011 yılları arasındaki enerji tüketim verileri kullanılmıştır.

3.2. Altın Geri Dönüş Tahmini

Bu çalışma kapsamında altın fiyatlarının tahmin edilmesine yönelik bir çalışma yapılmıştır. Bununla beraber altın fiyatlarını etkileyen faktörler ile ilgili de çok sayıda araştırma mevcuttur. Bu çalışmalar yapılırken zaman serilerinden ve farklı tekniklerden faydalanılmıştır. Altının; Dolar, Euro, enflasyon verileri, faiz oranları, ekonomik krizler, siyasi krizler vb. ile ilişkisi araştırılmıştır.

Adem ve diğ. (2017) yapmış oldukları çalışmada altın getiri tahmininde bulunmak için çeşitli bir veri seti kullanmışlardır. Brent petrol fiyatı, ABD doları, BIST100 endeksi, TCMB haftalık faiz oranı, gümüş ve bakır fiyatları kullanılmıştır. %98,17’lik bir başarımla altın fiyatları için gelecek tahminlerde bulunmuş, çok katmanlı yapay sinir ağı yöntemi kullanılmıştır.

Çelik ve Başarır (2017) altın fiyat tahmini ile ilgili yaptıkları çalışmada girdi verilerini daha da zenginleştirip araştırmalar yapmışlardır. Brent Petrol, doğal gaz, 30 yıllık bono,

10 yıllık bono, 5 yıllık bono, S&P 500, Nasdaq, Dow Jones, FTSE100, DAX, CAC40, SMI, NIKKEI, HANH, SEND ve Euro/Dolar verilerini kullanarak altın, gümüş, platin ve paladyum için gelecek fiyat tahmininde bulunmuşlardır. Yapay sinir ağları modellerinden faydalanmışlardır. Altın fiyatlarında doğruluk oranı yüksek tahminlerin yapılabileceğine dair tespitlerde bulunmuşlardır.

Smith (2001, 2002) yapmış olduğu çalışmalarda altın fiyatı ile hisse senetleri arasındaki ilişkiyi incelemiştir. Yapmış olduğu her iki çalışmada, altın fiyatları ile hisse senetleri arasında kısa vadede negatif bir ilişki olduğunu, uzun vadede nötr ilişki olduğunu tespit etmiştir.

Erer (2011) çalışmasında altın verilerindeki hareketliliği incelemiştir. Külçe altın (TL/gr) haftalık satış verilerini kullanmıştır. Veri setini 05.01.2001 – 04.02.2011 tarihleri arasından seçmiştir. ARCH(1), GARCH(1,1), EGARCH(1,1), TARARCH(1,1) ve TARARCH(2,2) modellemesi yapılmıştır. En başarılı sonucu TARARCH(2,2) modelinde elde etmiştir.

3.3. Değer Tahminlemede Sosyal Medya Kullanımı

Teknolojinin gelişmesi ve teknolojiye erişilebilirliğin günden güne kolaylık kazandığı günümüzde, insanlara yeni olanaklar sunulmuştur. Bu olanakların en önemlilerinden birisi de sosyal medya olmuştur. Sosyal medyanın gelişimi, internet ortamında sunulan içerikleri tek taraflı olmanın dışına çıkarmış, çift yönlü bir etkileşim imkanı getirmiştir.

Sosyal medya insanların çevrimiçi etkileşim içerisinde oldukları, birbirleri ile yazılı, sesli, görüntülü içerik paylaştıkları, birbirlerinin paylaşımlarına karşılıklı etkileşim içerisinde oldukları, yorumlarda bulunabildikleri internet tabanlı uygulamaların genel ifadesi şeklinde tanımlanabilir. Bu karşılıklı etkileşim sosyal medya uygulamaları üzerinden sağlanmaktadır.

Günümüzde hem internet kullanımı hem de akıllı cep telefonu kullanımı oldukça yaygınlaşmıştır. Cep telefonları sayesinde internete dolayısı ile sosyal medya uygulamalarına her an erişim sağlanabilmektedir. Bu sayede sosyal medya uygulamalarına erişim oldukça kolaylık kazanmıştır. Her geçen gün sosyal medyayı aktif

kullanan kiři sayısı artmaktadır. Sosyal medya hem gnlk hayatın bir parçası hem de bir kazanç kapısı haline gelmiştir.

Sosyal medyaya erişimin kolaylaşması, kullanıcıların birbiri ile etkileşim içerisinde olmaları, bilginin çok sayıda kişiye aynı anda ve oldukça hızlı şekilde ulaşabilmesi sayesinde sosyal medya, gazete, televizyon, radyo gibi iletişim kanallarının yerini almaya başlamıştır ve popüleritesi her geçen gün artmaktadır. Günümüzde en önemli sosyal medya uygulamalarına Twitter, Facebook, İnstagram, Tiktok, Youtube örnek gösterilebilir.

Birçok karar alıcı, karar verme noktasında sosyal medyadaki algıyı dikkate almaktadır. Bu sosyal medyadaki verilerin analizi ihtiyacını ortaya çıkarmıştır.

Sosyal medya uygulamalarında paylaşılan içeriklerinin kaç kişiye ulaştığı, bu paylaşım üzerindeki etkileşim biçimi (örneğin beğeni) ve sayısı, bu paylaşım yapılan yorum sayısı ve içeriği ile ilgili yapılan analiz çalışmalarına sosyal medya analizi denilmektedir. Sosyal medya kullanımının artması ile birlikte sosyal medya analizi de popüler hale gelmiştir. Çalışma kapsamında yapılan literatr taramasında Twitter platformunda oluşan veriler üzerinde çok sayıda analiz çalışması yapıldığı görlmştr.

Elkin ve diğ. (2017), yapmış oldukları çalışmada, Amerikada'ki grip hastalıkları ile ilgili yapılan paylaşımları analiz etmişler. Bu analiz ile mevcut salgınların tespitini ve salgınının coğrafik olarak hangi tarafa yöneleceğini tahmin etmeye çalışmışlar. Bu tahminlerini yapmak için Twitter üzerinden bu salgın ile ilgili atılan tweet'leri, bu tweet'leri atan kullanıcıları ve kullanıcıların yerlerini tespit edip salgınları ve salgınların yönünü belirlemişler. Amerikan Sağlık Merkezi (CDC) tarafından paylaşılan veriler incelendiğinde, mevcut salgınlar ve salgınların yönü ile ilgili Elkin ve diğ.'nin yapmış oldukları tahminlerin başarılı olduğu, Amerikan Sağlık Merkezi (CDC) verileri ile örtüştüğü görlmştr. Hastalıklara karşı önlem alma noktasında, coğrafik olarak nerede, hangi hastalıkların olduğu ve hangi taraf yöneleceğini tahmin etmek zaman kazandıracaktır. Salgınlara karşı erken alınan önlemler sayesinde büyük felaketlerin önüne geçilebilecektir.

Wang, Can (2012) çalışmalarında, 2012 ABD seçimleri için Twitter verileri üzerinde duygu analizi yaparak anlık ve sürekli güncellenen anket sonucu ortaya çıkarmışlardır. 2012 ABD seçimlerinde açıklanan başkan adayları için atılan tweet'lerde insanların siyasi yaklaşımları ve başkan adayları hakkındaki görüşlerini duygu analizi yöntemi ile bu sonuçları elde etmişleridir.

Pant ve diğ. (2018) yapmış oldukları çalışmada bitcoin ile ilgili gelecek fiyat tahmininde bulunmuşlardır. Bitcoin ile ilgili atılan tweet'lerin duygu analizini yapmışlardır. Ardından zaman serilerini kullanarak tahminlerde bulunmuşlardır. Yaptıkları çalışmada duygu analizi yönteminde yaptıkları tahminler, zaman serilerini kullanarak yaptıkları tahminlerden, gerçeğe daha yakın sonuçlar verdiğini görmüşlerdir.

Young (2015) yapmış olduğu çalışmasında sosyal medya üzerinde oluşan büyük veriden ve mobil teknolojilerden faydalanarak HIV salgını ile ilgili incelemeler yapmıştır. Bu çalışmada HIV salgınının belirlenmesinde büyük veriyi kullanmanın öneminden bahsetmiştir. Büyük veriden faydalanarak salgınların tespit edilmesi ve bu sayede önlenmesi konularını ele almıştır. Büyük veri üzerinde hastalık modelleme konusunun önemini ifade etmiştir.

Xiang ve diğ. (2015), büyük veri analizi ile turistik konaklama mekanları için önemli bir konu olan müşteri deneyimi hakkında, çıkarımlar elde edilip edilemeyeceği konusunu ele almışlardır. Bu çalışma kapsamında, Expedia.com isimli web sitesi üzerinden konaklama rezervasyonu yaptırmış/satın alanların yorumları ve beğen sayılarına ait veriler toplanmıştır. Bu veriler üzerinde sınıflandırma modelleri uygulanmış ve metin analizi yapılmıştır. Bu çalışmada müşterilerin yaşamış oldukları deneyim ile konaklama mekanına karşı tatminleri arasında bir ilişki olduğu çıkarımı elde edilmiştir.

4. FİNANSAL VERİLER ÜZERİNDE ALTIN GERİ DÖNÜŞ TAHMİNİ

4.1. Veri Seti Hazırlama

Veri seti satır ve sütunlardan oluşur. Gözlem birimleri satırlara, değişken birimler ise sütunlara yerleştirilir. Satır ve sütunların kesiştiği yere hücre denir. Gözlem sonucunda elde edilen değerler hücrelere yerleştirilir. Veri seti ses, video, resim, sayısal veri ya da metinsel verilerden oluşabilir.

Bu çalışma kapsamında Altın, Gümüş, Ham Petrol, S&P500, Russel 2000 Endeksi, 10 Yıllık ABD T-Note vadeli işlemleri, 2 Yıllık ABD T-Note vadeli işlemleri, Platin, Bakır, Dolar Endeksi, Volatilite Endeksi, MSCI EM ETF, Euro, Euronext100, Nasdaq enstrümanlarına ait günlük veriler toplanmıştır. Veriler Ocak 2013 – Aralık 2022 tarih aralığı dikkate alınarak toplanmıştır.

4.2. Çalışma Ortamı Hazırlıkları

Bu çalışma kapsamında toplanan veriler üzerinde uygulanan ön işleme, zaman serisi ve duygu analizi yöntemleri için gerekli kodlamalar Python programlama dili ile yapılmıştır. Kodların çalıştırılabilmesi için Python 3.x veya daha üst versiyonu ve NumPy modülü kullanılmıştır. Çalışma kapsamındaki işlem hacminin yüksek olması nedeni ile güçlü bir donanıma sahip makineye ihtiyaç duyulmuştur. Bu gereksinimin karşılanması için Google Colab Pro üyeliği yapılarak Google tarafından sağlanan yüksek RAM ve GPU ile çalışmalar tamamlanmıştır.

4.3. Keşifsel Veri Analizi

Veri setine ait temel özellikleri açıklamak ve veri setinde gizli kalan bilgileri açığa çıkarmak için veriler üzerindeki keşif sürecine keşifsel veri analizi denir. Akademik dünyada Exploratory Data Analysis olarak bilinir ve EDA şeklinde kısaltma adı da akademik anlamda kullanılır. EDA sürecinin işletilmesinde temel amaçlardan birisi veri setleri üzerinde yapılacak analizlerde güvenilirliği arttırmaktır. Bu anlamda öncü bir süreç olarak veri setindeki anormal durumları ortaya çıkarmakta ve asıl analiz süreçlerine rehberlik etmektedir. Aslında elde edilmiş olan veri setini tanımada ve analiz sürecine yön vermede çok önemli bir ön aşamadır. EDA sürecinde regresyon analizi, korelasyon analizi gibi istatistiksel yöntemlerin yanı sıra kümeleme gibi makine öğrenmesi yöntemleri, görselleştirme teknikleri gibi çok sayıda farklı teknik kullanılmaktadır.

4.4. Zaman Serisi Tahmini Performans Sonuçları ve Tartışma

Bu bölümde zaman serisi tahmin modelleri ile bilgi verilmektedir. Finansal veri seti üzerinde bu modeller işletilip sonuçları açıklanmaktadır

4.4.1. Süreklilik Modeli ile Tahmin

Süreklilik modeli (Persistence Modeli), zaman serisinin gelecek değerleri ile ilgili tahmin yaparken, şimdiki zaman ile gelecekteki tahmin zamanı arasında geçen sürede hiçbir şeyin değişmeyeceğini varsayarak tahminde bulunur.

4.4.2. ARIMA Modeli ile Tahmin

Durağan olmayan zaman serileri ile çalışma yapabilmek için durağanlığın sağlanması gereklidir. Durağanlığın sağlanması demek, ortalama ve varyans değerlerinin sabit kalması demektir.

ARIMA modelleri, fark alma işlemi uygulanarak durağan hale getirilen, ancak aslında durağan olmayan serilere uygulanır. Bu modeller “durağan olmayan doğrusal stokastik modeller” olarak da isimlendirilir. Bu modellerin gösterimi ARIMA (p, d, q) şeklindedir. Bu gösterimde yer alan parametreler ;

- p; Otoregresif (AR) Modelin derecesini
- d; fark alma derecesini
- q; Hareketli Ortalama (MA) Modelinin derecesini ifade etmektedir

Finansal veri seti üzerinde ARIMA Model işletilerek Şekil 4.1’deki gibi bir grafik elde edilmiştir. ARIMA ile yapılan tahminlerin başarılı olduğu anlaşılmaktadır.

4.4.3. Hareketli Ortalama Modeli ile Tahmin

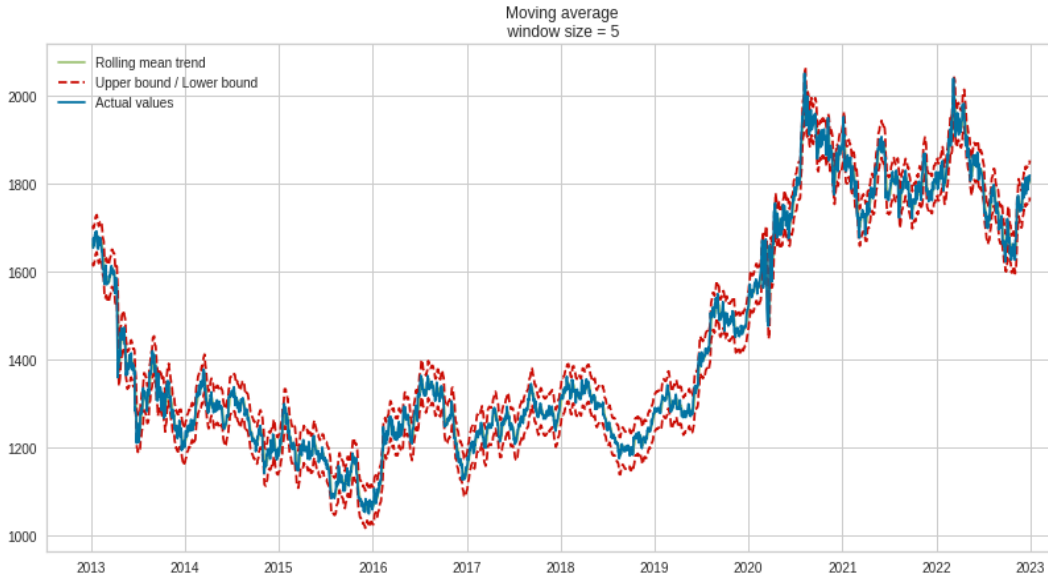
Zaman serileri için gelecek tahminleri yapılırken geçmiş değerler dikkate alınmaktadır. Hareketli ortalama, zaman serileri için uzun vadeli tahminlerde kullanılan istatistiksel bir yöntemdir. Zaman serisi üzerinde aralık hareket ettirmeye dayalı olarak, belirli aralıktaki verilerin ortalaması alınır. Bir aralık dönemi için hesaplama yapılırken, önceki aralığa ait ortalama değeri hesaplama dahil edilir.

Örneğin bir şirketin, 2010 – 2015 yılları arasındaki satış verileri için Hareketli Ortalama Modeli ile 3'er yıllık hareketli ortalama hesaplaması yapılacak olsun. Bu durumda 2010 – 2012, 2011 – 2013, 2012 – 2014, 2013 – 2015 dönemlerine ait ortalamalar hesaplanır.

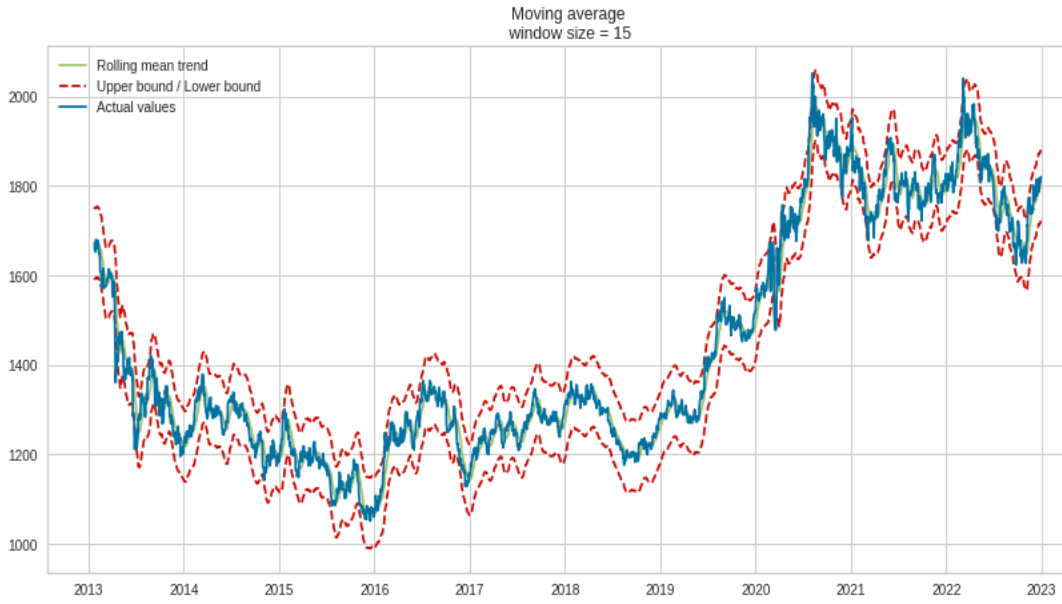
Finansal Veri Seti kullanılarak Hareketli Ortalama Modeli ile aralık değeri 5 günlük seçilerek tahmin işlemi yapıldığında Şekil 4.2'deki gibi sonuçlar elde edilmiştir. Hareketli Ortalama Modeli ile aralık değeri 15 günlük seçilerek tahmin işlemi yapıldığında Şekil 4.3'teki gibi sonuçlar elde edilmiştir. Hareketli Ortalama Modeli ile aralık değeri 30 günlük seçilerek tahmin işlemi yapıldığında Şekil 4.4'teki gibi sonuçlar elde edilmiştir.



Şekil 4.1. ARIMA Model Tahmini Gerçek Değer Karşılaştırma



Şekil 4.2. Hareketli Ortalama Modeli ile Tahmin -5 Gün Aralıklı



Şekil 4.3. Hareketli Ortalama Modeli ile Tahmin -15 Gün Aralıklı

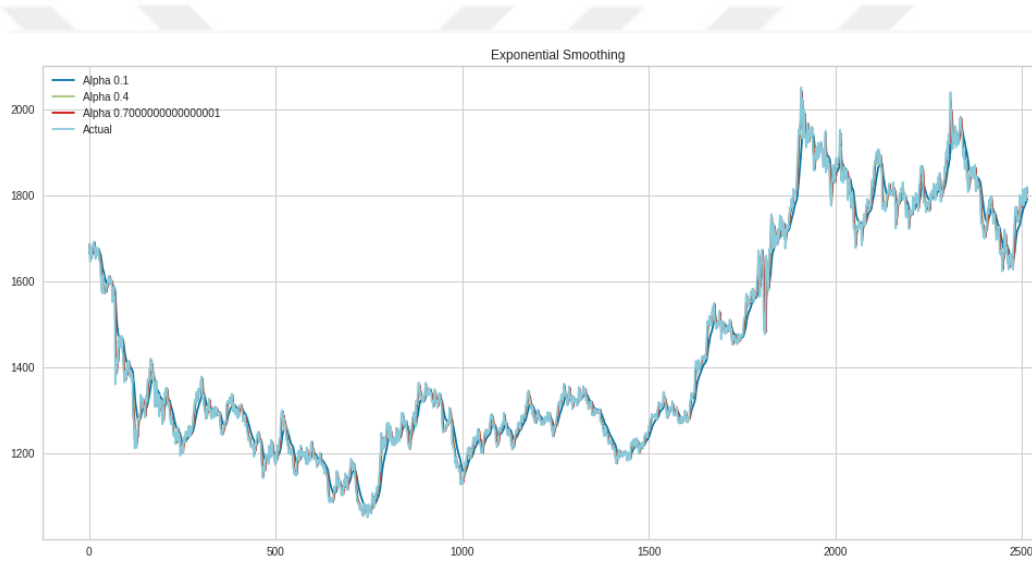


Şekil 4.4. Hareketli Ortalama Modeli ile Tahmin -30 Gün Aralıklı

Grafiklere bakıldığında aralık değerinin küçülmesi ile daha doğru sonuçları alınmıştır. Grafiklerde 5 günlük aralık değeri 15 günlük ve 30 günlük aralık değerlerinden daha iyi sonuçlar vermiştir.

4.4.4. Tek Üstel Düzeltme Modeli ile Tahmin

Bir ortalama etrafında devam eden zaman serileri için kullanılan tahmin yöntemidir. Burada önemli olan, zaman serisinin bir trend eğiliminde olmaması ve mevsimsel dalgalanmalar sergilememesidir. Zaman serisindeki verilere farklı ağırlıklar verilerek hesaplama yapılır. Tüm dönemlerdeki verilere aynı ağırlık değeri verilmez, zaman serisinin son dönemlerine daha büyük ağırlık değeri verilir. Çünkü zaman serisinin son dönemlerinin gelecek tahmini üzerinde daha fazla etkiye sahip olduğu varsayılmaktadır. Bu modelin avantajlarından birisi de az sayıda veri üzerinde de uygulanabilir olmasıdır. Finansal veri seti üzerinde Tek Üstel Düzeltme Modeli ile Tahmin yapıldığında Şekil 4.5'teki gibi sonuçlar elde edilmiştir.



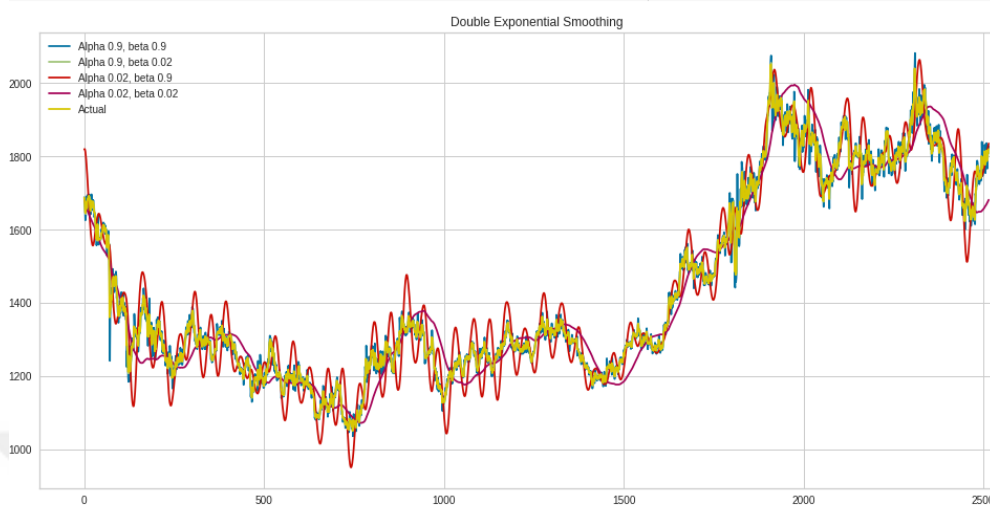
Şekil 4.5. Tek Üstel Düzeltme Modeli ile Tahmin

4.4.5. Çift Üstel Düzeltme Modeli ile Tahmin

Trend eğiliminde olan, mevsimsel dalgalanma göstermeyen zaman serileri için kullanılan bir tahmin yöntemidir. Tek Üstel Düzeltme Modelinde olduğu gibi zaman serisinin son dönemleri daha büyük oranlara sahip olacak şekilde verileri farklı değerlerde ağırlıklandırır. Tek Üstel Düzeltme Modelinden farkı ise iki düzleştirme uygulanmasıdır. Bu durumda iki düzleştirme sabitine ihtiyaç duyulur. Her iki sabit de 0 ile 1 arasında değer alır.

- α düzleştirme sabiti, orijinal verinin tahmini için kullanılır
- β doğrusal düzleştirme sabiti, trendin düzleştirilmesi için kullanılır.

Finansal veri seti üzerinde Çift Üstel Düzeltme Modeli işletildiğinde Şekil 4.6'daki gibi bir grafik elde edilmiştir.



Şekil 4.6. Çift Üstel Düzeltme Modeli ile Tahmin

4.4.6. Zaman Serisi Tahmini Modelleri Karşılaştırma

Zaman serisi tahmininde, finansal veri seti üzerinde işletilen 5 model için Ortalama Karekök Sapması, Ortalama Mutlak Hata ve tahmin değerleri Tablo 4.1.'de verilmiştir.

Tablo 4.1. Zaman Serisi Tahmini Modelleri Karşılaştırma

Model	MAE	RMSE	Tahmin
Süreklilik (Persistence)	28.26	36.48	1819.69
Arima	28.30	36.55	1817.99
Hareketli Ortalama	29.72	38.14	1819.00
Tek Üstel Düzeltme	28.29	36.49	1819.00
Çift Üstel Düzeltme	28.96	37.47	1825.00

4.5. Danışmanlı Öğrenme Performans Sonuçları ve Tartışma

4.5.1. Özellik Mühendisliği İşlemleri

Veri setlerinin makine öğrenmesinde kullanılabilir duruma gelmesi için bir takım ön işlemlerden geçmesi gerekir. İlgili modelin daha iyi sonuçlar vermesi bu ön işlemlerin başarısına bağlıdır. Özellik mühendisliği ise veri madenciliği tekniklerini kullanarak, işlenmemiş verileri temizleyip, bu verilerden yeni özellikler elde edilmesini amaçlar.

Özellik mühendisliği işlemleri dört başlık altında incelenebilir.

- Missing Values (Eksik Değerler)
- Outliers (Aykırı Değerler)
- Feature Scaling (Özellik Ölçeklendirme)
- Encoding (Label Encoding, One-Hot Encoding, Rare Encoding)

Veri setimizdeki verilerin göstermiş olduğu eğilimin çok uzağında kalan verilerdir. Aykırı değer analizi için standart sapma, z-skoru, boxplot yöntemleri kullanılır. Farklı nedenlerden dolayı veri setinde eksik değerler oluşabilmektedir. Bu eksik değerler analiz hesaplamalarında hatalı hesaplama yapılmasına neden olabilmektedir. Bu sorunu çözmek için eksik veriler için silme yapılabilir, bir değer atanabilir ya da tahmine dayalı bir değer belirlenip doldurulabilir.

Makine öğrenmesinde kullanılan algoritmalar için girdi değerleri sayılardan oluşur. Bu nedenle metinsel verilerin uygun yöntemler ile sayısal verilere dönüştürülmesi gereklidir. Bu işleme encoding denir. Label Encoding, One-Hot Encoding, Rare encoding yöntemleri ile encoding işlemleri yapılır.

Encoding işlemleri sonrası bazı değerler arasında çok büyük farklar oluşabilmektedir. Aralarında büyük farklar olan bu veriler ölçeklendirilerek veri setine ait sayısal değerler dar bir aralığa sıkıştırılır. Bu işleme özellik ölçeklendirme denir.

4.5.2. Doğrusal Regresyon ile Tahmin

Doğrusal Regresyon (Linear Regression), değerler arasında fonksiyonel bir ilişki varsayımı ile bağımsız değişkenleri kullanarak bağımlı değişkenlerin değerini tahmin etme yöntemidir. Tanımından da anlaşılacağı gibi bu yöntemde bağımlı ve bağımsız olmak üzere iki değişken kullanılır. Değeri tahmin edilmek istenen değişken bağımlı değişken olarak tanımlanır. Bağımlı değişkenin değerini tahmin etmek için kullanılan ve bağımlı değişken ile ilişkili olan değişkene de bağımsız değişken denir.

4.5.3. Rassal Orman Regresyonu ile Tahmin

Rassal Orman Regresyonu (Random Forest Regressor), çok sayıda karar ağacını oluşturup, bu karar ağaçlarını eğiterek, eğitilmiş karar ağaçlarını kullanarak tahmin

oluşturmaya yarayan denetimli bir öğrenme algoritmasıdır. Tercih edilmesindeki etkili bazı nedenler;

- Regresyon problemlerine uygulanabildiği gibi sınıflandırma problemlerine de uygulanabilmesi
- Hızlı eğitilebilmesi
- Tahmin hızının yüksek olması
- Çok boyutlu problemlere doğrudan uygulanabilmesi

4.5.4. Karar Ağacı Regresyonu ile Tahmin

Karar ağaçları (Decision Tree), verileri sınıflandırmak için uygulanan bir tekniktir. Bunun için bir veri setinde, sınıfı belirli olan verilerden yararlanır. Bu sayede sınıfı belirli olmayan veriler sınıflandırılır. Karar ağaçlarının tercih edilmesinin bazı nedenleri şu şekilde sıralanabilir.

- Veri madenciliğinde tanımlayıcı ve tahminleyici niteliklere sahip olmaları
- Kolay uygulanabilir olmaları
- Kolay değerlendirilebilir olmaları
- Yüksek güvenilirlik oranına sahip olmaları
- Depolama sistemleri ile kolaylıkla entegre olmaları başlıca tercih edilme nedenleridir.

4.5.5. Danışmanlı Öğrenme Modelleri Karşılaştırma

Danışmanlı Öğrenme modelleri finansal veri seti üzerinde yapılan tahminler, Ortalama Karekök Sapması ve Ortalama Mutlak Hata değerleri Tablo 4.2.'de verilmiştir. Decision Tree Regressor modeli en başarılı tahmini yapmıştır.

Tablo 4.2. Danışmanlı Öğrenme Modelleri Karşılaştırma

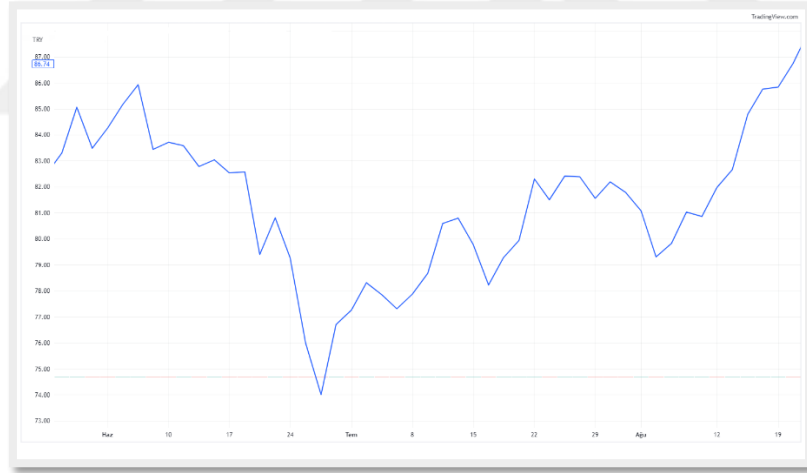
Model	MAE	Tahmin
Doğrusal Regresyon	73.55	1790.98
Rassal Orman Regresyonu	108.38	1812.67
Karar Ağacı Regresyonu	94.02	1839.25

4.6. Kritik Olaylar ile İlgili Değerlendirme

Bu bölümde, Tablo 4.3'te tarih aralığı ile gösterilmiş olayların altın fiyatları nasıl etki oluşturdukları incelenecektir. Belirtilen dönemlere ait Altın-TL grafikleri yorumlanacaktır.

Tablo 4.3. Piyasayı Etkileyen Kritik Dönemler

Kritik Olay	Tarih Aralığı
Gezi Parkı Eylemleri	28 Mayıs – 20 Ağustos, 2013
15 Temmuz Darbe Girişimi	20 Mart – 20 Ağustos, 2016
Papaz Andrew Brunson Olayı	1 Ağustos – 10 Kasım, 2018
Covid-19 Pandemisi	1 Haziran 2019 – 30 Haziran 2020
Türkiye Hükümeti Doğrudan Döviz Müdahalesi	1 Aralık – 31 Aralık, 2021

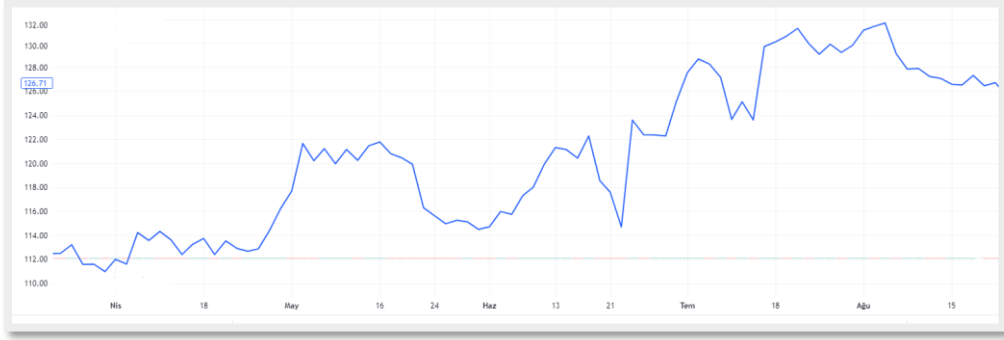


Şekil 4.1. Gezi Parkı Eylemleri Dönemi Altın Fiyatı (₺)

Şekil 4.1'deki grafiğe bakıldığında Gezi Parkı Eylemleri başladıktan sonra altın fiyatlarında düşüş yaşanmıştır. Dönemin hükümeti, olayları kontrol altında aldıktan sonra altın fiyatları, olayların başladığı zamanki seviyelerine geri dönmüştür. Gezi parkı olayları altın fiyatlarını olumsuz yönde etkilemiştir.

Şekil 4.2'de 15 Temmuz Darbe Girişimi Dönemine ait altın fiyatları Türk Lirası cinsinden grafik ile gösterilmiştir. 15 Temmuz öncesinde de dalgalanma eğiliminde olan altın

fiyatları, darbe girişimi sonrasında artış eğilimine girmiştir. Darbe girişiminin yaşandığı dönemde altının güvenli liman olarak görüldüğü söylenebilir. Bu dönemde yaşanan olaylar altın fiyatlarını olumlu yönde etkilemiştir.



Şekil 4.2. 15 Temmuz Darbe Girişimi Dönemi Altın Fiyatı (₺)



Şekil 4.3. Papaz Andrew Brunson Olayı Dönemi Altın Fiyatı

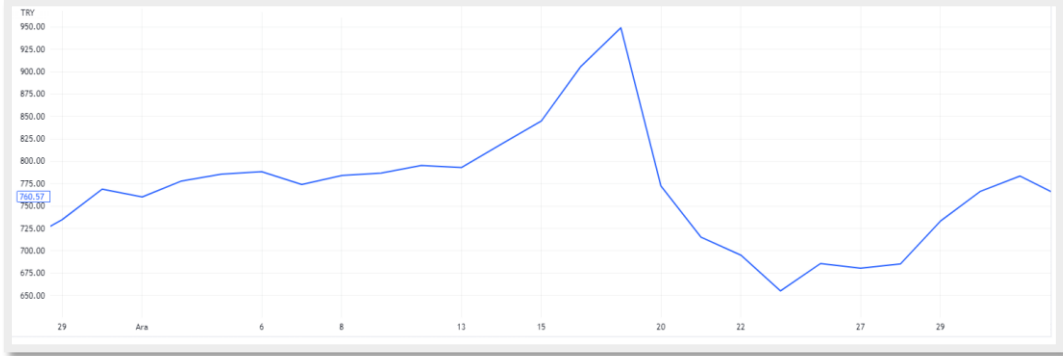
Şekil 4.3'te Papaz Andrew Brunson olayının yaşandığı döneme ait Altın/₺ fiyat grafiği gösterilmiştir. Bu dönemde altın fiyatlarında bir düşüş trendi yaşanmıştır. Yatırımcıların Türkiye ekonomisine karşı yaşadıkları güven sorunu nedeni ile altın fiyatları Papaz Andrew Brunson olayından olumsuz şekilde etkilenmiştir.

Şekil 4.4'te Covid-19 Pandemisi Dönemine ait Altın/₺ fiyat grafiği gösterilmiştir. Bu dönemde altın fiyatlarında bir artış terendi görülmektedir. Tüm dünyayı etkisi altına almış bir salgın yatırımcıda güven problemi ortaya çıkarmıştır. Bu durum güvenli yatırım aracı

olarak görülen altına olan talebi arttırmıştır. Bu dönemde altına fiyatlarında bir artış trendi yaşanmıştır. Covid-19 Pandemisi altın üzerinde pozitif etki oluşturmuştur.



Şekil 4.4. Covid-19 Pandemisi Dönemi Altın Fiyatı



Şekil 4.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi

2021 yılının sonlarında Türkiye Hükümeti tarafından doğrudan döviz müdahalesi yapılmıştır. Şekil 4.5'te bu döneme ait Altın/₺ grafiği verilmiştir. Art arda yapılan müdahalelerden sonra (5. Müdahale sonrası) altın fiyatlarında keskin bir düşüş görülmüştür. Ancak altın kısa sürede toparlanarak tekrar artış eğilimi göstermiştir. Bu müdahaleler altına fiyatlarını düşürmüş ancak artış eğilimini tam olarak durduramamıştır.

5. SOSYAL MEDYA VERİSİ ÜZERİNDE ALTIN GERİ DÖNÜŞ KONU MODELLEMESİ

5.1. Veri Toplama

Bu çalışma kapsamında metin tabanlı veri setini oluşturmak için Twitter verisi kullanılmıştır. Twitter üzerinden veri çekmek için BeautifulSoup4¹ ve Twitter API5² kullanılmıştır. Twitter’da atılan tweet’leri kategorilere ayırabilmek için etiketler (hashtag) kullanılır. Etiket oluşturmak için bir kelimenin başına diyez işareti (#) eklenir. Bir tweet içerisinde birden fazla etiket olabilir. Etiketlerin kullanımı ile belirli bir konu ile ilgili çok sayıda kişi tek bir başlık altında fikirlerini bir araya getirmiş olur. Etiketlenmiş olan konu ile ilgili hızlı bir bilgilendirme ve haberleşme ağı kurulmuş olur. Kullanıcılar her tweet için beğeni ya da retweet (aynı tweet’i paylaşma) yapabilir, tweet altına yorum ekleyebilirler.

Metin tabanlı veri seti oluşturulurken belirlenmiş olan kritik dönemlere ait tweet’ler kullanılmıştır. Altın getiri tahmini ile ilgili fikir verebilecek tweet’leri filtrelemek için etiketler kullanılmıştır. Bu etiketler: gold, ons, gold/sell, selling, coins, bitcoin, exchange, buy, buying, brent, euro, dollar, silver, pay, paying, yellow, forrest, expensive, cheap, coin, metal, collect, creditcards, goldjewelry, goldexchange, electronicgoldreceipts, goldinvestment, buyinggold, goldprice, diamond, xauusd, xaueur, xautr, bestwaybuygold, bestwaysellgold, goldpriceusd, carat, goldpricehourly.

Tablo 5.1. Kritik Dönemlere ait Tweet Sayıları ve Veri Boyutu

Kritik Olay	Veri Boyutu	Tweet Sayısı
Gezi Parkı Eylemleri	7 MB	17184
15 Temmuz Darbe Girişimi	13 MB	28350
Papaz Andrew Brunson Olayı	13 MB	21139
Covid-19 Pandemisi	45 MB	76609
Türkiye Hükümeti Doğrudan Döviz Müdahalesi	2 MB	3997

¹ <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

² <https://developer.twitter.com/en/docs/twitter-api>

Veri setlerinin toplandığı olaylar ve tarih aralıkları Tablo 2.1’de verilmiştir. Belirtilen dönemlere ait toplanan tweet sayıları ve veri büyüklüğü Tablo 5.1’ de verilmiştir.

5.2. Veri Ön İşleme

Twitter’den elde edilmiş verilerin analiz edilebilmesi için bir takım ön işlemden geçmesi gerekir. İlk olarak, dize türüne dönüştürülmeden önce, gereksiz sütunlar, NaN ve NaT satırları, tweet’lerin gövde sütunundan kaldırılmıştır. Ardından belirli bir ayraç ile tokenize edilip yani kelimelere ayırma işlemi uygulanmıştır. Genellikle boşluk karakteri ayraç olarak kullanılır. Bu sayede bir kelime dizisi elde edilir. Daha sonra metinler sırasıyla spaCy ile lemmatize edilmiş ve Gensim³ ile temizlenmiştir. Bir sonraki işlem, bağlaçlar, boşluklar, rakamlar, noktalama işaretleri, tek başına anlama bir katkısı olmayan kelimeler ve özel karakterler temizlenmiştir. Tüm kelimeler küçük harfle yazılmıştır. Ön işleme yapılmasının nedeni, verilerin sınıflandırıcı model ve sistemlerin kolaylıkla sınıflandırabileceği kelimeler cinsinden sunulmasıdır. Daha sonra verileri temizlemek ve yapılandırmak için Python’daki NLP araçları kullanılmıştır. Ardından duyarlılık analizi ve NMF konu modellemesi yapılmıştır.

5.3. Konu Modelleme

Konu modelleme, makine öğrenmesinde ve doğal dil işlemede (NLP) veri seti içerisindeki gizli konuları ortaya çıkarmak için kullanılan bir yöntemdir. Önceden belirlenmiş kategoriler veya etiketler yardımı olmadan, girdinin sınıflandırdığı denetimsiz bir öğrenme türüdür. Konu modelleme, yapılandırılmamış büyük veri setlerini düzenlemeye yardımcı olur. Konu modelleme biyoinformatik, NLP, genetik bilgi gibi alanlarda kullanılmaktadır. Konu modelleme için kullanılan bazı yöntemler vardır. Gizli Dirichlet Ayrımı(Latent Dirichlet Allocation -LDA) en çok kullanılan yöntemdir. LDA denetimsiz bir öğrenme algoritması olduğu için modelleme öncesinde tanımlanmış kelimelere ihtiyaç duyulmaz. Algoritma, her belgenin konu dağılımlarını ve her konunun kelime dağılımlarını tahmin etmeye çalışır. Her konu, kelime dağarcığındaki kelimeler üzerinden bir olasılık dağılımı olarak tanımlanır.

³ <https://pypi.org/project/gensim/>

Bu çalışmada, bir belge koleksiyonundaki temel konuları veya temaları bulma süreci olan Negatif Olmayan Matris Faktoring (NMF) tekniği kullanılmıştır. NMF kullanarak konu modellemede, her satırın bir metin belgesini ve her sütunun sözlükten bir kelimeyi gösterdiği bir metin belgeleri külliyyatından oluşan bir matris oluşturulur. Kelime frekansları negatif olmayan sayılar olduğundan, matristeki girişler her kelimenin her sayfada ne sıklıkla geçtiğini gösterir (Lee, 1999).

5.4. Duygu Analizi

Doğal dil işleme yöntemleri kullanılarak, dijital metinlerin duygusal olarak sahip oldukları ifadeleri pozitif, negatif veya nötr olarak sınıflandırmaya duygu analizi denir. Örneğin, insanların altın hakkında ne hissettiğini belirlemek için haberleri ve sosyal medyayı incelemek için duygu analizi kullanılabilir. Duygu analizi, insanların altına yönelik tutumlarının tipik olarak olumlu olduğunu tespit ederse, bu, değerli metale olan talebi artırabilir ve emtia fiyatını yükseltebilir. Öte yandan, duygu analizi altına karşı olumsuz bir tutum ortaya koyarsa, bu durum altına olan talebin azalmasına ve fiyatının düşmesine neden olabilir. GoldCriticalTweets veri seti üzerinde veri ön işleme adımlarından sonra metin sınıflandırma işlemleri gerçekleştirilmiştir. SVM (Vapnik, 1998), CNN (Albawi, 2017), LSTM (Hochreiter, 1996), MLP (Ramchoun, 2016), BERT (Devlin, 2018), ELECTRA (Clark, 2020) ve RoBERTa (Liu, 2019) gibi birçok farklı sınıflandırma algoritmasının performansı analiz edilmiştir.

5.4.1. Destek Vektör Makineleri

Destek Vektör Makineleri, sınıflandırma tekniklerinde kullanılan en etkili denetimli öğrenme tekniklerinden birisidir. Bir düzlemde yer alan noktalar arasında bir ayırma hiper düzlemi oluşturarak maksimize edilmiş margin değerine dayalı istatistiksel bir sınıflandırma tekniğidir.

Hiper düzlem, düzlemi bölen sınır olarak tanımlanabilir. Veri setini 2 sınıfa ayırmaya yarayan bir karar eşiğidir. Hiper düzlem, düz bir çizgi olabileceği gibi 2D bir düzlem de olabilir. Destek Vektör Makinelerinde, amaç gelecek verilerin hangi sınıfta yer alacağına karar vermektir. Bu karar mekanizmasını oluştururken, veri seti üzerindeki örnekler ile ayırıcı hiper düzlem arasındaki maksimum margin değeri elde edilmek istenmektedir.

ile karşılaştırılarak hata hesaplaması yapılır. Bu hata oranını azaltmak için model üzerindeki ağırlık değerleri güncellenir ve yeniden çıktı elde etme işlemi uygulanır. Model ilk işletildiği zaman ağırlık değerleri rastgele verilmektedir. Geriye doğru hesaplama işlemi ile elde edilen hata değerleri, ağırlık değerlerine dağıtılır ve model yeniden işletilerek hata oranı azaltılmaya çalışılır.

5.4.3. Evrişimli Sinir Ağları

Evrişimli sinir ağları (Convolutional neural network), derin öğrenmenin bir alt dalıdır ve genellikle görsel bilgileri analiz etmek için kullanılır. Yaygın kullanımlar arasında görüntü ve video tanıma, öneri sistemleri, görüntü sınıflandırma, tıbbi görüntü analizi ve doğal dil işleme yer alır. Evrişimli sinir ağı, sinir ağının katmanlardan en az birinde genel matris çarpımı yerine evrişim adı verilen matematiksel bir işlemi (özel bir doğrusal işlem türü) kullandığını belirtir.

CNN'deki katmanlardan ilki olan Evrişimli Katman, orijinal giriş boyutlarından daha küçük bir filtre matrisi üzerinde gezinir ve bu matristen belirli özellikleri yakalamaya çalışır. Özellik haritası, girdi ile filtre çarpılarak oluşturulur. Stride, her adımda evrişim katmanının adım boyutunu belirler. Adım boyutu arttıkça özellik haritası küçülür. Girişle aynı boyutta bir özellik matrisi oluşturmak için girişe fazladan 0'lar eklenir. Bu işleme dolgu denir. Evrişim katmanından sonra havuzlama işlemi uygulanır. Boyut küçültme yapılır. En yaygın kullanım, global maksimum havuzlamadır.

5.4.4. Uzun Kısa Süreli Bellek

Uzun Kısa Süreli Bellek, akademik alanda bilinen adı ile Long-Short Term Memory (LSTM), bir Yinelemeli Sinir Ağı (RNN) mimarisidir. RNN mimarisinde veriler arasındaki bağlam boşluklarına bir çözüm ihtiyacı duyulması ile LSTM ağları ortaya çıkmıştır. LSTM mimarisi, Hochreiter ve Schmidhuber tarafından 1997 yılında tanıtılmıştır.

LSTM hücreleri, kapılardan ve ağırlıklardan oluşur. Bir LSTM katmanı, bellek blokları olarak bilinen, tekrar tekrar birbirine bağlanan bir dizi bloktan oluşur. Her biri, bir veya daha fazla tekrarlayan bağlı bellek hücreleri ve bu hücreler için sürekli yazma, okuma ve sıfırlama işlemlerini sağlayan giriş, çıkış ve unutma kapıları olan üç çarpımsal birim

içerir. Ağ, hücrelerle ancak kapılar aracılığıyla etkileşime girebilir (Rothman, 2018: 376). En basit durumda, ağa bir seferde bir diziden bir gözlem gösterilir ve daha önce gördüğü hangi gözlemlerle alakalı olduğunu, bunların bir tahmin ile ilişkisini hatırlayabilir. Yani LSTM modeller, verilerden zamansal bağımlılığı öğrenebilir ve kullanabilir. Bir dizideki uzun vadeli korelasyonları öğrenme yeteneği nedeniyle LSTM ağları karmaşık çok değişkenli dizileri doğru bir şekilde modelleme yeteneğine sahiptir. Ayrıca birden fazla paralel girdi dizisini ayrı ayrı modelleyebilirler (Chollet, 2019: 196).

5.4.5. Transformatörlerden Çift Yönlü Enkoder Gösterimleri

Kısaca BERT diye isimlendirilen bu model 2018 yılında Google tarafından duyurulmuştur. Modelin işletilmesi esnasında, tüm katmanlarında, hem sağdan sola hem de soldan sağa değerlendirme yaparak kurduğu anlamsal bağlantı ile diğer modellerden ayrılmaktadır. Çift yönlü olması sayesinde veriler arasında daha iyi bağlar kurulabilmektedir. Örneğin Ahmet'ten bahsedilen bir metnin devamında "onu" kelimesinin Ahmet'i işaret ettiğini bilmektedir.

BERT önceden eğitilmiş bir modeldir. BERT'in eğitilmesi esnasında, içerisinde 800 milyon kelime bulunan BookCorpus ve içerisinde 2,5 milyar kelime bulunan Wikipedia veri setleri kullanılmıştır. BERT modeli eğitilirken kelimeler arasındaki ilişki için Masked Language Modeling (MLM) tekniği ile eğitilir. Benzer şekilde cümleler arasında doğru ilişkinin kurulması için Next Sentence Prediction (NSP) tekniği ile eğitilir.

MLM tekniğinde, kelime dizisinin %15'lik kısmı [MASK] belirteci ile maskelenir. Maskelenmemiş kelimeler arasında kurulan anlamsal bağlar ile maskelenmiş kelimeler tahmin edilmeye çalışılır.

NSP tekniğinde ise, veri setindeki ikinci cümlelerin %50'si rastgele yerleri değiştirilir. Kalan %50'lik kısım orijinal hali ile bırakılır. Rastgele değiştirilen ikinci cümlelerin ilk cümlelerin devamı olup olmadığı tahmin edilmeye çalışılır.

5.4.6. M-BERT

M-BERT modeli 104 farklı dil için, Wikipedia'dan elde edilmiş veri setlerinin BERT modeli üzerinde eğitilmesi ile ortaya çıkmıştır. Bu sayede M-BERT 104 dilden gelen tüm veriler arasına anlamsal bağlar kurabilir.

5.4.7. RoBERTa

Facebook geliştiricileri tarafından eğitilmiş bir modeldir. BERT'in Masked Language Modeling (MLM) tekniği ile eğitilmiştir. BERT'ten daha fazla veri üzerinde daha uzun süre eğitilmiş bir modeldir. BERT'in kullanmış olduğu BookCorpus Wikipedia veri setleri de dahil olmak üzere CommonCrawl News veri seti, Stories from Common Crawl veri seti ve Web text corpus veri seti üzerinde eğitilmiştir. Eğitim prosedürünü iyileştirmek için, RoBERTa, Next Sentence Prediction (NSP) tekniğini BERT'in ön eğitiminden kaldırır ve dinamik maskeleye sunar, böylece eğitimler sırasında maskelenmiş belirteç değişir.

5.4.8. Distilled BERT

Doğal Dil İşleme'de (NLP) büyük ölçekli önceden eğitilmiş büyük modellerin kısıtlı donanımlar üzerinde çalıştırılması oldukça zordur. Bu zorluğa karşı DistilBERT model ile daha küçük bir genel amaçlı dil temsil modeli önerilmiştir. DistilBERT 6 katmanlı, 768 gizli katmana ve 66M parametreye sahiptir (URL-1). Uygulamada, DistilBERT ktrain kütüphanesi ile gerçekleştirilir. Transformatör, 'distillbert-base-uncased' adıyla Huggingface kitaplığından elde edilir. Test ve eğitim verileri daha sonra ön işleme tabi tutulur.

5.4.9. OpenAI GPT-2

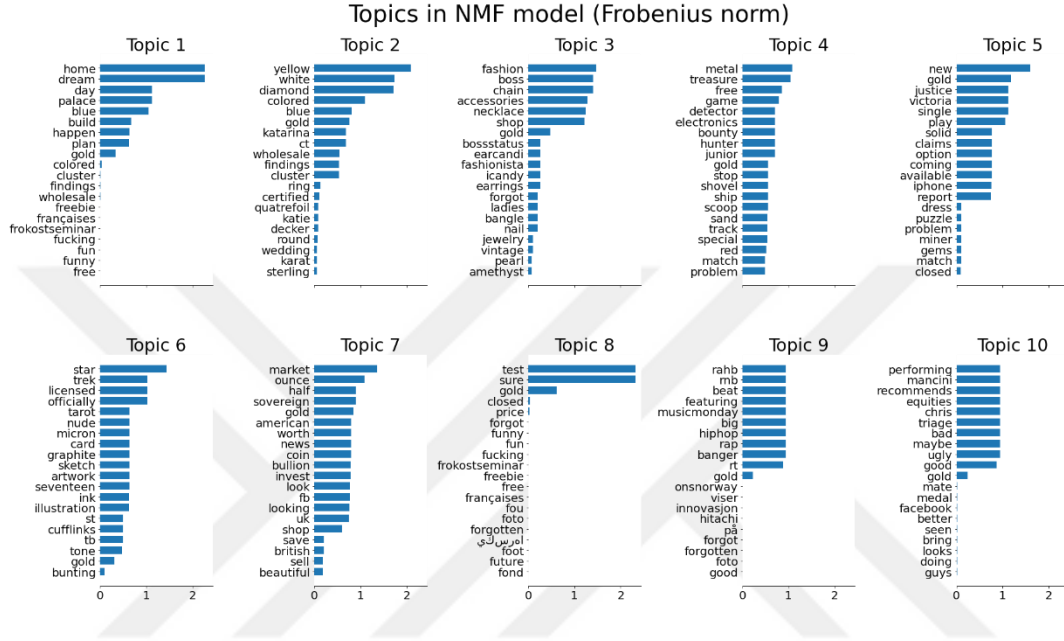
GPT-2, 8 milyon web sayfasından oluşan bir veri kümesinde eğitilmiş 1,5 milyar parametreye sahip büyük, dönüştürücü tabanlı bir dil modelidir. GPT-2, bir metinde verilen önceki kelimelerden sonraki kelimeyi tahmin etmeyi amaçlar. GPT-2'de BERT'den farklı olarak, tahmin için son 'belirteç' kullanılır. Bu nedenle padding işlemleri BERT'e göre sola yapılmaktadır.

5.5. Konu Bazlı Kritik Olaylar ile İlgili Değerlendirme

Bu bölümde elde edilen sosyal medya verisi üzerinde, her bir kritik olaya ait veri seti için ayrı ayrı konu modelleme ve duygu analizi algoritmaları işletilerek elde edilen sonuçlar gösterilmiştir.

Şekil 5.1'de Gezi parkı olaylarının yaşandığı döneme ait Twitter verisi kullanılarak, NMF yöntemi ile elde edilmiş konu modellemesi sonuçları gösterilmiştir. Bu konular ile bu

döneme ait farklı duygu analizi yöntemleri kullanılarak elde edilen sonuçlar Tablo 5.2’de gösterilmiştir. Tablo 5.2’deki sonuçlara bakıldığında CNN ve MLP modellerinin en iyi performans ölçütlerine sahip olduğu anlaşılmaktadır. Batılı yatırımcıların ETF (Exchange-traded Funds) portföyleri tükendikçe altın talebi yıldan yıla azalmıştır. Bu durumun başlıca nedenlerinden birisi Gezi Parkı olayları olmuştur.

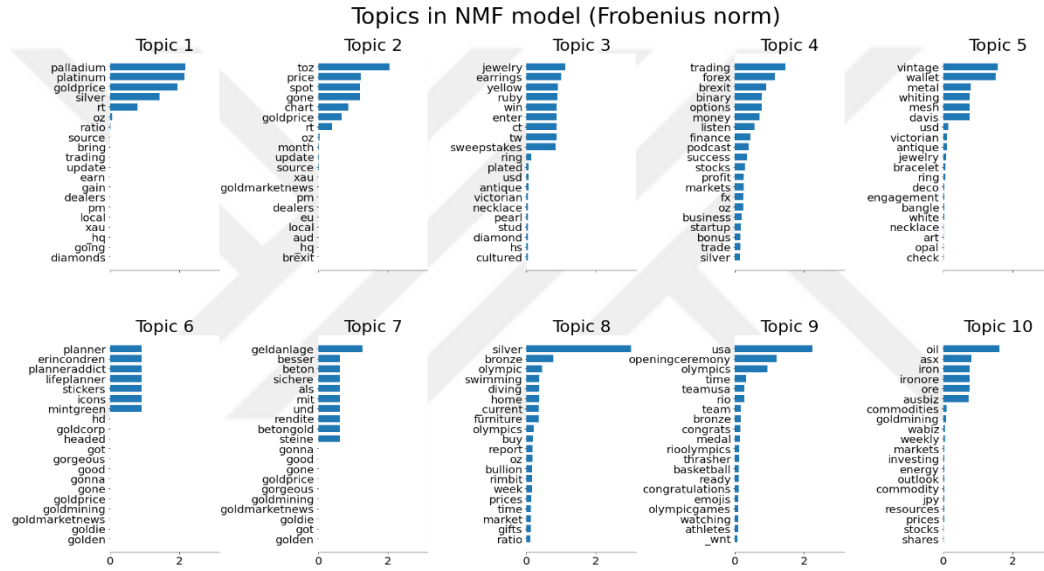


Şekil 5.1. Gezi Parkı Olayları Konu Modelleme

Tablo 5.2. Gezi Parkı Olayları Duygu Analizi Sonuçları

Model	Accuracy	F-Measure	Precision	Recall
SVM	0.99	0.99	0.99	0.99
CNN (1D)	1	1	1	1
MLP	1	1	1	1
LSTM	0.84	0.58	0.54	0.64
BERT	0.81	0.73	0.82	0.72
MBERT	0.7	0.5	0.46	0.57
RoBERTa	0.87	0.81	0.88	0.79
DistilBERT	0.92	0.91	0.91	0.91
GPT-2	0.68	0.6	0.63	0.59
ELECTRA	0.94	0.93	0.93	0.94

Şekil 5.2’de 15 Temmuz Darbe girişiminin yaşandığı döneme ait Twitter verisi kullanılarak, NMF yöntemi ile elde edilmiş konu modellemesi sonuçları gösterilmiştir. Bu konular ile bu döneme ait farklı duygu analizi yöntemleri kullanılarak elde edilen sonuçlar Tablo 5.3’te gösterilmiştir. Tablo 5.3’teki sonuçlara bakıldığında CNN modelinin en iyi performans ölçütüne sahip olduğu anlaşılmaktadır. WGC verilerine bakıldığında altın talebi 2016 yılının 3. çeyreğinde %10 düşüş ile 992.8 tona gerilemiştir. 2015 yılının 2. çeyreğinden bu yana en düşük seviye olmuştur. Bu dönemde, Türkiye’de enflasyon ve faiz oranlarında artış görülmüş, istihdam verileri kötüleşmiş ve büyüme yavaşlamıştır.

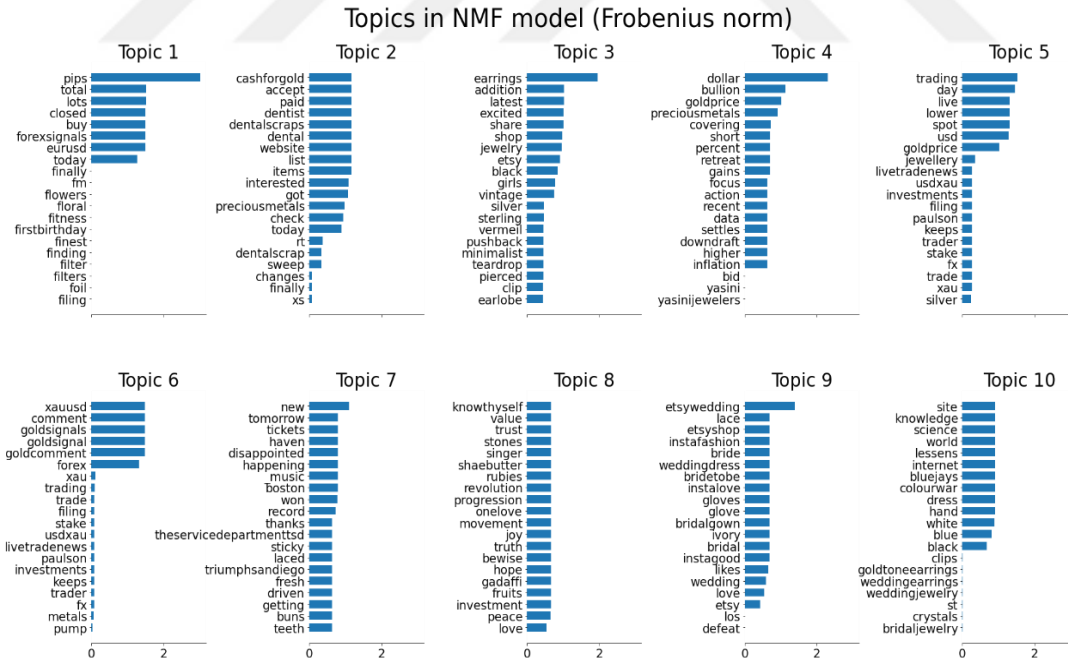


Şekil 5.2. 15 Temmuz Darbe Girişimi Konu Modelleme

Şekil 5.3’te Papaz Andrew Brunson Davasının yaşandığı döneme ait Twitter verisi kullanılarak, NMF yöntemi ile elde edilmiş konu modellemesi sonuçları gösterilmiştir. Bu konular ile bu döneme ait farklı duygu analizi yöntemleri kullanılarak elde edilen sonuçlar Tablo 5.4’te gösterilmiştir. Tablo 5.3’teki sonuçlara bakıldığında CNN ve MLP modellerinin en iyi performans ölçütüne sahip olduğu anlaşılmaktadır. Bu dönemde görülen büyük EFT çıkışları, Merkez Bankası tarafından yapılan alımlar ve tüketici talebinde görülen %13’lük artışla dengelenmiştir. Davanın 12 Ekim 2018’de görülen ikinci duruşması sonucunda Papaz Brunson’ın tahliye olması ile birlikte altın ve döviz kuru düşüş trendine girmiştir.

Tablo 5.3. 15 Temmuz Darbe Girişimi Olayları Duygu Analizi

Model	Accuracy	F-Measure	Precision	Recall
SVM	0.99	0.99	0.99	0.99
CNN (1D)	1	0.99	1	0.99
MLP	0.99	0.99	0.99	0.99
LSTM	1	0.99	0.99	0.99
BERT	0.95	0.93	0.94	0.93
MBERT	0.53	0.23	0.18	0.33
RoBERTa	0.88	0.69	0.79	0.68
DistilBERT	0.97	0.97	0.97	0.97
GPT-2	0.85	0.84	0.87	0.82
ELECTRA	0.92	0.78	0.95	0.75

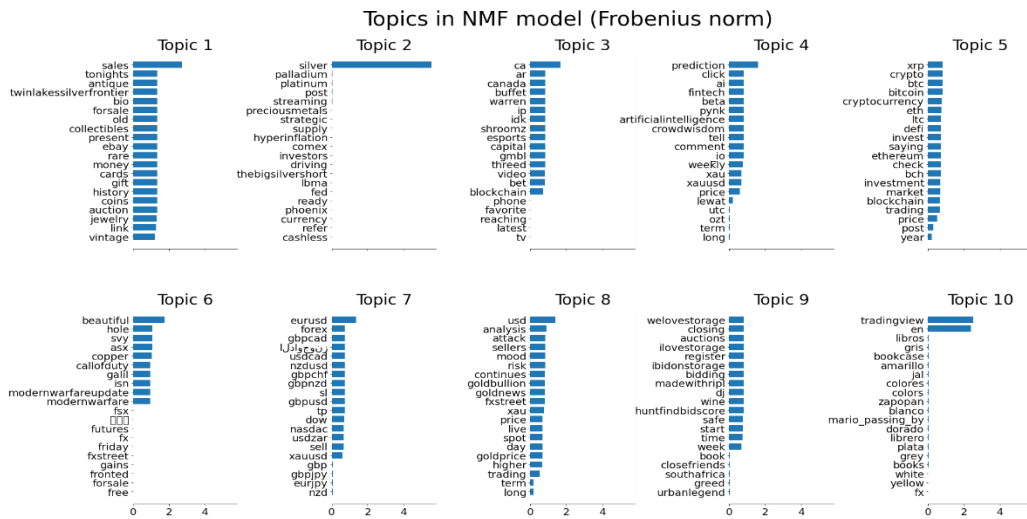


Şekil 5.3. Papaz Andrew Brunson Davası Konu Modelleme

Tablo 5.4. Papaz Andrew Brunson Davası Duygu Analizi

Model	Accuracy	F-Measure	Precision	Recall
SVM	0.99	0.99	0.99	0.99
CNN (1D)	1	1	1	1
MLP	1	1	1	1
LSTM	0.99	0.99	0.99	0.99
BERT	0.84	0.81	0.80	0.83
MBERT	0.47	0.21	0.16	0.33
RoBERTa	0.85	0.83	0.85	0.82
DistilBERT	0.88	0.86	0.87	0.86
GPT-2	0.74	0.72	0.71	0.74
ELECTRA	0.92	0.89	0.89	0.90

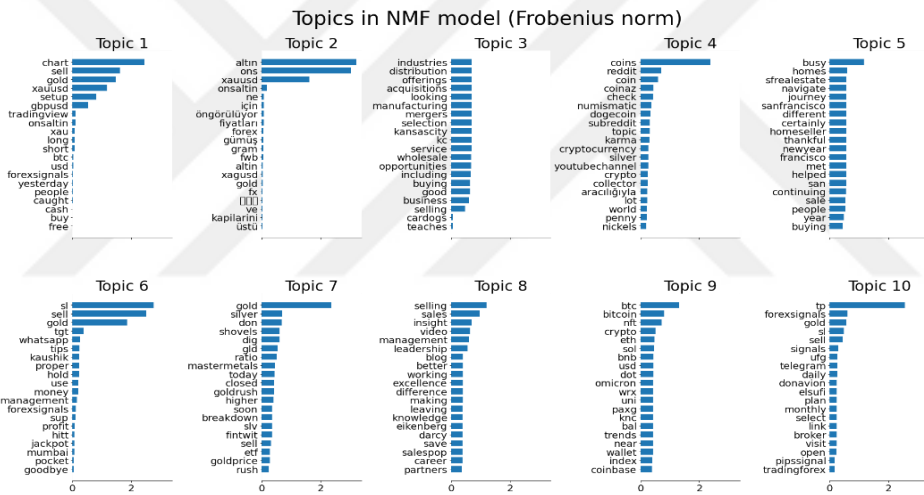
Şekil 5.4'te Covid-19 Pandemisinin yaşandığı döneme ait Twitter verisi kullanılarak, NMF yöntemi ile elde edilmiş konu modellemesi sonuçları gösterilmiştir. Bu konular ile bu döneme ait farklı duygu analizi yöntemleri kullanılarak elde edilen sonuçlar Tablo 5.5'te gösterilmiştir. Tablo 5.5'teki sonuçlara bakıldığında CNN, MLP ve LSTM en iyi performans ölçütlerine sahip olduğu anlaşılmaktadır. WGC'ye göre, altın talebi 2019 yılının 3. çeyreğinde hafif bir artışla 1.107,9 tona yükselmiş ve yıllık altın talebi 4,355,7 tona düşmüştür. Bu küresel salgın döneminde güvenli liman olarak değerlendirilen altına olan talep atmıştır. Pandemi dönemindeki olumsuz haber ve yorumlar hisse senetleri, altın ve faiz oranları üzerinde oldukça etkili olmuştur.



Şekil 5.4. Covid-19 Pandemisi Konu Modelleme

Tablo 5.5. Covid-19 Pandemisi Duygu Analizi Sonuçları

Model	Accuracy	F-Measure	Precision	Recall
SVM	0.99	0.99	0.99	0.99
CNN (1D)	1	1	1	1
MLP	1	1	1	1
LSTM	0.99	0.99	0.99	0.99
BERT	0.84	0.81	0.8	0.83
MBERT	0.47	0.21	0.16	0.33
RoBERTa	0.85	0.83	0.85	0.82
DistilBERT	0.88	0.86	0.87	0.86
GPT-2	0.74	0.72	0.71	0.74
ELECTRA	0.92	0.89	0.89	0.90



Şekil 5.5. Türkiye Hükümeti Doğrudan Döviz Müdahalesi Konu Modelleme

Şekil 5.5'te Türkiye Hükümetinin doğrudan döviz müdahalesinde bulunduğu döneme ait Twitter verisi kullanılarak, NMF yöntemi ile elde edilmiş konu modellemesi sonuçları gösterilmiştir. Bu konular ile bu döneme ait farklı duygu analizi yöntemleri kullanılarak elde edilen sonuçlar Tablo 5.6'da gösterilmiştir. Tablo 5.6'daki sonuçlara bakıldığında RoBERTa modelinin en iyi performans ölçütlerine sahip olduğu anlaşılmaktadır. WGC verilerine göre 2021'de bir bütün olarak (OTC hariç) altın talebi, 4. çeyrekte yaklaşık %50'lik artış göstermiştir. Art arda yapılan döviz müdahalelerinde 5. Müdahalenin ardından altın fiyatlarında sert düşüşler yaşanmıştır.

Tablo 5.6. Türkiye Hükümeti Doğrudan Döviz Müdahalesi

Model	Accuracy	F-Measure	Precision	Recall
SVM	0.99	0.99	0.99	0.99
CNN (1D)	1	1	1	0.99
MLP	1	1	1	0.99
LSTM	0.98	0.97	0.98	0.95
BERT	0.76	0.61	0.63	0.61
MBERT	0.72	0.59	0.6	0.59
RoBERTa	1	1	1	1
DistilBERT	0.76	0.58	0.65	0.57
GPT-2	0.7	0.55	0.58	0.55
ELECTRA	0.79	0.61	0.66	0.61

6. SONUÇLAR VE ÖNERİLER

Devletlerin izlemiş olduğu para politikaları, devletler arası ilişkiler, devletlerin ekonomik koşulları, sosyal hayatta yaşanan çok kritik olaylar, afetler vs. altın getirilerini etkilemektedir. Altın yatırımcılarının altın getirilerini etkileyen bu durumları dikkate almaları önemlidir.

Bu çalışmada kritik kabul edilecek önemli olayların yaşandığı dönemlere ait veri setleri kullanılarak altın getirileri ile ilgili araştırma yapılmıştır. Belirlenen kritik dönemlere ait finansal veriler ve sosyal medyadan toplanan metin tabanlı veriler kullanılmıştır. Finansal veriler üzerinde, zaman serisi tahmin modelleri ve danışmanlı öğrenme modelleri uygulanarak altın getirileri ile ilgili tahminler yapılmıştır. Sosyal medya platformu Twitter üzerinden aldığımız metin tabanlı veriler üzerinde konu modellemesi ve duygu analizi yöntemleri kullanılarak altın getiri tahmini yapılmıştır. Altın getirilerini tahmin etmek için duygu analizinin tek başına kullanılması tavsiye edilmez. Daha doğru ve güvenilir altın getiri tahmini için enflasyon verileri, faiz oranları, arz ve talep gibi ek temel teknik unsurlar da dikkate alınmalıdır.

6.1. Çalışmanın Kısıtları

Altın fiyatlarını etkileyen çok sayıda faktör var. Ülkeler arası ilişkiler, siyasi iklim, ekonomik krizler, faiz oranları, sosyal yaşamdaki çeşitli gelişmeler, döviz kurları vs. altın üzerinde olumlu ya da olumsuz etki oluşturabilmektedir. Altın fiyatlarında isabetli tahminler için tüm faktörler dikkate alınarak bir veri seti hazırlanmalı. Daha büyük veri setleri kullanılarak daha fazla modelleme ve tahmin yöntemi ile çalışmalar yapılmalıdır.

6.2. Gelecek Çalışmalar

Altın için yapılan gelecek fiyat tahminlerinin daha güvenilir olabilmesi için veri setinin zenginleştirilmesi ve daha büyük veri setleri üzerinde çalışma yapılması gerekmektedir. Bu nedenle altın üzerinde etki oluşturabilecek her konuda büyük bir veri seti oluşturulmalı. Veri setleri üzerindeki çalışmaların sağlıklı sonuçlar verebilmesi için daha fazla sınıflandırma modeli ve analiz/tahmin yöntemi kullanılmalıdır. Bu sayede her bir veri seti için en güvenilir sonucu veren yöntemler belirlenebilir. Gelecek çalışmalarda siyasi, ekonomik, sosyal yaşamdaki gelişmeler de dikkate alınarak daha fazla platformdan ve farklı kaynaklardan metin tabanlı veri toplayarak, faiz oranlarının, enflasyon ve

istihdam verilerinin kullanılarak veri seti oluşturulması planlanmaktadır. Daha güvenilir sonuçlar elde etmek için oluşturulacak bu zengin veri seti üzerinde daha fazla sınıflandırma ve analiz yöntemi ile tahminler yapılması planlanmaktadır.



KAYNAKLAR

- A.G. Malliaris ve Mary Malliaris. (2015), What drives gold returns? *A decision tree analysis, Finance Research Letters*, 13, 45–53.
- Akgul, H.G. (2019), Fake news as a tool of populism in turkey: The pastor andrew brunson case. *Polish Political Science Review* 7(2), 32–51
- Albawi, S., Mohammed, T.A., Al-Zawi, S. (2017), Understanding of a convolutional neural network. *In: 2017 international conference on engineering and technology (ICET)*. pp. 1–6. IEEE
- Al Amrani, Y., Lazaar, M., & El Kadiri, K. E. (2018). Random Forest and support vector machine based hybrid approach to sentiment analysis. *Procedia Computer Science*, 127, 511-520.
- Baur, D. G. ve G. Löffler (2013), Predicting the Equity Premium with the Demand for Gold Coins and Bars, *FIRN Research Paper*
- Benlagha, N., El Omari, S. (2022), Connectedness of stock markets with gold and oil: New evidence from covid-19 pandemic. *Finance Research Letters* 46, 102373
- Brian M. Lucey, Susan Sunila Sharma and Samuel A. Vigne. (2017), Gold and inflation(s) A time-varying relationship, *Economic Modelling*, 67, 88–101.
- Bruns, A., & Stieglitz, S. (2013). Towards more systematic Twitter analysis: metrics for tweeting activities. *International Journal of Social Research Methodology*, 16(2), 91-108.
- Boser, B.E., Guyon, I.M., Vapnik, V.N. (1992), A training algorithm for optimal margin classifiers. *In: Proceedings of the fifth annual workshop on Computational learning theory*. pp. 144–152
- Chen, Y.C., Huang, W.C. (2021), Constructing a stock-price forecast cnn model with gold and crude oil indicators. *Applied Soft Computing* 112, 107760
- C. Wen, S. Liu, X. Yao, L. Peng, X. Li, Y. Hu, T. Chi. (2019), A novel spatiotemporal convolutional long short-term neural network for air pollution prediction, *Science of The Total Environment*, 654,1091-1099.
- Clark, K., Luong, M.T., Le, Q.V., Manning, C.D. (2020), Electra: Pre-training text encoders as discriminators rather than generators. *arXiv preprint arXiv:2003.10555*
- 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*
- Ding, X., Wang, M. (2022), The effectiveness of foreign exchange intervention: Empirical evidence from vietnam. *The Journal of Asian Finance, Economics and Business* 9(2), 37–47

- Ekinci, E., Omurca, S.I. (2020), Net-lda: a novel topic modeling method based on semantic document similarity. *Turkish Journal of Electrical Engineering & Computer Sciences* 28(4), 2244–2260
- Lee, D.D., Seung, H.S. (1999), Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755), 788–791
- Liu, Y. et al. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*.
- Ramchoun, H., Ghanou, Y., Ettaouil, M., Janati Idrissi, M.A. (2016), Multilayer perceptron: Architecture optimization and training
- URL-1: https://huggingface.co/docs/transformers/model_doc/distilbert (Ziyaret Tarihi: 20 Nisan 2023)
- URL-2: <https://covid19.saglik.gov.tr/> (Ziyaret Tarihi: 20 Nisan 2023)
- Xu, H., Van Durme, B., & Murray, K. (2021). Bert, mbert, or bibert? a study on contextualized embeddings for neural machine translation. *arXiv preprint arXiv:2109.04588*.
- Wang, H., Can, D., Kazemzadeh, A., Bar, F., & Narayanan, S. (2012, July). A system for real-time twitter sentiment analysis of 2012 us presidential election cycle. In *Proceedings of the ACL 2012 System Demonstrations. Association for Computational Linguistics*. (pp. 115-120).
- Wang, Yu Shan and Chueh, Yen Ling (2013), “Dynamic Transmission Effects Between The Interest Rate, The US Dollar, And Gold And Crude Oil Prices”, *Economic Modelling*, 30, 792–798.
- Vapnik, V.N. (1998), Adaptive and learning systems for signal processing communications, and control. *Statistical learning theory*
- Yavuz, M.H., Balci, B. (2018), Turkey’s July 15th Coup: What Happened and Why. *University of Utah Press Salt Lake City*
- Zhang, Yue-Jun and Wei, Yi-Ming. (2010), “The Crude Oil Market and The Gold Market: Evidence for Cointegration, Causality and Price Discovery”, *Resources Policy*, 35, 168–177.

KİŞİSEL YAYIN VE ESERLER

Yavuz, A., Eken, S., Sayar, A. (2016). Akademik Çalışmalara Hakem Ataması için Büyük Veri Altyapısı, *10. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS'16)*

Yavuz, A., Eken, S. (2023). Gold Returns Prediction: Assessment based on Major Events, *EAI Endorsed Transactions on Scalable Information Systems, 10(5)*, 1-10.



ÖZGEÇMİŞ

Ahmet YAVUZ; 2012 yılında başladığı Kocaeli Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü'nden 2016 yılında mezun oldu. 2021 yılında Kocaeli Üniversitesi Bilişim Sistemleri Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimine başladı. Şuanda bir bankada yazılım uzmanı olarak görev yapmaktadır.

