



T.C.
HATAY MUSTAFA KEMAL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKÂ YÖNTEMLERİYLE BORSA ENDEKSİNİN
YÖNÜNÜN TAHMİNİ ÜZERİNE BİR ÇALIŞMA:
KARŞILAŞTIRMALI ANALİZ

MUSTAFA YILDIZ

ENFORMATİK ANABİLİM DALI

YÜKSEK LİSANS TEZİ

HATAY
TEMMUZ - 2022



T.C.
HATAY MUSTAFA KEMAL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YAPAY ZEKÂ YÖNTEMLERİYLE BORSA ENDEKSİNİN
YÖNÜNÜN TAHMİNİ ÜZERİNE BİR ÇALIŞMA:
KARŞILAŞTIRMALI ANALİZ

MUSTAFA YILDIZ
ORCID:0000-0001-8230-868X

ENFORMATİK ANABİLİM DALI

YÜKSEK LİSANS TEZİ

Danışman
Dr. Öğr. Üyesi OĞUZ KILIÇOĞLU
ORCID: 0000-0001-7358-8034

HATAY
TEMMUZ - 2022

27/07/2022

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını ve tez üzerinde Yükseköğretim Kurulu tarafından hiçbir değişiklik yapılamayacağı için tezin bilgisayar ekranında görüntülendiğinde asıl nüsha ile aynı olması sorumluluğunun tarafıma ait olduğunu beyan ederim.

İmzası

MUSTAFA YILDIZ

ÖZET

YAPAY ZEKÂ YÖNTEMLERİYLE BORSA ENDEKSİNİN YÖNÜNÜN TAHMİNİ ÜZERİNE BİR ÇALIŞMA: KARŞILAŞTIRMALI ANALİZ

Bu çalışmada, Makine öğrenmesinin alt dallarından, Regresyon Analizi ve Destek vektör makineleri ile Teknik analiz göstergeleri kullanılarak Bist100 endeksinin yönü tahmin edilmeye çalışılacak ve bulunan sonuçlar karşılaştırılacaktır. Çalışmanın sonunda Bist100 endeksinin tahmini için en iyi yöntem ve tekniğin hangisi olduğuna karar verilecektir.

Uygulama aşamasında Bist100 endeksinin günlük ve 60 dakikalık zaman serisi veri setleri kullanılarak, Logistic Regresyon, Lasso Regresyon, Ridge Regresyon ve Destek Vektör Makinesi tahmin modellerinin her biri ile uygulamalar yapıp simüle edilecek, elde edilen sonuçlar her bir model için ayrı ayrı değerlendirilecektir. En iyi sonuç veren model, teknik analiz göstergeleri ile birlikte kullanılarak Bist100 endeks tahmini için en iyi yöntem elde edilmeye çalışılacaktır.

2022, 60 sayfa

Anahtar Kelimeler: Yapay zekâ, makine öğrenmesi, borsa tahmini, regresyon analizi, destek vektör makinesi

ABSTRACT

A STUDY on THE PREDICTION of THE DIRECTION of THE STOCK MARKET INDEX with ARTIFICIAL INTELLIGENCE METHODS: COMPARATIVE ANALYSIS

In this study, the direction of the Bist100 index will be tried to be estimated by using Regression Analysis and Support vector machines, which are sub-branches of Machine learning, and Technical analysis indicators and the results will be compared. At the end of the study, it will be decided which is the best method and technique for the estimation of the Bist100 index.

During the implementation phase, applications will be made and simulated with each of the Logistic Regression, Lasso Regression, Ridge Regression and Support Vector Machine prediction models using the daily and 60-minute time series data sets of the Bist100 index, and the results will be evaluated separately for each model. The model with the best results will be used together with technical analysis indicators to try to obtain the best method for Bist100 index estimation.

2022, 60 pages

Keywords: Artificial intelligence, machine learning, stock market prediction, regression analysis, support vector machine

TEŞEKKÜR

Yüksek Lisans eğitimim süresince değerli vaktini bana ayırıp bilgilerini paylaşan, Yüksek Lisans tez sürecinde, konusunun belirlenmesi, araştırılması ve yazımı sırasında sahip olduğu bilgi birikimi ve tecrübesi ile çalışmayı yönlendiren ve her türlü yardımı esirgemeyen danışman hocam Dr. Öğretim Üyesi Oğuz Kılıçoğlu'na sonsuz saygı ve teşekkürlerimi sunarım.

Yüksek Lisans eğitimim boyunca ilgi ve bilgilerinden yararlandığım kıymetli hocalarım, Prof. Dr. Bünyamin Yıldız, Dr. Öğretim Üyesi Hakan Yetişkin ve Dr. Öğretim Üyesi Fatih Balaman'a teşekkür ederim.

Bununla birlikte her zaman yanımda olup beni sonuna kadar destekleyen aileme, sonsuz şükranlarımı sunarım.

İÇİNDEKİLER

ÖZET.....	I
ABSTRACT.....	II
TEŞEKKÜR.....	III
İÇİNDEKİLER	IV
ÇİZELGELER DİZİNİ	VI
ŞEKİLLER DİZİNİ.....	VII
KISALTMALAR DİZİNİ.....	VIII
1. GİRİŞ.....	1
1.1. Bist100 Endeksi.....	2
1.2. Zaman Serileri	2
1.2.1. Zaman Serilerinin Tahmininde Girdiler	2
1.3. Yapay Zekâ.....	3
1.3.1. Makine Öğrenimi.....	3
1.3.2. Makine Öğrenimi Yöntemleri	3
1.3.2.1 Denetimli Öğrenme Yöntemi.....	4
1.3.2.2 Denetimsiz Öğrenme Yöntemi.....	4
1.3.2.3 Yarı Denetimli Öğrenme Yöntemi.....	4
1.3.2.4 Pekiştirmeli Öğrenme Yöntemi	4
1.3.3. Makine Öğrenmesinin Uygulama Adımları	5
1.4. Regresyon Analizi	5
1.4.1. Doğrusal Regresyon (Basit Lineer Regresyon)	6
1.4.1.1 Artıklar	7
1.4.1.2 Parametre Tahmini	8
1.4.1.3 En Küçük Kareler Yöntemi (Least Square Method).....	8
1.4.1.4 Basit Doğrusal Regresyon Modeli Varsayımları	8
1.4.1.5 Hata Terimi (ϵ_i) Varsayımları	9
1.4.1.6 Çoklu Bağlantı Problemi.....	9
1.4.1.7 Düzenleştirme	9
1.4.2. Ridge Regresyon Yöntemi	10
1.4.3. Lasso Regresyon.....	11
1.4.4. Lojistik Regresyon.....	12
1.4.5. Destek Vektör Makineleri (Support Vector Machine - SVM)	14
1.4.5.1 Verilerin Doğrusal Ayrılabilirlik Durumu	15
1.4.5.2 Lineer Ayrılabilirlik Durumu	16
1.5. Teknik Analiz Göstergeleri (İndikatörler).....	16
1.5.1. Genel Bilgiler	16
1.5.2. İndikatörleri Alım Satım Noktalarını Nasıl Belirlerler.....	18
1.5.3. ADX İndikatörü (Average Directional Movement Index)	18
1.5.4. Aroon İndikatörü	20
1.5.5. Bollinger Bandı	21
1.5.6. CCI İndikatörü.....	23
1.5.7. DEMA İndikatörü (Double Exponential Moving Average).....	24
1.5.8. Envelopes İndikatörü.....	25
1.5.9. Hareketli Ortalamalar	26
1.5.10. MACD İndikatörü	28
1.5.11. Momentum İndikatörü.....	29

1.5.12. Parabolic Sar İndikatörü.....	30
1.5.13. RSI İndikatörü (Relative Strength Index)	31
1.5.14. Stochastic İndikatörü.....	32
1.5.15. TRIX İndikatörü.....	34
2. ÖNCEKİ ÇALIŞMALAR.....	36
2.1.1. Sadece Yapay Zeka Yöntemleriyle Yapılan Çalışmalar	36
2.1.2. Yapay Zeka Yöntemlerinin Teknik Analiz Göstergeleriyle Birlikte Kullanıldığı Çalışmalar	37
3. MATERYAL ve YÖNTEM.....	40
3.1. Materyal.....	40
3.2. Yöntem	42
4. ARAŞTIRMA BULGULARI ve TARTIŞMA	44
4.1. Günlük Periyotta Simülasyon Bulguları.....	44
4.1.1. Logistic Regresyon Günlük Periyotta Simülasyon Bulguları.....	44
4.1.2. Lasso Regresyon Günlük Periyotta Simülasyon Bulguları	45
4.1.3. Ridge Regresyon Günlük Periyotta Simülasyon Bulguları	46
4.1.4. Destek Vektör Makinesi Günlük Periyotta Simülasyon Bulguları.....	47
4.2. 60 Dakikalık Periyotta Simülasyon Bulguları	48
4.2.1. Logistic Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları.....	48
4.2.2. Lasso Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları	49
4.2.3. Ridge Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları	50
4.2.4. Support Vector Machine 60 Dakikalık Periyotta Simülasyon Bulguları ..	51
4.3. Bulguların Karşılaştırılması.....	52
4.3.1. Günlük Periyotlardaki Bulguların Karşılaştırılması	52
4.3.2. 60 Dakikalık Periyotlardaki Bulguların Karşılaştırılması	52
4.3.3. Tartışma	53
4.3.3.1 En Başarılı Yöntemin Teknik Analiz Göstergeleriyle Kullanılması	54
4.3.4. Bulgular	54
5. SONUÇ ve ÖNERİLER.....	56
5.1. Önerilen Model.....	56
KAYNAKLAR	58

ÇİZELGELER DİZİNİ

Çizelge 4.1. Logistic Regression XU100_Day simülasyon bulguları.....	44
Çizelge 4.2. Lasso RegressionXU100_Day simülasyon bulguları	45
Çizelge 4.3. Ridge RegressionXU100_Day simülasyon bulguları	46
Çizelge 4.4. Support Vector Machine_XU100_Day simülasyon bulguları.....	47
Çizelge 4.5. Logistic Regression_XU100_60 Minute simülasyon bulguları.....	48
Çizelge 4.6. Lasso Regression_XU100_60Minute Simülasyon bulguları.....	49
Çizelge 4.7. Ridge Regression_XU100_60 dakikalık simülasyon bulguları.....	50
Çizelge 4.8. Support Vector Machine_XU100_60 Dakikalık simülasyon bulguları.	51
Çizelge 4.9. Günlük periyotta simülasyon bulgularının karşılaştırılması	52
Çizelge 4.10. 60 dakikalık periyotta simülasyon bulgularının karşılaştırılması	53
Çizelge 4.11. Teknik Analiz göstergeleriyle kullanılan al-sat robotlarının simülasyon bulguları	54
Çizelge 5.12. Lasso ile birlikte Kullanılan Göstergelerin Sonuçları.....	57

ŞEKİLLER DİZİNİ

Şekil 1.1. Lineer ayrılabilen veriler (Irmak, 2019)	15
Şekil 1.2. Giriş uzayını özellik uzayına taşıma (Eray, 2008).....	16
Şekil 1.3. +DI ve –DI değerleri.....	19
Şekil 1.4. ADX Göstergesi.....	19
Şekil 1.5. Aroon İndikatörü.....	20
Şekil 1.6. Bollinger Bandı.....	22
Şekil 1.7. CCI İndikatörü	23
Şekil 1.8. Dema İndikatörü	25
Şekil 1.9. Envelopes İndikatörü	26
Şekil 1.10. MACD İndikatörü.....	28
Şekil 1.11. Momentum İndikatörü	29
Şekil 1.12. Parabolic Sar İndikatörü	31
Şekil 1.13. RSI İndikatörü.....	31
Şekil 1.14. Stochastic Slow İndikatörü	33
Şekil 1.15. Trix İndikatörü	34
Şekil 3.1. Günlük veri setinden örnek bir kısım.....	40
Şekil 3.2. Günlük periyottaki test veri setinin grafiği	41
Şekil 3.3. 60 dakikalık veri setinden örnek bir kısım.....	41
Şekil 3.4. 60 dakikalık periyotta test veri setinin grafiği	42
Şekil 4.1. Logistic Regresyon XU100_Day yöne göre toplam getiri yüzdesi	44
Şekil 4.2. Lasso RegresyonXU100_Day yöne göre toplam getiri yüzdesi.....	45
Şekil 4.3. Ridge RegresyonXU100_Day yöne göre toplam getiri yüzdesi.....	46
Şekil 4.4. Support Vector Machine_XU100_Day yöne göre toplam getiri yüzdesi..	47
Şekil 4.5. Logistic Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi .	48
Şekil 4.6. Lasso Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi	49
Şekil 4.7. Ridge Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi	50
Şekil 4.8. Support Vector Machine_XU100_60Minute yöne göre toplam getiri yüzdesi.....	51

KISALTMALAR DİZİNİ

KISALTMALAR

ML	: Makine Öğrenmesi
HKT	: Hata Kareler Toplamı
AKT	: Artık Kareler Toplamı
EKK	: En Küçük Kareler Yöntemi
DVM	: Destek Vektör Makinesi
SVM	: Support Vector Machine (Destek Vektör Makinesi)
MA	: Moving Average (Hareketli Ortalama)
SMA	: Simple Moving Average (Basit Hareketli Ortalama)
EMA	: Exponential Moving Average (Üssel Hareketli Ortalama)
WMA	: Weighted Moving Average (Ağırlıklı Hareketli Ortalama)
YSA	: Yapay Sinir Ağları

1. GİRİŞ

Borsalar en basit anlatımla kıymetlerin alınıp satıldığı pazarlardır. Borsalarda denetleyiciler vardır ve alım satım işlemleri belirli kurallar içerisinde olur. Borsalarda gerçekleşen fiyat hareketleri çok oynaktır ve fazlaca belirsizlik içerir. Bu nedenle borsa endekslerinde gerçekleşebilecek fiyatlamaların modellenmesi için henüz mutlak bir yol bulunamamıştır (Kutlu ve Badur, 2009). Bununla birlikte borsada herhangi bir finansal enstrüman için gelecekte oluşabilecek fiyatın değerini ve yönünü tahmin etmek amacıyla çeşitli yöntemler kullanılmıştır. Yapay zekâ sistemleri yüksek belirsizlik durumunda model kurma ve karar verme yetenekleri sayesinde finansal enstrümanların gelecekteki değerlerinin tahmini uygulamalarında sıklıkla kullanılmaktadır.

Çalışma içerisinde yapay zekâ ve yapay zekâ uygulamalarının alt dallarından biri olan makine öğrenmesi yöntemleri kapsamlı bir biçimde incelenecektir. Bunun için öncelikle Regresyon Analizi ve Destek Vektör Makineleri hakkında bilgi verilecektir. Ayrıca bu yöntemlerle birlikte borsanın geleneksel tahmin yöntemlerinden Teknik Analiz göstergelerinin bazıları kullanılarak “Bist100 endeksinin tahmini için en iyi model hangisi olabilir?” sorusuna cevap aranacak ve çalışma sonunda model veya modeller önerilecektir.

Bu çalışmanın bir önemi de sadece günlük değil 60 dakikalık periyotlarda da işlem yapılması ve günlük veriler ile sonuçlarının karşılaştırılmasıdır. Bundan önce incelediğimiz çoğu çalışmada zaman periyodu olarak günlük veriler alınmıştır.

Çalışmada öncelikle yapay zekâ ve makine öğrenmesi arasındaki ilişkiler irdelenmiştir. Daha sonra makine öğrenmesi ve teknikleri incelenmiştir. Support Vector Machine ve Regresyon Analizi hakkında bilgi verilmiş ve Regresyon Analizi tekniğinin en çok bilinen tahmin modellerinden, Logistic Regresyon, Lasso Regresyonu, Ridge Regresyon modellerinin her birinin arka planda çözdüğü problemin yapısına örnek olacak teorik bilgiler sunulmuştur. Ayrıca borsanın geleneksel tahmin yöntemlerinden Teknik Analiz yöntemi anlatılmış ve çalışmamızda kullandığımız Teknik Analiz göstergeleri hakkında bilgiler verilmiştir.

1.1. BİST100 Endeksi

Çalışmada tahminini yapmaya çalışacağımız BİST100 Endeksi; İstanbul Menkul Kıymetler Borsası'nda piyasa değeri ve gün içinde gerçekleşen işlem hacmi açısından en büyük 100 hisse senedinden oluşan bir endekstir. BİST100 endeksi finansal piyasalarda XU100 sembolü ile ifade edilmektedir.

BİST100 endeksini oluşturan hisselerin durumu tüm borsayı etkilemektedir. Borsa ile ilgili düştü, yükseldi, kararsız vs. şeklinde yapılan tüm yorumlar BİST100 endeksi üzerinden yapılmaktadır. Bu nedenle BİST100 endeksi sadece Borsa İstanbul için değil, Türkiye'deki mevcut finansal piyasalar içinde bir gösterge niteliği taşımaktadır.

1.2. Zaman Serileri

Tahminini yapmaya çalışacağımız BİST100 endeksi matematiksel olarak bir zaman serisi belirtir. Bu veriler belli bir sırayla, ard arda ve eşit zaman aralıkları içerisinde (saniye, dakika, vb.) dizilmişlerdir.

1.2.1. Zaman Serilerinin Tahmininde Girdiler

Bir zaman serisi veri setini tahmin etmek için kullanılan değişkenlerin hangisi olacağına karar vermek modelin performansı açısından son derece önemlidir. Analizimize konu olan BİST100 endeks verilerinin değerini tahmin etmek için kullanılabilecek pek çok değişken bulunmaktadır. Bu değişkenler; tahmine konu olan endeksin açılış fiyatının değeri, kapanış fiyatının değeri, gün içi en yüksek değeri, gün içi en düşük değeri, piyasada oluşan hacim miktarı, endeksin ortalaması, endekse ait göstergelerin değeri vb. olabilir. Bunun yanında döviz kurları, günlük faiz, ekonomik ortam, ekonomik istatistikler, politik kararlar, yurtdışı borsaları vb. değişkenler de tahmin modelinde kullanılabilir.

İnce ve Sönmez Çakır (2017) tarafından bildirildiğine göre, Lawrence (1997) yaptığı çalışmada, Johannesburg Menkul kıymetler borsasının performansını modellemeye çalışmış ve bu çalışma için belirlenebilecek gösterge sayısının 63 tane olduğunu belirlemiştir. Ancak Lawrence, yaptığı uygulamalarda bu göstergelerin tamamının kullanılmasına gerek olmadığı sonucuna ulaşmıştır.

Biz bu çalışmada, BİST100 endeksinin günlük ve saatlik periyotlardaki açılış değeri, kapanış değeri, periyot içerisinde oluşan en yüksek değeri ve periyot içerisinde oluşan en düşük değeri ile popüler teknik analiz göstergelerinden bazılarının değerlerini girdi olarak kullanacağız.

1.3. Yapay Zekâ

Genel olarak yapay zekânın tanımı yoktur. Ancak yapay zekâ için insan beyninin, düşünme, hatırlama, karşılaştırma değerlendirme ve karar verme gibi bilişsel fonksiyonlarına sahip sistemleri modelleme çalışmasının genel adıdır diyebiliriz.

Yapay Zekâ da temel amaç, insanların zorlanarak yaptığı işleri yapabilecek sistemler üretmektir.

1.3.1. Makine Öğrenimi

Makine öğrenmesi (ML) için genel olarak, makinelerin verileri analiz ederek insanların öğrenme yöntemlerini taklid edip bunun için çeşitli algoritma ve teknikler geliştiren bir bilim dalıdır diyebiliriz. Bu yöntemde bilgisayar komutlar kullanmadan tahminler veya kararlar vermek için bazı matematiksel hesaplamalar ve modellemelerden faydalanır. Makine öğrenmesi(ML), yapay zeka yöntemlerinin alt dallarından bir tanesidir.

Makine öğrenmesinde öğrenim örneklerden gerçekleşir. Bu yöntemde, algoritmalar klasik yazılım geliştirmedeki gibi kodlanmaz. Örneğin bir insanı bilgisayara tanıtmak için insanı tanımlamak yerine bilgisayara milyonlarca insan resmi tanımlanır. Makine öğrenmesinde kullanılan algoritma bazı hesaplamalar ve modellemeler kullanarak bu veriler ile insanı tanımlamak için kalıplar oluşturur. Daha sonra verilerden oluşturduğu kalıpları kullanarak tahminde bulunan bir model elde edilmiş olur. Böylece yinelenen kalıplardan algoritma insanı ayırt edebilmektedir (Aytekin, 2021).

1.3.2. Makine Öğrenimi Yöntemleri

Makine öğrenme yöntemlerini, dört başlık altında inceleyebiliriz.

1.3.2.1 Denetimli Öğrenme Yöntemi

Bu öğrenme çeşidinde eğitim verilerinin tamamının sonucu belli yani etiketlidir. Denetimli öğrenmede, etiketli veriler kullanılarak bir fonksiyon oluşturulur. Eğitim verileri ile makine eğitilir ve makinenin tahmin için kullanacağı modelin kurulması sağlanır.

Denetimli öğrenmede algoritmasının başarısı etiketli eğitim verilerinin miktarı ile değişebilmektedir. Algoritmaya ne kadar çok örnek veri sağlanırsa başarısı o denli artacaktır. Ancak bu durum denetimli öğrenmenin önemli bir sorunudur. Zira fazla miktarda eğitim verisi elde etmek ve bunları işlemek hem zor hem de zaman alıcı olabilir. (Aytekin, 2021)

1.3.2.2 Denetimsiz Öğrenme Yöntemi

Bu öğrenme yönteminde veriler etiketli değildir. Makine öğrenme algoritması etiketli olmayan veriler kullanılarak çıkarımlarda bulunur. Denetimsiz öğrenmede eğitim verilerinin sonucu bilinmediği için model denetlenmez. Model etiketlenmemiş verilerden gizli yapılar veya kalıpları açıklamak için çıkarımlar yapar.

Denetimsiz öğrenmenin en büyük sorunu sonucun genellikle tahmin edilemez olmasıdır. Bu nedenle genellikle algoritma doğru çıkarımlar yapana kadar geliştirici tarafından yönlendirilir.

1.3.2.3 Yarı Denetimli Öğrenme Yöntemi

Bu öğrenme türünde eğitim verisi, etiketli ve etiketsiz eğitim verisinden oluşur. Bir öğrenme algoritmasını eğitmek için yeterli etiketlenmiş veri yoksa yarı denetimli öğrenme etkili olacaktır. Eğitim sırasında, daha geniş, etiketsiz bir veri kümesinden daha küçük, etiketli bir veri kümesi kullanır.

1.3.2.4 Pekiştirmeli Öğrenme Yöntemi

Bu modelde yapay zekâ, problemi çözmek için deneme yanılma yöntemini kullanır. Ancak tıpkı oyun oynar gibi her bir çıkarımda ödül ya da ceza alır. Burada makinenin görevi oyunda alacağı toplam ödülü en yüksek düzeye çıkarmaktır.

Geliştirici ödül ve cezayı belirler ancak bu modelde oyunun nasıl çözüleceğine dair bilgi ya da ipucu vermez.

1.3.3. Makine Öğrenmesinin Uygulama Adımları

Genel olarak sorunları çözmek için kullanılan bir makine öğrenme süreci aşağıdaki gibi işler.

a. Verilerin Toplanması ve Hazırlanması

Mevcut verilerin türlerine göre hangi makine algoritmasını kullanacağımız belirlenir. Bazen verileri işlemek için dönüştürmek gerekebilir. Verilerde anormallikler, eksiklikler varsa düzenlenir ve veri bütünlüğü sorunu çözülür.

b. Modelin Eğitilmesi

Hazırladığımız verileri eğitim veri kümesi ve test veri kümesi şeklinde iki gruba ayırırız. Burada eğitim kümesi test kümesinden büyük olmalıdır. Modelin eğitilmesi için eğitim veri kümesindeki veriler kullanılır.

c. Modeli Doğrulama

Eğitilen modelin performansını ve doğruluğunu değerlendirmek için test kümesi kullanılır.

d. Sonuçlardan Çıkarımda Bulunma

Sonucu inceleyip modelin performansı incelenir ve çıkarımlarda bulunulur (azure.microsoft.com/tr, 2022).

1.4. Regresyon Analizi

Regresyon analizinin ekonomi, sosyal bilimler, bilişim ve tıp gibi daha birçok uygulama alanı vardır. Regresyon analizi, birden fazla değişken arasında var olan bir ilişkiyi tahmin edebilmek için matematiksel hesaplamalar kullanarak açıklamaya çalışan genel olarak istatistik temelli bir analiz yöntemidir (Vural, 2007).

Regresyon genellikle ilişkileri nicel olarak tanımlar. Bunun için önce bağımlı ve bağımsız değişkenler arasında var olan matematiksel model bulunur. Bu model yardımıyla bağımlı olan değişken değeri tahmin edilebilir. Buna örnek olarak doğrusal regresyon verilebilir.

Doğrusal regresyon, regresyon analizinin en çok kullanılan yöntemidir. Bu yöntemde tahminler yapabilmek için veriye en çok uyan doğruyu bulmak esastır. Bu doğruyu bulabilmek için ilerde de bahsedeceğimiz en küçük kareler yöntemi adı verilen matematiksel hesaplamalardan faydalanılır.

Verilerin türü ve yapısına göre kullanılabilecek birden çok regresyon tekniği vardır. Regresyon analizi yaparken verilerin türü ve yapısı dikkate alınıp doğru regresyon modeli kullanılmalıdır. Aksi taktirde modelin başarısı beklenen değerin altında olacaktır (Arı ve Önder, 2013).

Regresyon analizinin genel olarak iki türlü kullanım şekli vardır. Bunlardan ilki tahmin yapmak için kullanımıdır. Buradaki kullanımı makine öğrenme alanı ile önemli bir ölçüde örtüşür. İkinci olarak bazen bağımlı değişken ile bağımsız değişkenin birbirleriyle olan ilişkisini bulmak için regresyon analizi kullanılabilir.

1.4.1. Doğrusal Regresyon (Basit Lineer Regresyon)

Doğrusal regresyon birden çok değişkenin birbirleriyle var olan ilişkisini bir doğru yardımıyla modelleme yöntemidir. Burada tahmin edilmek istenen değişken bağımlı değişken, tahmin için kullanılan değişken ise bağımsız değişken olarak adlandırılır. Basit lineer regresyonda ana düşünce bağımlı değişken ile bağımsız değişkenin birbirleriyle bağlantısını anlatan bir doğru denklemi elde etmektir (Nacar ve Erdebilli, 2021).

Lineer Regresyon değişken sayısına göre sınıflandırılır. Eğer regresyon modeli kurulurken sadece bir tane bağımsız değişken kullanılıyorsa basit lineer regresyon modeli birden çok sayıda değişken kullanılıyorsa bu regresyona çoklu lineer regresyon adı verilir.

Basit lineer regresyon Formül (1.1) de gibi modellenir.

$$Y = \beta_0 + \beta_1 X_{1i} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.1)$$

Burada Y bağımlı değişkendir. X_1 bağımsız değişken ve β_0 ile β_1 bu bağımsız değişkenlerin parametreleridir. ε_i denklemin hata terimidir.

Eğer bağımsız değişken sayısı birden çoksa bu durumda, p adet bağımsız değişken ve n adet gözlem değeri için çoklu lineer regresyon Formül (1.2) deki gibi modellenmiştir (Arı ve Önder, 2013).

$$Y = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1.2)$$

Burada ε_i hata terimi, regresyon doğrusunun i 'inci gözlem değeri ile arasındaki mesafedir. ε_i 'nin ortalaması 0 olarak alınır. ε_i 'nin varyansının da σ^2 olarak normal dağılıma sahip olduğu varsayılır.

Doğrusal regresyon modelinde varsayım olarak X_i 'lerin ölçümünün hatasız olduğu, Y 'nin ise bir miktar hata (ε_i) ile ölçüldüğü varsayılır.

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (1.3)$$

Y 'nin hedef değişken olduğu ve yukarıdaki gibi ifade edilen doğrusal regresyon modelinde β_1 katsayısı doğrunun eğimini verecektir. β_0 katsayısı ise y eksenini kestiği noktadır. ε doğrunun hata terimi olarak ifade edilir.

Regresyon modeli tahmin edildiğinde ise Y 'nin tahmin denklemi şu şekilde gösterilir.

$$\hat{Y} = E(\hat{Y}) = \hat{\beta}_0 + \hat{\beta}_1 X \quad (1.4)$$

Bu denklemde $\hat{\beta}_0$, y eksenini kesim noktasının (β_0) tahmincisi, $\hat{\beta}_1$, doğrunun eğiminin (β_1) tahmincisidir.

1.4.1.1 Artıklar

Regresyon modelinde tahmin edilen değer ile gerçek değer arasındaki uzaklığa (residual) artık değeri denir. Kurulan regresyon modelinde i . artık;

$$\varepsilon_i = y_i - \hat{y}_i \quad (1.5)$$

Formülü ile hesaplanır. Başarılı tahminler yapan bir modelin teorik olarak beklenen değerinin $E[\varepsilon_i] = 0$ olması gerekir. Beklenen değer E harfi ile gösterilir.

Aşağıda formül (1.6)'da ise Artık (Hata) Kareler Toplamı (HKT veya AKT) verilmiştir.

$$AKT = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N \varepsilon_i^2 \quad (1.6)$$

Bir modelin başarılı tahminlerde bulunabilmesi için hata kareler toplamının küçük olması gerekir. Hata kareler toplamının değerinin düşük çıkması modelin tahmin ettiği $\hat{y}$ değeri ile gerçekte olan y değerinin farklarının düşük çıkması yani tahmin değerinin gerçek değere oldukça yakın olduğu anlamına gelir. Hata kareler toplamı için bir modelin veriyi açıklama ölçüsüdür dersek yanlış olmaz. Hata kareler toplamına aynı zamanda tahmin modelinin maliyet fonksiyonu da denir. Burada temel

felsefe maliyet fonksiyonunu minimum yapan en uygun parametrelerin bulunması üzerine kurgulanmıştır. Uygun parametreleri bulma işlemi bir optimizasyon problemidir.

1.4.1.2 Parametre Tahmini

Regresyon analizinde hedef; değişkenler arasındaki ilişkinin matematiksel biçimin tespit edilmesidir. Çoğu durumda bu ilişkinin bilinmesi kesin olarak mümkün olmadığı için ilişki yaklaşık olarak tahmin edilmeye çalışılır. Gözlenen veri yardımıyla bunların tahmin edilmesi gerekir. Parametre tahmini için gerek hesaplama kolaylığından ötürü gerekse doğruya en yakın sonuç vermesinden dolayı en küçük kareler yöntemi en çok kullanılan regresyon tahmin yöntemidir.

1.4.1.3 En Küçük Kareler Yöntemi (Least Square Method)

Bu yöntem lineer regresyonun temel kavramıdır denilebilir. Bu yöntemi ünlü matematikçi C. F. Gauss bulmuştur. En küçük kareler yöntemi hata kareler toplamını minimum yapmak üzerine kuruludur. Bu yöntemde hata kareler toplamı fonksiyonunu minimum yapmak için hata kareler toplamının türevi alınır ve sıfıra eşitlenir. Bu yöntem yardımıyla $\hat{\beta}_0$ ve $\hat{\beta}_1$ katsayıları tahmin edilir.

Her bir tahmin değerinin artık kareler toplamı hangi modelde en küçük ise o model seçilir. En küçük kareler yönteminin amacı hatayı minimize eden modelin seçimidir. $\hat{\beta}_0$ ve $\hat{\beta}_1$ katsayıları aşağıdaki denklemlerde gösterilmiştir.

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1.7)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (1.8)$$

Burada $\bar{x}$, x gözlem değerinin ortalaması, $\bar{y}$, y gözlem değerinin ortalamasıdır.

1.4.1.4 Basit Doğrusal Regresyon Modeli Varsayımları

Basit doğrusal regresyonun tahminlerinin güvenilir ve başarılı olması için aşağıdaki varsayımların doğruluğu gerekmektedir

1. X_i 'lerin (bağımsız değişkenler) değerinin sabit olduğu varsayılmalıdır.
2. Her bir değişkenin hatasız bir şekilde ölçüldüğü varsayılır.
3. Bağımlı değişkenlerin değerlerinin birbirleriyle ilişkisi yoktur.

4. Bağımlı değişkenler normal dağılıma sahiptir.
5. Bağımlı değişkenlerin her birinin varyansı eşittir.
6. Bağımlı ve bağımsız değişkenler arasındaki ilişki doğrusaldır.

1.4.1.5 Hata Terimi (ϵ_i) Varsayımları

1. Hata teriminin (ϵ_i) beklenen değeri sıfır olmalıdır.
2. Hata terimlerinin (ϵ_i) hata değerleri arasında ilişki yoktur.
3. Hata terimleri (ϵ_i) normal bir dağılıma sahiptir.
4. Hata terimleri(ϵ_i) ile bağımlı ya da bağımsız değişkenler arasında bir ilişki yoktur.

1.4.1.6 Çoklu Bağlantı Problemi

Çoklu regresyonda birden fazla bağımsız değişkenin aralarında yüksek bir korelasyonun olması durumuna çoklu doğrusal bağlantı problemi denir. Böyle bir durumda en küçük kareler tekniğinin modellemedeki başarısı düşük çıkar. Bu durumda EKK'ya alternatif yöntemler kullanılır (Karakaş, 2008).

Regresyon analizinde EKK tekniğinin uygulayabilmek için bir takım varsayımların geçerli olması gerekmektedir. Bu varsayımların geçerli olması halinde EKK yöntemi başarılı ve güvenilir sonuçlar verecektir (Yavuz, 2017).

1.4.1.7 Düzenleştirme

Bir modelde çoklu bağlantı problemi varsa EKK yöntemi başarılı sonuç vermeyeceğinden bu yöntem yerine başka yöntemler kullanılır. Böyle bir durumda model verisine yeni değişkenler eklenebilir ya da çoklu bağlantıya sahip değişkenlerden en az bir tanesi modelden çıkarılarak hesaplamaya alınmayabilir. Fakat çıkarılan değişkenin modele gerçek katkısı bilinmediğinden bunun sağlıklı olacağını söylemek pek mümkün değildir.

Genel olarak düzenleştirme, model içerisinde bulunan değişkenleri çıkarmadan regresyon modeline bir cezalandırma teriminin eklenmesi metodudur. Düzenleştirme yöntemleri çoklu bağlantı sorunlarında, modelin aşırı öğrenme problemi yaşaması veya öznitelik sayısının örnek sayısından çok büyük olduğu durumlarda kullanılırlar.. En çok bilinen düzenleştirme yöntemleri Ridge, Lasso ve Elastic Net yöntemleridir.

Ridge yönteminde L_2 normu, Lasso yönteminde L_1 normu ve Elastic Net yönteminde ise L_1 ve L_2 normu kullanılır.

1.4.2. Ridge Regresyon Yöntemi

Ridge regresyon için temel olarak modifiye edilmiş EKK yöntemidir diyebiliriz. Ancak bu yöntem EKK yöntemine göre daha güvenilir sonuçlar verir. Bunun nedeni Ridge regresyon yöntemi sorunlu verilere EKK yöntemine göre daha fazla hoşgörü tanıyan bir yöntemdir (Topdağ ve Acar, 2021).

Ridge regresyon yönteminde model kurulurken verilerdeki her bir değişken kullanılır. Modelde bulunan sorunlu olan değişkenleri modelden çıkarmak yerine bunların katsayıları küçültülerek modeldeki ağırlığı azaltılır (Aslan ve Yıldız, 2022).

Bir modelde aralarında çoklu bağlantı sorunu olan değişkenler çıkarılırsa bu aslında değişkenlerin katsayısının sıfır alınmış olması ile eşdeğerdir. Oysa Ridge regresyonda bu değişkenlerin katsayıları çıkarılmayıp, sıfıra yaklaştırılarak modelde bırakılır. Bu şekilde aralarında çoklu bağlantı sorunu olan değişkenler dahi modelde tutulup bunlar arasındaki ilişki incelenir (Topdağ ve Acar, 2021).

Ridge regresyon EKK'nin en küçük yapmaya çalıştığı fonksiyona L_2 düzenleme normunun eklenmesiyle elde edilir.

Daha önce vermiş olduğumuz Hata kareler toplamı(HKT) Formül 1.9'da hatırlatılmıştır (Aslan ve Yıldız, 2022).

$$HKT = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1.9)$$

HKT'na eklenen L_2 normu Formül (1.10) da verilmiştir. Formülde verilen λ , L_2 normunun katsayısıdır. Bu katsayı L_2 normunun cezalandırma katsayısıdır.

$$L_2 = \lambda \sum_{j=1}^p \beta_j^2 \quad (1.10)$$

Eşitlik 1.11 de Ridge regresyon modeli aşağıdaki gibi ifade edilmiştir:

$$SSE_{L_2} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p \beta_j^2 \quad (1.11)$$

Dikkat edilirse formüldeki $\lambda = 0$ olduğunda Ridge regresyon modelinin En Küçük Kareler metoduna eşit olacağı görülmektedir. λ değeri arttıkça model

katsayıları düzleşmeye başlar. λ değerinin artmasıyla beraber katsayılar üzerinde cezalandırma artmaktadır.

1.4.3. Lasso Regresyon

En küçük kareler yöntemi ile Ridge regresyon yönteminin tahminlerdeki başarı oranı yüksektir. Ancak bu iki yöntemde varyanslarının bir üst sınırı yoktur. Çoklu bağlantı sorununda bu yöntemlerin ikisi de yüksek varyansa sahip olup düşük verim gösterebilmektedirler. En küçük kareler yöntemindeki bu soruna çözüm olarak tahminin doğruluğunu arttırmak için parametrelerin katsayıları daraltılabilir veya bazı parametrelerin katsayıları sıfır alınabilir. Ridge regresyonda parametrelerin katsayıları küçültülerek parametrelerin katsayılarını daraltma işlemi gerçekleşir. Ancak parametrelerin katsayısı hiçbir zaman sıfır olmadığı için model içerisinde ilgisiz parametreler sürekli vardır. Bu durum elde edilen modelin yorumunu kimi zaman zorlaştırmaktadır. Tibshirani 1996'da buna benzer bazı problemler için Lasso metodunu önermiştir. Bu metoda göre, parametrelerin katsayılarının mutlak değerleri toplamı sabitten küçük ise HKT' ni minimum değere indirme temeli üzerine kurulmuştur (Küçük, 2019).

Lasso Regresyon, Ridge Regresyonuna benzer en temel fark Ridge regresyonda parametrelerin düzenlenmesi için L_2 ceza fonksiyonu kullanılırken Lasso regresyonda ise parametrelerin düzenlenmesi için L_1 ceza fonksiyonu kullanılır. Bunu şöyle anlatabiliriz. Ridge katsayıların karesini alırken Lasso katsayıların mutlak değerini alır. Lasso yönteminde kullanılan L_1 ceza fonksiyonu ve katsayısı Formül (1.12)'de verilmiştir.

$$L_1 = \lambda \sum_{j=1}^p |\beta_j| \quad (1.12)$$

Burada L_1 ceza fonksiyonudur. λ ise pozitif bir değer olup ceza parametresidir. λ büzülme miktarını kontrol eder. Lasso düzenleştirici yönteminin optimizasyon hedefi aşağıda belirtilen Formül 1.13'de gösterilmiştir:

$$SSE_{L_1} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^p |\beta_j| \quad (1.13)$$

Ridge regresyonda kullanılan L_2 düzenlestirmesi modelde bulunan parametrelerin katsayılarını küçültürken, Lasso regresyonda kullanılan L_1

düzenleştirmesi ise modelde bulunan parametrelerin katsayılarını sıfırlayarak modelin katsayılarının daha seyrek bir yapıda olmasını sağlamaya çalışır (Bardak, 2016).

Lasso ve Ridge regresyonlarından hangisinin daha başarılı olduğunu söylemek mümkün değildir. Frank ve Friedman (1993) çalışmalarında kimi zaman Ridge kimi zaman ise Lasso regresyonun daha başarılı tahminler ürettiklerini göstermişlerdir (Aslan ve Yıldız, 2022).

Lasso regresyon yöntemi hem değişkenleri seçme hem de düzenleme görevini yapar. Yöntemde katsayıların mutlak değerleri toplanarak sonucun sabit bir değerden küçük olması sağlanır. Bu şekilde hem kullanılacak değişkenler belirlenmiş olur hem de model yorumu daha iyi olur. Bu yöntemde kimi katsayılar sıfıra yaklaştırılırken kimi katsayılar da sıfır değerini alır. Lasso modeli aralarında yüksek ilişki bulunan katsayılardan bir tanesini alır. Lasso regresyon bu yönüyle değişken seçimini kendi yapabilmektedir. Lasso regresyonun başarılı tahminler üretmesi en büyük avantajıdır. Çünkü katsayılar sıfıra yaklaştırılırken veya sıfır olurken varyans doğal olarak azalacaktır. Bilhassa klasik tahmin metotlarının iyi sonuç vermediği değişken sayısının çok, örneklem genişliğinin az olduğu durumlarda çok iyi sonuçlar verir (Yıldırım ve Çiftçi, 2021).

1.4.4. Lojistik Regresyon

Lojistik regresyon, belirli bir olayın meydana gelme olasılığını tahmin etmek için kullanılan bir tür makine öğrenimidir. Bağımlı değişken (tahmin etmeye çalıştığımız veri) ve bağımsız değişkenler (etkileyebileceğini düşündüğümüz veri) arasındaki ilişkiyi modellemek için regresyon analizini kullanmanın bir yoludur.

Lojistik Regresyon her ne kadar bir regresyon çeşidi de olsa, sınıflandırma işlemi için kullanılan bir makine öğrenimi yöntemidir. Kategorik ya da sayısal veriler bu regresyonla sınıflandırılabilir. Lojistik regresyonda bağımlı değişken verileri (sonuç) sadece iki farklı değer alabilir. (Doğru/Yanlış, Evet/Hayır, Başarılı/Başarısız...) Lojistik ve lineer regresyon modelleri birbirlerine benzemektedirler. Ancak Lineer regresyonda doğrusal fonksiyon kullanılırken, Lojistik regresyonda Sigmoid fonksiyon kullanılır. Lojistik regresyonun temel felsefesi Sigmoid fonksiyonun $[0,1]$ aralığında değerler alması üzerinedir. Sigmoid fonksiyon formülü (1.14)'te verilmiştir.

$$f(x) = \frac{1}{1 + e^{-x}} = \frac{e^x}{1 + e^x} \quad (1.14)$$

Lojistik regresyon değişken sayısına göre isimlendirilir. Sadece bir değişken ile kurulan modellere tek değişkenli, birden çok değişken ile kurulan modellere çok değişkenli regresyon adı verilir (Parlak, 2019).

Lojistik regresyon ile Lineer regresyon arasındaki farklar şunlardır.

1. Lineer regresyonda tahmini yapılacak değişken (bağımlı değişken) sürekli. Oysa Lojistik regresyonda bu değer kesikli olabilmektedir.

2. Lineer regresyonda bağımlı değişkenin alabileceği değer tahmin edilirken, Lojistik regresyonda ise bağımlı değişkenin hangi değeri alabileceğinin olasılığı tahmin edilir.

3. Lineer regresyonda bağımsız değişkenin normal bir dağılıma sahip olması gerekirken lojistik regresyon için bir koşul yoktur (Bircan, 2004).

Lojistik regresyon modelinde bağımlı değişkenler ile bağımsız değişkenlerin arasındaki ilişki doğrusal olmayabilir. Ancak bağımlı değişkenler ile bağımsız değişkenler arasında doğrusal, üstel, polinomal bir ilişki olsa dahi lojistik regresyon modeli bu ilişkiyi logit bir ilişki olarak kabul eder. Bu nedenle lojistik regresyon modelleri doğrusal değildir (Kara, 2015).

Lojistik regresyon analizinde parametrelerin katsayılarını tahmin ederken “En Çok Olabilirlik Yöntemi (maximum-likelihood estimation)” adı verilen bir yöntem kullanılır. Bu yöntem Lineer regresyonda kullanılan EKK yöntemine benzemektedir (Gök, 2010).

Bir veri grubunda bağımlı değişken normal bir dağılıma sahipse EKK yöntemi ve en çok olabilirlik yöntemlerinin ikisi de aynı sonucu verir. Ancak EKK yöntemi sadece doğrusal modeller için kullanılabilirken en çok olabilirlik yöntemi için doğrusallık koşulu yoktur. Yani hem doğrusal olan hem de doğrusal olmayan modellerde kullanılabilir. Lojistik regresyon modeli doğrusal olmadığından EKK yöntemi yerine en çok olabilirlik yöntemi kullanılır. Bu yöntemde önce en çok olabilirlik fonksiyonu oluşturulur. Daha sonra bu fonksiyonu maximum yapan parametreler seçilir (Akçay, 2009).

Lojistik regresyonun ana fikri olasılık oranı üzerine kurgulanır. Olasılık oranı, bir olayın gerçekleşmesinin olasılığının aynı olayın gerçekleşmesinin olasılığına

orandır. Lojistik regresyon modeli için olasılık oranının doğal logaritması alınır. Parametreleri tahmin etmek için de en çok olabilirlik yöntemi kullanılır (Şahin, 2017).

Lojistik regresyon modelinde tahmin edilecek değişkenin gerçekleşme olasılığı tahmin edilir. Bu durumda, $E(Y|x) = \pi(x)$ eşitliği x 'in bilinmesi durumunda Y 'nin x 'e bağlı koşullu olasılığını göstermek üzere, koşullu olasılığı verecek olan Lojistik regresyon fonksiyonu Formül (1.15) de gösterilmiştir.

$$\pi(x) = \frac{e^{(\beta_0 + \beta_1 x_1)}}{1 + e^{(\beta_0 + \beta_1 x_1)}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1)}} \quad (1.15)$$

Logistik regresyon fonksiyonunun modelde yer alan β_0 ve β_1 parametreleriyle doğrusal hale dönüştürülmesi için bir transformasyona tabi tutulması gerekir. Bu dönüşüm logit dönüşüm olarak adlandırılır. Bunun için olasılık oranının doğal logaritması alınacaktır. Olasılık oranı bir olayın gerçekleşme olasılığının gerçekleşmeme olasılığına oranı olduğundan; gerçekleşme olasılığı $\pi(x)$ ise gerçekleşmeme olasılığı $1 - \pi(x)$ olacaktır.

Bu durumda logit dönüşümü Formül (1.16)'da olduğu gibi olacaktır.

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta_0 + \beta_1 x \quad (1.16)$$

Logit dönüşümü $g(x)$, sürekli ve modeldeki β_0 ve β_1 parametreleriyle doğrusal bir ilişkiye sahiptir. Bu parametreler genellikle en çok olabilirlik yöntemi ile tahmin edilir. Bunun için en çok olabilirlik fonksiyonu oluşturulur.

Olabilirlik fonksiyonu Formül (1.17)'da verilmiştir.

$$L(\beta) = \sum_{i=1}^n \{y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))\} \quad (1.17)$$

Olabilirlik fonksiyonundaki β_0 ve β_1 parametrelerini bulabilmek için fonksiyonu maximum yapan değerler bulunmalıdır. Bunun için β_0 ve β_1 'e göre türevleri alınıp 0'a eşitlenir. (Bircan, 2004)

1.4.5. Destek Vektör Makineleri (Support Vector Machine - SVM)

Destek vektör makineleri (DVM) farklı sınıflara ait olan verileri uygun olarak ayırabilmek için kullanılan istatistik temelli bir denetimli öğrenme algoritmasıdır. Destek vektör makinelerinde temel hedef eğitimde kullanılan veriyi uygun sınıflara ayıran en iyi fonksiyonu bulmaktır.

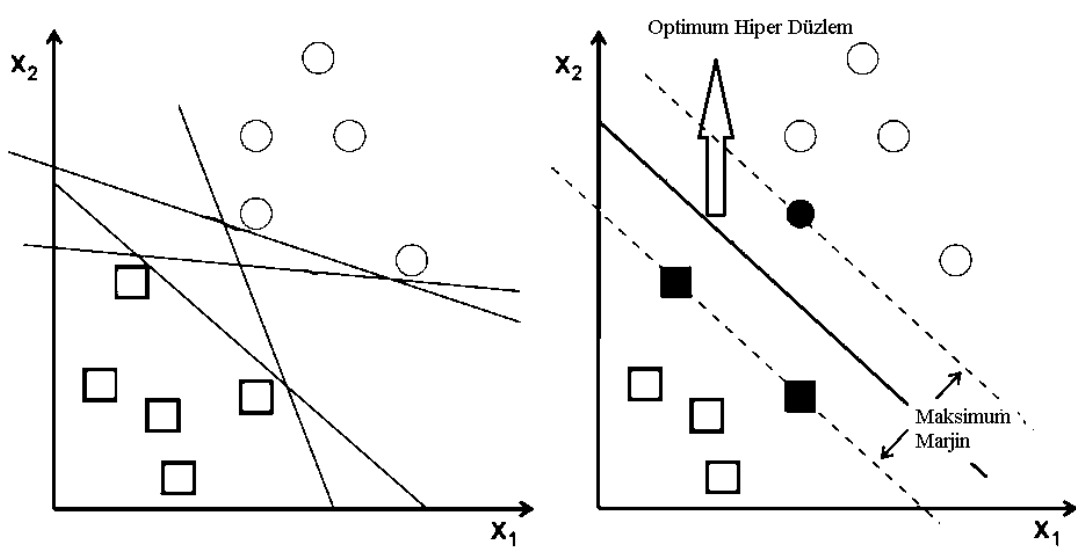
Destek vektör makineleri Yapay sinir ağlarına(YSA) çok benzerler. Çünkü DVM ileri beslemeli ve çift katmanlı olan YSA'ya sahiptirler. Destek vektör makineleri doksanlı yıllarda geliştirildiğinden beri çok popüler bir algoritma olarak kullanılan, sınıflandırma yapabilmek için kalıpları tanıyan ve verileri analiz edebilme yeteneği olan modellerdir. Destek vektör makineleri deneysel verilerin arasında bulunması çok güç olan bağlantıları yakalayabildikleri için finansal zaman serilerinin sınıflandırmasında ve modellenmesinde sıklıkla kullanılırlar (Kartal, 2020).

DVM'lerinin ana felsefesi eğitim verisindeki düzlemi ayırabilen en uygun aralığın bulunmasıdır. Böylece gelecek yeni verilerde sınıflandırılmış olacaktır. DVM'nin bir avantajı da eğitim verilerinin özelliklerinden etkilenmemeleridir. Bu özellikleri sayesinde çok büyük verileri modelleyebilir ve sınıflandırabilmektedir. Ancak DVM nin eğitim ve sınıflandırma algoritmaları oldukça karmaşıktır. Bu nedenle çalışmaları için güçlü bir bellek gerekir (Irmak, 2019).

Destek vektör makinelerinin sınıflandırıcıları verileri doğrusal ya da doğrusal olmayacak şekilde iki türlü ayırabilir.

1.4.5.1 Verilerin Doğrusal Ayrılabilme Durumu

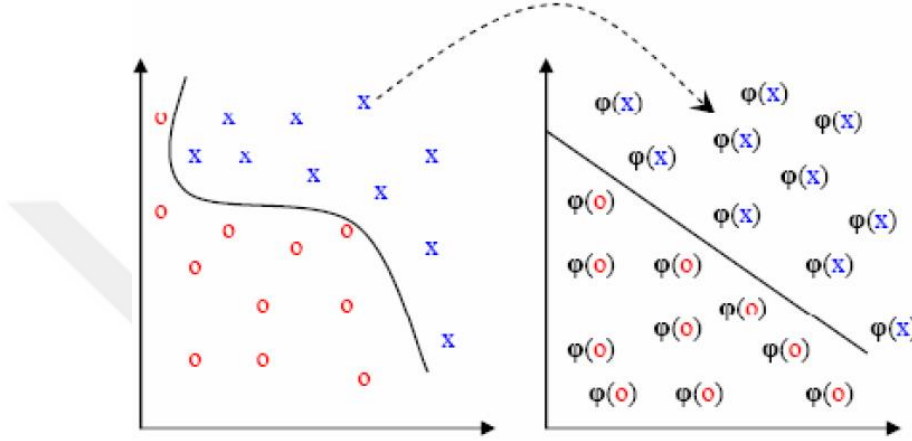
Eğitim verileri Destek vektör makinesi sınıflandırıcısı tarafından doğrusal olarak iki farklı gruba ayrılabilir. Eğer bir aşırı düzlem ile ayrılır. Bu düzlemin özelliği her iki sınıfa eşit mesafede olmasıdır (Eray, 2008).



Şekil 1.1. Lineer ayrılabilen veriler (Irmak, 2019)

1.4.5.2 Lineer Ayrılamama Durumu

Doğrusal olarak iki farklı gruba ayrılmayan problemlerde ana düşünce tüm veriyi doğrusal olacak şekilde ayırılabilen daha büyük boyutlara sahip bir özellik uzayına taşımaktır. Bu şekilde optimal ayırıcı aşırı düzlem çizilebilir. Başlangıçta giriş uzayında verilen eğitim verisinin tamamı yeni tanımlanan özellik uzayına taşınır. (Eray, 2008)



Şekil 1.2. Giriş uzayını özellik uzayına taşıma (Eray, 2008)

1.5. Teknik Analiz Göstergeleri (İndikatörler)

Teknik Analiz göstergelerine finansal piyasalarda İndikatör adı verilir.

1.5.1. Genel Bilgiler

İndikatörler, geçmişteki fiyat hareketlerinden faydalanarak geleceği öngörmeye çalışan göstergelerdir. İndikatörler gelecek tahmini için matematiksel hesaplamalardan faydalanırlar. Teknik Analizin temeli geçmiş fiyat hareketlerine bakarak geleceği öngörmektir. Bu nedenle İndikatörler bir nevi teknik analizin temelini oluştururlar diyebiliriz. Bununla birlikte, finans piyasalarında fiyat hareketleri çok oynaktır. Bu nedenle indikatörlerin kesin sonuç vermesini beklemek doğru bir yaklaşım olmayacaktır.

Bir çok indikatör olmasına rağmen, genel anlamda indikatörleri iki başlık altında toplayabiliriz. Şöyle ki kimi indikatörler gelecek için belli hesaplamalar ile fiyatın ne şekilde olacağını bulmak isterken, kimi indikatörlerde piyasada trendi ölçüp, trendin gücü ve varlığıyla ilgili öngörülerde bulunurlar.

a. Öncü İndikatörler

Öncü indikatörler, işlev olarak finansal piyasalarda fiyat hareketi gerçekleşmeden hareketin yönünü tahmin etmeye yarayan göstergeler olarak bilinirler. Diğer bir deyişle, mevcut bir trendin geri dönme olasılığını önceden belirten göstergelerdir. Öncü indikatörlerin çoğu, sabit bir yeniden-inceleme dönemi boyunca fiyat momentumunu ölçerler. En popüler Öncü indikatörlerden bazıları şunlardır; RSI (Göreceli Güç Endeksi), Stochastic Fast ve Stochastic Slow, CCI (Emtia Kanal Endeksi), , William's %R (William'ın yüzde aralığı) ve Stokastik Osilatör

b. Gecikmeli İndikatörler

Gecikmeli indikatörler, finansal piyasalarda fiyat hareketine göre karar veren, fiyatta değişimler gerçekleştikten sonra al ya da sat sinyali veren indikatörlerdir. Bu yönüyle gecikmeli indikatörler trend takip eden indikatörler olarak bilinirler. Gecikmeli indikatörlerde trend takibi için ortalamalar en sık kullanılan hesaplama yöntemleridir. Gecikmeli indikatörlerin en çok bilinenleri olarak, MA (Hareketli ortalamalar) ve MACD (Hareketli Ortalamaların mesafesi) göstergeleri sayılabilir.

Öncü İndikatörler, her ne kadar menkul kıymete ilişkin fiyat hareketinin muhtemel yönünü daha ortaya çıkmadan belirtmesinden dolayı, Gecikmeli indikatörlerden daha üstün olduğu düşünülse dahi; ortaya koydukları öngörüler, bu indikatörlerin doğası gereği kesinlik veya tam geçerlilik içermez. Dolayısıyla, bir indikatörden alınan sinyal onayının tam doğru olduğunu düşünmemek gerekir.

İndiktörler, öncü veya gecikmeli olup olmadıklarına bakılmaksızın piyasa hareketinin farklı yönlerini ölçerler. Ayrıca, nasıl hesaplandıklarına bağlı olarak, indikatörler sıfır çizgisinin üstünde ve altında salınabilirler. Bu tip indikatörlere, “Salınım Göstergeleri” ya da daha sık kullanımıyla “Osilatörler” denir. Diğer çeşit indikatörler ise şunlardır; momentum indikatörleri, oynaklık (volatilite) indikatörleri, piyasa gücü indikatörleri ve döngü indikatörleri. Bu indikatörlerden bazıları trend piyasalarda kullanıma uygunken, bazıları da trend olmayan ya da değişken piyasalarda kullanıma daha uygun olduğu için, farklı türlerdeki indikatörler genellikle biri biriyle çelişebilirler.

1.5.2. İndikatörleri Alım Satım Noktalarını Nasıl Belirlerler

1. Yöntem: Alım satım noktalarını belirlemek için finansal veri grafiği üzerine indikatör grafiği atılır. Tahmin edilecek sembolün fiyatı ile İndikatörün kesişim noktalarına göre alım satım noktaları belirlenir.

Örneğin fiyatın hareketli ortalamayı yukarı kesmesi al, aşağı kesmesi sat olarak kabul edilir.

2. Yöntem: Bazı indikatörler içerisinde 2 tane çizgi barındırırlar. Çizgilerin birbirlerini kesmesi bir alım satım sinyali olarak kabul edilir.

Macd, Trix, Stochastic Fast veya Stochastic Slow göstergeleri buna örnek olarak gösterilebilir.

3. Yöntem: Bazı indikatörler yatay iki çizgi arasında bulunurlar. İndikatör değeri üst çizgiye yaklaştığında veya yukarı kıldığında aşırı alım, İndikatör değeri alt çizgiye yaklaştığında veya aşağı kıldığında aşırı satım olduğu anlamına gelir. Burada indikatörün yatay iki çizgi arasındaki seviyeside önemlidir. Kimi İndikatörlerin alt ve üst çizgiye olan mesafesi trendin gücü hakkında bilgi verir.

CCI, RSI indikatörleri buna örnek olarak verilebilir.

1.5.3. ADX İndikatörü (Average Directional Movement Index)

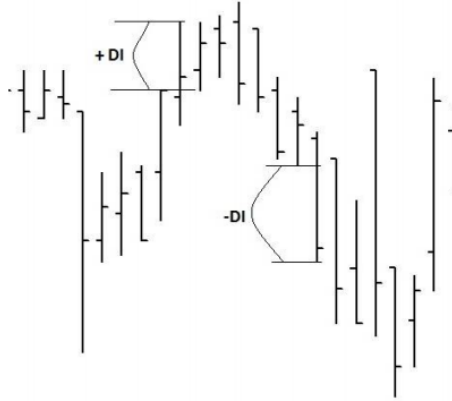
ADX göstergesi, fiyatın kendisi değil, yönü hakkında öngörude bulunmayı sağlayan bir göstergedir. Trendin varlığı ve yönünü belirleme açısından oldukça kullanışlı olan bir göstergedir. J. W. Wilder tarafından bulunmuştur. ADX göstergesinin ön tanımlı periyod değeri genellikle 14'tür.

Hesaplanması :

ADX(Average Directional Movement)'i anlatabilmek için önce DM (Directional Movement) göstergesini anlatmak gerekir. Çünkü Directional movement sadece ADX değil, ADXR, DIS, DX, DI+, DI- gibi bir çok göstergeninde ana fikrini oluşturur.

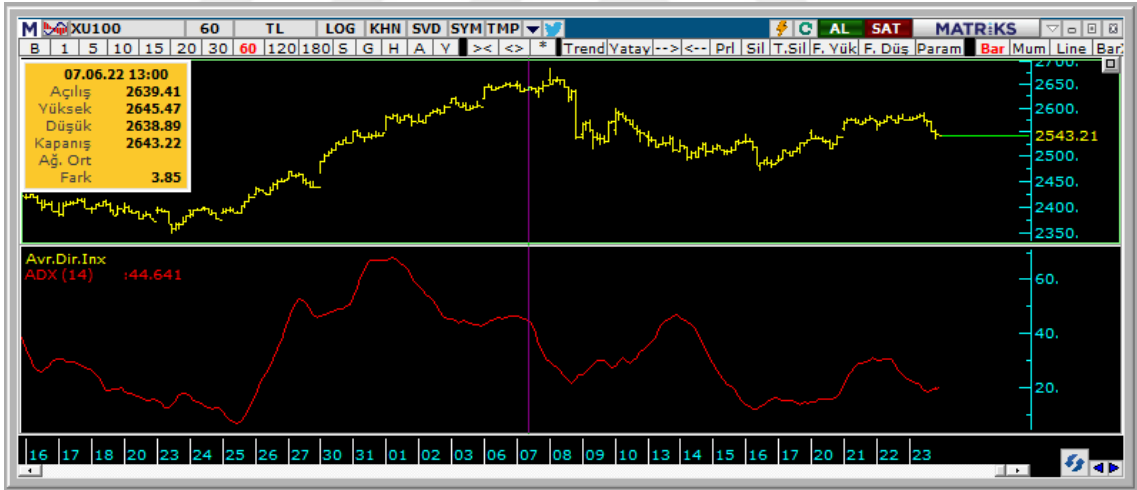
DM'yi hesaplayabilmek zor olmasa da epey zahmetlidir. DM değerini bulmak için bugünkü en yüksek ve en düşük değerlerin birleştirilmesi ile elde edilen doğru parçasından, dünkü en yüksek ve en düşük değerlerin birleştirilmesi ile elde edilen doğru parçasının kesişimi çıkarılır. Piyasanın yönü yukarı doğru ise DM değeri pozitif, aşağı

yönlü ise DM değeri negatif olur. Şekil 1.3 te görüleceği üzere, piyasada yön yukarı ise bu değer pozitif, yön aşağı ise bu değer negatif olur.



Şekil 1.3. +DI ve -DI değerleri

Bulunan +DM ve -DM değerlerinin toplamı, farklarına bölünerek DX değerini bulmuş oluruz. DX bir göstergedir ve oldukça hızlı hareket eder. DX göstergesinin basit hareketli ortalaması alındığında ADX indikatörünün değeri bulunmuş olur.



Şekil 1.4. ADX Göstergesi

Yorumlanması :

ADX(Average Directional Movement Index) indikatörü al ya da sat şeklinde sinyal üretmez, trend ile ilgili bilgiler verir. ADX, 0-100 aralığında değerler alan bir indikatördür. Bu değerler, trendin varlığı ve trendin gücü hakkında bilgiler içerir. Fiyatın yönünün yukarı ya da aşağı olması ADX açısından bir anlam taşımaz. Örneğin piyasada bir düşüş trendi varken ADX değeri yüksek olacaktır.

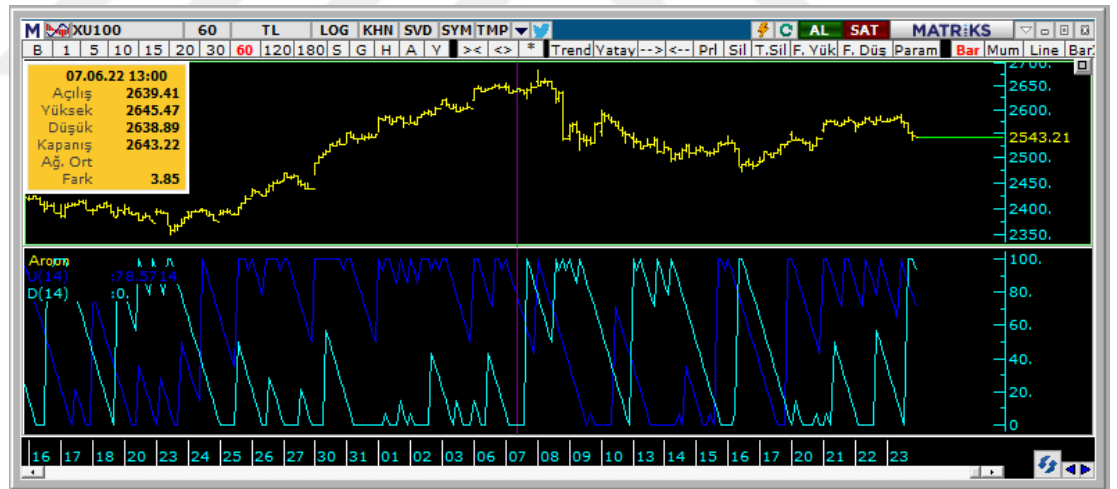
ADX değerinin yükselmesi piyasada bir trend başlangıcı olduğunu anlatır. Ancak trendin yönü ADX tarafından bilinemez. Bu trend iki yönde de olabilir. ADX değerinin düşmesi var olan trendin sonlanabileceği anlamı taşır. ADX değeri sıfır bölgesine yakın değerler alıyorsa bu mevcutta bir trend bulunmadığı anlamına gelir. Bu bize piyasanın kararsız bir yapıda olduğunu anlatır.

ADX göstergesinin alım satım noktaları üretmeyip, sadece trendi ölçmesinden dolayı bu gösterge başka indikatörler ile beraber kullanılır.

1.5.4. Aroon İndikatörü

Aroon indikatörünün temel işlevi piyasada trend var ise bu trendi erkenden bulmaktır. Aroon indikatörü 0 ile 100 arasında değerleri olan 2 ayrı çizgiden oluşur. Bu çizgilere, Aroon Up (Üst çizgi), Aroon Down (alt çizgi) adı verilir.

Bu göstergeyi 1995 yılında Tuschar Chande geliştirmiştir. Analistler genellikle Aroon göstergesini 14 periyotluk bir zaman diliminde kullanmayı önermektedirler. Ancak bu periyot kullanım amacına ve piyasaya göre değiştirilebilir.



Şekil 1.5. Aroon İndikatörü

Bu indikatörün amacı, fiyat yön değiştirdiğinde bu değişimin bir trend oluşturup oluşturmadığını bulmaktır. Aroon göstergesi, Aroon Up ve Aroon Down göstergelerinin birlikte aldıkları değerler ve konumlarına göre yorumlanır. Göstergenin kullanıldığı periyot boyunca fiyatlar en yüksek yeni bir değer almış ise Aroon Up göstergesi 100 değerini almamış ise 0 değerini alacaktır. Yine göstergenin

kullanıldığı periyot boyunca fiyatlar yeni bir en düşük değer almışsa Aroon Down göstergesi 100 değerini, almamış ise 0 değerini alacaktır.

Hesaplanması :

Aroon indikatörünün iki çizgiden oluşmasından dolayı, hesaplama, birbirine benzer adımları izleyen iki farklı parça halinde gerçekleşir.

$$\text{AroonUp}_i = \left[\frac{n - \max(H_i)}{n} \right] \times 100 \quad (1.18)$$

n = Hesaplama için seçilen geri inceleme dönemidir.

$\max(H_i)$ = Fiyatın max. olduğu andan itibaren geçen periyot sayısı

$$\text{AroonDown}_i = \left[\frac{n - \min(L_i)}{n} \right] \times 100 \quad (1.19)$$

$\min(L_i)$ = Fiyatın min. olduğu andan itibaren geçen periyot sayısı

Yorumlanması :

Aroon Up göstergesinin değerinin 100 olması, güçlü bir yukarı yönlü fiyat momentumuna işaret etmektedir. Bu durumda, Aroon Up göstergesinin değeri 70 ile 100 arasında iken, aynı anda Aroon Down göstergesinin değeri 0 ile 30 arasında ise yukarı doğru güçlü bir trendin başladığının sinyali alınmış demektir. Bu durumun tersi aşağı yönlü bir trend başlangıcıdır diyebiliriz.

Aroon Down göstergesinin değerinin 100 olması, güçlü bir aşağı yönlü fiyat momentumuna işaret etmektedir. Aroon Down göstergesinin değeri 70 ile 100 arasında değerler alırken, aynı anda Aroon Up göstergesinin değeri 0 ile 30 arasında değerler alıyorsa aşağı yönlü bir trendin varlığından bahsedebiliriz.

Bu yorumun dışında Aroon Up ve Aroon Down göstergelerinin birbirleriyle kesişmesi al sat sinyalleri olarak kullanılabilir. Bu durumda Aroon Up göstergesinin değerinin, Aroon Down göstergesinin değerinden büyük olduğu ilk anda alım yapılır ve pozisyona girilir. Aroon Down göstergesinin değerinin Aroon Up göstergesinin değerinden büyük olduğu ilk anda ise satış yapılır pozisyondan çıkılır.

1.5.5. Bollinger Bandı

Bollinger bandı birbirleriyle bağlantılı 3 ayrı eğriden oluşur. Bu eğriler alt bant, üst bant ve orta bant şeklinde isimlendirilirler. Bollinger bandında orta bant, fiyat

çizgisinin 20 günlük basit hareketli ortalamasından oluşur. Diğer iki bant ise orta bandın standart sapma değerinin 2 katı kadar aşağıya ve yukarıya doğru ötelenmesiyle elde edilirler. Bu şekilde fiyat alt ve üst band aralığında hareket eder. Fiyatın alt ve üst banda olan yakınlığına göre yorumlar yapılır. Alt ve üst bantlar kimi zaman destek ve direnç bölgesi şeklinde de kullanılabilir.



Şekil 1.6. Bollinger Bandı

Hesaplanması :

Orta Bollinger Bandı = 20 Günlük Basit Hareketli Ortalama

Üst Bollinger Bandı = Orta Bollinger Bandı + (2 × Standart Sapma)

Alt Bollinger Bandı

$$= \text{Orta Bollinger Bandı} - (2 \times \text{Standart Sapma}) \quad (1.20)$$

Bollinger Bandında orta bant genellikle 20 günlük basit ortalama ve standart sapma çarpanı 2 olarak alınmış olmasına rağmen, yatırım yapılacak finansal enstrümana ve piyasa şartlarına göre, ortalamanın türü ve periyodu ile birlikte standart sapma çarpanının değeri de değiştirilebilir.

Yorumlanması :

Bollinger bandında fiyatlar üst banda ulaştığında veya üst bandı geçerse yükseliş olacağı beklenmelidir. Fiyatlar üst banda değdiği sürece yükseliş trendinin devam edeceği beklenir.

Eğer fiyatlar alt banda değerse ya da alt bandı aşağı kırarsa bu durumda fiyatlarda düşüş olması beklenmelidir. Fiyatlar alt banda temas ettiği sürece düşüşün devam etmesi beklenir.

Bununla birlikte fiyatlar orta banttı üst banda doğru çıkarsa yükseliş, alt banda doğru inerse düşüş beklenenecektir.

Bir başka yorumda şöyle olabilir. Fiyatlar Bollinger Bandının dışına taşarsa tekrar kanal içerisine döneceği düşünülür.

Bollinger Bandının daralması da piyasanın kararsız bir yapıda olduğu anlamına gelir.

1.5.6. CCI İndikatörü

CCI göstergesi D. Lambert tarafından bulunmuştur. Bu gösterge asıl olarak mal piyasalarında işlem yapmak için geliştirilmiştir. Ancak hisse senedi piyasalarında da çok iyi sonuç verdiği için oldukça kullanışlı ve popüler bir göstergedir. CCI göstergesi mevcut fiyatın ortalamadan sapma miktarını bulmak için tasarlanmıştır.

Ortalamadan sapma miktarı değişen piyasa eğilimlerinin ve tercihlerinin tespit edilebilmesi açısından önemlidir. CCI trendin olmadığı yatay piyasalarda daha başarılıdır. Genellikle kısa vade için kullanılır. CCI göstergesi +100 ile -100 arasında değerler alır. CCI göstergesi, +100 e yaklaştıkça aşırı alım, -100 e yaklaştıkça aşırı satım bölgesine giriliyor demektir.



Şekil 1.7. CCI İndikatörü

Hesaplanması :

CCI hesaplanırken fiyat yerine ortalama fiyat kavramı kullanılır. Ortalama fiyat(O.F), gün içi max. değer, gün içi min. değer ve kapanış değerinin toplamının 3'e bölünmesiyle bulunan değerdir.

$$O.F = \frac{\text{En düşük} + \text{En yüksek} + \text{Kapanış}}{3} \quad (1.21)$$

$$CCI = \frac{O.F - O.F. Basit Hareketli ortalaması}{\text{Ortalama Fiyatın Mutlak sapma ortalaması} * 0.015} \quad (1.22)$$

Yorumlanması :

Genel olarak CCI göstergesini iki şekilde yorumlayabiliriz. Birinci yorumda CCI nin aşırı alım ve aşırı satım bölgesini kullanabiliriz. CCI göstergesinin +100 bölgesinde olması piyasanın aşırı satım bölgesinde olduğu ve mevcut yükselişin buralarda bitip piyasanın satışa geçeceği anlamına gelir. Bunun tersinde yani CCI değerinin -100 bölgesinde olması ise piyasanın aşırı alım bölgesinde olduğunun ve bundan sonra piyasadaki düşüşün bitip yukarı doğru bir hareketin olacağının habercisi olacaktır.

CCI göstergesinin ikinci yorumu fiyat uyumsuzluğu ile ilgilidir. Bunu kısaca şöyle açıklayabiliriz. Gösterge değeri ile fiyatın hareketi aynı yönde değil ise piyasada bir uyumsuzluk söz konusudur. Bu durumda fiyatlar yukarı doğru çıkıyorken CCI aşağı düşüyorsa fiyatlarda aşağı yönlü bir düzeltme beklenebilir. Bunun terside söz konusu olabilir. Yani fiyatlar düşerken, CCI değeri yukarı doğru artıyorsa fiyatlarda yukarı yönlü bir düzeltme beklenebilir.

1.5.7. DEMA İndikatörü (Double Exponential Moving Average)

DEMA(Double Exponential Moving Average), için üssel hareketli ortalamanın gecikmesini azaltmak için yine üssel ortalamaya bazı işlemler uygulanarak elde edilen bir ortalama türüdür diyebiliriz. Dema indikatörü genellikle 5 periyotluk standart değerle kullanılır. Ancak bu değer piyasa koşullarına ve ihtiyaca göre değişkenlik gösterebilir.



Şekil 1.8. Dema İndikatörü

Hesaplanması :

n gün EMA = n Günlük Üssel Hareketli Ortalama

$$\text{Dema} = (n \text{ gün EMA} * 2)$$

$$- (n \text{ gün EMA'nın, } n \text{ gün EMA sı}) \quad (1.23)$$

Yorumlanması :

Dema göstergesi bir nevi hareketli ortalama türü olduğundan bu göstergenin yorumlanmasında hareketli ortalamaların yorumlanmasıyla aynı olacaktır. En basit haliyle Dema göstergesi fiyat grafiği üzerine atılır. Fiyat Dema göstergesinin üstüne çıktığında alım, fiyat Dema göstergesinin altına indiğinde satım yapılabilir.

İkinci bir yorum iki Dema göstergesinin fiyat grafiği üzerinde kesişim noktaları alım ve satım noktaları olarak kullanılır. Bunun için uzun ve kısa periyotlu iki Dema grafiği fiyat grafiği üzerinde çizdirilir. Kısa periyotlu olan Dema, uzun periyotlu Dema'nın üstüne çıktığında alım, altına düştüğünde satım yapılır.

Dema ile ilgili önemli bir hususu şöyle açıklayabiliriz. Dema göstergesi işleme girerken en yüksek ve en düşükleri yakalamaya çalışmaz. Burada amaç trend dönüşlerinde işleme girerek trend boyunca işlemde kalmaktır.

1.5.8. Envelopes İndikatörü

Envelopes göstergesi alt ve üst bantlardan oluşan bir göstergedir. Fiyatlar bu iki bant arasında hareket eder. Bu yönüyle Bollinger bandına benzer. Envelopes göstergesinde önce fiyatın belli bir periyotta hareketli ortalaması alınır. Daha sonra bu

hareketli ortalama belirli bir oranda aşağıya ve yukarıya kaydırılarak alt ve üst bant değerleri bulunmuş olur.



Şekil 1.9. Envelopes İndikatörü

Hesaplanması :

$$\text{Üst Band} = Ema(n) + Ema(n) * 0,0x \quad (1.24)$$

$$\text{Alt Band} = Ema(n) - Ema(n) * 0,0x \quad (1.25)$$

$Ema(n) = n$ Günlük Üssel Hareketli Ortalama

$x = \text{Yüzde Miktarı}$

Yorumlanması :

Envelopes indikatörü yorumlanırken fiyat hareketlerinin alt ve üst bant içinde hareket etmesi beklenir. Burada alt ve üst bantların birer sınır olduğu kabul edilir. Yani fiyat bir banta yaklaştığında buradan döneceği varsayılır. Bu durumda fiyat üst banda değdiğinde satım, alt banda değdiğinde alım yapılır.

1.5.9. Hareketli Ortalamalar

Hareketli ortalamalar, teknik analiz göstergelerinin birçoğunun temelini oluştururlar. Fiyatların oynaklığını azaltmak için sıklıkla kullanılırlar.

Hesaplanması :

Birçok hareketli ortalama hesaplama şekli vardır. Hesaplanma yöntemine göre hareketli ortalama türleri basit, üssel, üçgensel, değişken ve ağırlıklı hareketli

ortalamlardır. Çalışmamızda basit ve ağırlıklı ve üssel ortalama kullanıldığı için sadece bu ortalama türlerinin hesaplanması üstünde duracağız.

1. Basit Hareketli Ortalama (SMA)

Basit hareketli ortalama (SMA) bildiğimiz aritmetik ortalama ile eşdeğerdir. Teknik Analiz de çok kullanılır. Bir kıymetin n periyotluk hareketli ortalamasını hesaplamak için kıymetin n periyot içindeki değerleri toplanıp n sayısına bölünür.

$$SMA_n = \frac{\sum_{i=1}^n C_i}{n} \quad (1.26)$$

2. Ağırlıklı Hareketli Ortalama (WMA)

Bir menkul kıymetin n periyotluk kapanış değerinin ağırlıklı ortalamasını hesaplamak için, n periyot içindeki her bir kapanış değeri 1'den n'e kadar doğal sayılarla sırasıyla çarpılıp toplanır. Bulunan toplam 1'den n'e kadar olan doğal sayıların toplamına bölünür.

$$WMA_n = \frac{\sum_{i=1}^n iC_i}{\sum_{i=1}^n i} = \frac{C_1 + 2C_2 + 3C_3 + \dots + nC_n}{1 + 2 + 3 + \dots + n} \quad (1.27)$$

3. Üssel Hareketli Ortalama (EMA)

Üssel hareketli ortalamanın (Ema), ortalaması alınacak periyodun son dönemlerine daha fazla ağırlık veren bir yapısı vardır. Bu nedenle finansal piyasalarda sıkça kullanılır. Üssel ortalama fiyat değişimlerine basit hareketli ortalama göre daha hızlı tepki verir.

n günlük üssel ortalamanın hesaplanması Formül 5.11 de verilmiştir. Üssel hareketli ortalama hesaplanırken başlangıç için basit hareketli ortalama kullanılabilir.

$$EMA_{bugün} = EMA_{dün} + \frac{2}{n+1} \times (Kapanış Fiyatı_{bugün} - EMA_{dün}) \quad (1.28)$$

Yorumlanması :

En yalın haliyle hareketli ortalama göstergesi fiyat grafiği üzerine atılır. Fiyat hareketli ortalamanın üstüne çıktığında alım, fiyat hareketli ortalamanın altına düştüğünde satım yapılabilir.

İkinci bir yorum iki hareketli ortalama göstergesinin fiyat grafiği üzerinde kesişim noktaları alım ve satım noktaları olarak kullanılır. Bunun için uzun ve kısa periyotlu iki hareketli ortalama grafiği fiyat grafiği üzerinde çizdirilir. Kısa periyotlu olan hareketli ortalama, uzun periyotlu hareketli ortalamanın üstüne çıktığında alım, altına düştüğünde satım yapılır.

1.5.10. MACD İndikatörü

Macd göstergesi üssel iki hareketli ortalamanın farklarından oluşan bir göstergedir. Bu gösterge standart kullanımda 26 günlük üssel ortalama 12 günlük üssel ortalama çıkarılarak elde edilir. Ancak bu sayılar piyasa koşullarına göre istenirse değiştirilebilir. Seçilen bu iki üssel ortalama göre Macd indikatörü piyasadaki mevcut eğilimin yönünü tespit etmeye çalışır.



Şekil 1.10. MACD İndikatörü

Macd göstergesi negatif veya pozitif değerler alabilir. 26 periyotluk üssel ortalamanın 12 periyotluk üssel ortalama büyük olduğu durumlarda macd değeri pozitif, küçük olduğu durumlarda Macd değeri negatif eşit olduğu durumlarda Macd değeri sıfır olacaktır.

Yorumlanması :

Macd bir trend indikatörüdür. Macd değerinin pozitif olduğu durumlarda yukarı doğru trendin başladığı, Macd değerinin negatif olduğu durumlarda ise aşağı yönlü trendin başladığı kabul edilir. Bu durumda Macd değeri pozitif olduğunda alım, Macd değeri negatif olduğunda ise satım yapılabilir.

Macd indikatörü ayrıca kendi üssel ortalaması alınarakta kullanılabilir. Macd nin genellikle 9 periyotluk üssel ortalaması alınarak macd grafiği ile birlikte kullanılır. Macd nin üssel ortalaması alınarak elde edilen bu ortalama trigger çizgisi adı verilir. Bu kullanımda trigger çizgisi Macd göstergesini yukarı doğru keserse alım, aşağı doğru keserse satım yapılacaktır. Macd göstergesi uyumsuzlukların tespitinde de kullanılabilir.

1.5.11. Momentum İndikatörü

Momentum göstergesinde amaç seçilen periyodun başlangıcındaki kapanış fiyatı ile bugünkü kapanış fiyatının karşılaştırılıp artış veya azalış oranı hakkında yorumda bulunmaktır. Momentum fiyatların artış yönünü ve miktarını bulurken aynı zamanda artış ya da azalış şiddeti hakkında yorumda bulunmamızı sağlar.



Şekil 1.11. Momentum İndikatörü

Momentum göstergesinin kısa vadede daha başarılı sonuçlar vereceği düşünülür. Momentum göstergesinin 12-14 periyotluk zaman diliminde de başarılı sonuçlar verdiği görülür. Ancak piyasa koşulları ve işlem yapılacak kıymete göre periyot değiştirilebilir. Ancak fiyat hareketlerinin hızlı olduğu piyasalarda işlem yaparken

göstergenin periyodu arttırılması, fiyat hareketlerinin yavaş olduğu piyasalarda göstergenin periyodunun azaltılması olumlu sonuçlar verecektir.

Hesaplanması :

n periyotluk momentum göstergesinin değeri aşağıdaki gibi bulunur.

$$Momentum_n = \frac{\text{Son Gün Kapanış Fiyatı}}{n \text{ Gün önceki kapanış Fiyatı}} * 100 \quad (1.29)$$

Momentum göstergesi hesaplanırken Son gün kapanış fiyatı ile ilk gün kapanış fiyatının eşit olduğu durumda 100 çıktığı için bu değer göstergenin referans değeridir.

Yorumlanması :

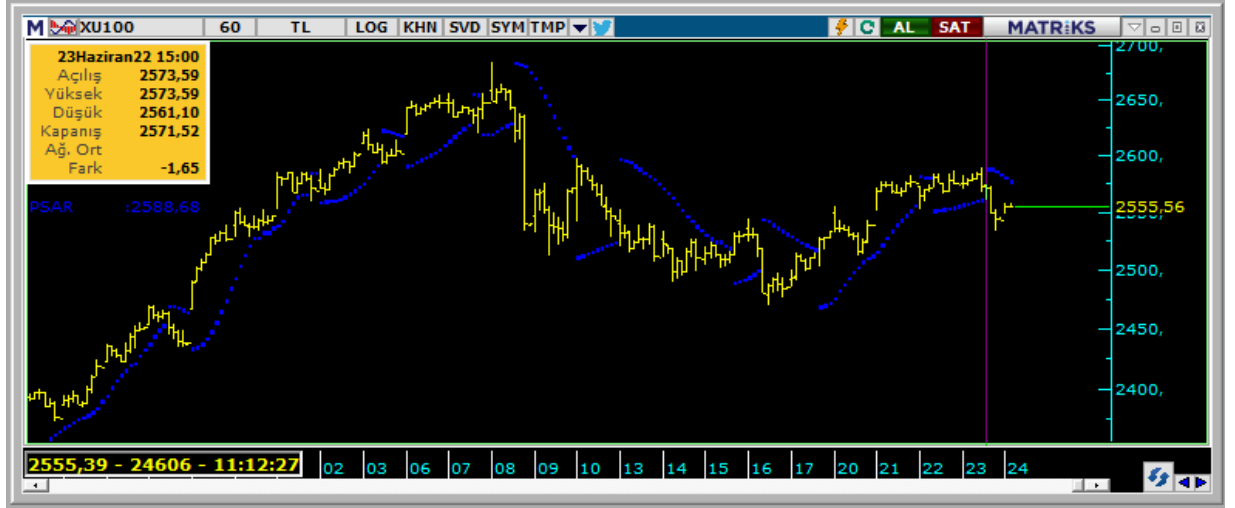
Momentum göstergesinde 100 referans değeri eşik kabul edilir. Momentum göstergesinin değeri 100 üzerine çıktığında alım yapılır, momentum değeri 100 altına düştüğünse ise satım yapılır.

Referans değerini kısa vade için aşırı alım ve aşırı satım sınırı olarak kabul eden bir yaklaşımda benimsenebilir. Bu durumda momentum göstergesinin değeri 100 üzerine çıkarsa fiyatlar aşırı alım bölgesinde olacağı için satım yapılır. Gösterge değeri 100 değerinin altına indiğinde ise alım yapılır. Ancak bu strateji yatay piyasalar için uygundur.

1.5.12. Parabolic Sar İndikatörü

Parabolic Sar göstergesini J.Welles Wilder bulmuştur. Parabolic Sar göstergesi grafik üzerinde çizildiğinde alım ve satım noktalarını grafik üzerinde belirleyen bir göstergedir. Sar, İngilizcede “Dur ve geri dön” anlamına gelen Stop And Reversal kelimelerinin baş harflerinin alınmasından elde edilen bir kısaltmadır.

Göstergenin pozisyonlara giriş ve çıkış noktalarını otomatik olarak tespit etmesi yanında kolay kullanımı göstergenin popüler olmasında etkili olmuştur. Hesaplanması oldukça uzun olan Parabolic sar değeri fiyatın altındaysa alım, fiyatın üstünde ise satım yapılır.



Şekil 1.12. Parabolic Sar İndikatörü

Yorumlanması :

Parabolic Sar göstergesi bir trend göstergesidir. Bu nedenle yatay piyasalarda başarılı sonuçlar üretmez. Bu göstergenin başka göstergelerle birlikte kullanılması halinde daha başarılı sonuçlar vermesi beklenir.

1.5.13. RSI İndikatörü (Relative Strength Index)

RSI göstergesi J. Welles Wilder tarafından bulunmuş bir göstergedir. Türkçeye göreceli güç endeksi olarak çevrilen RSI göstergesi, fiyatların aşırı alım veya aşırı satım bölgesinde olup olmadığını tespit eder. RSI göstergesi 0-100 aralığında değerler alan bir göstergedir. RSI göstergesinin değeri bizlere var olan trendin gücü hakkında bir fikir verir.



Şekil 1.13. RSI İndikatörü

Hesaplanması :

n periyotluk bir süre için herhangi bir senedin RSI değeri şu şekilde hesaplanır.

Son n günde senedin değerinin arttığı günlerin aritmetik ortalaması bulunur. (yukarı yönlü hareketin ortalaması). Yine son n günde senedin değerinin azaldığı günlerin aritmetik ortalaması bulunur. (aşağı yönlü hareketin ortalaması). Bu değerleri bulduktan sonra RSI değerinin bulunması Formül 5.13 te gösterilmiştir

$$RSI = 100 - \frac{100}{1 + \frac{\text{yukarı yönlü hareketin ortalaması}}{\text{aşağı yönlü hareketin ortalaması}}} \quad (1.30)$$

Yorumlanması :

RSI göstergesinin değeri 100'e yaklaştıkça fiyatlar aşırı alım bölgesinde, gösterge değeri 0'a yaklaştıkça fiyatlar aşırı satım bölgesindedir şeklinde düşünülür. Bu sebeple 30 ve 70 değerleri referans değerleri olarak kabul edilirler. Yani gösterge değeri 70 üstünde iken işlem yapılan kıymet aşırı alınmıştır, gösterge değeri 30 un altında iken işlem yapılan kıymet aşırı satılmıştır şeklinde yorumlanır.

RSI göstergesinin diğer bir kullanım şekli, gösterge 30 ve 70 değerlerini yukarı kırdığı anda alım işlemi, aşağı kırdığı anda ise satım işlemi yapılmasıdır.

RSI göstergesinin ortalama değeri 50'dir. RSI göstergesi 50 üstünde iken yukarı yönlü bir trend, gösterge değeri 50 altında iken aşağı yönlü bir trendin varlığından bahsedilir.

1.5.14. Stochastic İndikatörü

Stochastic göstergesi, Stochastic Slow ve Stochastic Fast şeklinde ikiye ayrılır. Adındanda anlaşılacağı üzere Stochastic Fast hızlı ve fiyat değişimlerine karşı çok duyarlı olduğundan dolayı bu gösterge yumuşatılarak Stochastic Slow göstergesi elde edilmiştir. Sochastic göstergesi 0 ile 100 arasında iki çizgiden oluşan bir göstergedir. Stochastic göstergesinin arkasındaki ana düşünce, mevcut fiyatın belirlenen periyot aralığındaki en yüksek ve en düşük değere olan uzaklığının yüzdesel bir ifadesi olarak özetlenebilir. Yani mevcut fiyatın, belirlenen periyot arasında aldığı en büyük ve en küçük değerler arasındaki konumuna göre yorumlanmasıdır.



Şekil 1.14. Stochastic Slow İndikatörü

Hesaplanması :

Stochastic göstergesi %K ve %D adı verilen iki çizgiden oluşur. %K çizgisinin hesaplanması aşağıdaki gibidir. Standart değerlerle kullanıldığında %K eğrisinin periyodu 5 olarak kullanılır. Bir senedin 5 günlük periyot için %K değeri Formül 5.14 te verilmiştir.

$$\%K = \frac{\text{Bugünün Kapanışı} - \text{Periyot içindeki min. değer}}{\text{Periyot içindeki max. değer} - \text{Periyot içindeki min. değer}} \times 100 \quad (1.31)$$

%D'nin değerini bulmak için yukarıda bulduğumuz %K'nın 3 günlük basit ortalaması alınır. %D ve %K eğrilerinden oluşan bu göstergeye Stochastic Fast göstergesi adı verilir. Ancak bu gösterge çok hızlıdır ve yavaşlatılması daha iyi sonuçlar verecektir. Bu göstergenin yavaşlatılmış hali Stochastic Slow göstergesidir.

Stochastic Slow göstergesini hesaplamak için Stochastic Fast'te bulduğumuz %D eğrisini kullanırız. Bu eğrinin bir kez daha 3 günlük basit ortalamasını alıp daha yavaş bir eğri elde ederiz. Stochastic Slow göstergesi için elde ettiğimiz bu iki eğriden yavaş olan %D, hızlı olan %K olacaktır.

Yorumlanması :

Stochastic Fast ve Stochastic Slow göstergeleri 0-100 arasında değerler alan göstergelerdir. Burada gösterge değerleri 100 e yaklaştığında senedin aşırı alım bölgesinde olduğu ve senedin mevcut fiyatının artık pahalı olduğu, gösterge değerleri 0'a yaklaştığında da senedin aşırı satıldığı ve bu bölgenin senet için artık ucuz olduğu yorumu yapılabilir. Bu yorumu yapmak için Stochastic göstergelerinin referans

değerleri genel olarak 20 ve 80 olarak alınır. Şöyle ki göstergenin değeri 80-100 aralığında ise bu bölgeden satım, göstergelerin değerleri 0-20 aralığında bu bölgeden alım yapılır.

Stokastik Fast ve Stochastic Slow göstergelerinin diğer bir kullanım şekli %K ve %D eğrilerinin kesişimi ile ilgilidir. Burada %K eğrisi %D eğrisini yukarı doğru keserse alım, aşağıya doğru keserse satım yapılır.

1.5.15. TRIX İndikatörü

Trix göstergesi hareketli ortalamaların fiyata olan duyarlılığını azaltmak için tasarlanmış bir göstergedir. Trix göstergesinin kurgusu belli bir periyotta bir hisse senedinin 3 defa üssel hareketli ortalaması alındıktan sonra bu ortalamanın yüzdesel olarak değişim miktarının bulunması üzerine kuruludur. Trix göstergesi hareketli ortalamaların kararlılığı arttırıp ve fiyattaki küçük değişimlere karşı bir nevi filtre görevi görür. Bu nedenle Trix göstergesi orta ve uzun vadede daha başarılı sonuçlar veren bir göstergedir.

Trix göstergesinde 0 referans değeridir. Göstergenin değeri sıfırdan küçük ise piyasada satış yönünde bir trend olduğu, göstergenin değeri 0 üzerinde ise piyasada alım yönünde bir trend olduğu varsayılır.



Şekil 1.15. Trix İndikatörü

Yorumlanması :

Trix göstergesinin en yaygın kullanımı bir hareketli ortalama ile kullanmaktır. Bu kullanımda Trix göstergesinin hareketli ortalaması alınır. Elde edilen göstergeye trigger çizgisi adı verilir. Trigger çizgisi ile Triks göstergesinin kesişim noktaları alım

satım noktaları olarak kullanılır. Trigger çizgisi Trix göstergesini yukarı doğru keserse alım, aşağı doğru keserse satım yapılır. Bu anlamıyla Trix göstergesinin kullanımı Macd göstergesinin kullanımına benzer.

Trix göstergesi aynı zamanda uyumsuzlukların tespitinde de kullanılabilir. Piyasa yükselirken Trix göstergesinin de değeri yükselmeli, Piyasa düşerken Trix göstergesinin de değeri düşmelidir. Aksi takdirde fiyatların mevcut trendin tersi yönünde bir düzeltme yapması beklenir.



2. ÖNCEKİ ÇALIŞMALAR

Literatür incelendiğinde yapay zeka yöntemleriyle endeks tahminin konusunda birçok çalışma mevcuttur. Bu çalışmaları, bir veya birden fazla Yapay Zeka modelinin birlikte kullanıldığı çalışmalar (Hibrit modeller) veya Yapay Zeka yöntemleriyle teknik analiz yöntemlerinin birlikte kullanılarak yapıldığı çalışmalar olarak 2 kategoriye ayırabiliriz.

2.1.1. Sadece Yapay Zeka Yöntemleriyle Yapılan Çalışmalar

- Pai ve Lin, (2004), Hisse senetlerinin tahmininde bulunmak için yaptıkları çalışmada ARIMA modeli ile Destek vektör makineleri modellerini hibritlemişlerdir. Sonuçta hibrid modelin ARIMA ve DVM modellerine göre tahmin öngörüsünün daha başarılı olduğu sonucuna ulaşmışlardır.
- Sutheebanjard ve Premchaiswadi, (2010), Geri yayımlı yapay sinir ağı yöntemiyle Tayland hisse senedi piyasasını tahmin etmeye çalışmışlardır.
- Aghababaeyan ve ark., (2011), Tahran borsa endeksinin (TSE) yönünü tahmin etmek için ileri beslemeli ve geri yayımlı bir YSA modeli kurmuşlardır. Bu çalışmada Borsa endeksinin yönünü %81 olasılıkla tahmin edebilmişlerdir.
- Güresen ve ark., (2011), NASDAQ 100 endeksini tahmin edebilmek için YSA ve GARCH yöntemlerini hibritleyerek bir model oluşturmuşlardır.
- Aygören ve ark., (2012), Çalışmalarında İMKB 100 endeksini tahmin etmek için Yapay sinir ağı modellerinden geri yayılım algoritması, nümerik arama modellerinden Newton modeli ve geleneksel zaman serileri yöntemlerinden ARMA(p,q) modeli ile tahmin etmeye çalışmışlardır. Sonuç olarak; Yapay sinir ağı modeli ile yaptıkları tahminin geleneksel zaman serileri ve nümerik arama modeline göre daha başarılı olduğunu gözlemlemişlerdir.
- Yavuz (2012) Yapay sinir ağı modelini kullanarak, Bist100 Sınai endeksindeki hisselerle risk ve getiri tahmini yapmış ve optimum portföy elde etmeye çalışmıştır.
- Yakut ve ark., (2014), Yapay Sinir Ağları ile birlikte ve DVM kullanarak Borsa Endeksinin tahmin edilebileceğini göstermişlerdir. Çalışmalarında Gecelik faiz oranı, Brezilya-İngiltere-Fransa-Almanya-Japonya borsalarının endeks

değerleri, Dolar kuru ve Bist100 endeksinin 1-2 ve 3 gün öncesinin değerlerini kullanmışlardır.

- Özçalıcı ve Ayriçay, (2016), Yaptıkları çalışmada yapay sinir ağları ile bir sonraki günün fiyat tahmini gerçekleştiren uzman bir sistem tasarlamaya çalışmışlardır. Bu çalışmada hangi değişkenlerin kullanılacağını ve kullanılacak değişkenlerin parametrelerini belirlemek için genetik algoritma tabanlı bir uzman sistem de tasarlanmıştır. Uzman sistemin diğer modellere göre daha başarılı sonuçlar verdiği gözlemlenmiştir.
- İnce ve Sönmez Çakır, (2017), çalışmalarında, Yapay Sinir Ağlarını ARIMA modeli ile birlikte kullanarak Nasdaq endeksinin değerlerini tahmin etmeye çalışmışlardır. Çalışmalarında Melez modellerin tek başına kullanılan modellerden daha başarılı oldukları sonucuna varmışlardır. (ARIMA; zaman serilerinin tahmini amacıyla kullanılan istatistik yöntemlerden biridir.)
- Pabuçcu (2019) Bist100 borsa endeksinin yönünü Yapay sinir ağı, destek vektör makinesi ve Naive Bayes makine öğrenme algoritmaları ile tahmin etmeye çalışmış ve yaptığı analizlerde destek vektör makineleri algoritmasının en iyi sonucu verdiğini gözlemlemiştir. (Naive Bayes olasılık tabanlı makine öğrenim algoritmasıdır.)
- Sarı (2019) Çalışmasında, hisse senedi getirilerinin yatırımcı duyarlılığı aracılığı ile tahmin edilmesini amaçlamış ve elde edilen bulgularda, destek vektör makinesi yöntemi ile dolaylı değişkenler kullanılarak yapılan tahminlerin daha başarılı sonuçlar verdiği görülmüştür.

2.1.2. Yapay Zeka Yöntemlerinin Teknik Analiz Göstergeleriyle Birlikte Kullanıldığı Çalışmalar

- Bahadır (2008) Çalışmasında, New York Borsası (NYSE) bünyesinde bulunan ETF'lerin değerlerini tahmin etmeye çalışmış bunun için Bayes teoremi ve yapay sinir ağlarını (YSA) kullanmıştır. Çalışmada iki yöntem kullanmış, birinci yöntemde Teknik analiz yöntemleri kullanılarak Bayes Karar modeli ile bir uzman sistem oluşturmuş. İkinci yöntemde yapay sinir ağı kullanılarak bir model uygulamıştır. Sonuçta yapay sinir ağının oluşturulan uzman sisteme

daha iyi sonuç verdiğini ancak her iki yönteminde tatmin edici sonuçlar verebileceğini gözlemlenmiştir.

- Şahin ve Özbayoğlu, (2014), çalışmalarında teknik analiz indikatörlerinden RSI kullanılmıştır. RSI için standart 14 periyodu ve 30-70 değerlerini Al-Sat olarak kullanmak yerine genetik algoritma kullanılarak her hisse için optimum RSI periyodu ve değeri keşfedilmeye çalışılmıştır. Bu çalışmada sadece genetik algoritma ile RSI optimizasyonu yapılmış ve ileri beslemeli yapay sinir ağı kullanılmamıştır. Değerlendirme yapmak için S&P 500 endeksi kullanılmıştır. Eğitim için 2001-2007 arası veriler, test için ise 2008-2012 arası veriler kullanılmıştır. Sonuçlar incelendiğinde ise borsanın yukarı ve aşağı yönde hareket ettiği zamanlarda optimize edilmiş RSI'nın, standart RSI 14'den hafif bir şekilde iyi performans gösterdiği görülmüştür. Bununla beraber finansal piyasanın iyi olduğu durumlarda ise RSI 14'ün optimize RSI'dan daha iyi performans gösterdiği görülmüştür. Fakat RSI 14'ün borsadaki yön değişimlerine karşı savunmasız olduğu vurgulanmıştır.
- Irmak (2019) yaptığı çalışmada RSI ve MACD göstergelerini öz düzenleyici haritalar kullanarak optimize etmiş ve yapay sinir ağları kullanarak al-tut stratejisinden daha başarılı sonuçlar veren bir model yapmayı başarmıştır.
- Kara (2019) çalışmasında Yapay Zeka yöntemlerini Teknik Analiz göstergeleriyle kullanılarak Borsa endekslerinin tahmininin yapılabileceğini ortaya koymuş aynı zamanda Yapay Zeka Yöntemlerinin tahmin performanslarını karşılaştırmıştır. Bunun için on farklı Teknik Analiz göstergesinin değerlerini Bist30 ve Bist100 endeksini tahmin için girdi olarak kullanmıştır. Çalışma sonunda Destek Vektör Makineleri, Lojistik Regresyon, Naive Bayes, Yapay Sinir Ağları, Karar Ağaçları, k-En Yakın Komşuluk yöntemleri içerisinde en başarılı sonucu Yapay Sinir Ağlarının verdiği görülmüştür.
- Kartal (2020) Bist100 ile birlikte DAX, S&P 500 ve NIKKEI 225 endekslerini Destek Vektör Makineleri kullanarak tahmin etmeye çalışmış elde ettiği modelin fiyatların artış yönlü olduğu trendlerde başarılı, azalış yönlü olduğu trendlerde ise başarılı sonuçlar veremediğini ortaya koymuştur.

- Arslankaya ve Toprak, (2021), yaptıkları çalışmada Polinom Regresyon, Random Forest Regresyon, Tekrarlayan Sinir Ağları (RNN) ve Uzun-Kısa Süreli Bellek (LSTM) yöntemlerini kullanarak hisse senetlerinin fiyatlarını tahmin etmeye çalışmış ve en iyi sonucu Random Forest Regresyon modelinin, en kötü sonucu ise Polinom Regresyon modelinin verdiğini gözlemlemişlerdir.



3. MATERYAL VE YÖNTEM

Bu çalışmada, denetimli öğrenme teknikleri kullanılarak çok değişkenli bir tahmin probleminin çözümü için makine öğrenimi teknikleri ile problemi en iyi modelleyecek yöntem belirlenmeye çalışılacaktır.

3.1. Materyal

Çalışma için Matriks IQ borsa işlem platformu kullanılacaktır. Matriks IQ programı C# programlama dili ile yazılmış bir borsa işlem platformudur.

Araştırmada girdi verisi olarak Bist100 endeksine ait 60 dakikalık ve günlük olmak üzere, 2 adet zaman serisi veri seti kullanılacaktır. Bu veri setleri Matriks IQ programından elde edilecektir.

Günlük veri setinde her bir gün için açılış değeri, kapanış değeri, gün içindeki en yüksek değeri ve gün içindeki en düşük değeri bulunmaktadır. Günlük veri setinden örnek bir kısmı Şekil 3.1. de gösterilmiştir.

Date	Time	Open	High	Low	Close
02.01.2020	00:00:00	1149.03	1159.32	1145.66	1159.32
03.01.2020	00:00:00	1151.43	1155.71	1126.17	1136.84
06.01.2020	00:00:00	1125.32	1129.73	1112.12	1114.13
07.01.2020	00:00:00	1125.94	1133.87	1123.48	1126
08.01.2020	00:00:00	1109.61	1133.42	1107.46	1128.76
09.01.2020	00:00:00	1152.37	1181.45	1152.37	1179.91
10.01.2020	00:00:00	1182.51	1191.4	1177.01	1186.64
13.01.2020	00:00:00	1195.95	1209	1195.95	1202.49
14.01.2020	00:00:00	1203.93	1215.04	1195.75	1214.23

Şekil 3.1. Günlük veri setinden örnek bir kısım

Günlük veri seti Bist100 endeksinin 14.06.2012 tarihinden 14.06.2022 tarihine kadar olan 3010 barlık(günlük) veriyi içermektedir. Veri setinin 14.06.2012-15.06.2020 tarihleri arasındaki 2509 barı eğitim verisi olarak, 16.06.2020-14.06.2022 tarihleri arasındaki 501 barı test verisi olarak kullanılacaktır. Günlük periyotta tahminini yapacağımız test veri setinin grafiği Şekil 3.2.'de verilmiştir.



Şekil 3.2. Günlük periyottaki test veri setinin grafiği

Test yapacağımız zaman serisi verisi günlük periyotta 501 bar (501 iş günü) ve yaklaşık 2 yıllık bir veridir. Başlangıç anında alıp hiç satmazsak(al-tut) yöne göre toplam kar yüzdesi %129,01 ve yöne göre toplam getirisi 1.431,3500 puandır.

60 dakikalık veri setinde her bir saat için Bist100 endeksinin açılış değeri, kapanış değeri, her bir saat için en yüksek değeri ve her bir saat için en düşük değeri bulunmaktadır. 60 dakikalık veri setinden örnek bir kısmı Şekil 3.3. de gösterilmiştir.

Date	Time	Open	High	Low	Close
06.08.2018	16:00:00	937.65	938.78	935.79	936.76
06.08.2018	17:00:00	936.76	944.72	936.44	942.63
06.08.2018	18:00:00	942.73	942.73	941.74	941.74
07.08.2018	10:00:00	946.23	954.37	946.23	952.13
07.08.2018	11:00:00	952.16	954.22	950.85	952.46
07.08.2018	12:00:00	952.48	952.91	946.83	948.39
07.08.2018	14:00:00	945.94	951.66	943.69	949.37
07.08.2018	15:00:00	949.37	952.99	947.58	951.18
07.08.2018	16:00:00	951.18	957.07	950.86	956.47

Şekil 3.3. 60 dakikalık veri setinden örnek bir kısım

60 dakikalık veri seti Bist100 endeksinin 12.11.2018 tarihinden 03.06.2022 tarihine kadar olan 7761 barlık(saatlik) veriyi içermektedir. Veri setinin 12.11.2018-28.09.2021 tarihleri arasındaki 6209 barı eğitim verisi olarak, 28.09.2021-03.06.2022 tarihleri arasındaki 1552 barı test verisi olarak kullanılacaktır. 60 dakikalık periyotta tahminini yapacağımız test veri setinin grafiği Şekil 3.4. de verilmiştir.



Şekil 3.4. 60 dakikalık periyotta test veri setinin grafiği

Test yapacağımız zaman serisi verisi 60 dakikalık periyotta 1552 bar ve yaklaşık 8,5 aylık bir veridir. Başlangıç anında alıp hiç satmazsak yöne göre toplam kar yüzdesi %88,56 ve yöne göre toplam getirisi 1.222,8700 puandır.

3.2. Yöntem

Araştırmamızın uygulama safhasını 3 bölümde ele alınacaktır.

Uygulamanın birinci bölümü; Aşağıda verilen makine öğrenmesi modellerinin her birinden günlük ve 60 dakikalık periyotlarda kullanılmak üzere 2 şer tane al-sat robotu oluşturulacaktır.

- a. Logistic Regresyon
- b. Lasso Regresyon
- c. Ridge Regresyon
- d. Support Vector Machine

Oluşturulan her bir robot kendi periyodundaki zaman serisinin eğitim veri seti ile eğitilecek, test veri seti ile performansı simüle edilecektir.

Al-Sat robotlarının her birine kendi periyodunda Bist100 endeksinin, açılış, kapanış, gün içi en yüksek ve en düşük değerleri verilerek endeksin bir sonraki gün kapanış değerinin yönü yüzdesel olarak tahmin edilecektir. Ayrıca simülasyonlarda tahmin yönümüz yukarı olduğunda alış, aşağı olduğunda satış yapacağımızı

varsayarak puan anlamında her bir robotun kaç puan kazandırdığı bulunacaktır. Tüm robotların aldığı puanlar karşılaştırılıp en fazla puan alan robotun modeli en başarılı sayılacaktır.

Uygulamanın İkinci Bölümü; Teknik Analiz göstergelerinin popüler olanlarından bazıları en başarılı model ile birlikte kullanılarak, oluşturulan modellerin performansı ölçümlenecek ve çıkarımlarda bulunulacaktır.

Uygulamanın Üçüncü Bölümü ise bütün yöntemler hakkında yorum yapılacak ve tahmin için kullanılacak en iyi yöntem hakkında sonuca varılacaktır.



4. ARAŞTIRMA BULGULARI VE TARTIŞMA

4.1. Günlük Periyotta Simülasyon Bulguları

4.1.1. Logistic Regresyon Günlük Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin her bir günlük periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın(günün) kapanış değeri Logistic Regresyon kullanılarak tahmin edilmeye çalışılmıştır. Modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.1. deki grafikte verilmiştir.



Şekil 4.1. Logistic Regresyon XU100_Day yöne göre toplam getiri yüzdesi

Modelimizin adı Logistic Regresyon_XU100_Day, yöne göre toplam kar yüzdesi %161,57 ve yöne göre toplam getirisi 1.770,1100 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 1'de verilmiştir.

Çizelge 4.1. Logistic Regression XU100_Day simülasyon bulguları

Sistem Adı	Logistic RegressionXU100 Day
Sembol	XU100
Periyod	Günlük
Eğitim Veri Seti Aralığı	14.06.2012-15.06.2020(2509 Bar)
Test Veri Set	16.06.2020-14.06.2022(501 Bar)
Test Verisindeki Bar Sayısı	501
Yön Tahmini (Başarılı / Başarısız)	311 / 190
Yön Tahmin Yüzdesi	%62,08
Trade Sayısı	12
Yöne Göre Toplam Getiri	1.770,1100
Trade Başına Ortalama Getiri	147,5093
Yüzde getirisi	%161,57

4.1.2. Lasso Regresyon Günlük Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin her bir günlük periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın(günün) kapanış değeri Lasso Regresyon kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.2. deki grafikte verilmiştir.



Şekil 4.2. Lasso RegresyonXU100_Day yöne göre toplam getiri yüzdesi

Modelimizin adı Lasso Regresyon_XU100_Day, yöne göre toplam kar yüzdesi %183.04 ve yöne göre toplam getirisi 2.005,3310 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.2.'de verilmiştir.

Çizelge 4.2. Lasso RegressionXU100_Day simülasyon bulguları

Sistem Adı	Lasso RegressionXU100_Day
Sembol	XU100
Periyod	Günlük
Eğitim Veri Seti Aralığı	14.06.2012-15.06.2020(2509 Bar)
Test Veri Seti Aralığı	16.06.2020-14.06.2022(501 Bar)
Test Verisindeki Bar Sayısı	501
Yön Tahmini (Başarılı / Başarısız)	312 / 189
Yön Tahmin Yüzdesi	%62,28
Trade Sayısı	36
Yöne Göre Toplam Getiri	2.005,3310
Trade Başına Ortalama Getiri	55,7036
Yüzde getirisi	%183,04

4.1.3. Ridge Regresyon Günlük Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin her bir günlük periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın(günün) kapanış değeri Ridge Regresyon kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.3. deki grafikte verilmiştir.



Şekil 4.3. Ridge RegresyonXU100_Day yöne göre toplam getiri yüzdesi

Modelimizin adı Ridge RegresyonXU100_Day, yöne göre toplam kar yüzdesi %178.06 ve yöne göre toplam getirisi 1.950,7910 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.3. te verilmiştir.

Çizelge 4.3. Ridge RegressionXU100_Day simülasyon bulguları

Sistem Adı	Ridge RegressionXU100_Day
Sembol	XU100
Periyod	Günlük
Eğitim Veri Seti Aralığı	14.06.2012-15.06.2020(2509 Bar)
Test Veri Seti Aralığı	16.06.2020-14.06.2022(501 Bar)
Test Verisindeki Bar Sayısı	501
Yön Tahmini (Başarılı / Başarısız)	303 / 198
Yön Tahmin Yüzdesi	%60,48
Trade Sayısı	86
Yöne Göre Toplam Getiri	1.950,7910
Trade Başına Ortalama Getiri	22,6836
Yüzde getirisi	%178,06

4.1.4. Destek Vektör Makinesi Günlük Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin her bir günlük periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın(günün) kapanış değeri Destek Vektör Makinesi kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.4. deki grafikte verilmiştir.



Şekil 4.4. Support Vector Machine_XU100_Day yöne göre toplam getiri yüzdesi

Modelimizin adı Support Vector Machine_XU100_Day, yöne göre toplam kar yüzdesi %61,68 ve yöne göre toplam getirisi 1.413,3510 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.4. de verilmiştir.

Çizelge 4.4. Support Vector Machine_XU100_Day simülasyon bulguları

Sistem Adı	Support Vector Machine_XU100_Day
Sembol	XU100
Periyod	Günlük
Eğitim Veri Seti Aralığı	14.06.2012-15.06.2020(2509 Bar)
Test Veri Seti Aralığı	16.06.2020-14.06.2022(501 Bar)
Test Verisindeki Bar Sayısı	501
Yön Tahmini (Başarılı / Başarısız)	309 / 192
Yön Tahmin Yüzdesi	%61,68
Trade Sayısı	5
Yöne Göre Toplam Getiri	1.413,3510
Trade Başına Ortalama Getiri	282,67
Yüzde getirisi	%129,01

4.2. 60 Dakikalık Periyotta Simülasyon Bulguları

4.2.1. Logistic Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin 60 dakikalık her bir periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın kapanış değeri Logistic Regresyon kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.5. deki grafikte verilmiştir.



Şekil 4.5. Logistic Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi

Modelimizin adı Logistic Regresyon_XU100_60Minute, yöne göre toplam kar yüzdesi %105,06 ve yöne göre toplam getirisi 1.450,7100 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.5. te verilmiştir.

Çizelge 4.5. Logistic Regression_XU100_60 Minute simülasyon bulguları

Sistem Adı	Logistic RegressionXU100_60Minute
Sembol	XU100
Periyod	60 Dakika
Eğitim Veri Seti Aralığı	12.11.2018-28.09.2021(6209 Bar)
Test Veri Seti Aralığı	28.09.2021-03.06.2022(1552 Bar)
Test Verisindeki Bar Sayısı	1552
Yön Tahmini (Başarılı / Başarısız)	853 / 699
Yön Tahmin Yüzdesi	%54,96
Trade Sayısı	42
Yöne Göre Toplam Getiri	1.450,7100
Trade Başına Ortalama Getiri	34,5407
Yüzde getirisi	%105,06

4.2.2. Lasso Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin 60 dakikalık her bir periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın kapanış değeri Lasso Regresyon kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.6. deki grafikte verilmiştir.



Şekil 4.6. Lasso Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi

Modelimizin adı Lasso Regression_XU100_60Minute, yöne göre toplam kar yüzdesi %149,02 ve yöne göre toplam getirisi 2.057,8300 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.6. da verilmiştir.

Çizelge 4.6. Lasso Regression_XU100_60Minute Simülasyon bulguları

Sistem Adı	Lasso Regression_XU100_60Minute
Sembol	XU100
Periyod	60 Dakika
Eğitim Veri Seti Aralığı	12.11.2018-28.09.2021(6209 Bar)
Test Veri Seti Aralığı	28.09.2021-03.06.2022(1552 Bar)
Test Verisindeki Bar Sayısı	1552
Yön Tahmini (Başarılı / Başarısız)	864 / 684
Yön Tahmin Yüzdesi	%55,93
Trade Sayısı	130
Yöne Göre Toplam Getiri	2.057,8300
Trade Başına Ortalama Getiri	15,8295
Yüzde getirisi	%149,02

4.2.3. Ridge Regresyon 60 Dakikalık Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin 60 dakikalık her bir periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın kapanış değeri Ridge Regresyon kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.7. deki grafikte verilmiştir.



Şekil 4.7. Ridge Regresyon_XU100_60Minute yöne göre toplam getiri yüzdesi

Modelimizin adı Ridge Regression_XU100_60Minute, yöne göre toplam kar yüzdesi %138.76 ve yöne göre toplam getirisi 1.916,1100 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.7. de verilmiştir.

Çizelge 4.7. Ridge Regression_XU100_60 dakikalık simülasyon bulguları

Sistem Adı	Ridge RegressionXU100_60Minute
Sembol	XU100
Periyod	60 Dakika
Eğitim Veri Seti Aralığı	12.11.2018-28.09.2021(6209 Bar)
Test Veri Seti Aralığı	28.09.2021-03.06.2022(1552 Bar)
Test Verisindeki Bar Sayısı	1552
Yön Tahmini (Başarılı / Başarısız)	842 / 710
Yön Tahmin Yüzdesi	%55,86
Trade Sayısı	297
Yöne Göre Toplam Getiri	1.916,1100
Trade Başına Ortalama Getiri	6.4515
Yüzde getirisi	%138,76

4.2.4. Support Vector Machine 60 Dakikalık Periyotta Simülasyon Bulguları

Modelimizde girdi olarak BİST100 endeksinin 60 dakikalık her bir periyodu için; açılış, kapanış, yüksek ve düşük değerleri kullanılmış ve bir sonraki barın kapanış değeri Support Vector Machine yöntemi kullanılarak tahmin edilmeye çalışılmıştır.

Yaptığımız modelin yöne göre toplam getiri yüzdesi (Kâr yüzdesi) Şekil 4.8. deki grafikte verilmiştir.



Şekil 4.8. Support Vector Machine_XU100_60Minute yöne göre toplam getiri yüzdesi

Modelimizin adı Support Vector Machine_XU100_60 dakikalık, yöne göre toplam kar yüzdesi %77.58 ve yöne göre toplam getirisi 1.071,2700 puandır. Yaptığımız modelin simülasyon bulguları Çizelge 4.8. de verilmiştir.

Çizelge 4.8. Support Vector Machine_XU100_60 Dakikalık simülasyon bulguları

Sistem Adı	Support Vector Machine_XU100_60Minute
Sembol	XU100
Periyod	60 Dakika
Eğitim Veri Seti Aralığı	12.11.2018-28.09.2021(6209 Bar)
Test Veri Seti Aralığı	28.09.2021-03.06.2022(1552 Bar)
Test Verisindeki Bar Sayısı	1552
Yön Tahmini (Başarılı / Başarısız)	849 / 703
Yön Tahmin Yüzdesi	%54,70
Trade Sayısı	6
Yöne Göre Toplam Getiri	1.071,2700
Trade Başına Ortalama Getiri	178,5450
Yüzde getirisi	%77,7

4.3. Bulguların Karşılaştırılması

4.3.1. Günlük Periyotlardaki Bulguların Karşılaştırılması

Lojistic Regresyon, Lasso Regresyon, Ridge Regresyon, Support Vector Machine Modelleri ile oluşturulan al-sat robotlarının her birine günlük periyotta Bist100 endeksinin, açılış, kapanış, gün içi en yüksek ve en düşük değerleri verilerek endeksin bir sonraki gün kapanış değerinin yönü yüzdesel olarak tahmin edildi. Simülasyonlarda tahmin yönümüz yukarı olduğunda alış, aşağı olduğunda satış yapacağımızı varsayarak puan anlamında her bir robotun kaç puan kazandığı bulundu.

Oluşturulan robotların aldıkları puanlar en yüksekten en düşüğe Çizelge 4.9. da sıralanmıştır.

Çizelge 4.9. Günlük periyotta simülasyon bulgularının karşılaştırılması

Sıralama	Kullanılan Tahmin Yöntemi	Puan	Yüzde Getiri
1.	Lasso Regresyon	2.005,3310	%183,04
2.	Ridge Regresyon	1.950,7910	%178,06
3.	Lojistic Regresyon	1.770,1100	%161,57
4.	Support-Destek Vector	1.413,3510	%61,68

Günlük periyotta test yaptığımız zaman serisi verisine göre, başlangıç anında alıp hiç satmazsak yöne göre toplam kar yüzdesi %129,01 ve yöne göre toplam getirisi 1.431,3500 puandır. Bu durumda Lasso, Ridge ve Lojistic Regresyon yöntemleriyle yapılan robotlar başarılı sonuçlar vermiştir denilebilir.

Bu sıralamaya göre, günlük periyotta Lasso Regresyon modeli ile yapılan robot en başarılı sonucu vermiştir.

4.3.2. 60 Dakikalık Periyotlardaki Bulguların Karşılaştırılması

Lojistic Regresyon, Lasso Regresyon, Ridge Regresyon, Support Vector Machine Modelleri ile oluşturulan al-sat robotlarının her birine 60 dakikalık periyotta Bist100 endeksinin, açılış, kapanış, gün içi en yüksek ve en düşük değerleri verilerek endeksin bir sonraki bar için kapanış değerinin yönü yüzdesel olarak tahmin edildi. Simülasyonlarda tahmin yönümüz yukarı olduğunda alış, aşağı olduğunda satış yapacağımızı varsayarak puan anlamında her bir robotun kaç puan kazandığı bulundu.

Oluşturulan robotların aldıkları puanlar en yüksekten en düşüğe Çizelge 4.10. de sıralanmıştır.

Çizelge 4.10. 60 dakikalık periyotta simülasyon bulgularının karşılaştırılması

Sıralama	Kullanılan Tahmin Yöntemi	Puan	Yüzde Getiri
1.	Lasso Regresyon	2.057,4100	%149,02
2.	Ridge Regresyon	1.916,1100	%138,76
3.	Lojistik Regresyon	1.450,7100	%105,106
4.	Support Vector Machine	1.071,2700	%77,7

60 dakikalık periyotta Lasso Regresyon modeli ile yapılan robot en başarılı sonucu vermiştir.

4.3.3. Tartışma

Günlük periyotta test süresi 60 dakikalık periyoda göre daha uzun olmasına rağmen getiri oldukça düşük çıkmıştır. Bu nedenle Bist100 endeksini tahmin için Günlük periyot yerine 60 dakikalık periyodun daha iyi sonuç verdiğini söyleyebiliriz.

Günlük ve 60 dakikalık periyotlarda birer düzenleştirme regresyonu olan Lasso ve Ridge Regresyon modelleri birbirine yakın değerler alarak sıralamada ilk iki sırayı almıştır. Bu iki regresyon modeli birer düzenleştirme regresyonudur ve temelde yapıları oldukça yakındır. Birer sınıflandırma yöntemi olan Lojistik Regresyon ve Destek Vektör Makinesi yöntemleri sıralamada son iki sırayı almıştır.

Tüm robotların aldıkları puanlar karşılaştırıldığında Lasso Regresyon modelinin daha başarılı bir model olduğu görülmüştür. Sonuç olarak yapılan simülasyonlarda Bist100 endeksinin tahmini için Lasso Regresyon yöntemi 60 dakikalık periyotta kullanıldığında en başarılı yöntem çıkmıştır.

Günlük ve 60 dakikalık periyotlarda yapılan simülasyonlarda Bist100 endeksinin tahmini için Lasso Regresyon yöntemi 60 dakikalık periyotta kullanıldığında en başarılı yöntem çıktığına göre, “Lasso Regresyon yöntemi ile Teknik Analiz göstergeleri 60 dakikalık periyotta birlikte kullanılırsa nasıl sonuçlar çıkar?” Sorusuna cevap aranacaktır. Burada hedef en başarılı robotun performansını arttıracak doğru göstergeleri bulmak olacaktır.

4.3.3.1 En Başarılı Yöntemin Teknik Analiz Göstergeleriyle Kullanılması

Bunun için oluşturduğumuz Lasso Regresyon al-sat robotuna Bist100 endeksinin her bir 60 dakikalık periyodu için, açılış, kapanış, en yüksek ve en düşük değerler ile birlikte teknik analiz göstergelerinden en çok bilinenleri girdi olarak verilerek endeksin bir sonraki bar için kapanış değerinin yönü yüzdesel olarak tahmin edilecektir. Simülasyonlarda tahmin yönümüz yukarı olduğunda alış, aşağı olduğunda satış yapacağımızı varsayarak puan anlamında her bir robotun kaç puan kazandığı bulunacaktır. Elde edilen Al-Sat robotlarının simülasyon sonucu karşılaştırılarak Bist100 endeksinin tahmini için en uygun model hakkında fikir edinilecektir. Teknik Analiz göstergeleriyle kullanılan Al-Sat Robotlarının simülasyon bulguları Çizelge 4.11. de verilmiştir.

Çizelge 4.11. Teknik Analiz göstergeleriyle kullanılan al-sat robotlarının simülasyon bulguları

Kullanılan Yardımcı Gösterge	Yön Tahmini (Başarılı / Başarısız)	Yön Tahmin Yüzdesi	Trade Sayısı	Yöne Göre Toplam Getiri	Yöne Göre % Getiri
AROON	861/691	%55,48	455	2.658,9900	%192,56
RSI	869/683	%55,99	369	2.563,4500	%185,64
Parabolic Sar	858/694	%55,28	433	2.530,2300	%183,23
Trix	870/682	%56,06	339	2.479,6500	%179,57
Stochactic Slow	865/687	%55,73	331	2.452,6090	%176,61
CCI	866/686	%55,80	317	2.423.6300	%175,51
EMA	862/690	%55,54	351	2.317,9100	%167,86
MACD	861/691	%55,48	319	2.307,2700	%167,09
Dema	862/690	%55,54	214	2.274,3900	%164,71
Bollinger Bandı	856/696	%55,15	315	2.260,7300	%163,72
Envelopes	861/691	%55,48	359	2.259,8300	%163,65
ADX	857/695	%55,22	293	2.077,9300	%150,48

4.3.4. Bulgular

Lasso Regresyon modeliyle yaptığımız al-sat robotunun, yöne göre toplam kar yüzdesi %149,02 ve yöne göre toplam getirisi 2.057,4100 puan olduğundan, Çizelge 7.3.3 teki Teknik Analiz göstergelerinin model performansına katkısı sırasıyla aşağıdaki gibi olmuştur.

1. Aroon Göstergesi robota girdi olarak eklendiğinde, robotun performansı 601,58 puan ve %29,23 artmıştır.
2. RSI Göstergesi robota girdi olarak eklendiğinde, robotun performansı 506,04 puan ve %24,60 artmıştır.
3. Parabolic Sar Göstergesi robota girdi olarak eklendiğinde, robotun performansı 472,82 puan ve %22,98 artmıştır.
4. Trix Göstergesi robota girdi olarak eklendiğinde, robotun performansı 422,24 puan ve %20,52 artmıştır.
5. Stochastic Slow Göstergesi robota girdi olarak eklendiğinde, robotun performansı 395,20 puan ve %19,21 artmıştır.
6. CCI Göstergesi robota girdi olarak eklendiğinde, robotun performansı 366,22 puan ve %17,80 artmıştır.
7. Ema Göstergesi robota girdi olarak eklendiğinde, robotun performansı 260,50 puan ve %12,66 artmıştır.
8. Macd Göstergesi eklendiğinde, robota girdi olarak robotun performansı 249,86 puan ve %12,14 artmıştır.
9. Dema Göstergesi robota girdi olarak eklendiğinde, robotun performansı 216,98 puan ve %10,55 artmıştır.
10. Bollinger Bandı robota girdi olarak eklendiğinde, robotun performansı 203,32 puan ve %9,89 artmıştır.
11. Envelopes Göstergesi robota girdi olarak eklendiğinde, robotun performansı 202,42 puan ve %9,84 artmıştır.
12. Adx Göstergesi robota girdi olarak eklendiğinde, robotun performansı 20,52 puan ve %1 artmıştır.

Bu sonuca göre Adx göstergesi dışında tüm göstergelerin robot performansını arttırdığı gözlemlenmiştir. Ancak özellikle ilk 6 göstergenin yani Aroon, RSI, Parabolic Sar, Trix, Stochastic Slow ve CCI nin katkısı daha fazla olduğu gözlemlenmiştir.

5. SONUÇ VE ÖNERİLER

Günlük periyotta kullanılan test verisi eğitim seti 501 Bar (501 iş günü), 60 dakikalık periyotta kullanılan test verisi eğitim seti 1552 Bar (194 iş günü) dır. Günlük periyotta test süresi 60 dakikalık periyoda göre daha uzun olmasına rağmen getiri 60 dakikalık periyoda göre oldukça düşük çıkmıştır. Bunun yanında getiri eğrisinden de açıkça görülebileceği gibi günlük periyotta pozisyona geç girilip geç çıkıldığı için yaşanan kayıpların daha fazla olduğu görülmüştür. Bu nedenle çalışmamızın sonucu olarak Bist100 endeksini tahmini için günlük periyot yerine 60 dakikalık periyodun daha iyi sonuç verdiğini söyleyebiliriz.

Çalışmamızda birer düzenleştirme regresyonu olan Lasso ve Ridge Regresyon yöntemlerinin birer sınıflandırma yöntemi olan Lojistic regresyon ve Destek Vektör Makinesi yöntemlerine göre daha başarılı sonuçlar verdiği gözlemlenmiştir.

Yaptığımız çalışmada sonuçlar beklediğimiz değerden oldukça yüksek çıkmıştır. Bunun en önemli nedeni çalışmayı yaptığımız dönemde tesadüfi olarak borsada ciddi bir yükseliş trendi söz konusudur. Başka dönemlerde sonuçların bu kadar yüksek olmayabileceği dikkate alınmalıdır. Bir diğer neden çalışmada kullanılan Teknik Analiz göstergeleri standart değerlerde kullanılmasına rağmen kurulan yapay zeka modelleri bir çok kez eğitilmiştir. Bu durum aşırı öğrenmeye sebep olmuş olabilir. Ancak bu çekincelere rağmen bu modellerin al-tut stratejisine göre oldukça başarılı sonuçlar verdiği gözlemlenmiştir.

5.1. Önerilen Model

Tüm robotların aldıkları puanlar karşılaştırıldığında. Lasso Regresyon modeline girdi olarak Teknik Analiz göstergelerinden Aroon, RSI, Parabolic Sar, Trix, Stochastic Slow ve CCI den herhangi biri eklendiğinde sonucun anlamlı bir şekilde değiştiği gözlemlenmiştir. Bu modellerin kullanıldığı robotlar sıralamada en iyi sonucu veren ilk 6 modeldir. Bu göstergeler Lasso Regresyon robotunun performansını %17,80 ile %29,24 oranında arttırmıştır.

Bunlar içerisinde Aroon göstergesinin kullanıldığı model en yüksek getiriye sahip modeldir. Ancak, simülasyonun sonucu, göstergelerin kullanılan değerleri, robotun eğitilme sıklığı, eğitim ve test için seçilen zaman aralığı verilerine göre değiştiğinden, Bist100 endeksini tahmin etmek için önerilen model olarak en başarılı

sonucu veren tek bir model yerine, en başarılı sonuçları veren ilk 6 modelin olması uygun olacaktır.

Önerilen modellerin getirileri için şu bilgiyi dikkate almakta fayda var. Şöyle ki, Bist100 endeksine test süresi boyunca Al-Tut stratejisi uygulansa kazanç %88,56 olacaktı. Bu nedenle önerilen robotların getirileri al-tut stratejisi ile elde edilecek gelirin yüzdesel artışına göre hesaplanmıştır.

Çizelge 5.12. Lasso ile birlikte Kullanılan Göstergelerin Sonuçları

Girdi Olarak Kullanılan Veriler <i>Açılış, Kapanış, En Yüksek, En Düşük</i>	Kullanılan Tahmin Yöntemi	Al-Tut Stratejisine Göre Yüzdesel Artış
Aroon Göstergesi	Lasso Regresyon	%117,44
RSI Göstergesi	Lasso Regresyon	%109,63
Parabolic Sar Göstergesi	Lasso Regresyon	%106,91
Trix Göstergesi	Lasso Regresyon	%102,77
Stochastic Slow Göstergesi	Lasso Regresyon	%100,56
CCI Göstergesi	Lasso Regresyon	%98,19

***Kullanılan Periyotların hepsi 60 dakikalıktır.**

Sonuç olarak Bist100 endeksini tahmin etmek için önerilen modeller ve bu modellerin al-tut stratejisine göre getirileri yukarıdaki tabloda özetlenmiştir.

KAYNAKLAR

- Aghababaeyan, Siddiqui ve ark., 2011. Forecasting the Tehran Stock Market by Artificial Neural Network (IJACSA). **International Journal of Advanced Computer Science and Applications**, Vol. 6, No. 8
- Akçay, A., 2009. Broiler Coccidiosis'inde Risk Faktörlerinin Lojistik Regresyon Analizi ile Belirlenmesi. **Ankara Üniversitesi Sağlık Bilimleri Enstitüsü** Doktora Tezi
- Anonim, 2022. <https://azure.microsoft.com/tr-tr/resources/cloud-computing-dictionary/what-is-machine-learning-platform/#techniques>
- Arı, A. ve Önder, H., 2013. Farklı Veri Yapılarında Kullanılabilecek Regresyon Yöntemleri. **Anadolu Tarım Bilim Dergisi**, 28(3):168-174
- Arslankaya, S. and Toprak, Ş., 2021. Makine Öğrenmesi ve Derin Öğrenme Algoritmalarını Kullanarak Hisse Senedi Fiyat Tahmini. **Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi**, 13(1): 178-192.
- Aslan, S. ve Yıldız, T., 2022. Makine Öğrenmesinde Rastgele Oran ve Sıralı Küme Örneklemesi Yöntemlerinin Doğrusal Regresyon Modellerine Etkisi. **DEUFMD**, 24(70): 29-36.
- Aygören, H., Sarıtaş, H. ve Moralı, T., 2012. İMKB 100 Endeksinin Yapay Sinir Ağları ve Newton Nümerik Arama Modelleri İle Tahmini. **Uluslararası Alanya İşletme Fakültesi Dergisi**, 4(1): 73 88.
- Aytekin, H.T., 2021. Makine Öğreniminin Araştırmacıların Veri Analizi Bağlamında Potansiyel Önemi. **Ufuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, 10(19): 85-106.
- Bahadır, İ., 2008. Bayes Teoremi ve Yapay Sinir Ağları Modelleriyle Borsa Gelecek Değer Tahmini Uygulaması. **TOOB Ekonomi ve Teknoloji Üniversitesi Fen Bilimleri Enstitüsü** Yüksek Lisans Tezi
- Bardak, B., 2016. Hastalık salgınlarının internet erişim ve arama verisi kullanılarak tahmini. Ankara: **TOBB ETÜ Fen Bilimleri Enstitüsü** Yüksek Lisans Tezi
- Bircan, H., 2004. Lojistik Regresyon Analizi: Tıp Verileri Üzerine Bir Uygulama. **Kocaeli Üniversitesi Sosyal Bilimler Dergisi**, 2(8): 185-208. Retrieved from <https://dergipark.org.tr/tr/pub/kosbed/issue/25712/271314>
- Derici Yıldırım, D. ve Çiftçi, A.T., 2021. Etkili değişkenlerin cezalı regresyon yöntemleri ile belirlenmesi: Diyabet veri kümesi üzerine bir uygulama. **Mersin Üniversitesi Sağlık Bilimleri Dergisi**, 14(1): 105-112. DOI: 10.26559/mersinsbd.816561
- Eray, O., 2008. Destek Vektör Makineleri İle Ses Tanıma Uygulaması. **Pamukkale Üniversitesi Fen Bilimleri Enstitüsü** Yüksek Lisans Tezi
- Gök, A.C., 2010. İşletmelerin Tahminleme Sürecinde Bulanık Doğrusal Regresyon Analizi ve Lojistik Regresyon Analizinin Uygulanması. **Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü** Yüksek Lisans Tezi

- Güresen, E., Kayakutlu, G., and Daim, T.U., 2011. Using artificial neural network models in stock market index prediction. **Expert Systems with Applications**, 38(8): 10389-10397. doi:10.1016/j.eswa.2011.02.068
<https://dergipark.org.tr/tr/pub/ufuksbedergi/issue/63452/911897>
- Ince, H. and Sonmez Cakir F., 2017. Analysis of financial time series with model hybridization. **Journal of Economics, Finance and Accounting (JEFA)**, 4(3): 331-341. Permement link to this document: <http://doi.org/10.17261/Pressacademia.2017.700>
- Irmak, H., 2019. Yapay Zekâ Kullanılarak Borsa İstanbul (Bist) İçin Algoritmik İşlem Stratejilerinin Geliştirilmesi. **Hacettepe Üniversitesi Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans Tezi**
- Kara, İ., 2019. Borsa Endeksi Hareket Yönünün Tahmininde Sınıflandırma Yöntemlerinin Performanslarının Karşılaştırılması: Bist100 Örneği. **Afyon Kocatepe Üniversitesi Sosyal Bilimler Enstitüsü Doktora Tezi**
- Kara, Ö.S., 2015, Lojistik Regresyon Analizi ve Kadın İşgücü Üzerine Bir Uygulama. Uludağ Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi
- Karakaş, S., 2008. Çoklu doğrusal bağlantı problemi ve yanlı regresyon tahmincileri. **İstanbul Üniversitesi Sosyal Bilimler Enstitüsü Yüksek Lisans Tezi**
- Kartal, C. (2020). Destek Vektör Makineleri ile Borsa Endekslerinin Tahmini. **İnsan ve Toplum Bilimleri Araştırmaları Dergisi**, 9(2): 1394-1418. DOI: 10.15869/itobiad.673015.
- Kutlu, B. ve Badur, B., 2009, Yapay sinir ağları ile borsa endeksi tahmini. **İstanbul Üniversitesi İşleme İktisadi Enstitüsü Yönetim Dergisi**, 63: 25-40.
- Küçük, A., 2019. Doğrusal Regresyonda Ridge, Liu ve Lasso Tahmin Edicileri Üzerine Bir Çalışma. **Hacettepe Üniversitesi İstatistik Anabilim Dalı Yüksek Lisans Tezi**
- Nacar, E.N. ve Erdebilli, B., 2021. Makine Öğrenmesi Algoritmaları ile Satış Tahmini. **Endüstri Mühendisliği Dergisi**, 32 (2): 307-320.
- Özçalıcı, M. ve Ayriçay Y., 2016. Bilgi İşlemsel Zeka Yöntemleri İle Hisse Senedi Fiyat Tahmini: Bist Uygulaması. **Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi**, 25(1): 274-298.
- Pabuçcu, H., 2019. Borsa Endeksi Hareketlerinin Tahmini: Trend Belirleyici Veri. **Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi**, 22(1): 246-256.
- Pai, P.F. and Lin, C.S., 2004. A Hybrid Arima and Support Vector Machines Model in Stock Price Forecasting. **Department of Industrial Engineering and Technology Management**, Da-Yeh University, 112 Shan Jiau Road, Da-Tsuen, Chung-Hwa 515, Taiwan.
- Parlak, T., 2019, Lojistik Regresyonda Robust Tahmin Yöntemlerinin Kullanılması. **Ankara Üniversitesi İstatistik Anabilim Dalı Yüksek Lisans Tezi**

- Sahin, U. and Ozbayoglu, A.M., 2014. TN-RSI: Trend-normalized RSI Indicator for Stock Trading Systems with Evolutionary Computation. **Procedia Computer Science**, 36(2014): 240-245. DOI: 10.1016/j.procs.2014.09.086.
- Sarı, S., 2019. Borsa İstanbul Hisse Senedi Getirilerinin Yatırımcı Duyarlılığı Aracılığıyla Tahmin Edilmesi. **Erzincan Binali Yıldırım Üniversitesi Sosyal Bilimler Enstitüsü** Doktora Tezi
- Sutheebanjard P. and Premchaiswadi W., 2010. Forecasting The Thailand Stock Market Using Evolution Strategies. **Phaisarn Sutheebanjard And Wichian Premchaiswadi Graduate School Of Information Technology**, Siam University Bangkok 10163, Thailand.
- Şahin, O., 2017. Lojistik Regresyon ile Ayvalık'ı Turizm Amacıyla Tercih Etmede Önemli Risk Faktörlerinin Belirlenmesi. **Elektronik Sosyal Bilimler Dergisi**, 16 (61): 647-660. DOI: 10.17755/esosder.304700.
- Topdağ, D. ve Acar, T., 2021. Modernizasyon Teorisi Bağlamında Demokrasinin Sosyal Gereksinimleri: Ridge Regresyon Yaklaşımından Ampirik Bulgular. **Doğuş Üniversitesi Dergisi**, 23(1): 39-57. DOI: 10.31671/doujournal.921754.
- Vural, A., 2007. Aykırı Değerlerin Regresyon Modellerine Etkileri ve Sağlam Kestiriciler. **Marmara Üniversitesi Sosyal Bilimler Enstitüsü** Yüksek Lisans Tezi
- Yakut, E., Elmas, B., Yavuz, S., 2014. Yapay Sinir Ağları ve Destek Vektör Makineleri Yöntemleriyle Borsa Endeksi Tahmini. **Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi**, 19(1): 139-157.
- Yavuz, E., 2017. Çoklu Doğrusal Bağlantı Problemi ve Yanlı Regresyon Tahmincileri. **Kahramanmaraş Sütçü İmam Üniversitesi Fen Bilimleri Enstitüsü** Yüksek Lisans Tezi
- Yavuz, M., 2012. Yapay Sinir Ağları ile Portföy Optimizasyonu. **Balıkesir Üniversitesi Fen Bilimleri Enstitüsü** Yüksek Lisans Tezi