



NVIDIA HISSE SENEDİ FİYAT TAHMINİ

Yazılım Mühendisliği Ana Bilim Dalı

Dönem Projesi

Jordi Hako Nintcheu

ORCID 0000-0003-3115-5251

Proje Danışmanı: Prof. Dr. Aytuğ Onan

Haziran 2025

Nvidia Hisse Senedi Fiyat Tahmini

Öz

Bu projede, NVIDIA hisse senedi fiyatının bir sonraki gün için tahmin edilmesinde istatistiksel, makine öğrenimi ve derin öğrenme tabanlı modellerin başarıları karşılaştırılmıştır. Çalışmada 2004-2024 yıllarını kapsayan geniş bir veri seti kullanılmış ve ARIMA, Random Forest, Doğrusal Regresyon, RNN, CNN ve LSTM modelleri test edilmiştir. Sonuçlar, özellikle 30 günlük kayan pencere ARIMA'nın en iyi performansı gösterdiğini; RNN'in ise derin öğrenme modelleri arasında en etkili yöntem olduğunu ortaya koymuştur. Elde edilen bulgular, model ve özellik seçiminin finansal tahmin doğruluğu için kritik olduğunu göstermektedir.

Anahtar Sözcükler: Hisse senedi fiyat tahmini, derin öğrenme, makine öğrenimi, ARIMA, NVIDIA, zaman serisi, finansal analiz

Stock Price Prediction for Nvidia

Abstract

This project compares the effectiveness of statistical, machine learning, and deep learning models in predicting the next-day closing price of Nvidia's stock. Using an extensive dataset spanning 2004 to 2024, several models—including ARIMA, Random Forest, Linear Regression, RNN, CNN, and LSTM—were evaluated. The results show that ARIMA with a 30-day rolling window achieved the best overall performance, while RNN emerged as the most effective deep learning method. The findings highlight that model selection and feature engineering play a critical role in the accuracy of financial forecasting.

Keywords: Stock price prediction, deep learning, machine learning, ARIMA, Nvidia, time series, financial analysis

Teşekkür

Bu proje çalışmasının her aşamasında, değerli bilgi birikimi, yönlendirmeleri ve desteği ile katkı sağlayan danışmanım Prof. Dr. Aytuğ Onan'a en içten teşekkürlerimi sunarım. Kendisi tarafından verilen Yapay Zekâ, Makine Öğrenmesi ve ilgili alanlardaki dersler, hem bu projede hem de akademik yolculuğumda bana ilham kaynağı olmuştur. Ayrıca proje sürecindeki özverili danışmanlığı, yol göstericiliği ve paylaştığı güncel bilgiler sayesinde, bu çalışmayı başarıyla tamamlayabildim. Destekleri için hocama bir kez daha teşekkür ederim.

İçindekiler

Öz	i
Abstract.....	ii
Teşekkür	iii
Şekiller Listesi.....	vi
Tablolar Listesi	vii
Kısaltmalar Listesi	viii
Semboller Listesi	11
İlgili Çalışmalar	3
2.1Geleneksel İstatistiksel Modeller.....	3
2.2Makine Öğrenimi Modelleri	4
2.3Makine Öğrenimi Modelleri	4
2.4Hibrit ve Gelişmiş Modeller	5
2.5Literatür Taraması Sonuç.....	6
3.1Borsa Temelleri	7
3.2Zaman Serisi Verilerinin Özellikleri	8
3.3NVIDIA Hisse Senedinin Uygunluğu	8
3.4Tahmine Dayalı Modeller	9
3.4.1İstatistiksel Yöntemler	10

3.4.2Makine Öğrenimi Yöntemleri	10
3.4.3Derin Öğrenme Mimarileri	11
3.5Veri Kümesi ve Özellik Mühendisliği	13
3.6Veri Ön İşleme ve Dönüşümler	15
3.7Model Değerlendirme Yöntemleri.....	16
3.8Öznitelik (Feature) Seçimi ve Önemi	17
3.9Model Seçimi ve Hiperparametre Ayarlaması	17
4.1Deney 1: ARIMA Modeli	19
4.2Deney 2: Tam Özellik Kümesi ile Makine Öğrenimi ve Derin Öğrenme Modelleri	20
4.3Deney 3: Azaltılmış Özellik Kümesi ile Makine Öğrenimi ve Derin Öğrenme Modelleri	21
Sonuç	23

Şekiller Listesi

Şekil 3.1	Hisse Senedi Mum Çubuğu açıklaması	9
Şekil 3.2	Random Forest Model Akışı	13
Şekil 3.3	Derin Öğrenme Modelleri Mimarisi	14
Şekil 3.4	CNN Model Architecture	15
Şekil 3.5	Giriş, çıkış ve unutma kapılarına sahip bir gözetleme deliği LSTM birimi (Wikipedia, t.y.)	15
Şekil 3.6	Eğitim, Doğrulama, Test Verilerinin Görselleştirilmesi	17
Şekil 3.7	Veri ön işleme adımlarının görselleştirilmesi	11
Şekil 4.1	30 boyutunda yuvarlanan pencere ARIMA modeli Performans.....	22
Şekil 4.2	ML ve DL Modellerinin Performansı – Tam Özellikler	23
Şekil 4.3	ML ve DL Modellerinin Performansı - Azaltılmış Özellikler	24

Tablolar Listesi

Tablo 4.1	ARIMA modelleri değerlendirme metrikleri	21
Tablo 4.2	ML vs. DL (Tam Özellikler) değerlendirme metrikleri	22
Tablo 4.3	ML vs. DL (Reduced Features) evaluation metrics	24

Kısaltmalar Listesi

AI	Artificial Intelligence
ML	Machine Learning
NN	Neural Network
DL	Deep Learning
ARIMA	Autoregressive Integrated Moving Average
LSTM	Long Short-Term Memory
BiLSTM	Bidirectional Long Short-Term Memory
CNN	Convolutional neural network
RNN	Recurrent Neural Network
SVM	Support Vector Machine
SVR	Support Vector Regression
MLP	Multi Layer Perceptron
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
MAPE	Mean Absolute Percentage Error
RMSE	Root Mean Squared Error
MAG 7	Magnificent 7
EMA	Exponential Moving Average
MA	Moving Average
MACD	Moving Average Convergence Divergence
RSI	Relative Strength Index

Bölüm 1

Giriş

Finansal piyasalar, yatırımcılar ve analistler için önemli karar süreçlerinin merkezinde yer almakta ve doğru hisse senedi fiyat tahmini, bu süreçlerin temel taşlarından biri olmaktadır. Hisse senedi fiyatlarının doğru şekilde tahmin edilmesi, yatırımcılara avantaj sağlamanın yanında, finansal riskleri yönetmede ve etkin portföy oluşturulmasında kritik bir rol oynamaktadır. Özellikle teknoloji sektöründe öncü firmalardan biri olan NVIDIA, sahip olduğu yüksek piyasa hacmi ve dalgalı fiyat hareketleri ile hem akademik çalışmalar hem de yatırımcılar açısından dikkat çeken bir hisse senedi konumundadır.

Bu projede, NVIDIA'nın hisse senedi fiyatının ertesi gün için tahmin edilmesi amacıyla; istatistiksel yöntemler, makine öğrenimi ve derin öğrenme tabanlı modeller karşılaştırmalı olarak incelenmiştir. Projede kullanılan veri seti, 2004-2024 yılları arasındaki günlük kapanış fiyatlarını içermektedir. ARIMA, Random Forest, Doğrusal Regresyon, Tekrarlayan Sinir Ağları (RNN), Konvolüsyonel Sinir Ağları (CNN) ve Uzun Kısa Süreli Bellek (LSTM) gibi farklı model ve algoritmalar değerlendirilmiş; elde edilen sonuçlar doğruluk, hata oranı ve model verimliliği açısından karşılaştırılmıştır.

Bu çalışma; bilimsel yazım kurallarına, terim ve ifade bütünlüğüne ve akademik etik ilkelere uygun olarak hazırlanmıştır. Tüm kaynaklar ve alıntılar uygun şekilde referans verilmiş, intihalden kaçınılmıştır. Proje metni, kolay anlaşılır, sade ve akıcı bir Türkçe ile kaleme alınmış; anlatımda kısa ve öz cümleler tercih edilmiştir. Projede, bilimsel araştırma yöntemleri esas alınmış ve sunulan analizler nesnel verilere dayandırılmıştır.

Hazırlanan bu dönem projesinin, hem finansal veri analizi hem de yapay zekâ temelli modelleme tekniklerinin finansal tahmin alanındaki uygulamaları konusunda faydalı bir kaynak olması amaçlanmıştır.

Bölüm 2

İlgili Çalışmalar

Hisse senedi fiyat tahmini, finansal araştırmaların ve veri biliminin en yoğun ilgi gören alanlarından biri olmuştur. Zaman içinde, finansal verilerdeki karmaşık örüntüleri anlamak ve gelecekteki fiyat hareketlerini daha doğru tahmin edebilmek için birçok yöntem geliştirilmiştir. Bu kapsamda; geleneksel istatistiksel modeller, makine öğrenimi tabanlı algoritmalar ve derin öğrenme yaklaşımları finansal tahmin uygulamalarında yaygın olarak kullanılmıştır. Bu bölümde, literatürde yer alan ve hisse senedi fiyat tahmini alanında öne çıkan yöntemler ile bu yöntemlerin avantajları ve sınırlılıkları özetlenmektedir. Ayrıca, ilgili çalışmaların sonuçları doğrultusunda bu projede tercih edilen modellerin seçilme nedenleri açıklanmaktadır.

2.1 Geleneksel İstatistiksel Modeller

Geleneksel istatistiksel modeller, finansal zaman serisi analizinde uzun yıllardır güvenilir araçlar olarak kullanılmaktadır. Bu modellerin en yaygın örneklerinden biri olan ARIMA (Otoregresif Bütünleşik Hareketli Ortalama) modeli, özellikle finansal verilerdeki doğrusal eğilimleri ve kısa vadeli dalgalanmaları yakalamada etkili olmuştur. ARIMA ve benzeri yöntemler, geçmiş fiyat bilgileri üzerine inşa edilerek, gelecekteki fiyat tahminlerini matematiksel bir çerçevede gerçekleştirir.

Literatürde, ARIMA ve GARCH gibi modellerin finansal piyasalardaki volatilitayı ve risk düzeylerini analiz etmede başarılı oldukları gösterilmiştir. Özellikle dalgalı piyasa koşullarında ARIMA-GARCH kombinasyonlarının, geleneksel modellere kıyasla daha iyi performans sergilediği çeşitli çalışmalarda vurgulanmaktadır. Ancak, bu tür istatistiksel yöntemlerin temel sınırlılığı, karmaşık ve doğrusal olmayan veri yapılarında performanslarının azalması ve yalnızca geçmiş veriye dayalı doğrusal ilişkileri modele dahil edebilmesidir. Bu nedenle, finansal zaman serisi analizinde

geleneksel yöntemlerin yanında, daha gelişmiş makine öğrenimi ve derin öğrenme yaklaşımlarının geliştirilmesine ihtiyaç duyulmuştur.

2.2 Makine Öğrenimi Modelleri

Makine öğrenimi modelleri, finansal piyasalardaki doğrusal olmayan örüntüleri ve karmaşık değişken ilişkilerini tespit etme yetenekleriyle, geleneksel istatistiksel yöntemlere kıyasla önemli avantajlar sunmaktadır. Bu modeller, büyük veri setlerinden anlamlı sonuçlar çıkarabilmek için geliştirilmiş ve hisse senedi fiyat tahmini gibi öngörüye dayalı analizlerde yaygın olarak kullanılmaya başlanmıştır.

Random Forest, Doğrusal Regresyon ve Destek Vektör Makineleri (SVM) gibi makine öğrenimi tabanlı algoritmalar; geçmiş fiyat hareketleri, teknik göstergeler ve piyasa verileri arasında ilişkiler kurarak, gelecekteki fiyatları tahmin edebilmektedir. Özellikle Random Forest gibi topluluk tabanlı modeller, çoklu karar ağacı oluşturarak aşırı uyum riskini azaltırken, doğrusal ve doğrusal olmayan ilişkileri aynı anda modelleyebilme kapasitesine sahiptir. Doğrusal Regresyon ise basit yapısı sayesinde hızlı ve yorumlanabilir sonuçlar sunar, ancak finansal verilerdeki karmaşıklığı tam olarak yansıtamayabilir.

Yapılan akademik çalışmalar, makine öğrenimi modellerinin, geleneksel istatistiksel yöntemlere kıyasla doğruluk oranını artırdığını ve piyasa dinamiklerine daha hızlı adapte olabildiğini göstermektedir. Bununla birlikte, bu modellerde özellik seçimi ve hiperparametre ayarlaması gibi süreçler, nihai tahmin başarısını doğrudan etkilemektedir. Bu nedenle, finansal zaman serisi analizlerinde makine öğrenimi modellerinin dikkatli şekilde yapılandırılması ve değerlendirilmesi gerekmektedir.

2.3 Makine Öğrenimi Modelleri

Derin öğrenme modelleri, son yıllarda finansal zaman serisi tahmininde önemli bir dönüşüm yaratmıştır. Özellikle sıralı ve karmaşık veri yapılarında güçlü öngörü kabiliyetine sahip olan bu yöntemler, hisse senedi fiyatlarındaki zamansal bağımlılıkları ve uzun vadeli eğilimleri etkili şekilde analiz edebilmektedir (DiPietro, 2020; Cao, 2024).

Uzun Kısa Süreli Bellek (LSTM), Tekrarlayan Sinir Ağları (RNN) ve Konvolüsyonel Sinir Ağları (CNN) gibi derin öğrenme mimarileri, geçmiş fiyat verilerinden hem kısa

hem de uzun dönemli ilişkileri öğrenerek, gelecekteki fiyat hareketlerinin daha doğru tahmin edilmesine olanak tanır. Özellikle LSTM ve RNN, ardışık veri içindeki bağımlılıkları modelleyerek, finansal piyasaların değişken yapısını anlamada geleneksel modellere kıyasla daha üstün performans sergileyebilmektedir (Sherstinsky, 2020; Zhenhao Yang, 2024). CNN ise, genellikle görüntü verilerinde kullanılmakla birlikte, zaman serisi verilerinde kısa vadeli kalıpları tespit etme amacıyla finansal alana da başarıyla uyarlanmıştır (Chua, 1997; Pawitra, 2024).

Literatürdeki güncel çalışmalar, derin öğrenme modellerinin özellikle yüksek hacimli ve karmaşık veri setlerinde, geleneksel ve makine öğrenimi tabanlı yaklaşımlara kıyasla daha yüksek doğruluk sağladığını göstermektedir (Cao, 2024; Pawitra, 2024). Ancak, bu modellerin eğitiminde yüksek hesaplama gücü gereksinimi ve hiperparametre ayarlarına olan duyarlılık gibi bazı zorluklar da bulunmaktadır. Bu nedenle, derin öğrenme modellerinin finansal tahminlerde başarılı sonuçlar verebilmesi için dikkatli şekilde tasarlanması ve optimize edilmesi gerekmektedir.

2.4 Hibrit ve Gelişmiş Modeller

Hibrit ve gelişmiş modeller, geleneksel istatistiksel yöntemler ile derin öğrenme yaklaşımlarının güçlü yönlerini birleştirerek, finansal zaman serisi tahmininde doğruluk ve esneklik sağlamayı hedeflemektedir. ARIMA-GARCH, CNN-BiLSTM ve FB-GAN gibi hibrit modeller, hem doğrusal hem de doğrusal olmayan ilişkileri aynı anda modele dahil ederek piyasa dalgalanmalarına ve dışsal faktörlere karşı daha dayanıklı tahminler üretebilmektedir (Yiluan Xing, 2024; Pawitra, 2024; Jain, 2024).

ARIMA-GARCH gibi modeller, istatistiksel tahminin gücünü volatilité modelleme ile birleştirerek özellikle oynak piyasa koşullarında öne çıkmaktadır (Yiluan Xing, 2024). CNN ve BiLSTM'in bir arada kullanıldığı yaklaşımlar ise, hem kısa vadeli kalıpları hem de uzun vadeli bağımlılıkları aynı anda yakalayabilmektedir ve literatürde üstün tahmin başarısı ile dikkat çekmektedir (Pawitra, 2024). Ayrıca FB-GAN modeli gibi yeni nesil hibrit yapılar, finansal haberlerdeki duyarlılık analizini fiyat tahminine entegre ederek, yalnızca tarihsel verilere değil, dışsal etkilere de duyarlı çözümler sunabilmektedir (Jain, 2024).

Bu modellerin en önemli avantajı, finansal verilerin karmaşık ve çok boyutlu doğasına daha iyi uyum sağlamalarıdır. Ancak, artan model karmaşıklığı, hesaplama maliyetleri

ve ayrıntılı ayar gereksinimi gibi bazı zorlukları da beraberinde getirmektedir. Buna rağmen, hibrit ve gelişmiş modellerin finansal piyasalarda tahmin doğruluğunu artırma potansiyeli son dönemde yapılan birçok çalışmada vurgulanmaktadır (Xing, 2024; Pawitra, 2024; Jain, 2024).

2.5 Literatür Taraması Sonuç

Literatürde yer alan çalışmalar, finansal zaman serisi tahmininde farklı model ve yaklaşımların çeşitli avantajlar ve sınırlamalar sunduğunu göstermektedir. Geleneksel istatistiksel modeller, doğrusal eğilimleri ve volatilitiyi başarılı şekilde analiz edebilmekte, ancak karmaşık ve doğrusal olmayan yapıları yeterince temsil edememektedir (Yiluan Xing, 2024; Merlo, 2023). Makine öğrenimi tabanlı modeller, özellikle Random Forest, XGBoost ve SVM gibi algoritmalar ile doğrusal olmayan ilişkileri yakalayabilmekte ve genel doğruluk oranını artırabilmektedir (Wong, 2023; Jiaqi, 2024; Tilla IZSÁK, 2023).

Derin öğrenme modelleri ise, özellikle ardışık ve dinamik veri yapısında, hem kısa hem de uzun vadeli bağımlılıkları modelleyerek finansal tahminde yüksek başarı elde edebilmektedir (Cao, 2024; Pawitra, 2024; Zhenhao Yang, 2024). Bunun yanında, hibrit yaklaşımlar, istatistiksel ve yapay zeka tabanlı yöntemlerin güçlü yönlerini birleştirerek karmaşık piyasa koşullarında daha sağlam ve hassas tahminler üretebilmektedir (Xing, 2024; Jain, 2024).

Sonuç olarak, literatürdeki eğilim; model seçiminde veri özelliklerinin, problem yapısının ve hedeflenen doğruluk düzeyinin belirleyici olduğuna işaret etmektedir. Özellikle finansal tahminlerde başarı için, uygun model mimarisinin ve güçlü bir özellik mühendisliği sürecinin önemi vurgulanmaktadır. Bu nedenle, bu projede de farklı model ve yöntemlerin performansları bütüncül bir yaklaşımla karşılaştırılmıştır.

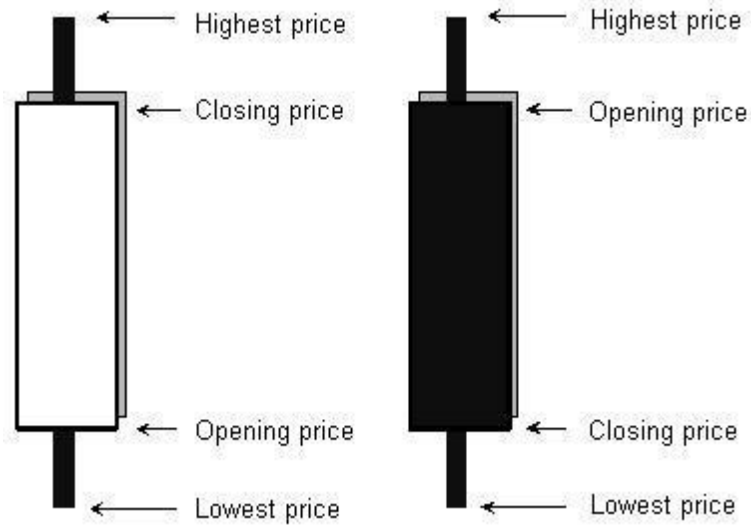
Bölüm 3

Hisse Senedi Fiyat Tahmininde Temel Kavramlar

3.1 Borsa Temelleri

Borsa, herkesin - ister bireysel yatırımcı, ister büyük kurum veya teknolojiyle yeni ilgilenen biri olsun - finansal sisteme katılım gösterebildiği kapsayıcı bir platformdur. Şirketler büyümek veya projelerini finanse etmek için burada yatırım ararken, yatırımcılar da kendi birikimlerini değerlendirme ve artırma fırsatı bulur. Modern borsalar, farklı yaş grupları, meslekler ve kültürel geçmişlerden gelen insanları buluşturarak küresel finansal sistemin kalbinde yer alır.

Borsada işlem gören hisse senetlerinin fiyatı, yalnızca ekonomik göstergeler ve finansal raporlarla değil, aynı zamanda toplumsal gelişmeler, yenilikler ve kolektif yatırımcı duygusuyla da şekillenir. Açılış fiyatı, bir hisse senedinin işlem gününe başladığı fiyatı; kapanış fiyatı ise günün sonunda ulaşılan değeri ifade eder. Yüksek ve düşük fiyatlar ise o gün yaşanan dalgalanmanın sınırlarını gösterir. Bu temel göstergeler, ister profesyonel yatırımcı ister yeni başlayan biri olsun, herkesin piyasa hareketlerini anlaması için bir temel oluşturur [(Wikipedia, n.d.)].



Şekil 3.1: Hisse Senedi Mum Çubuğu açıklaması

3.2 Zaman Serisi Verilerinin Özellikleri

Hisse senedi fiyat verileri, ister küçük yatırımcı ister büyük portföy yöneticisi olsun, herkesin analizinde kullandığı en temel veri türlerinden biridir. Zaman serisi, belirli aralıklarla kaydedilen ardışık gözlemlerden oluşur : her “nokta” bir tarihi, bir fiyatı, bir hacmi, bir duygu değişimini temsil edebilir. Böylece, herkes için piyasadaki eğilimleri, döngüleri ve ani dalgalanmaları gözlemlemenin bir yolu hâline gelir.

Bu verilerde genellikle trendler, mevsimsellik ve rastgele dalgalanmalar gibi örüntüler bulunur. Örneğin, bir hisse senedinin uzun vadede yükselmesi bir trend oluştururken, yılın belli dönemlerinde tekrar eden fiyat hareketleri mevsimselliğe işaret eder. Fakat hisse senedi piyasalarının en büyük zorluğu, bu istatistiksel özelliklerin (ortalama, varyans vb.) zaman içinde değişmesi ve öngörülemez dalgalanmalara açık olmasıdır. Yani, her yatırımcı için, hem kısa vadeli tepkileri hem de uzun vadeli eğilimleri modellemek özel teknikler gerektirir (Ariyo et al., 2015).

Zaman serisi analizi, finansal okuryazarlığı olan herkesin elindeki güçlü bir araçtır : doğru analizle, ister öğrenci ister profesyonel, herkes piyasa sinyallerini daha iyi yorumlayabilir.

3.3 NVIDIA Hisse Senedinin Uygunluğu

NVIDIA, yalnızca finans profesyonellerinin değil, teknolojiye ve geleceğe ilgi duyan herkesin dikkatini çeken bir şirkettir. Dünya çapında grafik işlemcileri (GPU) ve yapay zekâ altyapısı geliştiren NVIDIA, son yıllarda yapay zekâ devriminde oynadığı lider rol sayesinde borsada büyük bir görünürlük kazanmıştır. Her yaştan ve arka plandan yatırımcı, analist ya da öğrenci, NVIDIA'nın inovasyon gücünü ve hisse fiyatındaki hareketliliği yakından takip etmektedir.

NVIDIA'nın hisse senedi; hem yüksek hacimli işlemler, hem de fiyatındaki dalgalanmalar nedeniyle finansal tahmin projeleri için ideal bir örnek sunar. Ayrıca, teknolojiye meraklı gençlerden deneyimli yatırımcılara kadar herkesin ilgiyle takip ettiği bir şirket olması, bu çalışmayı daha geniş bir kitleye hitap eder hâle getirmektedir (Yiluan Xing, 2024).

Bu nedenle, NVIDIA'nın seçilmesi, hem günümüz teknolojisinin kalbindeki bir oyuncunun analiz edilmesini sağlamakta hem de farklı profildeki katılımcılara ilham vermektedir.

3.4 Tahmine Dayalı Modeller

Hisse senedi fiyatlarını tahmin etmek, hem profesyonel finansçılar hem de yatırım yapmaya yeni başlayan herkes için heyecan verici ve aynı zamanda karmaşık bir süreçtir. Çünkü piyasadaki fiyat hareketlerinin arkasında milyonlarca veri, duygu, haber ve toplumsal etki bulunur. Tüm bu karmaşayı anlamlandırmak için geliştirilen tahmine dayalı modeller, farklı bilgi düzeyindeki herkesin piyasayı analiz edebilmesine imkân tanır.

Yıllar içinde, araştırmacılar ve yatırımcılar tarafından üç ana model grubu öne çıkarılmıştır:

- istatistiksel Yöntemler (örneğin ARIMA, GARCH): Geçmiş fiyat verilerindeki doğrusal ilişkileri ve volatilitiyi modellemede kullanılır.
- Makine Öğrenimi Algoritmaları (örneğin Random Forest, Doğrusal Regresyon, SVM): Büyük veri kümelerinde gizli örüntüleri ve karmaşık ilişkileri yakalama konusunda avantajlıdır.

- Derin Öğrenme Mimarileri (örneğin LSTM, RNN, CNN): Sıralı ve karmaşık veri yapılarını analiz ederek, kısa ve uzun vadeli bağımlılıkları tespit eder.

Her yaklaşımın kendi güçlü ve sınırlı yönleri vardır. Bu modeller sayesinde, ister finans alanında deneyimli biri ister yeni öğrenmeye başlayan bir öğrenci olsun, herkes veriye dayalı tahminler üretme fırsatı bulabilir (Jiaqi, 2024; Cao, 2024).

3.4.1 İstatistiksel Yöntemler

İstatistiksel yöntemler, hisse senedi piyasalarını analiz etmeye ilgi duyan herkesin ilk başvurduğu temel araçlardır. Bu yaklaşımlar, geçmiş fiyat verilerindeki doğrusal ilişkileri ve kısa vadeli eğilimleri tespit etmeye odaklanır. Özellikle ARIMA ve GARCH gibi modeller, finansal verilerdeki trendleri ve oynaklıkları modellemede yaygın olarak kullanılır.

Örneğin, ARIMA modeli geçmiş kapanış fiyatlarına bakarak gelecekteki fiyatı tahmin etmeye çalışır. Böylece, piyasaya yeni adım atanlar için bile, fiyatların nasıl değişebileceği konusunda ilk izlenimler elde edilebilir (Ariyo et al., 2015).

Bu yöntemler, özellikle veri yapısı basit ve trendler belirgin olduğunda yüksek doğrulukla çalışır. Ancak, piyasalardaki karmaşık ve doğrusal olmayan hareketleri tam olarak yakalayamayabileceği için genellikle daha gelişmiş modellemelerle desteklenir.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (1)$$

Burada,

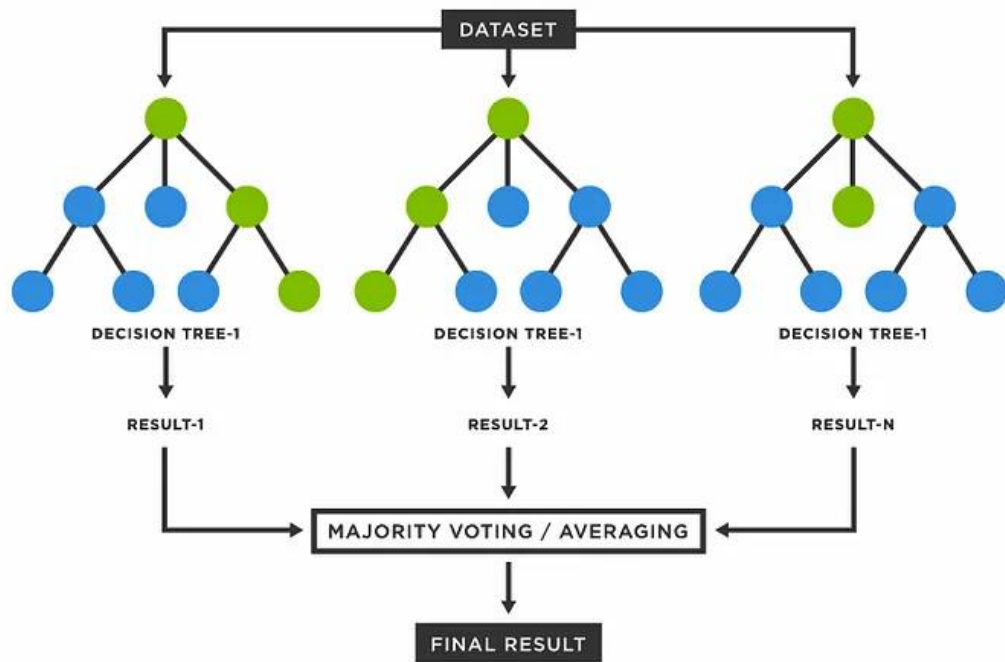
Y_t gerçek değerdir ve ϵ_t deki rastgele hatadır, i ve θ_j katsayılarıdır, p ve q tam sayılardır, genellikle sırasıyla otoregresif ve hareketli ortalama olarak adlandırılır.

3.4.2 Makine Öğrenimi Yöntemleri

Makine öğrenimi yöntemleri, son yıllarda veri bilimiyle ilgilenen herkesin gündemine girmiştir : finans uzmanlarından öğrencilere, hatta kodlama tecrübesi olmayan meraklılara kadar... Bu modellerin gücü, sadece geçmiş fiyatlara bakmakla kalmayıp, aynı zamanda teknik göstergeler, hacim verileri, haberler ve diğer pek çok faktörü birlikte analiz edebilmesinden gelir.

En yaygın kullanılan makine öğrenimi algoritmalarından bazıları Random Forest ve Doğrusal Regresyon'dur. Random Forest, çok sayıda karar ağacından oluşan bir topluluk modeliyle, karmaşık ve doğrusal olmayan ilişkileri başarılı şekilde modelleyebilir. Doğrusal Regresyon ise basitliği sayesinde özellikle hızlı ve kolay uygulanır, bu nedenle finans dünyasına yeni adım atanlar için iyi bir başlangıç noktasıdır.

Akademik çalışmalara göre, makine öğrenimi yöntemleri doğru özellik seçimi ve parametre ayarıyla, geleneksel istatistiksel modellere göre daha yüksek tahmin başarısı sağlayabilmektedir (Jiaqi, 2024; Wong, 2023).



Şekil 3.2: Random Forest Model Akışı

Bu yöntemler sayesinde, kodlama bilginiz ne olursa olsun, uygun araçlarla herkes kendi tahmin modellerini geliştirebilir.

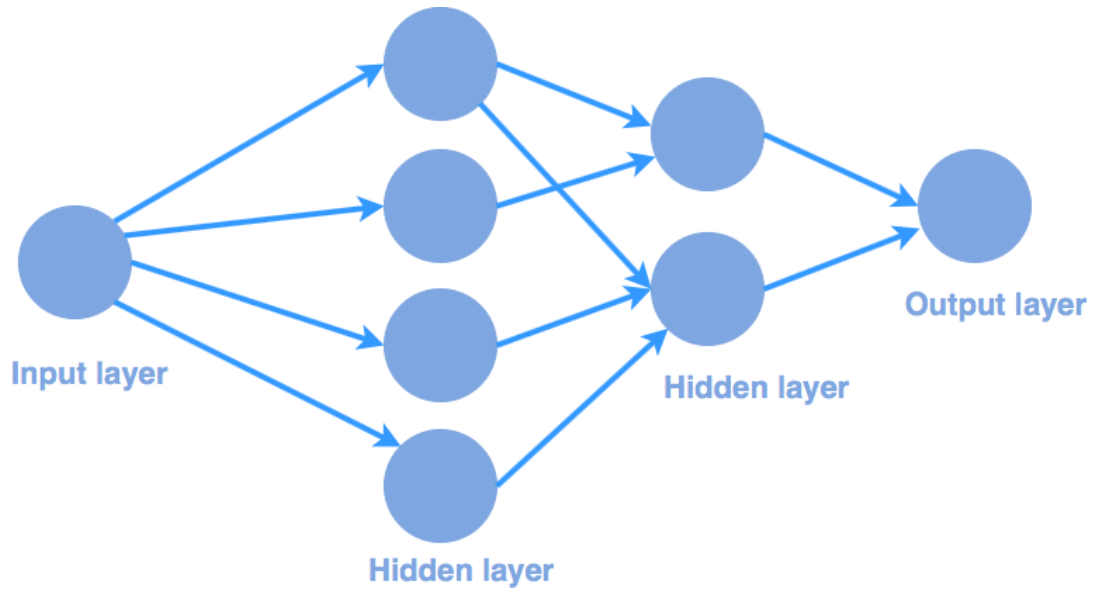
3.4.3 Derin Öğrenme Mimarileri

Derin öğrenme mimarileri, finansal veri analiziyle ilgilenen herkesin ilgisini çeken en güncel ve ileri düzey yöntemlerdendir. Bu yapılar, sıradan zaman serisi verilerinde dahi karmaşık ilişkileri ve uzun vadeli bağımlılıkları otomatik olarak öğrenme kapasitesine sahiptir. Özellikle büyük veriyle çalışan araştırmacılar, yatırımcılar, hatta

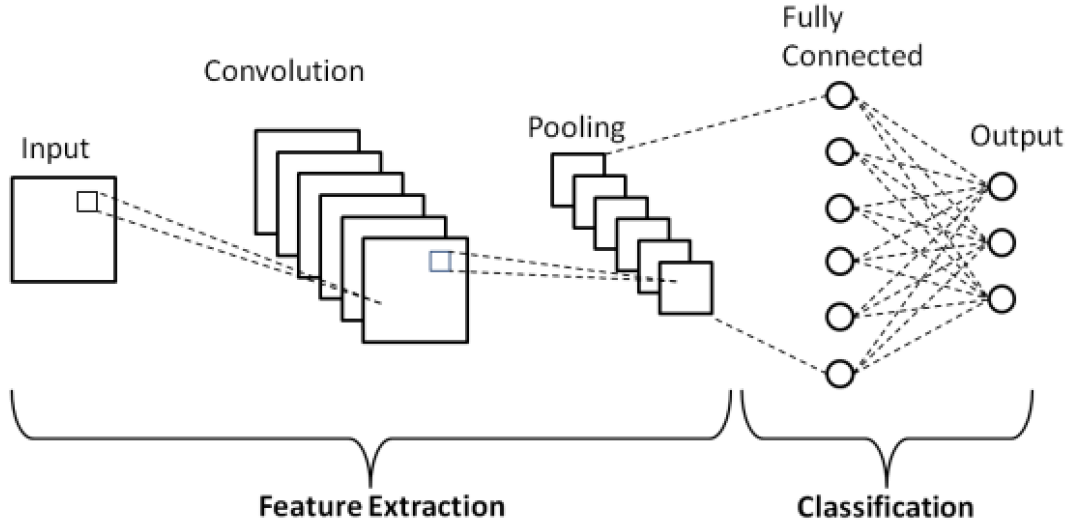
veri meraklısı öğrenciler için derin öğrenme, tahmin başarısını yeni bir seviyeye taşımıştır.

En çok kullanılan derin öğrenme mimarileri arasında Uzun Kısa Süreli Bellek (LSTM), Tekrarlayan Sinir Ağları (RNN) ve Konvolüsyonel Sinir Ağları (CNN) yer alır. LSTM ve RNN modelleri, sıralı verideki bağımlılıkları yakalama konusunda üstündür : örneğin, bir hisse senedinin geçmişteki davranışlarının, gelecekteki fiyatı üzerinde nasıl bir etkisi olabileceğini otomatik olarak öğrenebilirler. CNN ise, genellikle görüntü verilerinde kullanılsa da, finansal zaman serilerinde kısa vadeli kalıpları ve lokal örüntüleri tespit etmek için de başarıyla uygulanmaktadır (Cao, 2024; Pawitra, 2024).

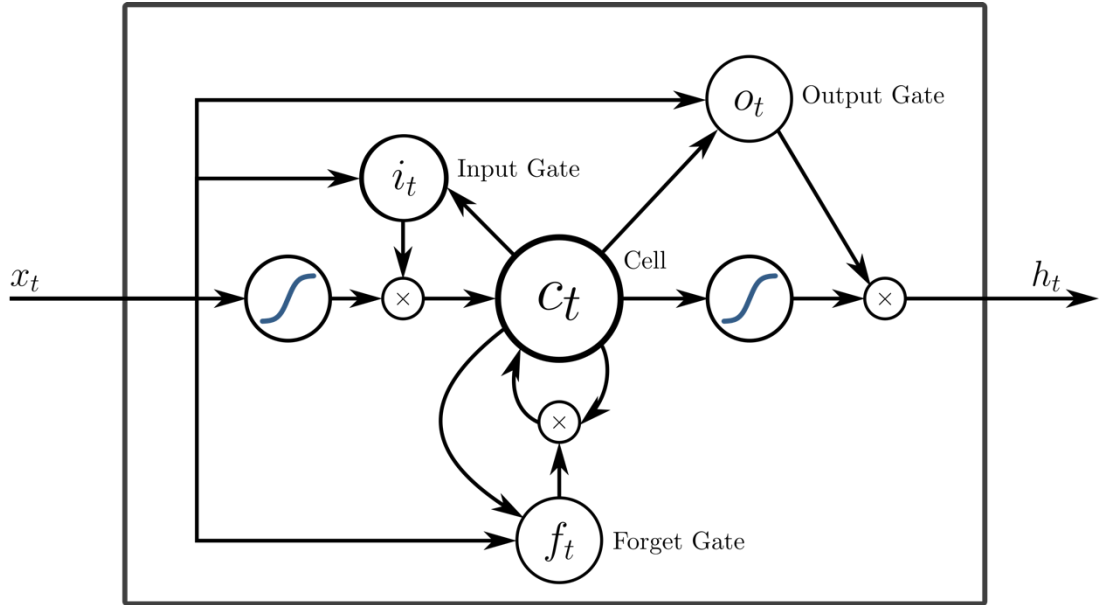
Bu modellerin uygulanması için yüksek hesaplama gücü ve veri hazırlama süreçleri gerekse de, günümüzde açık kaynaklı kütüphaneler sayesinde, teknik altyapısı olan herkes derin öğrenme ile kendi finansal tahmin projelerini geliştirebilir.



Şekil 3.3: Derin Öğrenme Modelleri Mimarisi



Şekil 3.4: CNN Model Architecture



Şekil 3.5: Giriş, çıkış ve unutma kapılarına sahip bir gözetleme deliği LSTM birimi (Wikipedia, t.y.)

3.5 Veri Kümesi ve Özellik Mühendisliği

Bu projede, NVIDIA hisse senedinin günlük kapanış fiyatlarını içeren kapsamlı bir veri kümesi kullanılmıştır. Bu veri seti, yalnızca fiyat bilgilerinden ibaret olmayıp, aynı zamanda işlem hacmi, en yüksek ve en düşük fiyat gibi birçok ek özelliği de kapsamaktadır. Böylece, ister finansal piyasalara yeni adım atanlar, ister deneyimli analistler olsun, herkes için çok boyutlu bir analiz ortamı sunulmuştur.

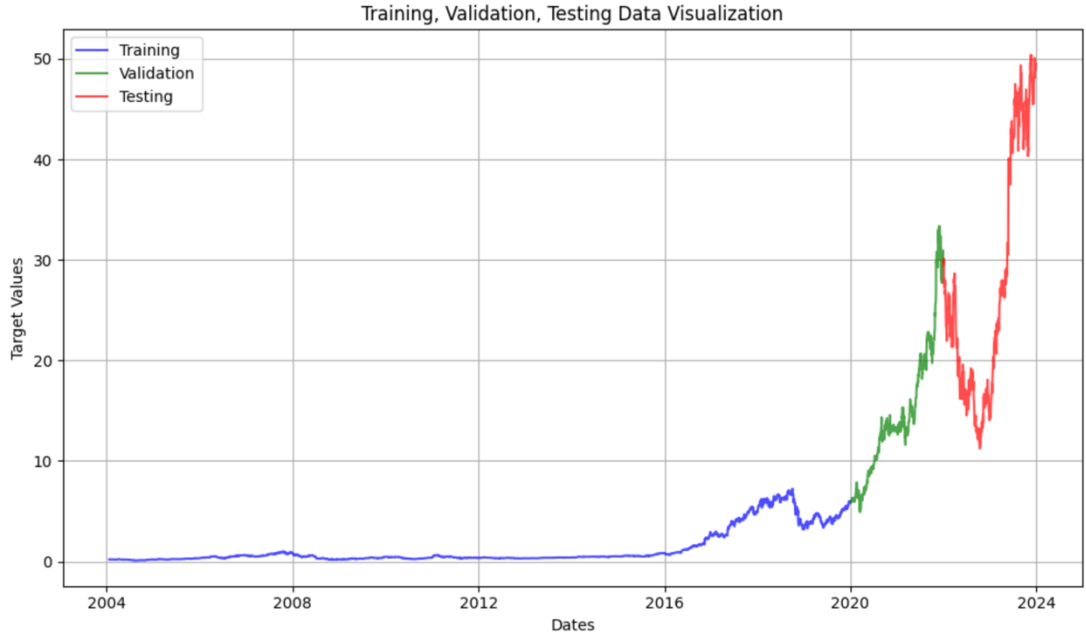
Özellik mühendisliđi, bir modelin başarıya ulaşmasında en kritik adımlardan biridir. Ham verilerden anlamlı özellikler üretmek, karmaşık ilişkileri daha net şekilde ortaya koymamıza olanak tanır. Örneđin, hareketli ortalama, volatilite veya fiyat değışim oranı gibi türetilmiş özellikler, hisse senedinin geçmiş davranışlarını daha iyi yansıtmakta ve modelin tahmin kapasitesini artırmaktadır.

Bu süreçte;

- “Lag” (gecikmeli) özellikleri,
- Hareketli ortalamalar,
- Günlük getiriler ve
- Teknik göstergeler (örneğin RSI, MACD gibi)

kullanılmıştır.

Bu sayede, ister teknik analiz ile ilgilenen bir yatırımcı, ister veri bilimi yolculuđuna yeni başlamış biri olsun, farklı seviyelerde analiz imkanı sağlanmıştır.



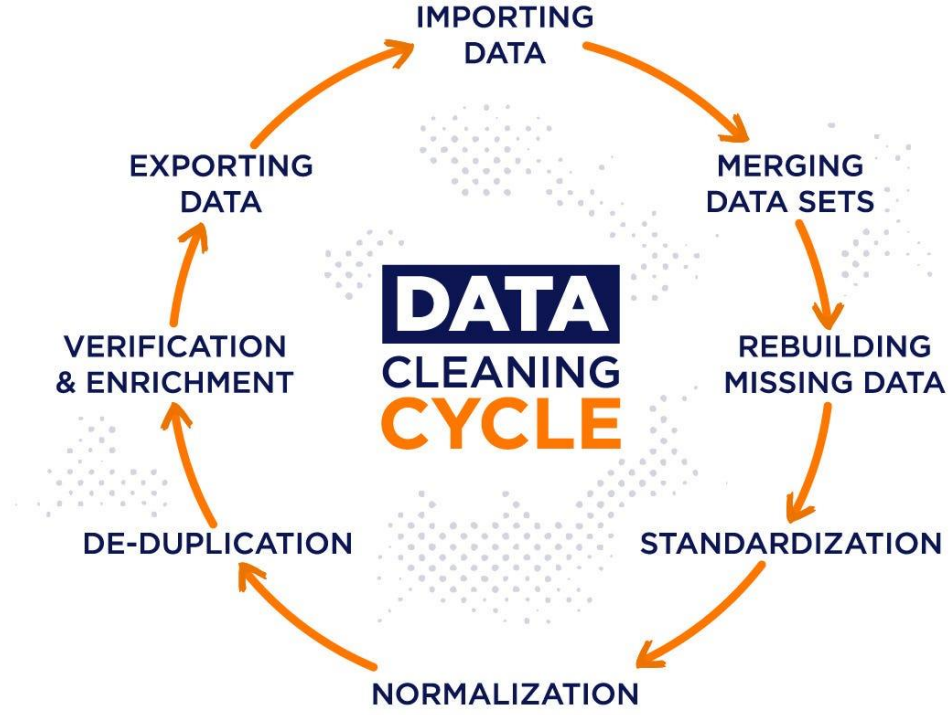
Şekil 3.6: Eğitim, Doğrulama, Test Verilerinin Görselleştirilmesi

3.6 Veri Ön İşleme ve Dönüşümler

Veri bilimi ve finansal modelleme ile ilgilenen herkes için veri ön işleme süreci, sağlam bir tahmin modeli oluşturmanın temel adımlarından biridir. Ham veriler genellikle eksik değerler, aykırı noktalar veya ölçek farklılıkları içerebilir. Bu nedenle, ister öğrenci, ister deneyimli bir veri bilimci olun, veriyi kullanıma hazır hale getirmek için sistemli bir şekilde temizlik ve dönüşüm işlemleri yapmak gereklidir.

Bu projede, ilk adım olarak eksik değerlerin tespiti ve uygun yöntemlerle doldurulması sağlanmıştır. Ardından, aykırı değerlerin etkisini azaltmak amacıyla çeşitli filtreleme ve sınırlandırma teknikleri uygulanmıştır. Ayrıca, tüm özellikler aynı ölçekte çalışabilsin diye, normalizasyon veya standartlaştırma gibi yöntemler kullanılmıştır. Bu tür dönüşümler, ister temel bir analiz ister karmaşık bir modelleme yapılacak olsun, herkesin daha güvenilir sonuçlar elde etmesine olanak sağlar.

Son olarak, zaman serisi yapısının korunmasına dikkat edilerek, veri seti eğitim ve test olarak ikiye ayrılmış ve çapraz doğrulama yöntemleriyle model performansı objektif biçimde değerlendirilmiştir.



Şekil 3.7: Veri ön işleme adımlarının görselleştirilmesi

3.7 Model Değerlendirme Yöntemleri

Model değerlendirme süreci, ister yeni başlayan bir veri bilimi öğrencisi olun, ister deneyimli bir finansal analist, doğru ve güvenilir tahminler üretebilmek için vazgeçilmez bir adımdır. Farklı modellerin karşılaştırılabilmesi ve en iyi sonuç veren yöntemin belirlenebilmesi için, genel kabul gören ölçütler ve testler kullanılır.

Bu projede, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) ve R-kare (R^2) gibi performans metrikleri temel alınmıştır.

- RMSE, tahmin edilen değerlerle gerçek değerler arasındaki farkın karelerinin ortalamasının kareköküdür; modelin hata düzeyini anlamak için kullanılır.
- MAE, tahminler ile gerçek değerler arasındaki mutlak farkların ortalamasıdır ve modelin genel doğruluğu hakkında fikir verir.
- R^2 ise modelin açıklayabildiği varyans oranını göstererek, modelin veriye ne kadar iyi uyduğunu anlamamızı sağlar.

Ayrıca, zaman serisi verilerinde “forward chaining” veya “rolling window” gibi özel doğrulama yöntemleri uygulanabilir. Bu yöntemler, farklı geçmiş dönemlere göre modelin başarısını test etmeye olanak tanır. Böylece, ister akademik, ister pratik bir uygulama olsun, her seviyedeki kullanıcı güvenilir ve nesnel sonuçlara ulaşabilir.

3.8 Öznitelik (Feature) Seçimi ve Önemi

Model başarısı, çoğu zaman kullanılan özniteliklerin (features) kalitesine ve seçimine bağlıdır. İster bir finans uzmanı, ister veri bilimi yolculuğuna yeni başlamış biri olun, doğru öznitelikleri belirlemek tahmin gücünüzü büyük ölçüde artırır. Ham veri çoğu zaman karmaşıktır ve doğrudan kullanılmak için uygun değildir. Bu nedenle, öznitelik seçimi sürecinde sadece en anlamlı, en bilgilendirici ve modelin performansına en çok katkı sağlayan değişkenlerin dahil edilmesi hedeflenir.

Bu projede, gecikmeli fiyatlar, hareketli ortalamalar, volatilité ve teknik göstergeler gibi pek çok farklı öznitelik değerlendirilmiş; istatistiksel testler ve model tabanlı önem skorları ile en etkili olanlar seçilmiştir. Gereksiz veya fazladan bilgi taşıyan öznitelikler çıkarılarak, modelin hem doğruluğu hem de işlem hızı artırılmıştır.

Ayrıca, öznitelik seçiminin doğru yapılması, modelin aşırı öğrenmesini (overfitting) önler ve genelleme başarısını yükseltir. Böylece, ister veri bilimiyle yeni tanışan biri ister deneyimli bir analist olun, daha güvenilir ve pratik sonuçlar elde etmek mümkün olur.

3.9 Model Seçimi ve Hiperparametre Ayarlaması

Model seçimi, ister yeni başlayan bir veri bilimci, ister deneyimli bir finansal analist için olsun, tahmin sürecinin en önemli aşamalarından biridir. Farklı modeller, farklı veri yapıları ve hedefler için çeşitli avantajlar sunar. Bu nedenle, projede çeşitli istatistiksel, makine öğrenimi ve derin öğrenme tabanlı modeller denenmiş, her birinin güçlü ve zayıf yönleri karşılaştırılmıştır.

Modelin başarısını maksimize etmek için yalnızca doğru modeli seçmek yeterli değildir; aynı zamanda her modelin hiperparametreleri de dikkatle ayarlanmalıdır. Hiperparametreler, modelin öğrenme sürecini ve esnekliğini belirleyen ayar

değerleridir (örneğin, ARIMA için p, d, q değerleri; Random Forest için ağaç sayısı; LSTM için katman sayısı gibi). Bu ayarların otomatik arama yöntemleriyle (grid search, random search gibi) veya deneysel olarak optimize edilmesi, modelin doğruluğunu ve genellenebilirliğini artırır.

Doğru model ve hiperparametre kombinasyonu sayesinde, ister küçük bir portföy yöneten biri, ister büyük ölçekli veriyle çalışan bir araştırmacı olun, herkes daha isabetli ve güvenilir tahminler elde edebilir.

Bölüm 4

Deneyler Ve Sonuçlar

Bu bölümde, NVIDIA hisse senedi fiyatlarının ertesi gün tahmini için uygulanan istatistiksel, makine öğrenimi ve derin öğrenme modellerinin performansları değerlendirilmiştir. Deneyler üç ana başlık altında yürütülmüş; model sonuçları RMSE, MAE ve R^2 gibi doğruluk ölçütlerine göre analiz edilmiştir.

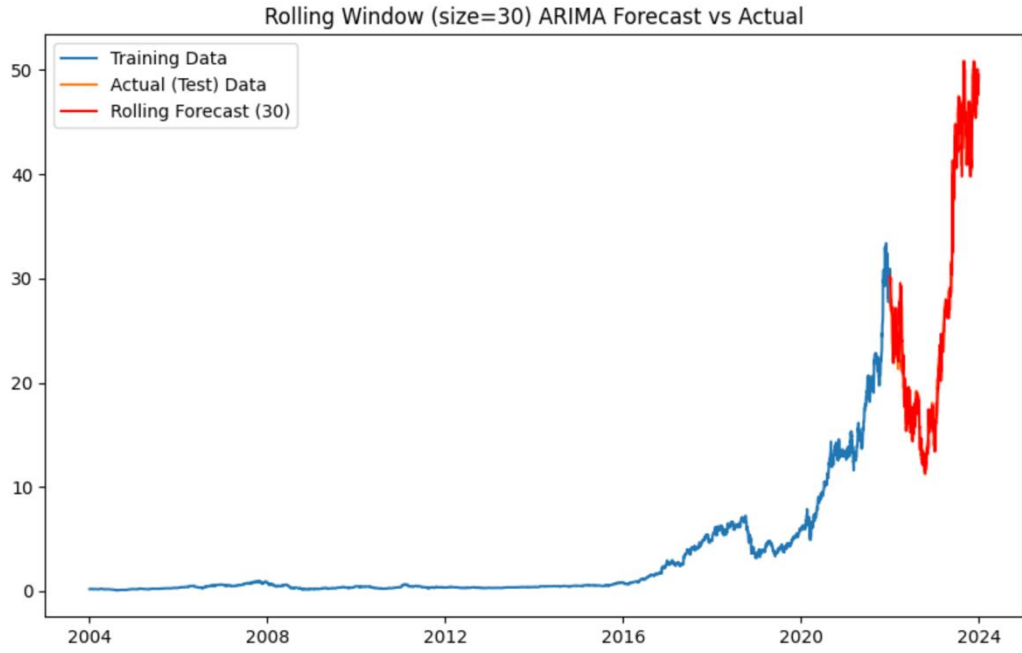
4.1 Deney 1: ARIMA Modeli

ARIMA modeli iki senaryo altında test edilmiştir: (i) kayan pencere kullanılmadan ve (ii) 30 günlük kayan pencere ile. Sonuçlar aşağıda sunulmuştur.

Tablo 4.1: ARIMA modelleri değerlendirme metrikleri

ARIMA TÜRÜ	SONUÇLAR		
	RMSE	MAE	R Square
ARIMA	12.49	11.41	-0.04
ARIMA 30 boyutunda yuvarlanan pencere ile	1.06	0.76	0.99

ARIMA modelinin kayan pencere ile kullanılması, performansını belirgin şekilde artırmış ve kısa vadeli tahminlerde yüksek doğruluk elde edilmiştir.



Şekil 4.1: 30 boyutunda yuvarlanan pencereli ARIMA modeli Performans

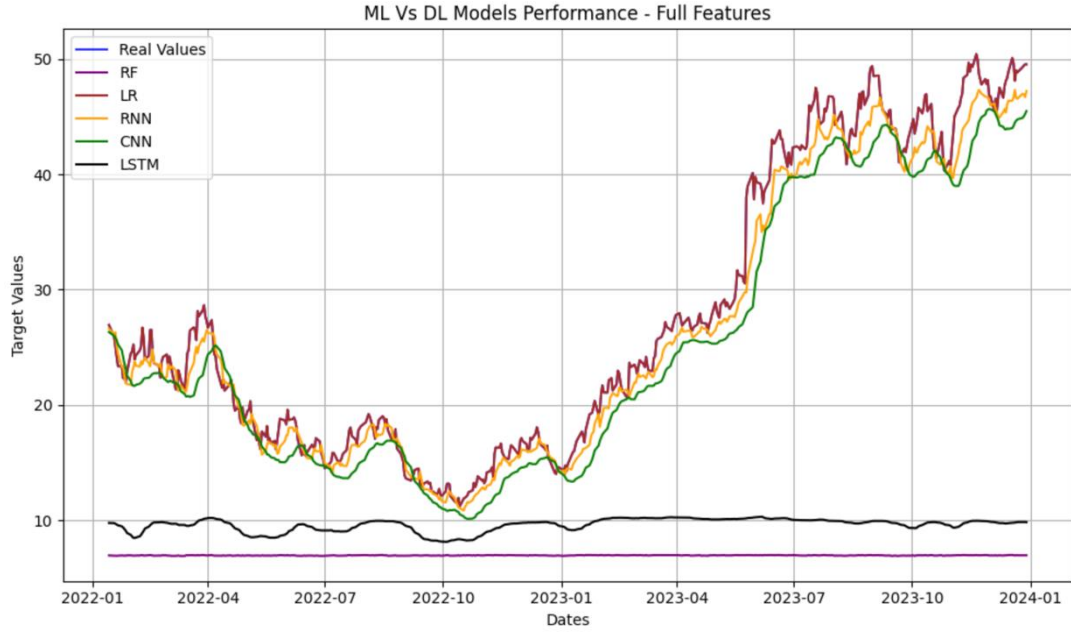
4.2 Deney 2: Tam Özellik Kümesi ile Makine Öğrenimi ve Derin Öğrenme Modelleri

Bu deneyde, gecikmeli değerler, hareketli ortalamalar, volatilité ölçümleri ve teknik göstergeler dahil olmak üzere 24 özellikten oluşan tam özellik seti ile çeşitli makine öğrenimi ve derin öğrenme modelleri test edilmiştir.

Tablo 4.2: ML vs. DL (Tam Özellikler) değerlendirme metrikleri

MODEL	SONUÇLAR		
	RMSE	MAE	R Square
LR	3.35	2.85	1.0
RF	23.92	20.57	-2.84
RNN	1.69	2.17	0.96
CNN	3.91	4.79	0.84
LSTM	17.27	20.95	-1.88

Sonuçlar, RNN'in en iyi performansı gösterdiğini; CNN'in onu takip ettiğini ve Doğrusal Regresyon'un ise veri kümesinin ötesinde genellemede zorluk yaşayabileceğini göstermiştir.



Şekil 4.2: ML ve DL Modellerinin Performansı – Tam Özellikler

4.3 Deney 3: Azaltılmış Özellik Kümesi ile Makine Öğrenimi ve Derin Öğrenme Modelleri

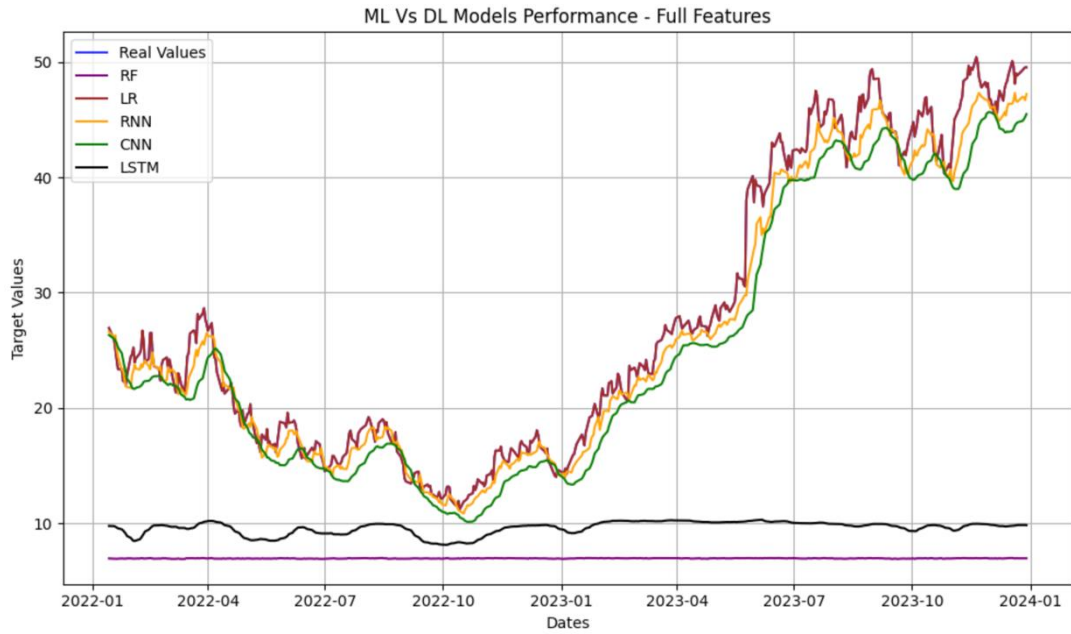
Bu deneyde, sadece gecikmeli özelliklerden ('t-1', 't-2', 't-3') oluşan azaltılmış bir özellik seti kullanılarak modeller yeniden test edilmiştir.

Tablo 4.3: ML vs. DL (Reduced Features) evaluation metrics

MODEL	SONUÇLAR		
	RMSE	MAE	R Square
LR	0.90	0.66	0.99
RF	23.83	20.48	-2.82
RNN	1.01	1.32	0.98

CNN	2.87	3.59	0.91
LSTM	6.37	9.45	0.41

Sonuçlar, RNN'in yine diğer tüm modellerden daha iyi performans gösterdiğini, CNN'in ise onu yakından takip ettiğini ortaya koymuştur. Doğrusal Regresyon, sade özellik setiyle güçlü performans sergilemiş, Rastgele Orman ve LSTM ise daha düşük başarı sağlamıştır.



Şekil 4.3: ML ve DL Modellerinin Performansı - Azaltılmış Özellikler

Bölüm 5

Sonuç

Bu projede, NVIDIA hisse senedinin ertesi gün kapanış fiyatının tahmini için istatistiksel, makine öğrenimi ve derin öğrenme tabanlı çeşitli modellerin performansları karşılaştırılmıştır. Yapılan deneylerde, 30 günlük kayan pencereyle kullanılan ARIMA modelinin doğrusal eğilimleri ve kısa vadeli dalgalanmaları başarıyla yakalayarak en iyi genel performansı sergilediği gözlemlenmiştir. Makine öğrenimi algoritmaları arasında Random Forest, karmaşık örüntüleri modellemede etkili olurken; Doğrusal Regresyon, özellikle azaltılmış özellik kümesiyle neredeyse mükemmel sonuçlar vermiştir. Derin öğrenme modellerinde ise, RNN'in sıralı bağımlılıkları en başarılı şekilde yakaladığı, CNN'in ise kısa vadeli trend tespitinde öne çıktığı görülmüştür. LSTM modeli ise, model hassasiyeti ve yüksek hesaplama gereksinimi nedeniyle diğer modellere kıyasla daha düşük başarı sağlamıştır.

Elde edilen bulgular, hisse senedi fiyat tahmininde model seçiminin ve özellik mühendisliğinin tahmin doğruluğu üzerinde kritik bir rol oynadığını göstermektedir. Ayrıca, hibrit ve gelişmiş model yaklaşımlarının, finansal piyasaların dinamik ve karmaşık yapısına daha iyi uyum sağlayabileceği ortaya konmuştur. Bu çalışma, finansal tahmin alanında yeni yöntemlerin geliştirilmesi ve mevcut yaklaşımların optimize edilmesi için temel bir kaynak olmayı amaçlamaktadır.

Kaynaklar

Dergi makalesi örnekleri:

Ariyo, A. A., Adewumi, A. O., & Ayo, C. K. (2015). Stock Price Prediction Using the ARIMA Model. IEEE, 2(2), 238.

Yiluan Xing, C. Y. (2024). Predicting NVIDIA's Next-Day Stock Price: A Comparative Analysis of LSTM, MLP, ARIMA, and ARIMA-GARCH Models. arXiv preprint arXiv:2405.08284.

Jiaqi, T. (2024). NVIDIA Stock Price Prediction by Machine Learning. Highlights in Business, Economics and Management, Volume 24 , 1072-1076.

Merlo, L. F. (2023). Risk Analysis of Nvidia Corporation Stock and Equally-Weighted US Portfolio. SSRN 4641295 .

Tilla IZSÁK, L. M. (2023). EVALUATION OF STOCK PRICE PREDICTION BASED ON THE SUPPORT VECTOR. Applied Computer Science, 19(3), 64-82.

Zhenhao Yang, Z. W. (24). The Research of NVIDIA Stock Price Prediction Based on LSTM And ARIMA Model. Highlights in Business, Economics and Management, Volume 24, 896-902.

Cao, R. (2024). Stock Price Prediction Using Deep-Learning Models: CNN, RNN, and LSTM. SHS Web of Conferences, 196, p. 02004.

Pawitra, M. T. (2024). Predicting Stock Market using CNN and BiLSTM Model. 2024 International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 267-272.

Pawitra, M. T. (2024). Predicting Stock Market using CNN and BiLSTM Model. nternational Conference on Computer, Control, Informatics and its Applications (IC3INA)(IEEE), pp. 267-272.

- Jain, J. K. (2024). FB-GAN: A Novel Neural Sentiment-Enhanced Model for Stock Price Prediction. In Proceedings of the Joint Workshop of the 7th Financial Technology and Natural Language Processing, the 5th Knowledge Discovery from Unstructured Data in Financial Services, and the 4th Workshop on Economics and Natural Language Processing, pp. 85-93.
- Wong, A. S.-W. (2023). Short-Term Stock Price Forecasting using Exogenous Variables and Machine Learning Algorithms. In 2023 3rd International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA) (IEEE), pp. 260-265.
- Xing, Y. C. (2024). Predicting NVIDIA's Next-Day Stock Price: A Comparative Analysis of LSTM, MLP, ARIMA, and ARIMA-GARCH Models. iv preprint arXiv:2405.08284.
- Duman, M. O. (2024). Performance Assessment of Temporal Frame Size in Stock Price Prediction. 2024 8th International Artificial Intelligence and Data Processing Symposium (IDAP)., pp. 1-5.
- Robert Alan Yaffee, M. M. (2000). An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS. Elsevier.
- Breiman, L. (2001). "Random forests. Machine learning 45.
- Sherstinsky, A. ((2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404, 132306.
- DiPietro, R. a. (2020). Deep learning: RNNs and LSTM. Handbook of medical image computing and computer assisted intervention Academic Press, 503-519.
- Chua, L. O. (1997). CNN: A vision of complexity. International Journal of Bifurcation and Chaos, 7.10, 2219-2425.
- Wikipedia. (tarih yok). Long short-term memory. (Wikipedia) 01 02, 2025 tarihinde https://en.wikipedia.org/wiki/Long_short-term_memory adresinden alındı.
- From Wikipedia, t. f. (tarih yok). Candlestick chart. (Wikipedia) 01 01, 2025 tarihinde https://en.wikipedia.org/wiki/Candlestick_chart adresinden alındı.

