

**T.C.**  
**BALIKESİR ÜNİVERSİTESİ**  
**FEN BİLİMLERİ ENSTİTÜSÜ**  
**BİLGİSAYAR VE BİLİŞİM MÜHENDİSLİĞİ ANABİLİM DALI**



**MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE KRİPTO VARLIK  
PİYASALARINDA FİYAT BALONLARININ TAHMİNİ**

**FATMA FEYZA ÖZEL**

**YÜKSEK LİSANS TEZİ**

**Jüri Üyeleri: Prof. Dr. Ayhan İSTANBULLU (Tez Danışmanı)**

**Doç. Dr. Hilmi Tunahan AKKUŞ (Eş Danışman)**

**Doç. Dr. Fatih AYDIN**

**Doç. Dr. İlyas ÖZER**

**Dr. Öğr. Üyesi Abdülgani KAHRAMAN**

**BALIKESİR, ŞUBAT – 2026**

## ETİK BEYAN

Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak tarafımda hazırlanan “**Makine Öğrenmesi Yöntemleri ile Kripto Varlık Piyasalarında Fiyat Balonlarının Tahmini**” başlıklı tezde;

- Tüm bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Kullanılan veriler ve sonuçlarda herhangi bir değişiklik yapmadığımı,
- Tüm bilgi ve sonuçları bilimsel araştırma ve etik ilkelere uygun şekilde sunduğumu,
- Yararlandığım eserlere atıfta bulunarak kaynak gösterdiğimi,

beyan eder, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ederim.

**Fatma Feyza ÖZEL**

(imza)

**Bu tez çalışması Balıkesir Üniversitesi Bilimsel Araştırma Koordinatörlüğü tarafından  
(BAP 2025/043) numaralı proje ile desteklenmiştir.**

## ÖZET

**MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE KRİPTO VARLIK  
PİYASALARINDA FİYAT BALONLARININ TAHMİNİ  
YÜKSEK LİSANS TEZİ  
FATMA FEYZA ÖZEL  
BALIKESİR ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ  
BİLGİSAYAR VE BİLİŞİM MÜHENDİSLİĞİ ANA BİLİM DALI  
(TEZ DANIŞMANI: PROF. DR. AYHAN İSTANBULLU)  
(EŞ DANIŞMAN: DOÇ. DR. HİLMİ TUNAHAN AKKUŞ)  
BALIKESİR, ŞUBAT - 2026**

Bu tez, kripto varlık piyasalarında, özellikle Bitcoin’de, ani değer kayıplarına yol açan fiyat balonlarının araştırılmasını ve nedenlerinin makine öğrenmesi yöntemleri ile tespit edilmesini amaçlamaktadır. Kripto varlık piyasalarının yüksek oynaklığı ve karmaşık yapısı, ilgili piyasada fiyat balonlarının varlığı konusunda şüphe uyandırmaktadır. Yapılan çalışmada ilk olarak 1 Ocak 2020 ile 5 Haziran 2024 tarihleri arasındaki veriler analiz edilmiş ve fiyat balonları GSADF testi ile tespit edilmiş ve etiketlenmiştir. Sonraki aşamada fiyat balonları, makine öğrenmesi ve derin öğrenme sınıflandırma algoritmaları kullanılarak tahmin edilmiştir. Sınıflandırma tahmini için finansal ve ekonomik değişkenlerin yanı sıra çeşitli teknik göstergeler de türetilerek kullanılmıştır. Metodolojik olarak, özellik mühendisliği aşamasında makroekonomik göstergelere ek olarak RSI, SMA ve ATR gibi teknik göstergeler kullanılmıştır. Veri sızıntısını önlemek amacıyla tüm değişkenler modele gecikmeli olarak dahil edilmiştir. Modellerin performansı, azınlık sınıfı başarımını ölçmeye uygun Test F1 Skoru, PR-AUC ve ROC-AUC metrikleri ile değerlendirilmiştir. Analiz sonuçlarına göre, makine öğrenmesi modellerine kıyasla RNN modeli, %70 eğitim – %30 test bölünmesinde 0,705 Test F1 skoru ve 0,797 PR-AUC değeri ile en yüksek performansı sergilemiştir. Elde edilen bulgular üç temel alanda somut katkı sağlamaktadır. İlk olarak yatırımcılar için spekülasyon balon dönemlerinde riskleri azaltmaya yardımcı olan bir erken uyarı mekanizması sunulmaktadır. İkinci olarak düzenleyici kurumlar açısından piyasa gözetimine yönelik uygulanabilir bir analitik çerçeve önermektedir. Son olarak kripto varlık fiyat balonlarını modelleyebilen RNN tabanlı yaklaşım sunulmuştur.

**ANAHTAR KELİMELELER:** Derin öğrenme, fiyat balonları, kripto varlık, makine öğrenmesi, tahmin

## **ABSTRACT**

### **PREDICTION OF PRICE BUBBLES IN CRYPTOCURRENCY MARKETS USING MACHINE LEARNING METHODS**

**MSc THESIS**

**FATMA FEYZA ÖZEL**

**BALIKESIR UNIVERSITY INSTITUTE OF SCIENCE**

**COMPUTER AND INFORMATION ENGINEERING**

**(SUPERVISOR: PROF. DR. AYHAN İSTANBULLU)**

**(CO-SUPERVISOR: ASSOC. PROF. DR. HİLMİ TUNAHAN AKKUŞ)**

**BALIKESIR, FEBRUARY - 2026**

This thesis aims to investigate price bubbles in crypto-asset markets—particularly in Bitcoin—that can trigger sudden value collapses, and to identify their underlying drivers using machine learning methods. The high volatility and complex structure of crypto-asset markets raise doubts regarding the presence and detectability of price bubbles in this domain. In this study, data spanning 1 January 2020 to 5 June 2024 are first analyzed, and price bubbles are detected and labeled using the GSADF test. In the subsequent stage, these bubble periods are predicted via machine learning and deep learning classification algorithms. For the classification task, a set of features is constructed that includes financial and macroeconomic variables as well as derived technical indicators. Methodologically, the feature engineering process incorporates technical indicators such as RSI, SMA, and ATR in addition to macroeconomic indicators. To prevent data leakage, all variables are included in the models using lagged values. Model performance is evaluated using metrics suitable for assessing minority-class performance, namely the Test F1-score, PR-AUC, and ROC-AUC. The empirical results show that, compared with conventional machine learning models, the RNN model achieves the best performance under a 70% training–30% testing split, attaining a Test F1-score of 0.705 and a PR-AUC of 0.797. The findings provide tangible contributions in three main areas. First, they offer an early warning mechanism that can help investors mitigate risks during speculative bubble periods. Second, they propose an actionable analytical framework for market surveillance from the perspective of regulatory authorities. Finally, the study presents an RNN-based approach capable of modeling crypto-asset price bubbles.

**KEYWORDS:** Deep learning, price bubbles, cryptocurrency, machine learning, forecasting

# İÇİNDEKİLER

## Sayfa

<b>ÖZET</b> .....	<b>i</b>
<b>ABSTRACT</b> .....	<b>ii</b>
<b>İÇİNDEKİLER</b> .....	<b>iii</b>
<b>ŞEKİL LİSTESİ</b> .....	<b>v</b>
<b>TABLO LİSTESİ</b> .....	<b>vi</b>
<b>KISALTMALAR LİSTESİ</b> .....	<b>vii</b>
<b>ÖNSÖZ VE TEŞEKKÜR</b> .....	<b>x</b>
<b>1. GİRİŞ</b> .....	<b>1</b>
<b>2. KRİPTO VARLIKLAR</b> .....	<b>3</b>
2.1 Kripto Para.....	3
2.2 Sanal Evren (Metaverse).....	4
2.3 Değiştirilemez Jeton (Non-Fungible Token - NFT) .....	6
2.4 Merkeziyetsiz Otonom Organizasyon (Decentralized Autonomous Organization - DAO).....	7
2.5 Merkeziyetsiz Finans (Decentralized Finance - DeFi) .....	8
<b>3. FİYAT BALONLARI</b> .....	<b>10</b>
3.1 Kavramsal Açıklama .....	10
3.2 Fiyat Balonlarının Önemi .....	11
3.3 Geçmişte Yaşanan Önemli Fiyat Balonları .....	12
3.3.1 Klasik dönem balonları .....	12
3.3.2 Modern dönem balonları.....	12
3.3.3 Dijital çağ balonları.....	13
3.3.4 Kripto varlık balonları.....	13
<b>4. KRİPTO VARLIKLARIN FİYATLARINI ETKİLEYEN ÖNEMLİ OLAYLAR</b> <b>15</b>	<b>15</b>
4.1 Bitcoinica Dolandırıcılığı .....	15
4.2 Silk Road Operasyonu .....	15
4.3 Çin Hükümetinin Kripto Para Piyasalarına Yönelik Düzenleyici Müdahaleleri.....	15
4.4 Mt. Gox Krizi.....	16
4.5 McAfee ve Kripto Manipülasyonu .....	16
4.6 Elon Musk ve Dogecoin Dalgalanmaları.....	16
4.7 Thodex Skandalı .....	17
4.8 FTX İflası.....	17
4.9 Terra (LUNA) ve UST Krizi .....	17
<b>5. LİTERATÜR İNCELEMESİ</b> .....	<b>19</b>
5.1 Kripto Varlıklarla İlgili Balon Çalışmaları.....	19
5.2 Kripto Varlıkların Diğer Varlıklarla İlişkisi Üzerine Çalışmalar .....	20
5.3 Fiyat Balonlarının Nedenlerinin Tahmini Üzerine Çalışmalar.....	21
<b>6. METODOLOJİ</b> .....	<b>23</b>
6.1 Yazılım ve Donanım Altyapısı .....	23
6.2 Akış Diyagramı.....	24
6.3 Veri Seti .....	25
6.4 Veri Ön İşleme.....	28
6.4.1 Veri temizleme işlemleri.....	28
6.4.2 Özellik çıkarma ve dönüştürme teknikleri .....	29
6.4.3 Veri setinin eğitim ve test olarak ayrılması (data splitting).....	30
6.5 Veri Ölçeklendirme Yöntemleri .....	31

## İÇİNDEKİLER(devam)

6.5.1 Min-Max ölçeklendirme .....	31
6.5.2 Z-skor ölçeklendirme (standardizasyon).....	31
6.6 GSADF Analiz Yöntemi.....	32
6.7 Optimizasyon Teknikleri .....	34
6.7.1 Hiperparametre ayarlama.....	34
6.7.2 Performans artırma stratejileri .....	34
<b>7. MAKİNE ÖĞRENMESİ MODELLERİ .....</b>	<b>37</b>
7.1 Rastgele Orman (Random Forest - RF).....	37
7.2 Aşırı Gradyan Artırma (Extreme Gradient Boosting - XGBoost).....	41
7.3 Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine - LightGBM).....	45
7.4 Uzun-Kısa Vadeli Bellek (Long Short-Term Memory - LSTM) .....	49
7.5 Evrişimli Sinir Ağları (Convolutional Neural Networks - CNN).....	52
7.6 Tekrarlayan Sinir Ağları (Recurrent Neural Networks - RNN) .....	58
7.7 Kapalı Tekrarlayan Ünite (Gated Recurrent Unit - GRU).....	60
7.8 Makine Öğrenmesi Modellerinin Performansları .....	62
7.8.1 Karmaşıklık matrisi (confusion matrix).....	62
7.8.2 Doğruluk (Accuracy) .....	63
7.8.3 Hassasiyet (Precision).....	63
7.8.4 Duyarlılık (Recall) .....	64
7.8.5 F1-Skoru .....	64
7.8.8 ROC-AUC ve Precision-Recall Eğrileri .....	64
7.8.9 Shapley katkılarına dayalı açıklayıcı yöntem (SHapley Additive Explanations - SHAP).....	65
7.9 Model Doğrulama ve Seçimi .....	66
7.9.1 Çapraz doğrulama yöntemleri.....	66
7.9.2 Öznitelik Seçimi (Feature Selection) .....	67
<b>8. ANALİZ VE BULGULAR .....</b>	<b>68</b>
<b>9. SONUÇ VE ÖNERİLER .....</b>	<b>78</b>
<b>10. KAYNAKLAR.....</b>	<b>80</b>
<b>ÖZGEÇMİŞ .....</b>	<b>Hata! Yer işareti tanımlanmamış.</b>

## ŞEKİL LİSTESİ

### Sayfa

Şekil 2.1: Metaverse oluşturan bileşenler .....	5
Şekil 2.2: Merkeziyetsizlikten merkezileşmeye analizi .....	8
Şekil 6.1: Akış diyagramı .....	24
Şekil 7.1: Karar ağaçlarının topluluk yaklaşımıyla birleştirilmesi .....	38
Şekil 7.2: XGBoost adım güçlendirme ve ağırlık güncelleme mekanizması .....	41
Şekil 7.3: Yaprak tabanlı büyüme .....	45
Şekil 7.4: LSTM çalışma modeli .....	50
Şekil 7.5: Ek katmanların özellik hiyerarşisi .....	53
Şekil 7.6: Konvolüsyon operasyonu ve çıktı hesaplama .....	54
Şekil 7.7: RNN çalışma prensibi.....	58
Şekil 7.8: Karmaşıklık matrisi değer gösterim tablosu .....	63
Şekil 8.1: Eviews sonuç dağılım grafiği .....	68
Şekil 8.2: Kripto varlık piyasalarında GSADF testi ile balon tespit sonuçları .....	69
Şekil 8.3: Korelasyon ısı haritası .....	70
Şekil 8.4: En iyi ağaç modeli karmaşıklık matrisi (XGBoost) .....	71
Şekil 8.5: En iyi ağaç modeli hassasiyet – duyarlılık eğrisi (XGBoost).....	72
Şekil 8.6: En iyi DL modeli karmaşıklık matrisi (RNN) .....	73
Şekil 8.7: En iyi DL modeli hassasiyet - duyarlılık eğrisi (RNN) .....	74
Şekil 8.8: Modellerin en iyi test F1 skorlarına göre karşılaştırılması .....	75

## TABLO LİSTESİ

### Sayfa

<b>Tablo 6.1:</b> Veri açıklama tablosu.....	26
<b>Tablo 6.2:</b> Tanımlayıcı istatistik tablosu .....	27
<b>Tablo 7.1:</b> Ağaç büyüme stratejilerinin karşılaştırılması .....	46
<b>Tablo 7.2:</b> LightGBM ve XGBoost arasındaki farklar .....	48
<b>Tablo 7.3:</b> RNN değer gösterim tablosu.....	60
<b>Tablo 8.1:</b> Modellerin farklı bölünmeleri ve ölçekleyicilere göre performans karşılaştırması .....	75

## KISALTMALAR LİSTESİ

<b>5AMLD</b>	: 5. Kara Para Aklamayı Önleme Direktifi
<b>ADF</b>	: Augmented Dickey-Fuller (Testi)
<b>AML</b>	: Kara Para Aklamayı Önleme (Anti-Money Laundering)
<b>AP</b>	: Average Precision (Ortalama Hassasiyet)
<b>AR</b>	: Artırılmış Gerçeklik
<b>ARDL</b>	: Autoregressive Distributed Lag (Otokorelasyonlu Dağıtılmış Gecikme)
<b>ATR</b>	: Ortalama Gerçek Aralık (Average True Range)
<b>AUC</b>	: Eğri Altındaki Alan (Area Under the Curve)
<b>BAP</b>	: Bilimsel Araştırma Projesi (Balıkesir Üniversitesi Projesi)
<b>BPTT</b>	: Backpropagation Through Time (Zaman İçinde Geri Yayılım)
<b>BTC</b>	: Bitcoin
<b>CEX</b>	: Merkezi (Merkezileşmiş) Borsalar (Centralized Exchanges)
<b>CFTC</b>	: Emtia Vadeli İşlemler Komisyonu (Commodity Futures Trading Commission)
<b>CMO</b>	: Siber Hareket Organizasyonları (Cyber Movement Organizations)
<b>CNN</b>	: Evrişimli Sinir Ağları (Convolutional Neural Networks)
<b>CSAD</b>	: Ortalama Mutlak Sapma (Cross-Sectional Absolute Deviation)
<b>DAO</b>	: Merkeziyetsiz Otonom Organizasyon (Decentralized Autonomous Organization)
<b>DeFi</b>	: Merkeziyetsiz Olmayan Finans (Decentralized Finance)
<b>DEX</b>	: Merkeziyetsiz Borsalar (Decentralized Exchanges)
<b>DJI</b>	: Dow Jones Industrial Average (Endeksi)
<b>DL</b>	: Derin Öğrenme (Deep Learning)
<b>EFB</b>	: Exclusive Feature Bundling (Ayrık Özellik Demetleme)
<b>EPU</b>	: Ekonomik Politika Belirsizliği (Endeksi)
<b>ETH</b>	: Ethereum
<b>EUR/USD</b>	: Euro-Dolar Döviz Kuru
<b>F1</b>	: F1-Skoru
<b>FBI</b>	: Federal Soruşturma Bürosu (Federal Bureau of Investigation)
<b>FN</b>	: Negatif Yanlış (False Negative)
<b>FNN</b>	: İleri Beslemeli Sinir Ağları (Feedforward Neural Networks)
<b>FP</b>	: Pozitif Yanlış (False Positive)
<b>FPR</b>	: Yanlış Pozitif Oranı (False Positive Rate)
<b>GDAXI</b>	: Alman DAX Endeksi
<b>GOSS</b>	: Gradient-Based One-Side Sampling (Gradyan Tabanlı Tek Taraflı Örnekleme)
<b>GSADF</b>	: Genelleştirilmiş Sup Augmented Dickey-Fuller (Testi)
<b>GRU</b>	: Kapılı Tekrarlayan Birim (Gated Recurrent Unit)
<b>H</b>	: Filtre/Kernel (CNN)
<b>HCI</b>	: İnsan-Bilgisayar Etkileşimi (Human-Computer Interaction)
<b>ICO</b>	: İlk Coin Arzı (Initial Coin Offering)
<b>KYC</b>	: Müşterini Tanı (Know Your Customer)
<b>LASSO</b>	: Least Absolute Shrinkage ve Selection Operator

## KISALTMALAR LİSTESİ (devam)

<b>LightGBM</b>	: Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine)
<b>LPPL</b>	: Log-Periyodik Güç Yasası (Log-Periodic Power Law Singularity)
<b>LSTM</b>	: Uzun-Kısa Vadeli Bellek (Long Short-Term Memory)
<b>MDI</b>	: Ortalama Safsızlık Azalması (Mean Decrease in Impurity)
<b>MinMax</b>	: Min-Max Ölçeklendirme (Ölçeklendirme Yöntemi)
<b>ML</b>	: Makine Öğrenmesi (Machine Learning)
<b>MR</b>	: Karma Gerçeklik (Mixed Reality)
<b>NASDAQ</b>	: National Association of Securities Dealers Automated Quotations (Endeksi)
<b>NFT</b>	: Değiştirilemez Jeton (Non-Fungible Token)
<b>OLS</b>	: Ordinary Least Squares (En Küçük Kareler Yöntemi)
<b>P</b>	: Padding (Doldurma - CNN)
<b>PR-AUC</b>	: Precision-Recall Eğrisi Altındaki Alan (Average Precision)
<b>PSY</b>	: Phillips, Shi ve Yu Yöntemi
<b>RF</b>	: Rastgele Orman (Random Forest)
<b>ReLU</b>	: Rectified Linear Unit (Doğrultulmuş Doğrusal Birim)
<b>RNN</b>	: Tekrarlayan Sinir Ağları (Recurrent Neural Networks)
<b>ROC</b>	: Receiver Operating Characteristic (Alıcı İşletim Karakteristiği)
<b>ROC-AUC</b>	: ROC Eğrisi Altındaki Alan
<b>RSI</b>	: Göreceli Güç Endeksi (Relative Strength Index)
<b>S</b>	: Stride (Adım Sayısı - CNN)
<b>SADF</b>	: Sup Augmented Dickey-Fuller (Testi)
<b>SEC</b>	: Menkul Kıymetler ve Borsa Komisyonu (Securities ve Exchange Commission)
<b>SHAP</b>	: SHapley Additive Explanations (Shapley Katkılarına Dayalı Açıklayıcı Yöntem)
<b>SMA</b>	: Basit Hareketli Ortalama (Simple Moving Average)
<b>S&amp;P 500</b>	: Standard & Poor's 500 (Endeksi)
<b>STAR</b>	: Smooth Transition Autoregressive
<b>SVECM</b>	: Structural Vector Error Correction Model (Yapısal Vektör Hata Düzeltme Modeli)
<b>SVM</b>	: Destek Vektör Makineleri (Support Vector Machines)
<b>TA-Lib</b>	: Teknik Analiz Kütüphanesi
<b>TN</b>	: Negatif Doğru (True Negative)
<b>TP</b>	: Pozitif Doğru (True Positive)
<b>TPR</b>	: Doğru Pozitif Oranı (True Positive Rate)
<b>U</b>	: Giriş verisinin gizli duruma etkisini belirleyen ağırlık matrisi
<b>VAR</b>	: Vektör Otoregresif Modeli (Vector Autoregression)
<b>VECM</b>	: Vektör Hata Düzeltme Modeli (Vector Error Correction Model)
<b>VIX</b>	: CBOE Volatilite Endeksi
<b>VR</b>	: Sanal Gerçeklik (Virtual Reality)

## **KISALTMALAR LİSTESİ (devam)**

<b>W</b>	: Önceki gizli durumu yeni gizli duruma taşıyan ağırlık matrisi
<b>XGBoost</b>	: Aşırı Gradyan Artırma (Extreme Gradient Boosting)
<b>XRP</b>	: Ripple
<b>Z-Skor</b>	: Z-Skor Ölçeklendirme (Standardizasyon)

## ÖNSÖZ VE TEŞEKKÜR

Bu tez çalışması hem akademik hem de kişisel gelişim sürecimde önemli bir dönüm noktası olmuştur. Bu yolculuk boyunca bana rehberlik eden, destek olan ve ilham veren birçok değerli insanın katkısı bulunmaktadır. Öncelikle, yüksek lisans eğitimim süresince bilgi ve tecrübeleriyle beni yönlendiren, araştırma sürecinde bilimsel bakış açımı geliştirmeme büyük katkı sağlayan değerli danışman hocam Prof. Dr. Ayhan İSTANBULLU'ya en içten teşekkürlerimi sunarım. Okul sürecim boyunca gösterdiği destek ve rehberlik benim için her zaman çok kıymetli olacaktır. Ayrıca tez sürecimdeki destekleri, sabrı ve öğrettiği kıymetli bilgilerle yanımda olan eş danışman hocam Doç. Dr. Hilmi Tunahan AKKUŞ'a da derin şükran borçluyum. Akademik disiplini, anlayışı ve yol göstericiliği sayesinde bu çalışmanın tamamlanmasına önemli katkılarda bulunmuştur. Bu süreçte bilgi birikimlerini paylaşan, akademik gelişimime destek olan tüm öğretim üyelerine ve üniversitemin kıymetli akademik kadrosuna teşekkür ederim.

Hayatım boyunca bana her koşulda destek olan, sevgisiyle ve duasıyla her zaman yanımda olan canım anneme minnettarım. Fedakârlıkları, sabrı ve yüreğindeki sonsuz sevgi olmasaydı, bugün burada olmam mümkün olmazdı. En büyük teşekkürü ise hayatımın en kıymetli yol arkadaşı, eşim Faruk Özel'e borçluyum. Bu zorlu süreçte gösterdiği anlayış, sabır ve motivasyon desteği sayesinde bu çalışmayı tamamlayabildim. Varlığı, bana hem huzur hem de güç verdi. Her zaman yanımda olduğun için sana sonsuz teşekkür ederim. Bu çalışmanın, alanında faydalı bir katkı sağlaması temennisiyle...

**Balıkesir, 2026**

**Fatma Feyza Özel**

# 1. GİRİŞ

Son yıllarda finans dünyasında kripto varlık piyasaları geleneksel finansal araçlardan farklı olarak merkeziyetsiz yapıları, yüksek getiri potansiyelleri ve sundukları teknolojik yeniliklerden dolayı ilgi odağı olmuştur. Bu durum giderek daha fazla yatırımcının dikkatini çeken popüler alan haline gelmiştir. Fakat bu piyasaların hızlı büyümesi, yüksek volatilité, meydana gelen fiyat balonları, ani düşüşler; yatırımcılar, akademisyenler ve düzenleyici kurumlar için önemli risk ve belirsizlikler oluşturmaktadır. Bu durum finansal istikrarın sağlanması için hızlı bir şekilde stratejik önlemlerin alınması gerektiğini ortaya koymaktadır.

Kripto paraların hızlı değer değişimleri ve sağladığı anonimlik, piyasada Bitcoinica, Silk Road ve Thodex gibi örneklerde görüldüğü üzere haksız kazanç olaylarına zemin hazırlamaktadır (Christin, 2012; Teiglve et al., 2013; Canbaz ve Berkun, 2021). Bu tür durumlar, kripto paraların hukuki altyapısı ve düzenleyici çerçevenin önemini artırmaktadır. Japonya gibi bazı ülkeler, kripto paraların yasal çerçeveye oturtulmasını teşvik ederken, Amerika Birleşik Devletleri'nde Menkul Kıymetler ve Borsa Komisyonu (Securities ve Exchange Commission - SEC) ile Emtia Vadeli İşlemler Komisyonu (Commodity Futures Trading Commission - CFTC) gibi düzenleyici kurumlar, yatırımcıların korunması amacıyla sıkı denetimler uygulamaktadır. Avrupa Birliği ise 5. Kara Para Aklamayı Önleme Direktifi (5th Anti-Money Laundering Directive - 5AMLD) kapsamında, finansal suçlarla mücadeleyi güçlendirmek adına Müşterini Tanı (Know Your Customer - KYC) ve Kara Para Aklamayı Önleme (Anti-Money Laundering - AML) ilkelerine dayanan kapsamlı düzenlemeler getirmiştir (Blockchain Türkiye Platformu, 2019; Carlson, 2024).

Bu çalışma, kripto varlık piyasalarında fiyat balonlarının tahmin edilmesi konusunda makine öğrenmesi temelli modellerin uygulanabilirliğini kapsamlı biçimde ortaya koymayı hedeflemektedir. Kullanılan veri seti, Bitcoin'in kendi fiyat dinamikleri (Kapanış Fiyatı, Açılış Fiyatı, En Yüksek Fiyat, En Düşük Fiyat, İşlem Hacmi) ile altın, petrol, uluslararası borsa endeksleri, volatilité göstergeleri gibi makroekonomik ve finansal değişkenleri içermektedir. Bu çok boyutlu yaklaşım, literatürde sıkça vurgulanan analiz yaklaşımlarını temel alarak, fiyat balonlarının daha erken ve güvenilir şekilde tespit edilmesini amaçlamaktadır. Araştırma sonuçları, mevcut ekonometrik testler ile makine öğrenmesi modellerinin birbirini tamamlayıcı nitelikte olduğunu ortaya koymaktadır. Böylelikle

yatırımcıların risk yönetimi ve düzenleyici otoritelerin önleyici politikalar geliřtirmesi aısından stratejik temel oluřturmayı hedeflemektedir.

Arařtırma toplam dokuz blmden oluřmakta olup giriř blmn takip eden ikinci blmde kripto varlık trlerinin temel zellikleri, kullanım alanları ve finansal dinamikleri hakkında detaylı bilgiler sunulmaktadır. nc blmde arařtırmanın kavram aısından btnlğnn korunması ve amacının daha net ifade edilmesi iin fiyat balonları hakkında detaylı bilgi verilmiřtir. Drdnc blmde kripto varlıkların fiyatlarında byk dalgalanmalara neden olan nemli olaylar, krizler aıklanmaktadır. Beřinci blmde kripto varlıklarda balonların varlıđını tespit eden, kripto varlıkların diđer deđiřkenlerle iliřkisini analiz eden ve ML yntemleriyle kripto paraların deđerini/ynn tahmin etmeye alıřan literatr alıřmalarına yer vermektedir. Altıncı blmde alıřmanın yntemine ve kapsamına yer verilmiř olup yedinci blmde makine đrenmesi modellerinden bahsedilmektedir. Sekizinci blmde de ise elde edilen bulgular sunulmuř ve dokuzuncu blmde sonu ve nerilere iliřkin genel deđerlendirmeler yapılmıřtır. Bylelikle yapılmıř olan tez alıřması kripto varlık piyasalarının fiyat balonlarının tespiti ve nedenlerine katkıda bulunurken, aynı zamanda yatırımcıların bilinli kararlar almasını destekleyecek, risklerin nceden tespit edilmesine ynelik stratejilerin geliřtirilmesinde nemli referans olmayı amalamaktadır.

## 2. KRİPTO VARLIKLAR

İlk olarak Bitcoin’le birlikte para olarak ortaya çıksa da sonradan geliştirilen diğer kripto varlıklarla beraber kripto varlıklar toplam beş sınıfa ayrılmaktadır. Kripto varlıklar, dijital teknolojilerin ve blokzinciri altyapısının sunduğu yenilikçi çözümlerle, finansal sistemlerin ötesinde yeni ekosistem oluşturmuştur. Bu bölümde, kripto paralar, sanal evren (metaverse), değiştirilemez jeton (NFT), merkeziyetsiz otonom organizasyonlar (DAO) ve merkeziyetsiz finans (DeFi) ele alınarak her varlık türünün temel özellikleri, kullanım alanları ve finansal dinamikleri incelenmektedir.

### 2.1 Kripto Para

Kripto para kavramı, finansal sistemlerdeki merkeziyetçiliği ortadan kaldırmak ve aracılara duyulan gereksinimi azaltmak amacıyla geliştirilen dijital para birimlerinin temelini oluşturmaktadır. Bu alandaki devrimsel adım, Bitcoin’in doğuşuyla başlamıştır. Kripto para fikrinin temelleri, 1980’lerin sonlarında ve 1990’ların başlarında atılmıştır (Aleksejeva et al., 2021). Ancak, 2008 yılında Satoshi Nakamoto takma adını kullanan kişi veya grup tarafından yayımlanan “Bitcoin: A Peer-to-Peer Electronic Cash System” başlıklı makale, modern kripto paraların başlangıç noktası olmuştur (Nakamoto, 2008). Bu makale, herhangi merkezi otoriteye ihtiyaç duymadan, kullanıcılar arasında doğrudan işlem yapabilmeyi mümkün kılan Bitcoin protokolünü tanıtmıştır. 2009 yılında hayata geçirilen Bitcoin’in ilk bloğu, madencilik yoluyla “Genesis Block” adıyla üretilmiştir. Bu bloğun içinde yer alan “The Times 03/Jan/2009 Chancellor on brink of second bailout for banks” ifadesi, geleneksel bankacılık sistemine karşı tepkiyi simgelemektedir (Caetano, 2015; Judmayer et al., 2022). Oluşumunun ardından bugün Bitcoin, finansal işlemlerde güvenli transfer sağlamayı amaçlayan, merkezi otoriteden bağımsız, kriptografi ve dağıtık hesaplama teknikleri üzerine inşa edilmiş dijital para birimidir (İnci ve Alper, 2018; Akgül vd., 2022). Kripto paralar merkeziyetsizliklerinin yanı sıra, geleneksel para birimlerine kıyasla daha hızlı ve daha ucuz transfer edebilme yapılarıyla öne çıkmaktadırlar. İşlemlerin doğrulanması ve saklanması, “blokzinciri” (blockchain) adı verilen halka açık, dağıtık defter teknolojisi sayesinde gerçekleşmektedir. Blokzincirinde, her blok, önceki bloğun hash değeriyle birbirine bağlanarak zincirin bütünlüğü ve güvenliği sağlanmaktadır. Bu nedenle herhangi müdahale veya değişiklik durumunda tüm zincir etkilenmektedir (Sompolinsky and Zohar, 2013; Ali et al., 2016). Bitcoin’in toplam arzının 21 milyon ile sınırlı tutulması, enflasyona karşı doğal koruma mekanizması oluştururken, düşük işlem ücretleri ve hızlı transfer süreleri de kullanıcılar için önemli avantajlar sağlamaktadır. Ayrıca işlemlerin yüksek derecede

anonimlik sağlanarak gerçekleştirilmesi, kullanıcıların dijital cüzdanlar aracılığıyla özel anahtarlarla varlıklarına erişebilmesi, ancak bu anahtarların kaybedilmesi durumunda erişimin tamamen imkânsız hale gelmesi gibi önemli sorumlulukları da beraberinde getirmektedir (Segendorf, 2014; Güven ve Şahinöz, 2018).

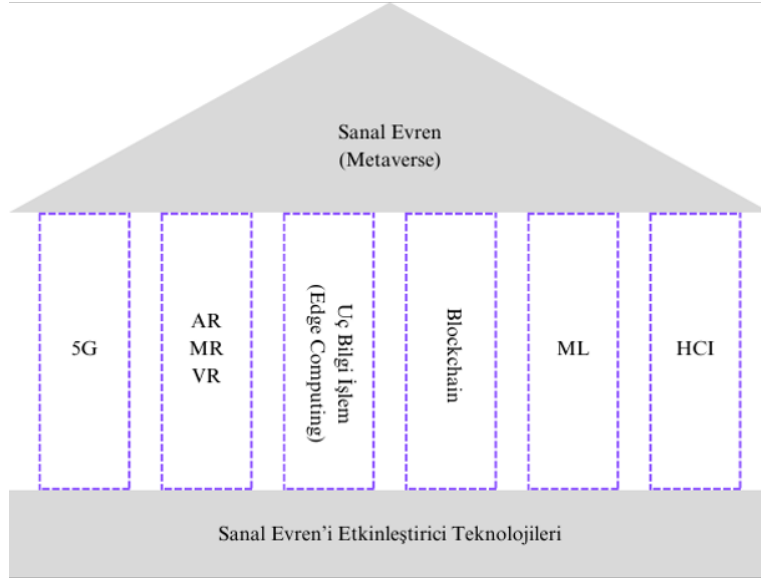
Kripto para birimlerinin yapısı, sadece işlem ve transfer mekanizmalarıyla sınırlı kalmayıp, kullanım amaçları ve teknik özellikler bakımından da çeşitlilik göstermektedir. Örneğin üç büyük kripto parayı incelediğimizde; Bitcoin, çoğunlukla “dijital altın” olarak anılmaktadır. Uzun vadeli değer saklama aracı olarak konumlanmaktadır. Bu durumun yanı sıra Ethereum, sunduğu akıllı sözleşme (smart contract) altyapısıyla merkeziyetsiz uygulamaların geliştirilmesine olanak tanıyan güçlü bir platformdur. Ripple (XRP) ise özellikle bankalar arası hızlı ve düşük maliyetli para transferleri için tasarlanmış, finansal kurumlara odaklanan çözüm sunmaktadır. Bu çeşitlilik, kripto ekosisteminin farklı sektörlerde kullanılabilmesine ve teknolojik inovasyonun yaygınlaşmasına önemli ölçüde katkı sağlamaktadır.

Sonuç olarak kripto para alanı merkeziyetsizlik, dağıtık defter teknolojisi ve özellikle bazı kripto paraların önceden belirlenmiş maksimum arz sınırına sahip olması gibi temel özellikleriyle, modern finansal sistemlerin dönüşümünü simgelemektedir. Bu yapısal özellikler, kripto paraların hem bireysel kullanıcılar hem de kurumsal yatırımcılar tarafından benimsenmesinde, finansal işlemlerin güvenliği ve maliyet etkinliğinde önemli rol oynamaktadır. Söz konusu teknolojik yenilik, blokzinciri altyapısı sayesinde, küresel ölçekte güvenli, şeffaf ve hızlı işlem imkânları sunarak finansal sistemlerin geleceğini şekillendirmektedir (Nguyen, 2016).

## **2.2 Sanal Evren (Metaverse)**

Metaverse, Neal Stephenson’ın (1994) Snow Crash romanında ortaya çıkmaktadır. 2020’lerden itibaren teknoloji dünyasında büyük ilgi gören bir terimdir. Gerçek insanların avatarlarıyla yani sanal temsilleriyle etkileşimde bulunduğu metaverse, kesintisiz ve paylaşılan dijital ortamları ifade etmektedir. Bu şekilde literatüre kazandırılan metaverse blokzinciri, sanal gerçeklik (Virtual Reality - VR), artırılmış gerçeklik (Augmented Reality - AR) ve karma gerçeklik (Mixed Reality - MR) gibi ileri teknolojilerin entegrasyonu ile günümüz dijital ekonomisinde önemli dönüşüm aracı haline gelmiştir. Özellikle 2020’lerden itibaren, Meta gibi önde gelen teknoloji devlerinin bu alana yaptığı büyük yatırımlar, metaverse’ün potansiyelini artırarak, dijital varlıkların benzersizliğini ve değiştirilemez

özelliklerini ön plana çıkarmaktadır. Dünyalar olarak tanımlanan bu kavram, 2021’de popülaritesini artırarak, kullanıcıların dijital ortamlarda yaşamını yeniden şekillendirmesine yönelik vizyon sunmaktadır (Kim, 2021).



**Şekil 2.1:** Metaverse oluşturan bileşenler

Şekil 2.1’de metaverse’ü oluşturan ana bileşenler gösterilmiştir. Bu bileşenler 5G, AR, VR ve MR, makine öğrenmesi (ML), insan-bilgisayar etkileşimi (HCI) teknolojileri yer almaktadır. Kullanıcıların avatarlarıyla üç boyutlu etkileşimde bulunabileceği, sosyal etkinliklere katılabileceği ve içerik oluşturabileceği platform sunan bu bileşenler, metaverse deneyimini şekillendirmektedir. Örneğin, Roblox ve Fortnite gibi oyun platformları, kullanıcıların kendi dünyalarını oluşturabilmelerine olanak tanırken, Nvidia’nın Omniverse platformu, 3D dünyaların sanal evrenler arasında bağlantı kurmasına yardımcı olmaktadır (Cheng et al., 2022).

Metaverse kullanıcılara eğlence, ticaret ve sosyal etkileşimin ötesinde, eğitimde de devrim niteliğinde fırsatlar sunmaktadır. VR destekli eğitim, öğrencilerin sanal alanlarda daha etkileşimli ve katılımcı şekilde öğrenmelerini sağlamakta, böylece fiziksel sınıf ortamlarının sınırlarını aşarak, global öğrenme deneyimi sunmaktadır (Mystakidis, 2022).

Metaverse’ü sadece geçici moda veya popülerlik akımı olarak görmekten ziyade, teknolojik evrimin doğal sonucu olarak değerlendirmek gerekmektedir. Smart ve ark. (2007), internetin sürekli açık bir dijital ekosisteme evrileceğini ve sanal dünyaların bu evrimde belirleyici

olacağını belirtmiştir. Günümüzde sanal ortamlar arası etkileşim ve geçiş, metaverse deneyiminin zenginleşmesinde önemli bir altyapı sunmaktadır.

Metaverse, karmaşık yapı ve çok yönlü deneyim sunma potansiyeli taşıırken, henüz tam anlamıyla şekillenmiş kavram değildir. Kullanıcıların fiziksel ve sanal dünyalarda araya gelerek etkileşimde bulunmasını sağlayan bu platformun, dijital ekonomileri ve sosyal etkileşimleri dönüştürmesi beklenmektedir. Ancak metaverse'ün tam potansiyelini gerçekleştirebilmesi için kullanıcı deneyiminin sorunsuz şekilde sağlanması, yüksek maliyetler, güvenlik ve gizlilik konularının ele alınması gerekmektedir (Mystakidis, 2022). Diğer kripto varlıklardan farklı olarak metaverse'de yazılım yanında donanıma da ihtiyaç duyulmaktadır.

### **2.3 Değiştirilemez Jeton (Non-Fungible Token - NFT)**

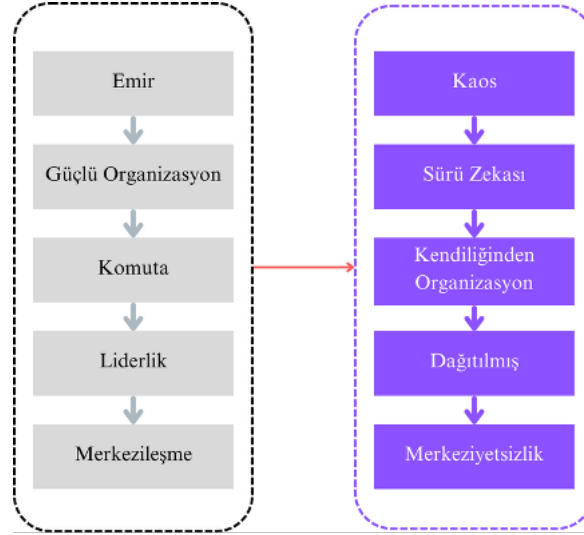
NFT'ler son yıllarda dijital ekonomide dikkat çeken dönüşüm aracı haline gelmiştir. NFT'ler teknolojinin sunduğu olanaklarla geliştirilen ve her biri benzersiz olan dijital varlık temsilidir. NFT'ler, değiştirilemez ve takas edilemez tokenlar olarak da anılmaktadır. Her NFT, kendine özgü meta veriler içermektedir. Böylece başka token ile birebir değiştirilememektedir. Bu yönüyle değiştirilebilir kripto paralardan (örneğin Bitcoin, Ethereum gibi) ayrılmaktadır (Chohan, 2021c). NFT'lerin en dikkat çekici özelliklerinden biri, dijital varlığın orijinalliğini ve mülkiyetini blokzinciri üzerinde şeffaf biçimde kayıt altına alabilmesidir. Bu tokenlar, dijital sanat eserlerinden müzik ve video içeriklerine, koleksiyon ürünlerinden oyun içi varlıklara kadar geniş yelpazede kullanılabilir. Her NFT, ilgili dijital içeriğin kim tarafından oluşturulduğunu, kimler tarafından sahiplenildiğini ve ne zaman el değiştirdiğini gösteren dijital sertifika işlevi görmektedir. Bu özellik, özellikle telif hakları, lisanslama ve mülkiyet transferi gibi konuların dijital ortama taşınmasını kolaylaştırmakla birlikte güvenilir hale de getirmektedir. Başlangıçta Ethereum blokzinciri üzerinde geliştirilen NFT yapıları, zamanla Solana, Binance Smart Chain ve Polygon gibi diğer blokzinciri platformlarında da benimsenmiştir. Bu durum NFT kullanım alanlarını genişletmiştir (Ante, 2022). NFT'ler, yalnızca teknik yenilik değil, aynı zamanda yeni endüstrilerde devrim oluşturan ekonomik model olarak da öne çıkmaktadır. NFT'lerin ilk büyük çıkışı, sanat ve koleksiyon dünyasında yaşanmıştır. Özellikle dijital sanatçı Beeple'in NFT eserinin 69,3 milyon dolara satılması (Nadini et al., 2021), bu teknolojinin potansiyelini küresel ölçekte görünür kılmıştır. Bu gelişme, sanatçıların eserlerini dijital ortamda hem koruma altına almasını hem de dünya genelindeki alıcılara doğrudan sunmasını

mümkün kılmıştır. Öte yandan NFT'ler oyun sektöründe de önemli kırılma noktası oluşturmuştur. Örneğin CryptoKitties gibi projeler, dijital varlıkların alınıp satılabildiği yeni ekonomik ekosistemin öncüsü olmuştur (Serada, 2020). Bugün birçok oyun ve sanal evren platformu, kullanıcıların oyun içi öğeleri NFT formatında edinmesine ve takas etmesine olanak tanımaktadır. NFT'lerin etkisi yalnızca sanat ve oyunla sınırlı kalmamış, müzik ve eğlence sektörlerine de yayılmıştır. Müzisyenler, içerik üreticileri ve markalar, kendi NFT koleksiyonlarını oluşturarak dijital platformlar üzerinden yeni gelir kaynakları elde etmekte böylelikle hayranlarıyla doğrudan etkileşim kurma fırsatı yakalamaktadırlar. Ancak NFT piyasasının hızlı büyümesi beraberinde bazı riskleri de oluşturmuştur. NFT'lerin piyasa değeri, çoğunlukla insanların onlara olan ilgisine ve gelecekte değer kazanacağına dair beklentilerine göre yükselip düşmektedir. Başka deyişle, hem arz-talep dengesi hem de “ileride daha çok değerlenir” düşüncesi fiyatlar üzerinde önemli rol oynamaktadır. Bu nedenle NFT'ler yatırım aracı olarak cazip görünse de yüksek volatilité ve değer belirsizliği riski barındırmaktadır (Nadini et al., 2021). NFT varlık süreci göz önüne alındığında tüm bu gelişmeler bize göstermektedir ki, NFT'ler yalnızca dijital varlıkların ticarileştirilmesinde değil, aynı zamanda dijital çağda mülkiyet kavramının yeniden tanımlanmasında önemli rol oynamaktadır.

#### **2.4 Merkeziyetsiz Otonom Organizasyon (Decentralized Autonomous Organization - DAO)**

DAO'lar, blokzincir teknolojisinin sunduğu imkanlarla merkezi otorite veya hiyerarşik yapıya ihtiyaç duymadan işleyen, akıllı sözleşmelerle kodlanmış organizasyonlardır. Bu yapılar, tüm yönetim ve operasyonel kuralları blokzincir üzerinde şeffaf ve değiştirilemez şekilde kaydetmektedir. Bu durum katılımcılar arasında eşitlik, gönüllülük ve iş birliğine dayalı yönetim modeli oluşturmaktadır. DAO'lar, bireylerin çevrimiçi platformlarda kendilerini koordine etmelerine ve organizasyonun işleyişine aktif olarak katkıda bulunmalarına olanak tanırken, şeffaflık ve hesap verebilirlik sağlamaktadır. Bu durum, yatırımcıların toplumsal fayda odaklı ortak hedefler doğrultusunda hareket etmelerini mümkün kılmaktadır (Beck et al., 2018; De Filippi and Wright, 2018). DAO'ların kökeni, Şekil 2.2'de gösterildiği gibi doğal ekosistemlerdeki kendiliğinden organizasyon olgularına kadar uzanmaktadır. Örneğin, balık sürüleri, bireysel balıkların basit hareketleriyle karmaşık yönelim ve savunma davranışları sergileyen yapı oluşturmaktadır. Bu tür organizasyonlar, merkezi kontrol olmaksızın, grup içindeki bireylerin koordinasyonu ile işlevsellik kazanmaktadır. Benzer şekilde, internet üzerindeki siber hareket organizasyonları (Cyber

Movement Organizations - CMO) ve Dağıtılmış Yapay Zekâ (Distributed Artificial Intelligence - DAI) sistemleri de kendiliğinden organizasyon yapılarının örnekleri olarak DAO'nun temellerini oluşturmaktadır (Wang et al., 2019).



**Şekil 2.2:** Merkeziyetsizlikten merkezileşmeye analizi

Blokzincir teknolojisinin sağladığı dağıtık veri depolama, değişmezlik ve şeffaflık özellikleri, DAO'ların uygulanabilirliğini artırmaktadır. 2015 yılında Ethereum platformunun bizleri akıllı sözleşmelerle tanıştırması, DAO'ların gelişimine önemli zemin oluşturmuştur. Akıllı sözleşmeler, organizasyonların kendi kendini yönetmesini ve önceden tanımlanmış iş mantığına göre otomatik olarak çalışmasını sağlamaktadır. 2016 yılında kurulan The DAO, token satışı yoluyla yaklaşık 150 milyon dolar fon toplayarak en büyük ölçekli erken dönem DAO örneklerinden biri olmuştur (Wang et al., 2019). Ancak DAO'lar, güvenlik, yasal belirsizlikler ve performans gibi çeşitli zorluklarla karşı karşıyadır. Örneğin, 2016'da yaşanan The DAO saldırısı, sistemdeki güvenlik açısından yararlanarak 50 milyon doların üzerinde Ether'in çalınmasına yol açmıştır. Ayrıca mevcut yasal düzenlemelerle uyum sağlama konusunda zorluklar da söz konusudur (Hassan and De Filippi, 2021).

## 2.5 Merkeziyetsiz Finans (Decentralized Finance - DeFi)

DeFi ekosistemi, blokzincir teknolojisi ve akıllı sözleşmelerden yararlanarak merkezi olmayan finansal hizmetlerin sunulmasını mümkün kılan ve bu yönüyle geleneksel finans sistemine alternatif oluşturan finansal teknoloji alanıdır. Bu ekosistem, borç verme, borç alma, ticaret ve sigorta gibi geleneksel finansal hizmetleri merkeziyetsiz yapıda sunmaktadır. Bu finansal işlemler, merkezi bir aracı kuruma ihtiyaç duyulmaksızın, Ethereum gibi blokzincir platformları üzerinde çalışan akıllı sözleşmeler aracılığıyla

yürütülmektedir (Chohan, 2021a). Ayrıca kullanıcılar varlıklarının kontrolünü tamamen ellerinde tutabilmekte ve diledikleri gibi kullanabilmektedir. Geleneksel finans sisteminde bu tür işlemler genellikle yavaş, maliyetli ve karmaşık süreçlerle yönetilmektedir. DeFi sayesinde kullanıcılar doğrudan birbirleriyle etkileşim kurarak daha basit işlemler gerçekleştirebilmektedir (Zetsche et al., 2020). Akıllı sözleşmeler, belirli koşullar sağlandığında otomatik olarak yürürlüğe girmektedir. Ayrıca DeFi kullanıcılar arasında şeffaf, güvenli, hızlı ve düşük maliyetli işlemler gerçekleştirilmesine imkan tanımaktadır. Bu da DeFi'yi, finansal hizmetlerin küresel ölçekte herkes için erişilebilir hale geldiği sistem olarak öne çıkarmaktadır (Zetsche et al., 2020). Böylelikle yenilikçi ve şeffaf finansal sistem oluşturmayı amaçlamaktadır (Chen and Bellavitis, 2020). DeFi ekosisteminin temelini oluşturan bu merkeziyetsizlik, dijital varlıkların işlem gördüğü platformlar olan kripto varlık borsalarında da temel ayrımı beraberinde getirmektedir.

Kripto varlık borsaları, dijital varlıkların alım satım işlemlerinin gerçekleştirildiği çevrimiçi platformlar olarak tanımlanmaktadır. Bu borsalar, merkezi (CEX) ve merkeziyetsiz (DEX) olmak üzere iki ana kategoriye ayrılmaktadır. Merkezi borsalar, belirli şirket ya da kurum tarafından yönetilmekte, yüksek likidite ve kullanıcı dostu arayüzler sunarken, hacklenme riskine karşı daha savunmasız olabilmektedir. Öte yandan, merkeziyetsiz borsalar tamamen blokzinciri teknolojisi üzerine inşa edilmektedir. Kullanıcıların varlıklarını kendi cüzdanlarında saklamalarına olanak tanıyarak güvenliği artırmaktadır. Binance, Coinbase, Kraken ve Bitfinex gibi platformlar, geniş işlem yelpazeleri, gelişmiş ticaret araçları ve çeşitli güvenlik önlemleri ile piyasada öne çıkmaktadır (Priya et al., 2023). Bu avantajların yanı sıra DeFi'nin karşılaştığı bazı zorluklar da bulunmaktadır. Akıllı sözleşmelerdeki hatalar, sistem açıkları veya dolandırıcılık girişimleri gibi riskler, kullanıcıların varlıklarını kaybetme olasılığını artırmaktadır. Ayrıca merkeziyetsiz yapısı nedeniyle DeFi ekosisteminde düzenleme ve denetim eksikliği bulunmaktadır. Bu durum DeFi'nin daha geniş kitleler tarafından benimsenmesi ve güvenilir finansal sistem olarak kabul edilmesi için önemli engel teşkil etmektedir (Chohan, 2021a).

### 3. FİYAT BALONLARI

Bu bölümde fiyat balonları olgusu, hem kavramsal hem de tarihsel yönleriyle ele alınmaktadır. Öncelikle balon kavramının ne anlama geldiği, hangi ekonomik ve psikolojik dinamiklerle ortaya çıktığı ve davranışsal finans literatüründe nasıl yorumlandığı açıklanmaktadır. Ardından fiyat balonlarının neden yalnızca bireysel yatırımcılar için değil, aynı zamanda finansal sistemin istikrarı ve makroekonomik dengeler açısından da kritik risk faktörü olduğu tartışılmaktadır.

Bölümün devamında, tarih boyunca yaşanan önemli balon örnekleri üzerinden kitlesel yanılgı ve piyasa dinamikleri arasındaki ilişki incelenmektedir. Bu kapsamda klasik dönemden lale çılgınlığı, Güney Denizi ve Mississippi balonları; modern dönemden 1929 Büyük Buhran ve Japonya varlık balonu; dijital çağdan dot.com balonu ve 2008 mortgage krizi ele alınmaktadır. Son olarak kripto varlık piyasalarında gözlemlenen güncel balon örnekleri değerlendirilmektedir.

#### 3.1 Kavramsal Açıklama

Genel anlamda fiyat balonu kavramı, varlık fiyatlarının temel değerlerden sapması şeklinde tanımlanmaktadır. Fiyat balonları üzerine yapılan araştırmalar, bu gelişimin farklı coğrafyalarda çeşitli biçimlerde ortaya çıktığını göstermektedir. Kindleberger ve Aliber (2005), finansal balonları "varlık fiyatlarındaki sürdürülemez artışlar" olarak tanımlamaktadır. Bu ham artışların genellikle piyasa katılımcılarının aşırı iyimserliği ile beslendiğini vurgulamaktadır. Bu nedenle balonlar yalnızca fiyat hareketleri değil, aynı zamanda ortak piyasa psikolojisinin yansıması olarak da değerlendirilmektedir. Diğer yaklaşıma göre ise fiyat balonu, fiyatların ani ve büyük artışlarının ardından gelen keskin düşüşleri ifade etmektedir. Bu durum özellikle yanlış fiyatlamaların sonucu olarak ortaya çıkmaktadır (Abreu and Brunnermeier, 2002). Santoni (1987), fiyat balonlarını, varlığın piyasada temel değerlerden sürekli ve sistematik olarak sapma göstermesi şeklinde açıklamaktadır. Blanchard ve Watson (1982) ise rasyonel balonlar kavramını geliştirerek, yatırımcıların gelecekteki fiyat artışlarına dair beklentilerinin kendi kendini besleyen döngü oluşturabileceğini göstermiştir. Bu teoriye göre, yatırımcılar varlığın temel değerinin üzerinde olduğunu bilseler bile, kısa vadeli kazanç beklentisiyle alım yapmaya devam edebilmektedirler (Blanchard and Watson, 1982).

Ekonomik modellerde fiyat balonlarını tam olarak anlamak oldukça güç olmuştur. 1990'lar boyunca finansal balonlar genellikle patladıkları aşamada fark edilmekteydiler. Geleneksel finans teorileriyle açıklanamayan bu durum, davranışsal finans modelleri aracılığıyla daha iyi anlaşılmaya başlanmıştır (Anavatan ve Kayacan, 2018). Shiller (2015), irrasyonel coşku (irrational exuberance) kavramıyla, piyasa katılımcılarının duygusal ve psikolojik faktörlerden etkilenecek şekilde rasyonel olmayan kararlar alabileceğini ve bunun balon oluşumuna zemin hazırlayabileceğini ortaya koymuştur.

### **3.2 Fiyat Balonlarının Önemi**

Fiyat balonları hem yatırımcılar hem de düzenleyici kurumlar açısından kritik role sahiptir. Balonlar, varlık fiyatlarının temel değerlerden sapması sonucu oluşan durumlar olarak sadece bireysel yatırımcıların zarar etmesine neden olmakla kalmamaktadır. Aynı zamanda finansal sistemlerde likidite sıkıntısı ve makroekonomik krizlerin habercisi olarak da görülmektedir. Bu nedenle fiyat balonlarının erken tespiti, piyasanın risk yönetimi stratejilerini geliştirebilmeleri ve düzenleyici kurumların önceden önlemler alabilmesi için büyük önem taşımaktadır (Santoni, 1987; Abreu and Brunnermeier, 2002; Oran, 2011).

Reinhart ve Rogoff (2009), sekiz yüzyıllık finansal kriz verilerini inceledikleri çalışmalarında, varlık fiyat balonlarının sistemik finansal krizlerin en önemli öncü göstergelerinden biri olduğunu tespit etmişlerdir. Buna göre balonların patlaması, yalnızca ilgili varlık sınıfında değil, kredi piyasaları, reel ekonomik aktivite ve istihdam üzerinde de domino etkisi oluşturabilmektedir. Ayrıca yatırımcı duyarlılığı ve spekülasyon davranışları, balon oluşumunun temel etmenleri arasında yer aldığı da çıkarılmaktadır (Shiller, 2000; Kristoufek, 2013; Wei, 2017). Özellikle sosyal medya çağında, bilgi yayılımının hızlanması ve sürü davranışının güçlenmesi, balon dinamiklerini daha karmaşık hale getirmektedir. Baker ve Wurgler (2007), yatırımcı duyarlılığının ölçülmesi ve bu ölçümlerin balon tahmininde kullanılması konusunda öncü çalışmalar yapmışlardır. Bu bağlamda, erken uyarı sistemleri ve gelişmiş modelleme yaklaşımlarının kullanılması, gelecekte oluşabilecek balonların öngörülmesi ve olası finansal krizlerin önlenmesinde stratejik avantaj sağlamaktadır. Phillips et al., (2011) tarafından geliştirilen SADF testi gibi ekonometrik yöntemler, balonların gerçek zamanlı tespitinde önemli ilerlemeler kaydetmiştir.

Kripto varlık piyasalarının ortaya çıkışıyla birlikte, fiyat balonlarının önemi daha da artmıştır. Geleneksel finansal varlıklardan farklı olarak, kripto varlıkların temel değerinin

belirlenmesindeki zorluklar ve piyasanın 7/24 işlem görmesi, balon dinamiklerini daha karmaşık hale getirmektedir (Corbet et al., 2019). Bu nedenle makine öğrenmesi gibi gelişmiş analitik yöntemlerin kullanılması, kripto varlık piyasalarındaki tespit edilen balonların tahmini için kritik öneme sahiptir.

### **3.3 Geçmişte Yaşanan Önemli Fiyat Balonları**

Tarih boyunca farklı dönemlerde ve coğrafyalarda pek çok spekülasyon balon yaşanmıştır. Bu balonların ortak özellikleri ve farklılıkları, günümüz piyasalarını anlamak için önemli bilgiler sunmaktadır.

#### **3.3.1 Klasik dönem balonları**

İlk örneklerden biri, 1634–1637 arasındaki Lale Çılgınlığı (tulipmania) olup, Hollanda'da nadir lale soğanlarına yönelik talebin astronomik fiyat artışlarına yol açmasıyla belirginleşmiştir (Garber, 2001). Fiyatlar o denli yükselmiştir ki, lale soğanı Amsterdam'da evle eşdeğer hâle gelmiştir. Ancak 1637'de balon patlamış ve pek çok yatırımcı zarara uğramıştır (Oran, 2011; Yıldırım, 2021). Mackay (2003), bu olayı "olağanüstü popüler yanılgılar ve kalabalıkların çılgınlığı" olarak nitelendirmiş ve kolektif irrasyonelitenin ilk örneklerinden biri olarak tarihe geçmiştir. Lale Çılgınlığı'ndan yaklaşık 80 yıl sonra, İngiltere'de Güney Denizi Balonu olarak bilinen olayda İngiltere'de, Güney Denizi Balonu olarak bilinen olayda, 1711'de kurulan Güney Denizi Şirketi'nin hisseleri büyük kâr vaatleri ve devlet borçlarını üstlenmesi nedeniyle aşırı değerlendirilmiş, 1720'de sert düşüşle sonuçlanarak yatırımcılara ağır kayıplar yaşatmıştır (Bali ve Büyükşalvarcı, 2011; Eren ve Saraçoğlu, 2017). Dale (2014), bu balonun oluşumunda politik manipülasyon ve asimetrik bilginin rolünü detaylı olarak incelemiştir. Aynı dönemde Fransa'da Mississippi Balonu, John Law tarafından kurulan Mississippi Şirketi'nin Kuzey Amerika'daki ticari girişimlerine dair abartılı beklentilerin oluşturduğu büyük fiyat artışlarının ardından hızla çökmesiyle patlak vermiş, ülke ekonomisi ciddi krize sürüklenmiştir (Schubert, 1988; Akıncı vd., 2014).

#### **3.3.2 Modern dönem balonları**

Daha yakın tarihlerde 1929 Büyük Buhran, hisse senedi piyasasındaki aşırı spekülasyonlar ve krediyle alım furyası sonucu ABD'de başlayan ve küresel çapta etkili olan kriz olarak kayıtlara geçmiştir (Buluş ve Kabaklarlı, 2010; Yavuz, 2021). Borsa çöküşüyle ABD ekonomisi büyük durgunluğa girmiş, yüksek işsizlik ve iflaslar dünya geneline yayılmıştır.

Galbraith (2009), bu dönemde marjin alımlarının ve kaldıraçlı işlemlerin balon oluşumundaki kritik rolünü vurgulamıştır.

Japonya'da 1980'lerin sonunda yaşanan varlık fiyatları balonu, hem gayrimenkul hem de hisse senedi piyasalarını kapsayan geniş çaplı balon olarak tarihe geçmiştir. 1989'da zirveye ulaşan Nikkei endeksi, sonraki yıllarda %60'tan fazla değer kaybetmiş ve Japonya "kayıp on yıl" olarak adlandırılan uzun durgunluk dönemine girmiştir (Okina et al., 2001).

### **3.3.3 Dijital çağ balonları**

Dot.com Balonu olarak anılan ve 2000'lerin başında internet şirketlerinin aşırı değerlenmesine dayalı balon, teknolojik yeniliklerin yatırımcılar tarafından fazlasıyla iyimser yorumlanmasıyla şişmiş, gerçekte sürdürülebilir iş modeli olmayan pek çok şirket iflas ederek borsalarda ciddi kayıplara neden olmuştur (Panko, 2008). Ofek ve Richardson (2003) yapmış oldukları çalışmada bu dönemde halka arz kısıtlamalarının kaldırılmasının balon dinamiklerini nasıl etkilediğini göstermişlerdir. Son olarak 2008 Mortgage Krizi, ABD konut piyasasında düşük faiz ve kolay kredi koşullarıyla fiyat balonu oluşturmuştur. Konut fiyatlarındaki düşüş ve kredi geri ödemelerinin aksamasıyla balonun patlaması sonucunda küresel finansal krize dönüşmüştür (Demir vd., 2008). Mian ve Sufi (2009), bu krizde subprime mortgage kredilerinin ve menkul kıymetleştirmenin rolünü detaylı olarak analiz etmişlerdir.

### **3.3.4 Kripto varlık balonları**

Yüzyılın en dikkat çekici balon örneklerinden biri, kripto varlık piyasalarında gözlemlenmektedir. Bitcoin'in 2017 yılında 20.000 dolara yaklaşması ve ardından %80'den fazla değer kaybetmesi, klasik balon dinamiklerinin dijital varlıklarda da geçerli olduğunu göstermiştir (Cheah and Fry, 2015). Baur et al., (2018), Bitcoin'in spekülasyon yatırım aracı olarak kullanıldığını ve geleneksel varlık sınıflarından farklı balon karakteristikleri sergilediğini ortaya koymuşlardır.

Bu örnekler, varlık fiyatlarının ekonomik gerçeklerden kopması, kitlesel psikoloji ve yatırımcı irrasyonelliği gibi faktörlerin fiyat balonu oluşumunda nasıl kritik rol oynadığını açıkça göstermektedir. Balonlar, sıklıkla ani ve keskin çöküşlerle sonlanarak hem yatırımcılar hem de reel ekonomi üzerinde ciddi tahribat oluşturacağını gözler önüne

sermiřtir. Tarihsel deneyimler, teknolojik yenilikler, denetimlerin gevřetilmesi, ařırı kredi geniřlemesi ve yatırımcı cořkusunun bir araya gelmesinin balon oluřunu iin ideal kořullar oluřturduęu gstermektedir (Kindleberger and Aliber, 2005).

## **4. KRİPTO VARLIKLARIN FİYATLARINI ETKİLEYEN ÖNEMLİ OLAYLAR**

Kripto para piyasası, geleneksel finansal sistemlere alternatif olarak gelişen, ancak yapısında barındırdığı risk odaklı davranışlar, düzenleyici eksiklikler ve teknik açıklar nedeniyle zaman zaman ciddi krizlerle gündeme gelen alandır. Bu bölümde, yalnızca kripto varlıklar üzerinde etkili olan krizler ayrıntılarıyla ele alınmakta ve piyasanın karşı karşıya olduğu risklere dikkat çekilmektedir.

### **4.1 Bitcoinica Dolandırıcılığı**

Bitcoinica, 2011 yılında Singapurlu genç girişimci Zhou Tong tarafından kurulan Bitcoin işlem platformudur. Kuruluşundan kısa bir süre sonra, Mart 2012'de Bitcoinica'nın web barındırma sağlayıcısı hacklenerek 43.554 Bitcoin çalınmıştır. Ardından çok geçmeden Bitcoinica tekrar hacklenmiştir. Bu seferde de 18.547 Bitcoin çalınmıştır. Tong, Bitcoinica'yı Intersango'ya satmasına rağmen siber saldırı olayları devam etmiştir. Belirtilen olayların önü alınamayınca site kapanmıştır. Bitcoinica dolandırıcılığı, kripto para borsalarının güvenliği konusunda önemli örnek olarak kripto para dünyasında yankı bulmuştur (Teigland et al., 2013).

### **4.2 Silk Road Operasyonu**

2011 yılında Ross Ulbricht tarafından kurulan Silk Road, Dark Web üzerinden yasadışı ticaret yapılmasına olanak tanıyan e-ticaret platformudur. İndirilebilir mobil uygulama (app) olmayan bu platform, Google Chrome gibi standart tarayıcılarla girilemeyen ve yalnızca Tor ağı üzerinden erişilebilen '.onion' uzantılı gizli web sitesi olarak faaliyet göstermiştir. Bu site üzerinden kullanıcılar anonim kalmak amacıyla Bitcoin kullanarak uyuşturucu, silah ve diğer yasadışı ürünlerin ticaretini gerçekleştirmişlerdir. 2013 yılında Federal Soruşturma Bürosu (Federal Bureau of Investigation - FBI) tarafından yapılan operasyonla site kapatılmış ve Ross Ulbricht tutuklanmıştır. FBI, Silk Road operasyonu sonucunda büyük miktarda Bitcoin ele geçirmiştir. Bu durum da kripto paraların yasadışı kullanımları hakkında geniş çaplı tartışmalar başlatmıştır (Christin, 2012).

### **4.3 Çin Hükümetinin Kripto Para Piyasalarına Yönelik Düzenleyici Müdahaleleri**

Çin, Bitcoin ve kripto para piyasalarına müdahale eden ilk ülkelerden biridir. 2013 yılında Çin Halk Bankası, Bitcoin'i yasa dışı kabul eden ve bankaların Bitcoin ile ilgili işlem yapmasını yasaklayan bildiri yayınlamıştır. Bildirinin ardından Bitcoin'in değeri %30

oranında düşüş yaşamıştır. Çin hükümetinin müdahaleleri, kripto paraların merkezi otoriteye bağlı olmamasından duyduğu rahatsızlık nedeniyle gerçekleştirmiştir. 2017 yılında ise Çin hükümeti, ilk coin arzını (Initial Coin Offering – ICO) yasaklamış ve kripto para borsalarının kapatılmasını istemiştir. 2019 yılında Bitcoin madenciliğini kısıtlayan düzenlemeler getirilmiştir. 2021’de ise Çin hükümeti kripto para ticaretini ve madenciliğini tamamen yasaklamıştır (Songur ve Ordu, 2023).

#### **4.4 Mt. Gox Krizi**

Japonya merkezli Mt. Gox, 2013–2014 döneminde dünya genelindeki Bitcoin işlemlerinin %70’inden fazlasını yönetmekteydi. Ancak şirket, 2014 yılında yaklaşık 850.000 adet Bitcoin’in kaybolduğunu duyurarak faaliyetlerini durdurmuş, ardından iflas başvurusunda bulunmuştur. Kriz, kripto tarihindeki ilk büyük sistemik çöküş olarak kabul edilmektedir. Mt. Gox krizi kripto varlık piyasalarında güvenlik ile denetim eksikliklerinin kritik olduğunu ortaya koymaktadır (Cheung et al., 2015).

#### **4.5 McAfee ve Kripto Manipülasyonu**

John McAfee, 1987 yılında kurduğu McAfee Associates ile dünyanın ilk ticari antivirüs yazılımını geliştiren isim olarak tanınmıştır. 1994 yılına kadar şirketini yöneten McAfee, aynı yıl istifa etmiş ve şirketten ayrılmıştır. 1996’da hisselerini satarak 100 milyon dolar gelir elde etmiştir. McAfee Associates 2010 yılında Intel tarafından satın alınmıştır. Son yıllarda kripto paralara olan ilgisi artan McAfee, 2017-2018 yıllarında Dogecoin, Reddcoin ve Verge gibi kripto paralar üzerinde manipülasyon yaparak yatırımcıların bu paraları satın almasını sağlamıştır. Ardından yükselen fiyatlardan satış yaparak büyük kazançlar elde etmiştir. Bu olaylar sonrasında ABD’de vergi kaçakçılığı suçlamasıyla karşı karşıya kalmıştır. İspanya’da da vergi kaçırma nedeniyle tutuklanan McAfee, 2021 yılında Barselona’daki hapisane hücrelerinde intihar etmiştir (Eğilmez, 2023).

#### **4.6 Elon Musk ve Dogecoin Dalgalanmaları**

Dogecoin, Billy Markus ve Jackson Palmer tarafından 2013 yılında şaka olarak oluşturulmuştur. Amacı kripto para piyasasının ciddiyetini kırmak ve daha erişilebilir kripto para birimi oluşturmaktır. Lansmanının ardından kısa sürede popülerleşen Dogecoin, topluluk desteğiyle önemli bir yere gelmiştir. Dogecoin Ocak 2014’te işlem hacmi diğer kripto paraları aşmış olup büyük kullanıcı kitlesi kazanmıştır (Chohan, 2021b). Elon Musk, Dogecoin’in en büyük destekçilerinden biri olarak bilinmektedir. Sık sık sosyal medya

üzerinden Dogecoin hakkında paylaşımlar yaparak fiyatını etkilemiştir. Musk'ın bu eylemleri Dogecoin'in fiyatında büyük dalgalanmalara yol açmıştır. Ancak aynı zamanda manipülasyon suçlamalarına da neden olmuştur. 2021 yılında Musk'ın tweetleri sayesinde Dogecoin büyük yükseliş yaşamış, ancak aynı yıl içinde fiyatında ani düşüşler meydana gelmiştir (Öncü, 2021).

#### **4.7 Thodex Skandalı**

Thodex, Türkiye'nin ilk lisanslı kripto para borsası olarak 2020 yılında uluslararası piyasalara giriş yapmış ve hızla büyümüştür. Ancak 2021 yılında Thodex'in sahibi Faruk Fatih Özer, yatırımcıların 2 milyar dolara yakın parasını alarak yurt dışına kaçmıştır. Kullanıcılar, borsanın kapandığını öğrenince paralarını çekememiş ve binlerce yatırımcı mağdur olmuştur. Olayın ardından Türkiye'de kripto para kullanımına ilişkin düzenleme ihtiyacı daha belirgin hale gelmiş ve kripto para ile ödeme yapma işlemleri yasaklanmıştır (Canbaz ve Berkun, 2021). Nitekim 16 Nisan 2021 tarihinde Türkiye Cumhuriyet Merkez Bankası tarafından "Ödemelerde Kripto Varlıkların Kullanılmamasına Dair Yönetmelik" yayımlanarak yürürlüğe girmiştir.

#### **4.8 FTX İflası**

2022 yılında dünyanın en büyük kripto para borsalarından biri olan FTX, kullanıcı fonlarının şeffaf olmayan biçimde kullanılması, iç denetim eksikliği ve yanlış yönetim uygulamaları sonucunda likidite krizine girmiştir. Ardından 11 Kasım 2022 tarihinde iflas başvurusunda bulunmuştur. Şirketin iflasının ardından kullanıcılar hesaplarındaki varlıkları çekememiş, panik satışlar başlamıştır. Bu olayın ardından Bitcoin dahil olmak üzere birçok kripto varlığın fiyatında %20'ye varan düşüşler gözlemlenmiştir. Olay, sektördeki yasal düzenlemelerin eksikliğini bir kez daha gündeme getirmiştir. Ayrıca kripto borsalarının kurumsal denetim altına alınması gerekliliği tartışmaya açılmıştır (The Guardian, 2022).

#### **4.9 Terra (LUNA) ve UST Krizi**

Do Kwon liderliği ile 2018 yılında piyasaya sürülen Terra ağı, özellikle stabil coini UST sayesinde büyük ilgi toplamıştır. Ancak Mayıs 2022'de Terra ekosistemi büyük çöküş yaşamıştır. 7 Mayıs 2022'de Terra'ya ait UST'nin fiyatı 1 dolar çıpasından saptmaya başlamıştır. Tarih 9 Mayıs'a geldiğinde 35 cente kadar düşmüştür. Terra'nın LUNA tokeni ise 12 Mayıs'ta birkaç dolardan birkaç sente kadar düşerek tarihi çöküş yaşamıştır. Bu

öküşün temel nedeni, UST'nin fiyatını sabitlemek için LUNA'nın arzının artırılmasıdır. Bu durum da LUNA'nın deęerinin hızla düşmesine neden olmuştur. Kripto para piyasalarındaki volatilité ve spekülasyon bu çöküşü daha da hızlandırmıştır. Terra'nın çöküşü, piyasa aktörleri tarafından “ölüm sarmalı” olarak adlandırılan süreçle ilişkilendirilmiştir (Liu et al., 2023).

## 5. LİTERATÜR İNCELEMESİ

Bu bölümde, kripto varlık piyasalarındaki balon oluşumu, ilgili değişkenler, yapay zeka tabanlı tahmin modelleri ve fiyat balonlarının öngörülmesine yönelik akademik çalışmalar sistematik çerçevede incelenmiştir.

### 5.1 Kripto Varlıklarla İlgili Balon Çalışmaları

Kripto varlık piyasalarında fiyat balonlarının tespiti, literatürde sıklıkla ele alınmaktadır. Cheung et al., (2015), Mt. Gox borsasındaki BTC fiyatlarını inceleyerek 2011 ile 2014 arasında 66 ila 106 gün süren üç büyük balon dönemi gözlemlemiştir. Zeren ve Esen (2018) ise 2010-2017 dönemi günlük BTC fiyatlarında GSADF testi uygulamış ve sonuçların %99 güvenilirlikle çoklu balonların varlığını desteklediğini bulmuştur. Landgraf (2016) da benzer şekilde GSADF testini kullanarak Bitcoin fiyatlarında balonların varlığını belirlemiştir. Wei (2017), Bitcoin'in piyasa dinamiklerini incelemiş ve avantajlarının abartılmasından kaynaklanan aşırı değerlenme ve manipülasyonların ortaya çıkmasına neden olduğunu belirtmiştir. Daha geniş kripto varlık yelpazesi incelendiğinde, Mete vd., (2019) SADF ve GSADF testlerini BTC, ETH ve Ripple üzerinde uygulamıştır. Analizler sonucunda BTC'de en uzun 295 gün süren dokuz farklı balon dönemi, ETH'da en uzun 134 gün süren dört farklı balon dönemi ve XRP'de en uzun 63 gün süren üç farklı balon dönemi tespit edilmiştir. Çalışma, BTC'de başlayan balonların gecikmeli olarak diğer kripto paralara yansıdığını, ancak bu balonların genellikle daha kısa sürdüğünü göstermiştir.

Bu çalışmaların yanı sıra bazı çalışmalar, piyasadaki balonların kavramsal niteliği konusunda farklı sonuçlar sunmuştur. Güleç ve Aktaş (2019), piyasada en çok işlem gören sekiz kripto para biriminde SADF testi uygulamış ve spekülasyon fiyat balonu varlığına dair bulguya rastlanmadığını belirtmiştir. Bu durum, kripto para birimi fiyatlarının ani çöküş yerine, aniden değil de daha kontrollü düşüş yaşayacağına dair işaretler olarak yorumlanmıştır. Ayrıca Cheah ve Fry (2015), BTC'nin spekülasyon balonlar sergilediğini ve temel fiyatının sıfır olduğunu iddia etmiştir. Fry ve Cheah (2016) ise, ani düşüşleri analiz etmek üzere literatüre "negatif balon" kavramını kazandırmıştır. Yeni nesil dijital varlıklarda da balon dinamikleri gözlenmiştir. Akkuş, Gursoy, Dogan ve Demir (2022), MANA tokeninde GSADF testi uygulayarak balon dönemlerinin varlığını belirlemiştir. Farklı şekilde Shu et al., (2021), Log-Periodic Power Law Singularity (LPPLS) modelini kullanarak balonları endojen (içsel spekülasyon davranışlar) ve egzogen (Tesla alımı gibi dışsal

şoklar) olarak sınıflandırmıştır. Bu ayırım, risk yönetimi stratejileri açısından kritik olup, egzojen balonların yönetimi için haber akışı ve düzenleyici ortamın sürekli izlenmesi gerekmektedir. Yapılmış olan literatür çalışması kripto varlık piyasalarındaki balon dinamiklerinin daha iyi anlaşılması, bu yenilikçi finansal araçların sürdürülebilir gelişimi için kritik öneme sahiptir.

## **5.2 Kripto Varlıkların Diğer Varlıklarla İlişkisi Üzerine Çalışmalar**

Kripto varlıkların klasik finans piyasaları ve ekonomik göstergelerle olan ilişkisi, yatırımcıların risklerini doğru yönetip portföylerini çeşitlendirmesi bakımından büyük önem taşımaktadır. Yapılan çalışmalar, bu ilişkilerin genellikle düşük, dönemsel ve piyasa koşullarına bağlı olduğunu göstermektedir.

Geleneksel piyasalar ve makroekonomik faktörler ile ilgili çalışmalar incelendiğinde, Akkaya (2022), BTC fiyatını etkileyen ekonomik ve finansal göstergeleri (EUR/USD, NASDAQ, NIKKEI, Petrol, Altın, vb.) incelemiş ve Johansen eşbütünleşme testi sonucunda, değişkenler arasında uzun dönemde eşbütünleşme denkleminin var olduğunu tespit etmiştir. Ancak, VECM sonuçları, BTC fiyatının kısa dönemde belirlenen ekonomik ve finansal değişkenlerden etkilenmediğini göstermektedir. Akkaya (2022) çalışması, BTC fiyatının NIKKEI ile en yüksek korelasyona (0.8371) ve ardından NASDAQ endeksi ile (0.7212) korelasyona sahip olduğunu bulmuştur. Diğer çalışma olarak Aghalibayli (2019), 2016-2018 dönemine ait haftalık verilerle Vektör Otoregresif Modeli (VAR) ve Granger nedensellik testini kullanarak, Bitcoin'in hem petrol fiyatından hem de altın fiyatından etkilendiğini ve bu iki değişkendeki artışların Bitcoin fiyatlarında yükselişe neden olduğunu tespit etmiştir. Haffar ve Le Fur (2021) yapmış oldukları çalışmada normal piyasa koşullarında kripto varlık getirilerinin geleneksel varlıklarla (S&P500, NASDAQ, VIX, Altın, Petrol) zayıf pozitif korelasyon gösterdiğini, ancak finansal stres dönemlerinde geleneksel varlıklardan ayrıştığını ortaya koymuştur. Aynı şekilde Corbet et al., (2018), kripto varlıkların kendi aralarında yüksek korelasyon gösterdiğini, ancak geleneksel varlıklarla zayıf ilişki sergilediğini, bu durumun portföy çeşitlendirmesinde "bağımsız varlık sınıfı" olarak değerlendirilebileceğini gösterdiğini belirtmiştir.

Başka bakış açısıyla BTC'nin güvenli liman statüsü ve piyasa dalgalanmaları araştırılmıştır. Bu konuda Bouoiyour ve Selmi (2015), Bitcoin'in altın ile dönemsel korelasyonunu

incelemiş ve zaman zaman güvenli liman işlevi görebileceğini öne sürmüştür. Bouri et al., (2017), Bitcoin'in küresel belirsizlik dönemlerinde kısa vadede güvenli liman özelliği sergileyebildiğini, fakat uzun vadede riskli doğasının baskın çıktığını vurgulamıştır. Yermack (2013), Bitcoin'in para biriminden ziyade, yatırım varlığı veya spekülasyon aracı olma eğiliminde olduğunu belirtmiştir. Akkaya (2022) bulguları da kısa dönemde Bitcoin ticaretinin spekülasyon eğilimlere açık olduğunu teyit etmektedir. Tüm bu çalışmaların yanı sıra kripto paraların değerlendirilmesi ve kullanıcı tabanları büyük ölçüde düzenleyici kurumların ve ülkelerin haberlerine duyarlıdır. Örneğin, Musk'ın sosyal medya paylaşımları Dogecoin'in fiyatında büyük dalgalanmalara yol açmıştır.

Son olarak bazı çalışmalarda teknik göstergeler ve duyarlılığın ön plana çıktığı görülmektedir. Bunlardan Çarkacıoğlu (2016), MACD, RSI, ATR ve OBV gibi teknik göstergelerin kripto varlık fiyat hareketlerini kısa vadede açıklamada etkili olduğunu, uzun vadede ise sınırlı kaldığını gözlemlemiştir. Panagiotidis, Stengos ve Vravosinos (2018), LASSO yaklaşımını kullanarak arama yoğunluğu (search intensity) ve altın getirilerini, Bitcoin getirileri için en önemli değişkenler olarak belirlemiştir. Kristoufek (2013) ve Garcia et al., (2014) gibi çalışmalar, arama hacimlerinin ve sosyal medya etkileşimlerinin kripto fiyatlarını nasıl etkilediğini inceleyerek piyasa duyarlılığının rolünün kritik olduğunu göstermiştir.

### **5.3 Fiyat Balonlarının Nedenlerinin Tahmini Üzerine Çalışmalar**

Kripto varlık piyasalarında fiyat balonlarının oluşum nedenlerini anlamak ve tahmin etmek, hem akademik araştırmalar hem de pratik risk yönetimi açısından kritik öneme sahiptir. Bu alandaki çalışmalar, geleneksel ekonometrik yöntemleri modern makine öğrenmesi teknikleriyle birleştirerek, balon dinamiklerinin daha derinlemesine anlaşılmasını sağlamaktadır.

Fiyat balonlarının tahmin edilmesi, yalnızca varlık fiyatı hareketlerini modellemekten öte, piyasa dinamiklerinin ve dışsal şokların karmaşık etkileşimini anlamayı gerektiren zorlu alanlardan biridir. Bu alandaki modern çalışmalar, geleneksel ekonometrik testlerin gücünü, makine öğrenmesi ve derin öğrenme tekniklerinin örüntü tanıma yetenekleriyle birleştiren hibrit yaklaşımlara ve ekonometrik entegrasyona odaklanmıştır. Sawhney et al (2022),

GSADF testinden elde edilen balon indeksini makine öğrenmesi modeline entegre ederek, potansiyel fiyat çöküşlerinin orta ve uzun vadeli tahmin performansını önemli ölçüde artırmıştır. Park ve Yang (2024), bu hibrit yaklaşımı daha da detaylandırarak, balon göstergeleriyle zenginleştirilmiş modelin, temel modellere kıyasla kesinlik-hatırlama eğrisi altındaki alanda (PR-AUC) anlamlı derecede üstün performans sergilediğini bulmuştur. Park ve Yang (2024) yapmış olduğu çalışmanın ampirik bulgularında, yaklaşık iki hafta (14 gün) veya daha uzun süre kesintisiz devam eden fiyat balonunun, gelecekte çöküş yaşanma ihtimalini belirgin şekilde artırdığını tespit etmiştir. Bu durum balon sürekliliğinin risk değerlendirmesinde kritik gösterge olduğunu göstermektedir.

## 6. METODOLOJİ

Bu bölümde, tezin uygulama aşamasında izlenen teknik yol haritası detaylandırılmaktadır. İlk olarak yazılım ve donanım altyapısının tanıtılmasıyla başlayan bölüm, önerilen modelin genel işleyişini özetleyen akış diyagramı ile devam etmektedir. Analizin temelini oluşturan veri setinin yapısal özellikleri, verilerin analize uygun hale getirilmesi için uygulanan ön işleme adımları ve ölçeklendirme yöntemleri sırasıyla açıklanmıştır. Ardından bu çalışmada kullanılan GASDF testi nedir sorusu cevaplanmıştır. Son olarak, model performansını maksimize etmek amacıyla kullanılan hiperparametre optimizasyon teknikleri ve doğrulama stratejileri ele alınmıştır.

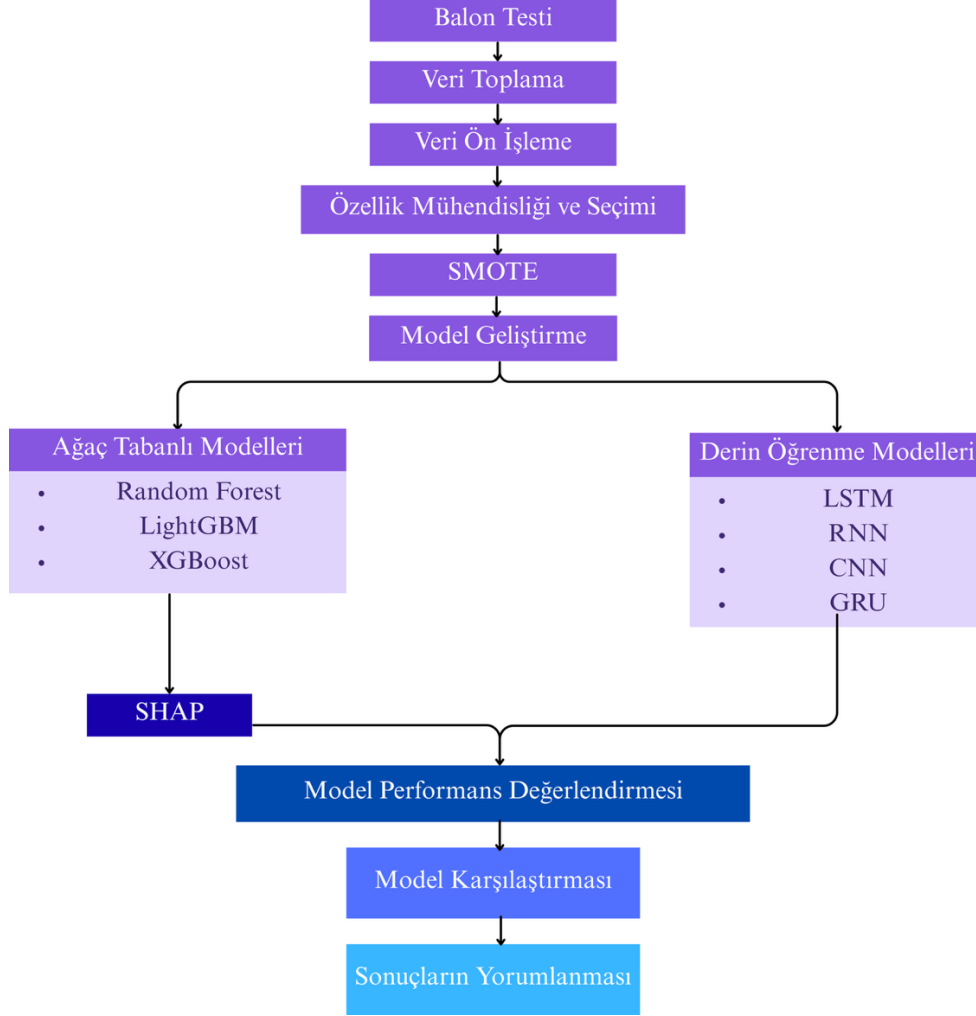
### 6.1 Yazılım ve Donanım Altyapısı

Bu çalışmanın teknik altyapısı, ekonometrik analizler ile yapay zekâ uygulamalarını birleştiren iki aşamalı düzende kurgulanmıştır. Analizin ilk ayağında, çalışmanın bel kemiğini oluşturan fiyat balonlarını tespit etmek amacıyla EViews 12 paket programından yararlanılmıştır. Burada uygulanan GSADF testi sonucunda elde edilen veriler etiketlenmiş ve asıl modellemenin yapılacağı Python ortamına aktarılmıştır. Makine öğrenmesi modellerinin eğitimi ve test süreçleri için ise donanım kısıtlarını aşmak ve hesaplama hızını artırmak adına Google Colab platformu tercih edilmiştir. Bulut tabanlı bu ortamın sağladığı Intel Xeon işlemciler ve özellikle derin öğrenme modelleri için kritik olan NVIDIA Tesla T4 GPU desteği, karmaşık hesaplamaların makul sürelerde tamamlanmasına olanak tanımıştır.

Yazılım tarafında Python ekosisteminin sunduğu araçlardan geniş ölçüde faydalanılmıştır. Verilerin düzenlenmesi ve matris işlemleri için Pandas ve NumPy kullanılırken, zaman serilerinin durağanlık testleri Statsmodels kütüphanesi üzerinden yürütülmüştür. Modele girdi olarak verilen teknik göstergelerin (RSI, SMA vb.) hatasız hesaplanması için “teknik analizler (technical analysis – ta)” kütüphanesi kullanılmıştır. Modelleme aşamasında ise Scikit-learn’ün yanı sıra XGBoost, LightGBM ve CatBoost gibi yüksek performanslı ağaç modelleri, derin öğrenme mimarileri için ise TensorFlow ve Keras kütüphaneleri devreye alınmıştır. Ayrıca en uygun modeli hızlıca belirlemek için LazyPredict kütüphanesinden, modellerin verdiği kararları yorumlayabilmek için ise SHAP kütüphanesinden

yararlanılmıştır. Kurulan bu bütünleşik yapı, çalışmanın hem istatistiksel güvenilirliğini hem de teknolojik derinliğini sağlamıştır.

## 6.2 Akış Diyagramı



Şekil 6.1: Akış diyagramı

Bu çalışmada, kripto varlık piyasalarında fiyat balonlarının tespiti amacıyla geliştirilen makine öğrenmesi tabanlı sınıflandırma sistemine ait işlem adımları Şekil 6.1'de akış diyagramı biçiminde sunulmuştur. Süreç, EViews 12 ile GSADF testinin ardından veri toplama ve ön işleme aşamalarıyla başlar ve özellik mühendisliği ile devam eder.

Model geliştirme aşamasında ağaç tabanlı modeller (Random Forest, LightGBM) ve derin öğrenme tabanlı modelleri (LSTM, RNN, CNN, GRU) ile iki ana kategoriye ayrılan yöntemler kullanılmıştır.

Ağaç tabanlı modellerin karar mekanizmalarının yorumlanabilirliğini artırmak için SHAP yöntemiyle değişken önem düzeyleri analiz edilmiştir. Her model için performans değerlendirmesi için karmaşıklık matrisi, azınlık sınıfı başarımını ölçmeye uygun Test F1 Skoru, PR-AUC ve ROC-AUC metrikleri ile değerlendirilmiştir istatistiksel metrikler ile gerçekleştirilmiştir.

Son aşamada ise tüm modellerin başarı oranları karşılaştırılarak en yüksek performans gösteren yöntem belirlenmiş ve bulgular detaylı olarak yorumlanmıştır.

### 6.3 Veri Seti

Bu çalışmada, kripto varlık piyasalarında oluşabilecek fiyat balonlarını tespit edebilmek amacıyla, 1 Ocak 2020 ile 5 Haziran 2024 tarihleri arasında kapsayan günlük frekansta veri seti oluşturulmuştur. Ayrıca Bitcoin fiyat hareketlerini daha derinlemesine analiz edebilmek amacıyla açılış, en yüksek, en düşük fiyatlar ile işlem hacmi gibi içsel değişkenler de dahil edilmiştir. Bitcoin fiyatlarının yalnızca kendi içsel dinamikleriyle değil aynı zamanda küresel ekonomik göstergeler ve yatırımcı davranışlarıyla da etkileşim içinde olduğu düşünülerek veri setine çeşitli dışsal finansal değişkenler de entegre edilmiştir. Bu kapsamda altın ve petrol fiyatları, yatırımcıların risk algısını ve piyasalardaki güvenli liman eğilimlerini yansıtan emtia göstergeleri olarak değerlendirilmiştir. Ayrıca geleneksel sermaye piyasalarının Bitcoin ile olan olası korelasyonunu ortaya koymak amacıyla ABD hisse senedi piyasalarını temsil eden S&P500, NASDAQ, Russell 2000 ve Dow Jones Industrial Average (DJI) endeksleri ile Avrupa piyasalarını temsilen Almanya DAX (GDAXI) endeksi veri setine dahil edilmiştir. Bu veriler, Investing.com, Yahoo Finance ve CoinMarketCap gibi finansal veri sağlayıcılarından temin edilmiştir.

Veri setinin daha öngörücü hale getirilmesi amacıyla çalışmaya üç temel teknik analiz göstergesi entegre edilmiştir: Göreceli Güç Endeksi (RSI\_14), 20 günlük Basit Hareketli Ortalama (SMA\_20) ve Ortalama Gerçek Aralık (ATR\_14). Bu göstergeler sırasıyla momentum, trend ve volatilité dinamiklerini temsil ederek modellerin balon dönemlerine ilişkin sinyalleri daha iyi yakalamasına katkı sağlamaktadır. Veri setindeki tüm değişkenlerin tanımları ve kaynakları Tablo 6.1'de sunulmaktadır.

**Tablo 6.1:** Veri açıklama tablosu

<b>Kısaltması</b>	<b>Açıklaması</b>	<b>Kaynak</b>
BTC_Kapanış	Bitcoin fiyatının kapanış değeri	Investing.com
BTC_Açılış	Bitcoin fiyatının açılış değeri	Investing.com
BTC_Yüksek	Bitcoin'in gün içindeki en yüksek fiyatı	Investing.com
BTC_Düşük	Bitcoin'in gün içindeki en düşük fiyatı	Investing.com
BTC_Hacim	Bitcoin'in toplam işlem hacmi	Investing.com
Altın	Altın fiyatlarının günlük kapanış değeri	CoinMarketCap
Petrol	Petrol fiyatlarının günlük kapanış değeri	CoinMarketCap
S&P500	S&P500 endeksinin günlük kapanış değeri	CoinMarketCap
NASDAQ	NASDAQ endeksinin günlük kapanış değeri	CoinMarketCap
Russell 2000	Russell 2000 endeksinin günlük kapanış değeri	Yahoo Finance
DJI	Dow Jones Industrial Average kapanış değeri	Yahoo Finance
CBOE	CBOE volatilité endeksinin kapanış değeri	Investing.com
GDAXI	Alman DAX endeksinin kapanış değeri	Yahoo Finance
SMA_20	20 günlük basit hareketli ortalama	Teknik Analiz, ta kütüphanesi
RSI_14	14 günlük göreceli güç endeksi	Teknik Analiz, ta kütüphanesi
ATR_14	14 günlük Ortalama Gerçek Aralık	Teknik Analiz, ta kütüphanesi

BTC kapanış, açılış, yüksek, düşük, hacim gibi sütunlara ait istatistiksel özetler eksik gözlem içermeyen biçimde, örneklem bütünlüğü korunarak elde edilmiştir. Buna karşın teknik göstergelere ilişkin analizlerde yalnızca eksiksiz değerlere sahip gözlemler analiz kapsamına alınmıştır. Bu amaçla, sadece teknik göstergelerin yer aldığı sütunlar seçilmiş ve dropna() yöntemiyle eksik veri içeren satırlar elenmiştir.

**Tablo 6.2:** Tanımlayıcı istatistik tablosu

Değişken	Gözlem Sayısı	Ortalama	Standart Sapma	Min	Medyan	Max	Çarpıklık	Basıklık
BTC Kapanış	1618.00	31785.96	17183.55	4970.79	29219.04	73083.50	0.43	-0.73
BTC Açılış	1618.00	31749.42	17170.51	5002.58	29206.01	73079.38	0.43	-0.73
BTC En Yüksek	1618.00	32467.96	17599.24	5331.83	29580.83	73750.07	0.42	-0.75
BTC En Düşük	1618.00	30988.87	16690.72	4106.98	28790.10	71334.09	0.43	-0.71
BTC Hacim	1618.00	32162843629.16	18428913440.97	5331172801	29098203848	350967941479	4.21	56.69
Altın	1618.00	1849.38	125.57	1477.30	1845.14	2081.90	-0.27	-0.42
Petrol	1618.00	70.11	20.89	12.34	72.19	123.70	-0.29	0.03
S&P500	1618.00	4037.82	541.21	2237.40	4135.10	4796.56	-0.79	-0.02
NASDAQ	1618.00	12643.40	1943.65	6860.67	13060.88	16057.44	-0.58	-0.35
Russell2000	1618.00	1880.96	277.40	991.16	1891.66	2442.74	-0.47	0.14
DJI	1618.00	32520.57	3690.61	18591.93	33592.94	37715.04	-0.98	0.64
CBOE	1618.00	1404.72	2948.46	1.19	3.50	9970.00	1.71	1.12
GDAXI	1618.00	14517.18	1656.98	8441.71	15097.78	16794.43	-0.84	0.41
SMA_20	1599	31708.8546	16798.2615	5953.1080	29046.6641	69130.1357	0.3906	-0.7710
RSI_14	1419	53.53	14.17	19.01	52.42	90.72	0.29	-0.40
ATR_14	1419	1621.63	1054.77	158.81	1359.48	5003.11	0.66	-0.57

Tablo 6.2’de sunulan istatistikler, kripto para piyasası ile geleneksel piyasalar arasındaki derin yapısal farklılıkları açıkça gözler önüne sermektedir. Rakamlar incelendiğinde Bitcoin’in ne denli sert dalgalanma (volatilite) sürecinden geçtiği görülmektedir. Analiz dönemi boyunca ortalama 40.141 USD seviyesinde işlem gören Bitcoin, en düşük 3.750

USD'yi, en yüksek ise 73.750 USD'yi görerek inanılması güç fiyat aralığında hareket etmiştir. Bu hareketliliğin boyutu, 17.645 yüksek standart sapma değeriyle istatistiksel olarak da kanıtlanmaktadır. Bitcoin'in bu agresif risk profili, güvenli liman olarak görülen Altın (101,78 standart sapma) ve Petrol (15,24 standart sapma) ile kıyaslandığında çok daha net anlaşılmaktadır. Benzer şekilde işlem hacmindeki oynaklık da piyasaya giriş-çıkışların ne kadar istikrarsız olduğunu doğrulamaktadır. Kısacası bu tablo, Bitcoin'in geleneksel varlıklara kıyasla tamamen farklı ve çok daha riskli bir sınıfta yer aldığını sayısal olarak ispatlamaktadır.

## **6.4 Veri Ön İşleme**

Bu bölümde, çalışmada kullanılan veri kümesinin analiz öncesinde nasıl işlendiği, hangi ön işleme tekniklerinin uygulandığı ve modelleme aşamasında nasıl hazır hale getirildiği detaylı olarak açıklanmaktadır. Veri temizleme, özellik çıkarma ve dönüştürme süreçleri, özellikle finansal zaman serisi verilerinde önemli yer tutmaktadır. Kullanılan makine öğrenmesi ve derin öğrenme modellerinin başarılı sonuçlar vermesi için veri setinin uygun şekilde düzenlenmesi ve hazırlanması kritik adımdır (Bishop, 2006).

### **6.4.1 Veri temizleme işlemleri**

Veri temizleme süreci, ham verinin analiz için uygun hale getirilmesini sağlayan ilk ve en kritik adımdır. Çalışmada kullanılan veri kümesi, Bitcoin fiyat hareketlerinin yanı sıra, altın ve petrol gibi emtia fiyatları, S&P 500, NASDAQ, Russell 2000 ve DJI gibi ABD borsa endeksleri, CBOE volatilité endeksi ile GDAXI (Almanya) endeksini kapsayan çok değişkenli zaman serisi verisidir. Finansal veriler, genellikle eksik, hatalı ya da gereksiz bilgileri barındırabildiğinden (Tsay, 2005), öncelikle veri kümesindeki eksik ve tutarsız değerler dikkatle incelenmiştir. Yapılan analizler sonucunda, veri kümesinde eksik değer bulunmadığı belirlenmiştir. Ancak zaman bağımlılığı göz önüne alınarak, orijinal CSV dosyasında yer alan "Date" sütunu pd.to\_datetime fonksiyonu ile doğru datetime formatına dönüştürülmüş ve analiz sürecinde kronolojik düzenin korunabilmesi amacıyla indekslenmiştir. Bu işlem, zaman serilerinin doğal akışını bozmayacak şekilde verinin yapılandırılmasını sağlamış ve verinin güvenilirliğini artırmıştır.

Bunun yanı sıra, veriye entegre edilecek teknik göstergelerin hesaplanması sürecinde ortaya çıkabilecek eksik ya da tutarsız değerler, dropna () metodu ile sistematik olarak temizlenmiştir. Bu yaklaşım, modelin yalnızca anlamlı ve etkili değişkenler üzerinde çalışmasını sağlayarak, öğrenme sürecinde hatalı ya da yanıltıcı bilgilerle karşılaşma riskini minimize etmiştir. Tüm bu işlemler sonucunda, temizlenmiş ve analiz için uygun hale getirilmiş, güvenilir ve tutarlı veri seti elde edilmiştir. Böylece sonraki analiz ve modelleme aşamalarının sağlam temeller üzerine inşa edilmesi mümkün kılınmıştır.

#### **6.4.2 Özellik çıkarma ve dönüştürme teknikleri**

Makine öğrenmesi ve derin öğrenme modellerinin yüksek doğrulukla öğrenebilmesi, yalnızca modelin yapısına bağlı değildir. Aynı zamanda modelin beslendiği verinin doğru biçimde işlenmesine de bağlıdır. Bu nedenle ham verinin uygun biçimde dönüştürülmesi, ölçeklendirilmesi ve zenginleştirilmesi model başarısı açısından kritik öneme sahiptir (Goodfellow et al., 2016).

İlk olarak, fiyat ve işlem hacmi gibi değişkenlerin birbirinden oldukça farklı ölçeklerde yer almasından dolayı tüm sayısal değişkenler Min-Max ölçeklendirme yöntemiyle [0,1] aralığına normalize edilmiştir. Bu yöntemin tercih edilmesinin başlıca nedenleri, gradyan tabanlı optimizasyon algoritmalarının, özellikle eğilim azalışı (gradient descent) gibi yöntemlerin, daha hızlı yakınsamasını sağlamaktır. Ayrıca farklı büyüklükteki özelliklerin neden olduğu ağırlık dengesizliklerini ortadan kaldırmak ve sigmoid ile tanh aktivasyon fonksiyonlarının tanımlı aralıkta daha etkin çalışmasını mümkün kılmak da bu tercihin önemli unsurlarıdır (Han et al., 2011). Tüm bunlar bir araya geldiğinde, modelin öğrenme süreci daha tutarlı ve kararlı hale gelmektedir.

Finansal zaman serisi verisi olarak kullanılan veri kümesinde, fiyat balonlarının tespiti için öncelikle EViews yazılımı kullanılarak GSADF testi uygulanmıştır. Bu testin çıktısı olan ikili sınıflandırma sonuçları (0: balon yok, 1: balon var), makine öğrenmesi modellerinin hedef değişkeni olarak kullanılmıştır. Modeller, Bitcoin fiyatı, işlem hacmi, teknik göstergeler ve diğer finansal değişkenleri girdi olarak alarak, GSADF testinin belirlediği balon dönemlerini tahmin etmeyi öğrenmektedir. Bu yaklaşımla, GSADF'nin objektif

sonuçları ile modern makine öğrenmesi yöntemlerinin örüntü tanıma yetenekleri birleştirilerek, balon tespitinde hibrit ve daha güçlü metodoloji oluşturulmuştur.

Özellikle LSTM ve CNN gibi derin öğrenme modellerinde, zaman penceresi yaklaşımı kullanılarak son 3 günlük (`time_steps = 3`) veri dizisi üzerinden tahmin yapılmaktadır. Bu düzenleme sayesinde modeller, sadece anlık değerlere değil, ardışık günler arasındaki ilişkilerin modellenmesine olanak tanımıştır. Böylece GSADF testinin belirlediği balon dönemlerini daha yüksek doğrulukla tahmin edebilmeleri sağlanmıştır.

Bu kapsamlı veri dönüşüm ve ölçeklendirme süreci, hem ham verinin modelin gereksinimlerine uygun hale getirilmesini hem de finansal zaman serisi verilerinin dinamik yapısının etkin biçimde yakalanmasını sağlamıştır. Böylece elde edilen modellerin performansının ve genelleme yeteneğinin artırılmasına önemli ölçüde katkıda bulunmuştur.

#### **6.4.3 Veri setinin eğitim ve test olarak ayrılması (data splitting)**

Makine öğrenimi ve derin öğrenme modellerinin gerçek dünya performansını objektif şekilde ölçebilmek, tezimizin en kritik aşamalarından biridir. Bu nedenle modellerin öğrenme sürecinde görmediği veriler üzerinde ne kadar başarılı tahminler yaptığını test etmek amacıyla, ham veri seti eğitim ve test olmak üzere iki temel parçaya ayrılmaktadır.

Finansal zaman serileri için veri ayrımında rastgele seçim yerine zamansal sırayı koruyan yöntem kullanılmaktadır. Böylece model, geçmiş verilerle eğitilirken, gelecekteki (test setindeki) verileri tahmin etmeye çalışmaktadır. Bu durum piyasa koşullarını simüle eden en gerçekçi yaklaşımdır. Bu durumun yanı sıra yapılmış olan çalışmada, güçlü ve genellenebilir modellere ulaşmak adına veri setini farklı senaryolarda denemek için eğitim ve test olarak %70 - %30, %75 - %25 ve %80 - %20 oranlarında ayrılmaktadır. Kod akışımızda ise %75 - %25 (75/25 Split) ayrımı, hızlı ve tutarlı başlangıç senaryosu olarak tercih edilmektedir.

Diğer durum ise veri setindeki balon tespiti (GSADF testi ile belirlenen 1 sınıfı) gibi nadir olaylar, doğal olarak sınıf dengesizliğine yol açmaktadır. Bu dengesizlik, özellikle eğitim setinde modelin yaygın olan 0 sınıfına (balon yok) eğilim göstermesine ve 1 sınıfını göz ardı etmesine neden olmaktadır. Bu sorunu gidermek adına ağaç modellerinde, eğitim seti

içerisindeki sınıf oranlarına göre dinamik olarak hesaplanan pozitif sınıf ağırlığı ( $scale\_pos\_weight / class\_weight$ ) kullanılmaktadır. Derin öğrenme modellerinde ise benzer şekilde, kayıp fonksiyonunun (loss function) 1 sınıfına daha fazla önem vermesini sağlayan sınıf ağırlıkları ( $class\_weight\_dict$ ) tanımlanmaktadır.

## 6.5 Veri Ölçeklendirme Yöntemleri

Makine öğrenmesi ve istatistiksel modelleme süreçlerinde, veri ön işleme adımı model başarımını doğrudan etkileyen kritik aşamadır. Bundan dolayı değişkenlerin farklı ölçeklerde olması, algoritmaların performansını olumsuz yönde etkilemektedir. Bu nedenle literatürde de görüldüğü üzere en çok kullanılan yöntemler olarak Min-Max Ölçeklendirici ve Z-Skor Ölçeklendirici kullanılmıştır.

### 6.5.1 Min-Max ölçeklendirme

Min-Max ölçeklendirme, değişkenin değerlerini belirli aralığa (genellikle 0 ile 1 arasında) dönüştürmeyi amaçlayan normalizasyon yöntemidir. Bu yöntem, her gözlem değerinin, değişkenin minimum ve maksimum değerleri arasındaki konumuna göre yeniden hesaplanmasını sağlamaktadır. Min-Max dönüşümü aşağıdaki formül ile ifade edilmektedir:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (6.1)$$

Burada orijinal gözlem değerini,  $X_{norm}$  ilgili değişkenin minimum değerini,  $X_{max}$  maksimum değerini ve  $X_{min}$  ' ise ölçeklenmiş değeri temsil etmektedir. Bu yöntem, özellikle yapay sinir ağları gibi bazı algoritmaların daha iyi performans göstermesi için veri aralığının daraltılması gerektiğinde tercih edilmektedir. Ancak Min-Max yöntemi, aykırı değerlere karşı oldukça duyarlıdır. Ayrıca ekstrem gözlem,  $X_{min}$  veya  $X_{max}$  değerini belirgin ölçüde etkilemektedir. Bu durum normalleşmiş değerlerin dağılımını bozmaktadır (Han et al., 2011; Géron, 2019).

### 6.5.2 Z-skor ölçeklendirme (standardizasyon)

Z-skor normalizasyonu, verilerin ortalamasının 0 ve standart sapmasının 1 olacak şekilde dönüştürülmesini sağlayan standartlaştırma yöntemidir. Bu teknik, özellikle değişkenlerin

farklı ortalama ve varyanslara sahip olduđu durumlarda daha eşit veri yapısı elde etmeye olanak tanımaktadır. Z-skor dönüşümü aşağıdaki formül ile tanımlanmak:

$$X_{\text{std}} = \frac{X - \mu}{\sigma} \quad (6.2)$$

Bu denklemde  $x$  orijinal değeri,  $\mu$  değişkenin aritmetik ortalamasını, ise standart sapmasını ifade etmektedir. Z-skor yöntemi, Min-Max yöntemine kıyasla aykırı değerlere daha az duyarlıdır. Özellikle regresyon, destek vektör makineleri (SVM) ve k-en yakın komşu (k-NN) gibi mesafe tabanlı algoritmalarda etkili ön işleme tekniğidir. Bununla birlikte, bu yöntemin varsayımlarından biri, verilerin yaklaşık normal dağılım göstermesidir. Normal dağılmayan verilerde, standardizasyonun etkinliği azaltmaktadır (James et al., 2013; Kuhn and Johnson, 2013).

## 6.6 GSADF Analiz Yöntemi

Fiyat balonlarının tespiti literatüründe, geleneksel birim kök (örneğin ADF testi) ve eşbütünleşme testleri piyasalarda periyodik olarak oluşturulup aniden çöken (periodically collapsing) balonları tespit etmekte yetersiz kalmaktadır. Bu eksikliği gidermek amacıyla Phillips, Wu ve Yu (2011) tarafından Supremum Artırılmış Dickey-Fuller (SADF) testi geliştirilmiştir. SADF testi, sağ kuyruk (right-tail) yaklaşımını kullanarak varlık fiyatlarındaki "hafif patlayıcı (mildly explosive)" davranışları tespit edebilmektedir (Phillips et al., 2015). Ancak SADF testi, uzun zaman serilerinde birden fazla balonun (multiple bubbles) bulunduğu durumlarda, ardışık balonların başlangıç ve bitiş noktalarını tutarlı şekilde belirlemede ve balonların varlığını kanıtlamada yetersiz kalmaktadır (Wei et al., 2022). Bu kısıtı aşmak ve aynı örneklem içinde birden fazla balonun varlığını tespit edebilmek amacıyla Phillips, Shi ve Yu (2015) tarafından Genelleştirilmiş Supremum Artırılmış Dickey-Fuller (GSADF) testi literatüre kazandırılmıştır (Phillips et al., 2015).

GSADF testinin çalışma mantığı ve matematiksel altyapısı temel olarak bir fiyat serisinde "birim kök (unit root)" olduğu yönündeki boş hipotezi, serinin "patlayıcı bir süreç (explosive process)" izlediği yönündeki alternatif hipoteze karşı test etmektedir. GSADF testinde hipotezler aşağıdaki gibi tanımlanmaktadır:

$$\begin{aligned} H_0: \delta &= 1 \text{ (balon yok / birim kök)} \\ H_1: \delta &> 1 \text{ (en az bir balon var / patlayıcı davranış)} \end{aligned} \quad (6.3)$$

Burada  $\delta > 1$  olması, serinin ilgili alt örnekleminde birim kök sınırının aşıldığını ve balon rejimi ile uyumlu patlayıcı dinamiklerin oluştuğunu ifade etmektedir. Bu testte boş hipotezin reddedilmesi, ilgili varlığın fiyatında rasyonel bir balon olduğuna dair ampirik bir kanıt olarak kabul edilmektedir (Caspi, 2017). Yöntemlerin pencereleme (windowing) yaklaşımları arasındaki fark, GSADF'nin gücünü ortaya koymaktadır. SADF testinde, örneklemin başlangıç noktası  $r_1 = 0$  sabit tutulup, bitiş noktası  $r_2$  ileriye doğru genişletilerek özyineli (recursive) regresyonlar hesaplanmaktadır. GSADF testinde ise bu pencere yapısı esnetilmiştir. GSADF, sadece bitiş noktasını ( $r_2$ ) ileriye doğru değiştirmekle kalmaz, aynı zamanda başlangıç noktasının da ( $r_1$ ) izin verilen aralıkta (0 ile  $r_2 - r_0$ ) arasında değişmesine olanak tanımaktadır (Phillips et al., 2015). Bu esnek pencere genişliği (flexible window width) sayesinde model, çok daha fazla alt örneklemini (subsample) kapsamaktadır. Aynı zaman serisi içindeki çoklu patlayıcı davranışları taramada geleneksel yöntemlerden çok daha üstün bir performans sergilemektedir. GSADF istatistiği, değişen başlangıç ( $r_1$ ) ve bitiş ( $r_2$ ) noktalarına sahip tüm geçerli alt örneklemler üzerinden elde edilen ADF istatistiklerinin en büyük (maksimum) değeri olarak aşağıdaki gibi tanımlanmaktadır (Wei et al., 2022):

$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1], r_1 \in [0, r_2 - r_0]} \{ADF_{r_1}^{r_2}\} \quad (6.4)$$

Kripto varlık piyasaları, doğası gereği yüksek volatiliteye sahip olan ve uzun bir zaman dilimi içinde ardışık birçok yükseliş ve çöküş döngüsü barındıran piyasalardır. Bu bağlamda, incelenen periyot boyunca kripto para fiyatlarında tek bir balon yerine çoklu balon dönemlerinin yaşanmış olması kuvvetle muhtemeldir. Test penceresinin hem başlangıç hem de bitiş noktalarını dinamik olarak değiştirerek çoklu patlayıcı davranışları başarılı bir şekilde saptayabilen ve zamanlarını damgalayabilen (date-stamping) GSADF yaklaşımı, kripto varlık piyasalarındaki irrasyonel fiyat hareketlerinin tespiti için oldukça uygun ve güvenilir bir yöntemdir (Caspi, 2017).

## 6.7 Optimizasyon Teknikleri

Optimizasyon teknikleri, makine öğrenmesi ve derin öğrenme modellerinin doğruluğunu artırmak ve modelin tahmin gücünü iyileştirmek için kullanılan yöntemlerdir (Sun et al., 2019). Çalışmada kullanılan modellerin en iyi performansı göstermesi için hiperparametre ayarlama (hyperparameter tuning) ve performans artırma stratejileri uygulanmıştır.

### 6.7.1 Hiperparametre ayarlama

Makine öğrenmesi ve derin öğrenme modellerinin başarısını artırmak için hiperparametre ayarlama süreci önemli rol oynamaktadır. Hiperparametreler, modelin eğitim sürecinden önce belirlenen ve modelin öğrenme kapasitesini doğrudan etkileyen değişkenlerdir. Bu parametrelerin en uygun şekilde belirlenmesi, modelin doğruluğunu artırmak ve genelleme yeteneğini iyileştirmek açısından büyük önem taşımaktadır (Bergstra and Bengio, 2012).

Bu çalışmada ağaç tabanlı modellerde Grid Search yöntemi tercih edilerek, her modelin ana hiperparametre alanı sistematik şekilde taranmıştır. Örneğin RF'de bulunan RandomForestClassifier'da ağaç sayısı olarak 100 ve 200, derinlik limiti (max\_depth) için 10 ve 20 değerleri ve yaprak için minimum örnek sayısı (min\_samples\_leaf) değerleri taranmıştır. LightGBMClassifier'da iterasyon sayısı (n\_estimators) 100 ve 200, öğrenme hızı (learning\_rate) 0.05 ve 0.1, yaprak sayısı (num\_leaves) ise 20 ve 31 aralıklarında denenirken XGBoostClassifier tarafında ise 100 ve 200 ağaç, 3–5 aralığında maksimum derinlik, 0.05 ve 0.1 öğrenme hızları ile Grid Search'le taranmıştır. Gerçekleştirilen bu optimizasyon süreci sayesinde, her model için en uygun hiperparametreler belirlenmiş, model doğruluğu artırılmış ve genelleme performansı iyileştirilmiştir. Böylece hem klasik makine öğrenmesi hem de derin öğrenme modellerinin, finansal zaman serisi verileri üzerindeki tahmin gücü en üst düzeye çıkarılmıştır.

### 6.7.2 Performans artırma stratejileri

Makine öğrenmesi ve derin öğrenme modellerinin başarısını artırmak için çeşitli stratejiler uygulanmaktadır. Bu stratejiler, modelin genelleme yeteneğini artırmak, aşırı öğrenmeyi (overfitting) önlemek, veri dengesizliklerini gidermek ve modelin tahmin gücünü iyileştirmek amacıyla geliştirilmiştir. Finansal zaman serileri ve özellikle fiyat balonu tespiti gibi dengesiz veri dağılımına sahip problemler, doğrudan sınıflandırma modellerinin

başarısını etkileyebileceğinden, performans artırma teknikleri bu tür veriler üzerinde kritik rol oynamaktadır (Fernández et al., 2018).

Özellik mühendisliği (feature engineering), modelin veri içindeki anlamlı özellikleri öğrenmesini kolaylaştırmak ve gereksiz özniteliklerden arındırarak performansını iyileştirmek için uygulanan diğer kritik adımdır. Bu çalışmada, Durağanlaştırma (differencing) yöntemi kullanılarak, finansal zaman serilerindeki değişkenlerin çoğunda durağanlık (stationarity) koşulu sağlanamadığından, modelin daha güvenilir sonuçlar üretmesi için Augmented Dickey-Fuller (ADF) testi ile durağan olmadığı tespit edilen tüm değişkenlerin birinci farkları alınmıştır. Bu işlem, değişkenleri durağan hale getirerek modelin uzun vadeli trendler yerine anlık değişimlere odaklanmasını sağlamıştır.

Teknik analiz göstergeleri ile fiyat ve hacim verilerine dayalı olarak RSI (Göreceli Güç Endeksi), SMA (Basit Hareketli Ortalama) ve ATR (Ortalama Gerçek Aralık) gibi endüstri standartlarında teknik analiz göstergeleri türetilerek, modelin karar verme yeteneği zenginleştirilmiştir. Son olarak gecikme ile de tüm bağımsız değişkenler hedef değişkene göre bir periyot geriye kaydırılarak, modelin tahmin yapacağı anda yalnızca geçmiş verileri kullanması sağlanmış ve böylece veri sızıntısı (data leakage) kesinlikle önlenmiştir. Derin öğrenme modellerinde yaygın olarak karşılaşılan problemlerden biri aşırı öğrenme olup, modelin eğitim verisine fazla uyum sağlaması nedeniyle yeni verilerde düşük doğruluk göstermesine neden olmaktadır (Goodfellow et al., 2016). Bu problemi önlemek için erken durdurma (early stopping) yöntemi uygulanmıştır. Erken durdurma tekniği, derin öğrenme sürecinde, modelin doğrulama (validation) veri setindeki hata oranını sürekli olarak izlenmiş ve hata belirli sabır (patience) süresi boyunca düşmeye devam etmediğinde eğitim durdurulmuştur. Bu teknik, modelin gereğinden fazla eğitilerek yalnızca eğitim verisine aşırı öğrenme sağlaması riskini azaltmıştır. Ardından derin öğrenme katmanlarında (LSTM, RNN, CNN, GRU) Dropout katmanları kullanılmıştır. Bu durum, eğitim sırasında rastgele nöronları geçici olarak devre dışı bırakarak nöronların birbirine aşırı bağımlı hale gelmesini engellemektedir. Böylelikle modelin farklı özellik kombinasyonlarını öğrenmeye zorlayarak genelleme yeteneğini artırmaktadır.

Son olarak, modelin tahmin performansını artırmak için sınıf ağırlıklandırma (class weighting) stratejisi uygulanmıştır. Dengesiz veri kümelerinde, model genellikle çoğunluk sınıfa daha fazla önem verirken azınlık sınıfı göz ardı edilebilmektedir. Bunu önlemek için ağaç tabanlı modellerde (XGBoost, RandomForest, LightGBM) ve derin öğrenme modellerinde (LSTM, RNN, CNN, GRU) azınlık sınıfının eğitimdeki önemi, çoğunluk sınıfa kıyasla ters orantılı olarak artırılmıştır. Bu matematiksel dengeleme sayesinde, modelin azınlık sınıfı üzerindeki hassasiyeti (precision) yükseltilmiş ve finansal balonları doğru tespit etme gücü artırılmıştır. Böylece model, yalnızca çoğunluk sınıfta değil, azınlık sınıfta da yüksek doğrulukla tahmin yapabilmektedir.

Gerçekleştirilen bu stratejiler sayesinde, modelin hata oranları minimize edilmiş, aşırı öğrenme önlenmiş ve fiyat balonlarının tahmin edilmesi için en iyi performansa sahip modeller belirlenmiştir. Özellikle finansal zaman serisi analizlerinde, bu yöntemlerin arada uygulanması modelin daha güvenilir ve sağlam tahminler yapması sağlanmıştır.

## 7. MAKİNE ÖĞRENMESİ MODELLERİ

Bu bölümde, Bitcoin'in GSADF balon durumunu tahmin etmeye çalışan makine öğrenmesi modellerinden bahsedilmektedir. Her model, finansal verilere eklenen teknik göstergelerden yararlanarak, GSADF değerinin temsil ettiği 'balon' (1) veya 'normal' (0) durumunu sınıflandırmayı amaçlamaktadır. İlk olarak, güçlü ve esnek yapılarıyla bilinen Random Forest, XGBoost ve LightGBM ağaç tabanlı algoritmaları incelenmektedir. Ardından, zaman serisi verilerinin kendine özgü karmaşık yapılarını öğrenmede başarılı olan LSTM, CNN, RNN ve GRU derin öğrenme modelleri ele alınmaktadır.

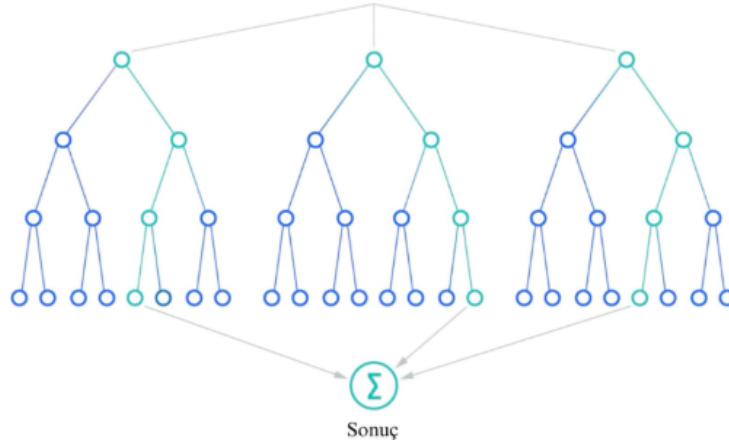
Her modelin ne kadar iyi çalıştığını görmek için çeşitli ölçütler kullanılmaktadır. Karmaşıklık Matrisi (Confusion Matrix) ile tahminlerin ne kadar doğru olduğunu, Doğruluk (Accuracy) ile genel başarı oranını, Hassasiyet (Precision) ile pozitif tahminlerimizin ne kadarının gerçekten doğru olduğunu, Duyarlılık (Recall) ile gerçek pozitiflerin ne kadarını yakalayabildiğimizi, F1-Skoru ile hassasiyet ve duyarlılık arasındaki denge ile daha dengeli başarı ölçütünü ve Brier Skoru ile de modelin olasılık tahminlerinin ne kadar güvenilir olduğunu değerlendirmekteyiz.

Çalışmada ayrıca modellerin sınıflandırma becerilerini daha net görebilmek için ROC-AUC ve Precision-Recall Eğrileri gibi görsel araçlardan faydalanılmaktadır. Buna ek olarak, modellerin karar süreçlerini şeffaflaştırmak için SHAP analizinden yararlanılmıştır. Böylece tahminlerde hangi değişkenlerin daha baskın rol oynadığı tespit edilmiştir. Son olarak, modellerimizin güvenilirliğini test etmek ve en iyi tahmini yapan modeli seçmek adına Çapraz Doğrulama Yöntemleri ve Özellik Seçimi (Feature Selection) gibi önemli stratejileri tartışılmaktadır.

### 7.1 Rastgele Orman (Random Forest - RF)

RF algoritması, Leo Breiman (2001) tarafından geliştirilen ve hem sınıflandırma hem de regresyon problemleri için yaygın olarak kullanılan güçlü topluluk (ensemble) öğrenme yöntemidir (Breiman, 2001). Bu algoritma, birçok karar ağacının tahminlerini araya getirerek modelin genelleme yeteneğini artırmakta ve aşırı öğrenme sorununu azaltmaktadır. RF, yüksek doğruluk sağlaması, değişken önemini belirleyebilmesi, genelleştirme hatasını değerlendirme kapasitesi ve sınıf dengesizliklerini yönetme yeteneği ile öne çıkmaktadır (Rodriguez et al., 2006). RF algoritmasının çalışma mekanizması Şekil 7.1.'de verilmiştir.

(Ljumović and Klar, 2015). Temel olarak algoritma, Bootstrap Aggregation (Bagging), rastgele özellik seçimi ve tahmin birleştirme süreçlerine dayanmaktadır.



**Şekil 7.1:** Karar ağaçlarının topluluk yaklaşımıyla birleştirilmesi

Random Forest algoritmasının çalışma prensibi üç temel aşamadan oluşmaktadır:

- i) **Bagging:** Bagging yöntemi, veri setinden rastgele örnekler seçilerek yeni alt veri kümelerinin oluşturulmasını sağlamaktadır. Her karar ağacı, bu alt veri kümeleri üzerinde bağımsız olarak eğitilmektedir. Bu süreç, modelin çeşitliliğini artırarak aşırı öğrenme sorununu azaltmaktadır. Her karar ağacının çıktısı regresyon problemlerinde ortalama alınarak, sınıflandırma problemlerinde ise çoğunluk oyu ile birleştirilmektedir. Bu yöntem, tek karar ağacının yüksek varyans sorununu minimize etmektedir (Breiman, 2001; Ljumović and Klar, 2015).

Formülasyon olarak,  $X = \{x_1, x_2, \dots, x_n\}$  ve  $Y = \{y_1, y_2, \dots, y_n\}$  eğitim veri seti olmak üzere B bootstrap örnekleme şu şekilde oluşmaktadır (Hastie et al., 2009):

$$\{X_b, Y_b\}_b = 1^b \quad (7.1)$$

Her bootstrap örnekleme üzerinde karar ağacı eğitilir. Yeni bir gözlem için tahmin formülü (Hastie et al., 2009):

$$f(x') = \frac{1}{B} \sum_{b=1}^B f_b(x') \quad (7.2)$$

Sınıflandırma problemlerinde çoğunluk oyu şu şekilde alınmaktadır:

$$\text{Mode}(f_1(x'), f_2(x'), \dots, f_B(x')) \quad (7.3)$$

$f_b(x')$ : b'inci karar ağacının, yeni gözlem  $x'$  için yaptığı tahminidir.

$B$  : Toplam karar ağacı sayısıdır

- ii) **Rastgele özellik seçimi:** Her düğümde, tüm değişkenler yerine rastgele seçilen alt küme değişken üzerinden bölünme işlemi gerçekleştirilmektedir. Bu yöntem, karar ağaçlarının birbirine bağımlı hale gelmesini önleyerek genelleme hatasını azaltmaktadır. Rastgele özellik seçimiyle, büyük boyutlu veri setlerinde bile yüksek performans elde edilmektedir (Dubey, 2018).

Matematiksel olarak, veri setinde toplam  $p$  değişken varsa, her düğümde rastgele seçilen  $m$  özellik üzerinden bölünme yapılmaktadır. Bu durum şu şekilde ifade edilmektedir:

$$m \ll p \quad (7.4)$$

Sınıflandırma problemlerinde genellikle  $m = \sqrt{p}$ , regresyon problemlerinde ise  $m = \frac{p}{3}$  seçilmektedir (Hastie et al., 2009).

- iii) **Tahminlerin birleştirilmesi:** Her ağaçtan elde edilen tahminler birleştirilerek nihai sonuç oluşmaktadır. Regresyon problemlerinde ağaçların ortalaması denklem (7.2) formülü ile alınmaktadır

Sınıflandırma problemlerinde ise çoğunluk oyu veya olasılıkların ortalaması ile sınıf tahmini denklem (7.3) ile yapılmaktadır.

Sınıflandırma bağlamında, bu tahminler genellikle her sınıf için olasılıklar veya doğrudan sınıf etiketleri (0 veya 1 gibi) şeklinde olabilmektedir. Eğer ağaçlar olasılık tahmini yapıyorsa, bu olasılıkların ortalaması alınarak nihai sınıf olasılığı elde edilmektedir. En yüksek olasılığa sahip sınıf seçilmektedir. Eğer doğrudan sınıf etiketi tahmini yapıyorlarsa, en çok oyu alan (yani, en çok ağacın tahmin ettiği) sınıf, Random Forest modelinin nihai tahmini olarak kabul edilmektedir. Bu yöntem, her ağacın "fikrini" alarak daha sağlam ve genellenebilir karar mekanizması oluşturulmasını sağlamaktadır (Hastie et al., 2009).

**Özellik önemi analizi:** Random Forest algoritması, her değişkenin modele katkısını değerlendirmek için Permütasyon Özellik Önemi ve Ortalama Safsızlık Azalması (Mean Decrease in Impurity, MDI) yöntemlerini kullanılmaktadır.

**Permütasyon özellik önemi:** Bir özelliğin modeli ne kadar etkilediğini anlamak için özelliğin değerleri rastgele karıştırılarak modelin performansı yeniden değerlendirilmektedir. Permütasyon öncesi ve sonrası performans farkı şu şekilde hesaplanmaktadır:

$$\Delta\text{Önem} = \text{Karıştırılmış Hata} - \text{Orijinal Hata} \quad (7.5)$$

Daha büyük fark, ilgili özelliğin model için daha önemli olduğunu göstermektedir. Ancak, yüksek fazla kategoriye sahip özellikler yanıltıcı sonuçlar üretmektedir (Parr et al., 2018).

**Ortalama safsızlık azalması (MDI):** Bu yöntem, karar ağaçlarında özelliğin bölünme sırasında safsızlığı ne kadar azalttığını ölçmektedir. Safsızlık ölçütü genellikle gini indeksi veya entropi gibi istatistiklerdir. Özelliğin önemi, şu formülle ifade edilmektedir:

$$\text{Önem}(x) = \frac{1}{n_T} \sum_{i=1}^{n_T} \sum_{j \in T_i} P_{(T_i)}(j) \Delta i_{(T_i)}(j) \quad (7.6)$$

$x$  : Modele beslenen her özelliği (feature) ifade etmektedir.

$n_T$  : Ormandaki toplam ağaç sayısını göstermektedir.

$T_j$ : Ormandaki -inci karar ağacını temsil etmektedir.

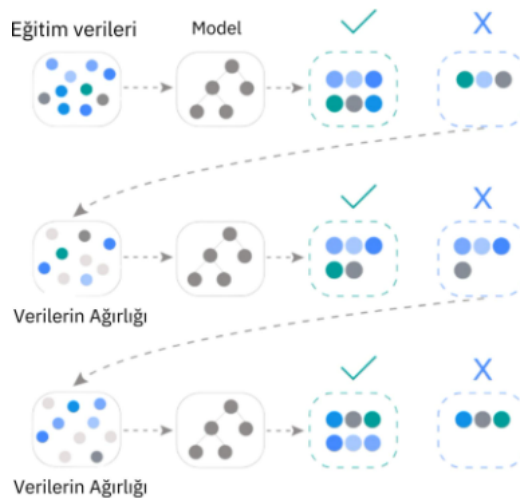
$P_{T_i}(j)$ : Ağacındaki numaralı düğüme düşen örneklerin oranıdır.

$\Delta i_{T_i}(j)$ : O düğüme gerçekleşen safsızlık azalmasını (örneğin Gini azalması veya entropi azalması) göstermektedir.

Bu yöntem, özelliğin model üzerindeki etkisini ölçmede etkili olmasına rağmen bazı durumlarda yanıltıcı sonuçlar verebilmektedir. Özellikle, yüksek kardinaliteli özellikler, bu yöntemde daha yüksek önem skorlarına sahip olma eğilimindedir (Parr et al., 2018). Ayrıca Scikit-learn gibi yaygın kütüphanelerdeki varsayılan uygulamalarda eğitim istatistiklerine dayalı ölçümler, özelliğin test setindeki tahmin başarısını doğru şekilde yansıtmayabilmektedirler (Pedregosa et al., 2011). Bu tür sınırlamaların üstesinden gelmek için çapraz doğrulama veya permütasyon testleri gibi yöntemler önerilmektedir.

## 7.2 Aşırı Gradyan Artırma (Extreme Gradient Boosting - XGBoost)

XGBoost, Gradient Boosted Decision Trees (GBDT) yönteminin optimize edilmiş ve geliştirilmiş versiyonudur. Tianqi Chen tarafından Washington Üniversitesi'nde geliştirilen bu algoritma, makine öğreniminde yüksek performans, hız ve ölçeklenebilirlik sağlamasıyla öne çıkmaktadır (Chen and Guestrin, 2016). Sınıflandırma ve regresyon problemlerinde etkin şekilde kullanılan XGBoost, büyük veri kümelerini işleme kapasitesi ve dağıtık eğitim desteği sayesinde Kaggle yarışmaları gibi zorlu uygulamalarda sıkça tercih edilmektedir. Şekil 7.2'de de gösterildiği gibi XGBoost, paralel ve dağıtık hesaplama özelliklerine sahiptir. Veri, bellekte "blok" adı verilen birimler halinde saklanarak makineler arasında dağıtılabılır veya harici bellekte işlenebilmektedir. Apache Spark, Dask ve Kubernetes gibi araçlarla entegrasyon sağlayarak dağıtık eğitim yapılmasına olanak tanımaktadır. Bu özellik, XGBoost'un büyük veri kümeleri üzerinde hızlı şekilde çalışmasını mümkün kılmaktadır. Ayrıca önbellek farkındalığı (cache-aware prefetching) algoritması ile veriyi hızlı şekilde işleyerek çalışma süresini önemli ölçüde azaltmaktadır. Tek makinede diğer yöntemlere kıyasla 10 kata kadar daha hızlı çalışabilen XGBoost, milyarlarca veri noktasını sınırlı kaynaklarla işleyebilecek kapasiteye sahiptir (Kavlakoglu and Russi, 2024).



Şekil 7.2: XGBoost adım güçlendirme ve ağırlık güncelleme mekanizması

Dahili düzenleme mekanizmaları, XGBoost'un model genelleme kapasitesini artırmasını ve aşırı öğrenmeyi önlemesini sağlamaktadır. Düzenleme parametreleri arasında, yaprak ağırlıkları için kullanılan  $\lambda$  (L2 norm düzenlemesi) ve yeni düğüm eklemek için gereken minimum bilgi kazancını belirleyen  $\gamma$  öne çıkarmaktadır. Bu düzenleme mekanizmaları, hiperparametrelerin ince ayarını yaparak daha iyi sonuçlar elde edilmesine olanak tanımaktadır (Chen and Guestrin, 2016; XGBoost Developers, 2024). Eksik veri yönetimi,

XGBoost'un önemli diğer özelliğidir. Seyrek veri farkındalığı (sparsity-aware) algoritması, eksik verileri varsayılan yöne sınıflandırır ve daha sonra bu verilerin model üzerindeki etkisini en aza indirmektedir. Bu yaklaşım, eksik veriler içeren veri kümelerinde XGBoost'un performansını artırmaktadır (Chen and Guestrin, 2016; XGBoost Developers, 2024). Esneklik, XGBoost'u güçlü kılan başka özelliktir. Kullanıcı tanımlı kayıp fonksiyonlarıyla çalışabilen XGBoost, özel veri kümeleri veya uygulamalar için özelleştirilebilir çözümler sunmaktadır. Bunun yanında, regresyon problemleri için ortalama karesel hata (MSE) ve sınıflandırma problemleri için çapraz entropi kaybı gibi standart kayıp fonksiyonlarını desteklemektedir. Bu özellikler, XGBoost'un geniş uygulama yelpazesinde kullanılmasını sağlamaktadır (Chen and Guestrin, 2016; XGBoost Developers, 2024). Tüm bu özellikler, XGBoost'u makine öğrenimi projelerinde performans, hız ve güvenilirlik açısından güçlü araç haline getirmektedir.

### **XGBoost algoritması:**

XGBoost, Gradient Boosting prensibine dayanmaktadır. Her ağaç, önceki ağacın tahmin hatalarını düzeltmeye odaklanmaktadır. Bu, ardışık öğrenme süreci ile gerçekleşmektedir. Aşağıdaki adımlar, XGBoost algoritmasının temelini oluşturmaktadır (XGBoost Developers, 2024):

- i) **Model başlatma:** İlk tahminler, genellikle hedef değişkenin ortalamasıyla ( $y_i(0)$ ) başlatılmaktadır.
- ii) **Artımlı tahmin:** Her yeni ağaç, önceki ağaçların hatalarını (kayıp fonksiyonu üzerinden) öğrenir. Model şu şekilde güncellenmektedir:

$$\begin{aligned}
 \widehat{y}_i^{(0)} &= 0 \\
 \widehat{y}_i^{(1)} &= f_1(x_i) = \widehat{y}_i^{(0)} + f_1(x_i) \\
 \widehat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \widehat{y}_i^{(1)} + f_2(x_i) \\
 &\vdots \\
 \widehat{y}_i^{(t)} &= \widehat{y}_i^{(t-1)} + f_t(x_i)
 \end{aligned} \tag{7.7}$$

$f_1(x_i)$ :  $t$ 'inci ağacın tahminini temsil etmektedir.

- iii) **Amaç fonksiyonu:** Model, kayıp fonksiyonu ( $l$ ) ve düzenleme ( $\Omega$ ) terimlerinden oluşan amaç fonksiyonunu optimize etmektedir:

$$\begin{aligned}
obj^{(t)} &= \sum_{i=1}^n l(y_i, \widehat{y_i^{(t)}}) + \sum_{i=1}^t \Omega(f_i) \\
&= \sum_{i=1}^n l(y_i, \widehat{y_i^{(t-1)}} + f_t(x_i)) + \Omega(f_t) + c \\
&= \sum_{i=1}^n \left( y_i - \left( \widehat{y_i^{(t-1)}} + f_t(x_i) \right) \right)^2 + \sum_{i=1}^t \Omega(f_i) \\
&= \sum_{i=1}^n \left( 2 \left( \widehat{y_i^{(t-1)}} - y_i \right) f_t(x_i) + f_t(x_i)^2 \right) + \Omega(f_t) + c
\end{aligned} \tag{7.8}$$

$l(y_i, \widehat{y_i^{(t)}})$ : Kayıp fonksiyonu (örneğin, kare hata veya log-loss).

$\Omega(f_i)$ : Düzenleme terimi, ağaç karmaşıklığını kontrol etmektedir.

$c$ : sabit sayıdır.

- iv) **Taylor serisi yaklaşımı:** Amaç fonksiyonu, Taylor serisi açılımı kullanılarak ikinci dereceye kadar genişletilmektedir:

$$obj^{(t)} = \sum_{i=1}^n \left[ l(y_i, \widehat{y_i^{(t-1)}}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) + c \tag{7.9}$$

Burada  $g_i$  ve  $h_i$  şu şekilde tanımlanabilir:

$g_i = \frac{\partial}{\partial \widehat{y_i^{(t-1)}}} l(y_i, \widehat{y_i^{(t-1)}})$ : Kayıp fonksiyonunun birinci türevi (gradyan)

$h_i = \frac{\partial^2}{\partial (\widehat{y_i^{(t-1)}})^2} l(y_i, \widehat{y_i^{(t-1)}})$ : Kayıp fonksiyonunun ikinci türevi (Hessian)

Sabiti kaldırarak formülün basitleştirilmiş hali:

$$\sum_{i=1}^n \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \tag{7.10}$$

- v) **Ağaç yapısı ve düzenleme:** Her ağaç, yaprak düğümleri ( $T$ ) ve yapraklara atanan ağırlıklar ( $w_j$ ) ile temsil edilmektedir:

$$f_t(t) = w_{q(x)}, w \in \mathbb{R}^T, q: \mathbb{R}^d \rightarrow \{1, 2, 3, \dots, T\} \tag{7.11}$$

- $w$ : Ağacın yapraklarındaki puanların vektörüdür,

- $q$  : Her veri noktasını karşılık gelen yaprağa atayan fonksiyondur
  - $T$  : Yaprak sayısıdır
- vi) **Düzenleme terimi:** Düzenleme, modelin genelleme kapasitesini artırmaktadır. Düzenleme terimi daha sonra şu şekilde tanımlanmaktadır:

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (7.12)$$

Amaç fonksiyonu şu hale gelir:

$$\begin{aligned} \text{obj}^{(t)} \approx \sum_{i=1}^n \left[ g_i w_{q(x_i)} + \frac{1}{2} h_i w_{q(x_i)}^2 \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \\ \sum_{j=1}^T \left[ \left( \sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left( \sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] \end{aligned} \quad (7.13)$$

Sadeleştirilmiş hali:

$$\text{obj}^{(t)} = \sum_{j=1}^T \left[ G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \quad (7.14)$$

- $\gamma$ : Yeni düğüm eklemek için gereken minimum bilgi kazancı.
  - $\lambda$ : Yaprak ağırlıklarını düzenlemek için ceza parametresi.
- vii) **Düzenleme terimi: Optimizasyon süreci:** Amaç fonksiyonu, her yaprağın ağırlığı ( $w_j$ ) etmektedir. İlgili ifadeler:

$$\text{Kazanç} = \frac{1}{2} \left[ \frac{G_L^2}{(H_L + \lambda)} + \frac{G_R^2}{(H_R + \lambda)} - \frac{(G_L + G_R)^2}{(H_L + H_R + \lambda)} \right] - \gamma \quad (7.15)$$

**Kazanç hesaplama:** Amaç fonksiyonu, her yaprağın ağırlığı ( $w_j$ ) için optimize edilmektedir:

$G_L, G_R$  :Sol ve sağ yapraklardaki Gradyan toplamları

$H_L, H_R$  :Sol ve sağ yapraklardaki Hessian toplamları.

**En iyi ağırlık ve hedef değer:**

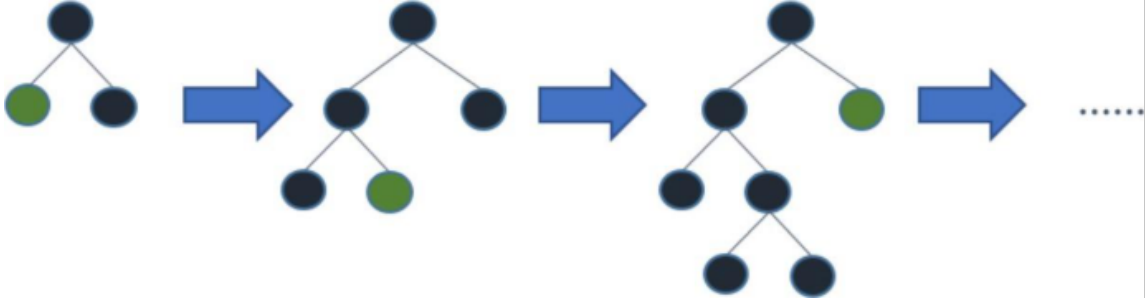
$$w_j^* = - \frac{G_j}{(H_j + \lambda)} \quad (7.16)$$

Modelin nesne fonksiyonundaki en iyi azalma:

$$obj^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{(H_j + \lambda)} + \gamma T \quad (7.17)$$

### 7.3 Hafif Gradyan Artırma Makinesi (Light Gradient Boosting Machine - LightGBM)

LightGBM, Microsoft tarafından geliştirilmiş, yüksek performanslı ve açık kaynaklı gradyan artırma çerçevesidir. Bu algoritma, hem büyük ölçekli veri setlerini etkili şekilde işleyebilmesi hem de hız ve bellek kullanımı açısından sunduğu avantajlarla dikkat çekmektedir. LightGBM, sıralama, sınıflandırma ve regresyon gibi makine öğrenimi görevlerinde yaygın şekilde kullanılmaktadır. Yenilikçi tasarımı, bellek kullanımını optimize ederken eğitim süresini önemli ölçüde kısaltmayı hedeflemektedir. Bu çerçevede, algoritma performansını artıran Gradient-Based One-Side Sampling (GOSS) ve Exclusive Feature Bundling (EFB) gibi özel teknikler içermektedir (Ke et al., 2017).



Şekil 7.3: Yaprak tabanlı büyüme

Şekil 7.3'te gösterildiği gibi LightGBM, ağaç yapılarında yaprak-tabanlı büyüme stratejisi ile çalışmaktadır. Bu strateji, geleneksel derinlik-tabanlı büyüme yerine her adımda kayıp fonksiyonunda en büyük azalmayı sağlayan yaprağı genişleterek asimetrik ve daha verimli ağaçlar oluşturmaktadır (LightGBM Development Team, 2024). İki farklı stratejinin karşılaştırması tablo 7.1'de verilmiştir. Bu tabloda da gösterildiği üzere yaprak-tabanlı büyüme stratejisi, daha düşük kayıp değerleri ve daha yüksek doğruluk sunarken, aşırı öğrenmeye yol açma riskine karşı derinlik sınırlamasıyla desteklemektedir. Ayrıca algoritma eksik değerlerin yönetiminde de son derece etkili yöntem olan seyrek veri farkındalığı algoritmasını kullanmaktadır. Bu sayede, eksik veriler varsayılan yöne atanarak en iyi çözüm öğrenerek eksik değerlerin olumsuz etkisi minimize etmektedir (LightGBM Development Team, 2024).

**Tablo 7.1:** Ağaç büyüme stratejilerinin karşılaştırılması

<b>Kriter</b>	<b>Geleneksel Derinlik Tabanlı Ağaç Büyümesi</b>	<b>Yaprak Tabanlı Ağaç Büyümesi</b>
<b>Genişleme Sırası</b>	Ağacı seviyeler halinde genişletmektedir; önce aynı seviyedeki tüm düğümleri bölerek sonraki seviyeye geçmektedir.	Ağacı yapraklar halinde genişletmektedir; her derinlikte en yüksek bölme kazancına sahip düğümü öncelikli olarak bölmektedir.
<b>Kayıp Azaltımı</b>	Küresel olarak en iyi bölme noktasını dikkate almadığından, kayıpları en aza indirme konusunda optimal olmayabilmektedir.	Her zaman en yüksek kazancı sağlayan bölme seçtiği için daha fazla kayıp azaltımı sağlayabilmektedir.
<b>Aşırı Öğrenme</b>	Dengeli ve sık ağaçlar oluşturduğu için aşırı öğrenmeye daha az yatkındır.	Daha karmaşık ve derin ağaçlar oluşturduğu için aşırı öğrenme daha yatkındır.
<b>Karmaşıklık Kontrolü</b>	Maksimum derinlik veya düğüm başına minimum örnek sayısı sınırı koyarak ağacın karmaşıklığını kontrol etmektedir.	Maksimum yaprak sayısı veya minimum bölme kazancı sınırı koyarak ağacın karmaşıklığını kontrol etmektedir.
<b>Algoritma Örnekleri</b>	C4.5, CART, XGBoost	LightGBM

LightGBM, günümüzde en çok tercih edilen gradyan artırma algoritmalarından biri olarak, özellikle büyük ölçekli veri setleri üzerinde sunduğu hız ve verimlilikle öne çıkmaktadır. Modelin temelinde yer alan histogram tabanlı öğrenme algoritması, sürekli değişkenleri ayrık aralıklara bölerek eğitim süresini önemli ölçüde kısaltmakta ve bellek kullanımını optimize etmektedir. Histogram yaklaşımı, bölünme kazançlarının hızlı hesaplanmasına imkân tanırken, bellek üzerindeki yükü de azaltmaktadır. Böylece LightGBM, sınırlı donanım kaynaklarında dahi büyük veri setlerini etkili biçimde işleyebilmektedir. Bunun yanı sıra, LightGBM'nin GPU hızlandırma ve paralel işlem desteği, modelin çok çekirdekli işlemciler ve grafik işlem birimleri üzerinde çalışmasına olanak tanımakta, böylece eğitim sürecini daha da hızlandırmaktadır (Ranka ve Singh, 1998; Jin ve Agrawal, 2003; Li et al., 2007). Bu özellikler, algoritmayı hem araştırma hem de endüstriyel ölçekte uygulamalar için son derece cazip hale getirmektedir. LightGBM'nin başarısında, aynı zamanda yenilikçi optimizasyon teknikleri de rol oynamaktadır. Gradient-based One-Side Sampling (GOSS) yöntemi, yüksek bilgi taşıyan örnekleri önceliklendirerek gereksiz hesaplamaları azaltırken, Exclusive Feature Bundling (EFB) tekniği ise seyrek ve birbirini dışlayan özellikleri araya getirerek bellek kullanımını ciddi ölçüde düşürmektedir. Bu iki yöntem sayesinde LightGBM, yüksek doğruluk ile düşük kaynak tüketimini arada sunarak diğer gradyan artırma algoritmalarından ayrılmaktadır.

LightGBM'nin düzenleme mekanizmaları, modelin genelleme kapasitesini artırırken aşırı öğrenmeyi önlemek için tasarlanmıştır. L1 ve L2 norm düzenlemesi, yaprak düğüm sayısı sınırlandırması ve minimum bölme kazancı gibi teknikler bu mekanizmalar arasında yer almaktadır. Bunun yanı sıra, hiperparametre ayarları için sunduğu geniş seçenekler, kullanıcıların modeli veri setine ve problemin gereksinimlerine göre optimize etmesine olanak tanımaktadır (Zhang et al., 2019).

Tüm bu avantajları sayesinde LightGBM; arama motoru sıralamaları, öneri sistemleri, tıklama oranı tahmini ve finansal modelleme gibi birçok alanda etkili sonuçlar vermekte ve modern makine öğrenmesi projelerinde güçlü araç olarak öne çıkmaktadır (LightGBM Development Team, 2024).

### LightGBM Algoritması:

Aşağıda LightGBM algoritmasının temel adımları açıklanmaktadır (Taha and Malebary, 2020):

#### Girdi:

- Eğitim verisi:  $x_i \in \mathbb{R}^d, y_i \in \{-1, +1\} \rightarrow D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ .
  - $L(y, \theta(x))$ : Kayıp fonksiyonudur
  - $M$ : İterasyon sayısıdır
  - $a$ : Büyük gradyan veri örnekleme oranıdır
  - $b$ : Küçük gradyan veri örnekleme oranıdır
- i) **Ayrık özellik demetleme (Exclusive Feature Bundling - EFB)**: Birbirine bağımsız olan özellikler (aynı anda sıfır olmayan değerleri nadiren kabul eden) birleştirilir. Bu özellikler, "Exclusive Feature Bundling (EFB)" yöntemi ile bir araya getirilir ve boyut indirgeme gerçekleştirilir.
- ii) **Başlangıç modelinin ayarlanması**: İlk model parametresi  $\theta_0(x) = \arg \min \sum_{i=1}^N L(y_i, c)$  olacak şekilde ayarlanmaktadır.
- iii)  $m = 1$  ile  $M$  arasında bir döngü gerçekleştirilmektedir:
- a) Gradyanların hesaplanması: Gradyanların mutlak değerleri hesaplanmaktadır.
  - b) Gradyan tabanlı tek taraflı örnekleme (Gradient-Based One-Side Sampling - GOSS):

Veriler büyük ve küçük gradyanlara göre örneklenir. Böylelikle  $topN = a \times len(D)$  en büyük kadar gradyan seçilmektedir. Denklem 7.18'de gösterilmektedir.

$$r_i = \left| \frac{\partial L(y_i, \theta^{(m)}(x_i))}{\partial \theta(x_i)} \right| \theta(x) = \theta_m(x) - 1(x) \quad (7.18)$$

- Kalan gradyanlar arasından  $randN = b \times len(D)$  kadar rastgele seçim yapılmaktadır.
  - Seçilen veri kümeleri:  $A$  ve  $B$
  - Nihai veri seti:  $D = A + B$
- c) Bilgi kazançlarının hesaplanması: Her bölünme adayı için bilgi kazancı hesaplanmaktadır.
- d) Yeni karar ağacının oluşturulması:  $\hat{D}$  veri seti kullanılarak yeni karar ağacı  $\theta_m(x)$  oluşturulmaktadır.
- e) Model Güncellemesi:  $\theta_m(x)$  önceki modelle birleştirilmektedir.

## Sonuç

$$\theta_m(x) = \theta_{m-1}(x) + \theta_m(x) \quad (7.19)$$

İterasyonlar tamamlandıktan sonra model  $\hat{\theta}(x) = \theta_M(x)$  olarak döndürülmektedir. Bu algoritma, GOSS ve EFB gibi yenilikçi tekniklerle bellek kullanımını optimize ederek ve eğitim süresini kısaltarak büyük veri setleri üzerinde etkili model oluşturmayı sağlamaktadır. Leaf-wise (yaprak bazlı) büyüme stratejisi ve histogram tabanlı öğrenme algoritması sayesinde, LightGBM diğer gradyan artırma yöntemlerinden üstün performans sergilemektedir (Taha and Malebary, 2020).

**Tablo 7.2:** LightGBM ve XGBoost arasındaki farklar

Özellik	LightGBM	XGBoost
Tanıtım Yılı	2017	2014
Kategorik Verilerin İşlenmesi	Bir-sıcak encoding ve doğrudan kategorik veri desteği	Ön işleme gerektirir
Ağaç Bölme Stratejisi	Yaprak bazlı bölme	Derinlik bazlı bölme
Hız ve Ölçeklenebilirlik	Büyük veri setleri için çok hızlı	Hızlı ve çok yönlü
Hafıza Kullanımı	Düşük	Orta

**Tablo 7.2 (devam)**

<b>Düzenleştirme</b>	Evet	Evet
<b>Paralel İşleme</b>	Evet	Evet
<b>GPU Desteği</b>	Evet	Evet
<b>Özellik Önem Sıralaması</b>	Mevcut	Mevcut
<b>Yorumlanabilirlik</b>	Özellik önem sıralamaları, bölme değer histogramları	Özellik önem sıralamaları, ağaç görselleri
<b>Kategorik Verilerde Performans</b>	İyi	Orta
<b>Hız ve Verimlilik</b>	Büyük veri setleri için çok verimli	Çok yönlü ve hızlı

Tablo 7.2’de, LightGBM ve XGBoost’un yeteneklerini, tasarım stratejilerini ve performans özelliklerini özetlemektedir (parthshukla211, 2023; akshisaxena, 2024). Kullanıcılar, bu karşılaştırmayı göz önünde bulundurarak projelerine en uygun kütüphaneyi seçilebilmektedir.

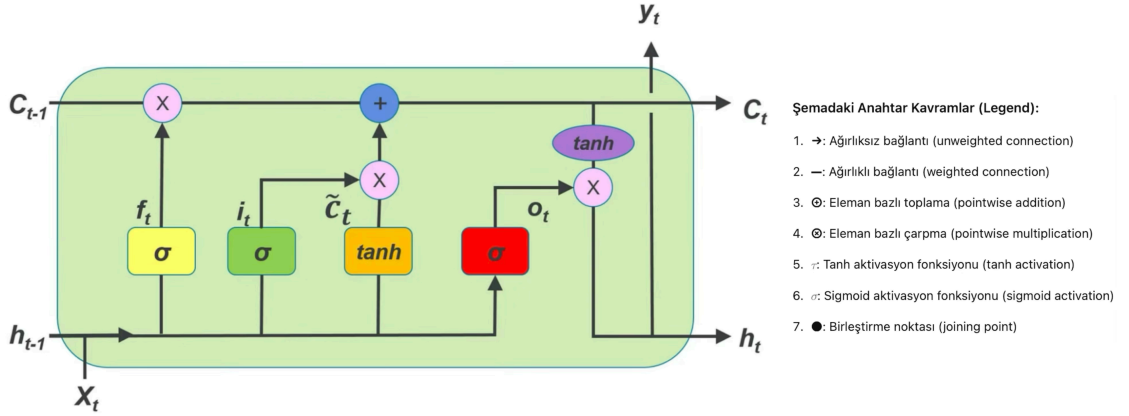
#### **7.4 Uzun-Kısa Vadeli Bellek (Long Short-Term Memory - LSTM)**

LSTM, Tekrarlayan Sinir Ağlarının (RNN) temel zorluklarını aşmak için geliştirilmiş mimaridir. RNN’ler, geçmiş girdileri hatırlayarak zaman serisi veya ardışık veri setleri üzerinde çalışabilirken, bu ağlar özellikle uzun zaman bağımlılıklarıyla başa çıkmakta zorlanmaktadır. LSTM, bu zorlukların üstesinden gelmek için tasarlanmış ve özellikle vanishing gradient (kaybolan gradyan) ve exploding gradient (patlayan gradyan) problemlerine karşı çözüm sunmaktadır (Hochreiter and Schmidhuber, 1997). LSTM’nin temel bileşenleri hücre durumu (cell state) ve kapılar (gates) mekanizmasıdır. Hücre durumu, bilgiyi uzun süreli olarak saklamayı sağlarken, giriş, çıkış ve unutma kapıları, bu bilginin güncellenmesi ve gereksiz bilginin elenmesi süreçlerini yönetmektedir. Bu mekanizma sayesinde, LSTM’ler uzun vadeli bağımlılıkları öğrenme ve işleme yeteneği kazanmaktadır (Gers et al., 2000). LSTM’nin başarısı, özellikle dil modelleme, konuşma tanıma, makine çevirisi ve video analizi gibi ardışık veri içeren uygulamalarda kendini göstermektedir (Graves et al., 2005). Geleneksel RNN’ler, yalnızca belirli zaman aralığındaki bağlama erişebilirken, LSTM’ler 1000’den fazla zaman adımı arasında bağlamları öğrenip aktarmaktadır. Bu özellik, özellikle uzun dönem bağımlılıkların önemli olduğu uygulamalarda büyük avantaj sağlamaktadır (Gers et al., 2000). Örneğin, konuşma tanıma gibi uygulamalarda bağlamın yalnızca geçmiş girdilerle değil, gelecekteki girdilerle de ilişkilendirilmesi gerektiği durumlarda bidirectional LSTM (çift yönlü LSTM) kullanılmaktadır. Bu yöntem, veriyi ileri ve geri yönlü işleyerek her iki yönden de bağlamsal bilgiyi optimize etmektedir (Graves et al., 2013). LSTM’lerin başka önemli kullanım alanı,

sequence-to-sequence (seq2seq) modellerdir. Bu modeller, diziyi girdiye alıp başka diziyeye dönüştürmek için encoder ve decoder iki farklı LSTM ağından yararlanmaktadır. Encoder, giriş dizisini sabit boyutlu vektöre dönüştürürken, decoder bu vektörü hedef çıktıya dönüştürmektedir. Seq2seq modeller, makine çevirisi ve otomatik altyazı oluşturma gibi alanlarda başarıyla uygulanmaktadır (Sutskever et al., 2014).

### LSTM algoritması:

Şekil 7.4'te LSTM hücresi, giriş kapısı (input gate), unutma kapısı (forget gate) ve çıkış kapısı (output gate) gösterildiği gibi bileşenler içermektedir. Bu kapılar, hücre durumunu (cell state,  $C_t$ ) ve gizli durumu (hidden state,  $h_t$ ) kontrol ederek uzun süreli bağımlılıkları öğrenme yeteneğini sağlamaktadır (Sagheer and Kotb, 2019).



Şekil 7.4: LSTM çalışma modeli

- $f_t$ : Unutma kapısı (forget gate)
  - $i_t$ : Giriş kapısı (input gate)
  - $o_t$ : Çıkış kapısı (output gate)
  - $\tilde{C}_t$ : Şimdiki hücre durumu (current cell state)
  - $C_{t-1}$ : Önceki hücre durumu (previous cell state)
  - $h_t$ : Şimdiki gizli durum (current hidden state)
  - $h_{t-1}$ : Önceki gizli durum (previous hidden state)
  - $X_t$ : Giriş vektörü (input vector)
- i) **Unutma kapısı ( $f_t$ ):** Unutma kapısı, hangi bilginin hücre durumundan atılacağını belirlemektedir. Geçmiş hücre durumundan ( $C_{t-1}$ ) hangi bilginin “unutulacağını” kontrol etmektedir. Sigmoid fonksiyonu ( $\sigma$ ), 0 ile 1 arasında çıktı

üretmek için bilginin ne kadarının tutulup ne kadarının silineceğini belirlemektedir (Graves et al., 2013; Cui et al., 2020).

$$f_t = \sigma(W_f \cdot [h_{(t-1)}, x_t] + b_f) \quad (7.20)$$

- $W_f$ : Geçmiş gizli durum ( $h_{t-1}$ ) ve yeni giriş ( $x_t$ ) arasındaki bağlantıyı öğrenen ağırlık matrisidir.
  - $b_f$ : Öğrenilen bias değeridir.
  - $\sigma$ : Her hücre için “unutulması gereken” bilginin oranını belirlemektedir.
- ii) **Giriş kapısı ( $f_t$ ) ve aday hücre durumu ( $\tilde{C}_t$ ):** Giriş kapısı, yeni bilginin ne kadarının hücre durumuna dahil edileceğini belirlenmektedir (Graves et al., 2013; Cui et al., 2020).

$$i_t = \sigma(W_i \cdot [h_{(t-1)}, x_t] + b_i) \quad (7.21)$$

Giriş kapısı, hangi bilginin yeni hücre durumuna eklenmesi gerektiğini kontrol etmektedir.

- $W_i$ : Giriş kapısı için ağırlık matrisi.
- $b_i$ : Giriş kapısı için bias.

Hücreye yeni bilgi eklenmeden önce aday hücre durumu ( $\tilde{C}_t$ ) oluşturmaktadır (Graves et al., 2013; Cui et al., 2020).

$$\tilde{C}_t = \tanh(W_c \cdot [h_{(t-1)}, x_t] + b_c) \quad (7.22)$$

- $W_c$ : Hücre durumu için ağırlık matrisi.
- $b_c$ : Hücre durumu için bias.
- $\tanh$ : Tanh aktivasyon fonksiyonu.

Aday hücre durumu, eklenmesi düşünülen yeni bilgiyi temsil etmektedir. Tanh fonksiyonu, -1 ile 1 arasında çıktı üretmek için veriyi normalleştirir ve stabilize sağlamaktadır. Bu iki bileşen, eski hücre durumunun üzerine yeni bilgi eklemek için birlikte çalışmaktadır.

- iii) **Hücre Durumunun Güncellenmesi ( $C_t$ ):** Hücre durumu ( $C_t$ ) geçmiş bilgi ( $C_{t-1}$ ) ve yeni bilgi ( $\tilde{C}_t$ ) arasında denge kurarak güncellenmektedir (Graves et al., 2013; Cui et al., 2020).

$$C_t = f_t \odot C_{(t-1)} + i_t \odot \tilde{C}_t \quad (7.23)$$

- $\odot$ : Eleman bazında çarpma işlemi.
  - Unutma kapısı ( $f$ ): Önceki hücre durumu ( $C_{t-1}$ ) çarpılarak gereksiz bilgiler silinmektedir.
  - Giriş kapısı ( $i_t$ ): Aday hücre durumu ( $C_t$ ) çarpılarak yeni bilgi eklenmektedir.
- iv) Çıkış Kapısı ( $o_t$ ): Çıkış kapısı, hangi bilginin hücreden gizli duruma aktarılacağını belirlemektedir.

$$o_t = \sigma(W_o \cdot [h_{(t-1)}, x_t] + b_o) \quad (7.24)$$

- v) Gizli Durumun Güncellenmesi ( $h$ ): Güncellenmiş hücre durumu kullanılarak gizli durum hesaplanmaktadır (Graves et al., 2013; Cui et al., 2020).

$$h_t = o_t \odot \tanh(C_t) \quad (7.25)$$

- $W_o$ : Çıkış kapısı için ağırlık matrisi.
- $b_o$ : Çıkış kapısı için bias.

Bu mekanizma, LSTM'nin uzun dönemli bağlamları öğrenmesini ve kısa dönemli gereksiz bilgileri filtrelemesini sağlamaktadır.

## 7.5 Evrişimli Sinir Ağları (Convolutional Neural Networks - CNN)

CNN, görsel veriler ve zaman serisi analizleri gibi yapılandırılmış verilerin işlenmesinde devrim niteliğinde derin öğrenme algoritmasıdır. CNN'ler, manuel özellik mühendisliği gereksinimini ortadan kaldırarak, verilerden doğrudan desen ve özellik öğrenebilme kapasitesine sahiptir. Bu özellikleriyle modern yapay zeka sistemlerinin temel yapı taşlarından biri haline gelmiştir. Görüntü tanıma, nesne algılama, metin analizi ve medikal teşhis gibi çeşitli alanlarda kullanılmaları, CNN'leri yapay zekanın en etkili araçlarından biri yapmaktadır (IBM, 2025).

Şekil 7.5'de görüldüğü gibi, eklenen her konvolüsyonel katman veriyi daha karmaşık ve yüksek seviyeli özelliklere dönüştürmektedir. Bu süreç, nesneyi daha basit yapı taşlarına bölerek her seviyede daha gelişmiş desenlerin öğrenilmesine olanak tanımaktadır (IBM,

2025). Örneğin, bisikletin çerçeve, tekerlek ve pedal gibi parçalarını algılayan katmanlar, daha üst seviyelerde bu parçaları birleştirerek bisikletin varlığını tanımlayabilmektedir. Bu CNN'lerin hiyerarşik şekilde öğrenmesini sağlamaktadır.

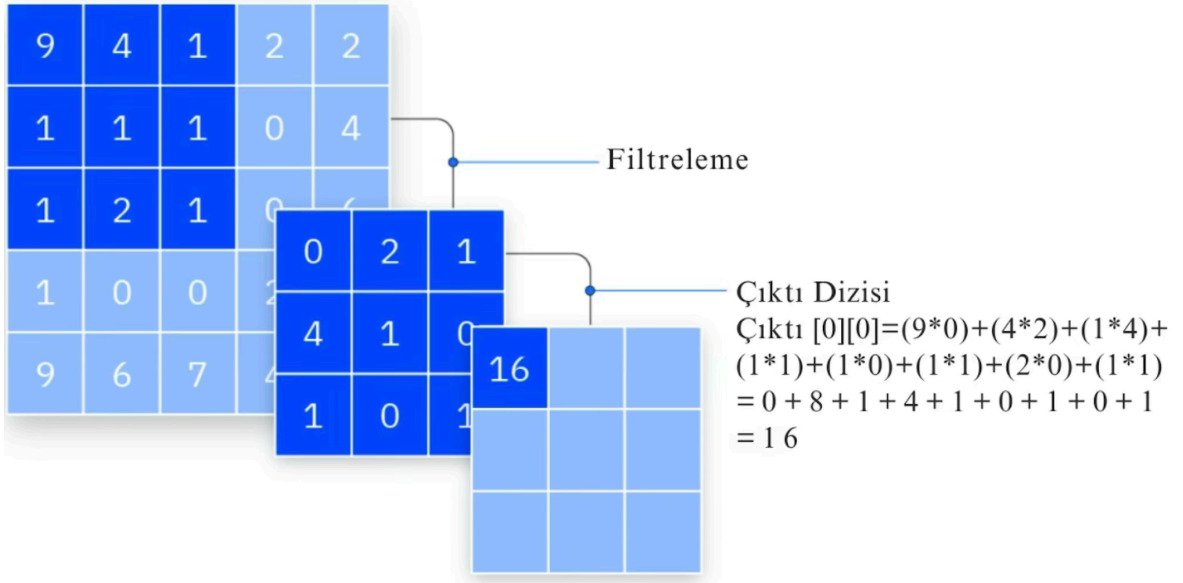


**Şekil 7.5:** Ek katmanların özellik hiyerarşisi

LeCun ve ekibi tarafından geliştirilen LeNet-5, CNN'lerin tarihindeki ilk başarılı uygulamalardan biri olarak öne çıkmıştır. Bu model, el yazısı karakter tanımda yüksek performans göstermiş ve gradyan tabanlı öğrenmenin gücünü vurgulamıştır (LeCun et al., 1998). Derin öğrenmenin popülerleşmesi ise Krizhevsky, Sutskever ve Hinton tarafından geliştirilen AlexNet ile gerçekleşmiştir. AlexNet, ImageNet yarışmasındaki başarısıyla büyük veri setlerinde CNN'lerin üstün performansını kanıtlamış olup yapay zeka araştırmalarında yeni dönemin kapısını aralamıştır (Krizhevsky et al., 2012). Daha sonra, Simonyan ve Zisserman tarafından geliştirilen VGGNet, daha derin ağ mimarilerinde küçük filtre boyutları kullanarak görsel veriler üzerinde daha yüksek doğruluk elde edilmesini sağlamıştır (Simonyan and Zisserman, 2015).

Şekil 7.6'da gösterildiği gibi, CNN'lerde giriş matrisine uygulanan filtreler önemli özelliklerin çıkarılmasını sağlamaktadır. Filtre, giriş görüntüsü üzerinde kaydırılarak her konumda çarpım ve toplam işlemi gerçekleştirilmektedir (IBM, 2025). Bu işlem sonucunda, anlamlı özelliklerin korunduğu ve boyutların azaltıldığı "çıkıktı matrisi" elde edilmektedir. Bu süreç, CNN'lerin görsel verilerdeki temel desenleri tanımasını sağlamaktadır.

Giriş Resmi



Şekil 7.6: Konvolüsyon operasyonu ve çıktı hesaplama

CNN mimarilerindeki diğer önemli ilerleme, He ve ekibi tarafından geliştirilen ResNet ile sağlanmıştır. ResNet, derin ağlarda yaygın sorun olan kaybolan gradyan problemini çözmek için “skip connections” (atlama bağlantıları) yöntemini tanıtmıştır. Bu sayede çok daha derin ağların etkin şekilde eğitilmesine olanak sağlamıştır (He et al., 2016). Bununla birlikte, Huang ve ekibinin geliştirdiği DenseNet, ağ katmanları arasındaki bağlantıları yoğunlaştırarak hem bilgi akışını iyileştirmiş hem de parametre verimliliğini artırmıştır (Huang et al., 2017).

CNN’lerin uygulama alanları görsel verilerin ötesine geçmiştir. Örneğin, O’Shea ve Nash tarafından yapılan çalışma, CNN’lerin temel prensiplerini açıklayarak yeni başlayanlar için kapsamlı giriş kaynağı sunmuştur (O’Shea and Nash, 2015). Benzer şekilde, Stanford Üniversitesi’nin ünlü CS231n kursu, görsel tanıma için CNN’lerin teorik ve uygulamalı yönlerini detaylı şekilde ele almaktadır (Stanford University, 2024-2025).

Bu mimarilerin gücünü artıran önemli gelişmelerden biri de Inception ağlarıdır. Szegedy ve ekibi tarafından geliştirilen bu mimari, hesaplama verimliliğini artırarak daha karmaşık görevlerde üstün sonuçlar elde edilmesini sağlamıştır (Szegedy et al., 2015). Ayrıca TensorFlow gibi çerçeveler, CNN uygulamalarını araştırmacılar ve geliştiriciler için daha erişilebilir hale getirerek bu teknolojinin hızla yaygınlaşmasına katkıda bulunmuştur

(TensorFlow, 2022). Son olarak, Gu ve ekibi tarafından yazılan kapsamlı derleme makalesi, CNN'lerdeki son gelişmeleri, karşılaşılan zorlukları ve gelecekteki potansiyel uygulamaları detaylandırarak bu alanda değerli rehber sunmaktadır (Gu et al., 2018).

### **CNN Algoritması:**

Convolutional Neural Networks (CNN) içinde kullanılan formüller, görüntü işleme sürecinde birbirini takip eden aşamalar oluşturur ve birbirleriyle doğrudan ilişkilidir.

#### **i) Konvolüsyon işlemi ve stride, padding etkisi:**

- a) Konvolüsyon operasyonu: CNN'deki temel işlem konvolüsyondur. Konvolüsyon işlemi, giriş görüntüsüne ( $X$ ) filtre ( $H$ ) uygulayarak belirli özellikleri çıkarmaya çalışmaktadır.

$$y(i, j) = \sum_{u=1}^l \sum_{s=1}^l H(u, s)X(i + u, j + s) \quad (7.26)$$

Bu işlem, görüntüdeki kenarları, dokuları ve diğer önemli detayları algılamaktadır. Ancak, bu işlem sırasında bazı ek parametreler çıktıyı doğrudan etkilenmektedir.

- b) Stride ( $S$ ) ve Padding ( $P$ ) etkisi: Konvolüsyon işleminin ardından elde edilen çıktının boyutları şu formülle belirlenmektedir:

$$\text{OutputSize} = \frac{\text{InputSize} + 2P - \text{KernelSize}}{S} + 1 \quad (7.27)$$

- c) Eğer (stride) büyükse, konvolüsyon filtresi görüntü üzerinde daha büyük adımlarla hareket eder ve çıktı boyutu küçülür. Eğer (padding) eklenirse, giriş görüntüsü büyütülür ve kenar bilgileri kaybolmaz. Eğer Kernel Size büyütülürse, filtre daha büyük alanı kapsar ve daha karmaşık özellikleri öğrenebilir, ancak çıktı boyutu küçülür. Yani konvolüsyon işlemi, stride ve padding gibi hiperparametrelerle doğrudan bağlantılıdır. Bunlar çıktının boyutunu belirlemektedir.

- ii) **Çıktı boyutu ve Max-Pooling etkileşimi:** Konvolüsyon katmanından sonra genellikle max-pooling uygulanmaktadır. Max-pooling, görüntünün önemli özelliklerini koruyarak boyutlarını küçültmek için kullanılmaktadır.

$$y(i, j) = \max\{X(2i, 2j), X(2i, 2j + 1), X(2i + 1, 2j), X(2i + 1, 2j + 2)\} \quad (7.28)$$

Pooling işlemi çıktı boyutunu küçülterek hesaplama maliyeti azaltmaktadır. Max-pooling işlemi doğrudan konvolüsyon sonrası elde edilen çıktıyı etkilemektedir, yani konvolüsyon çıktısı büyükse pooling işlemi ile boyutları küçültebilmektedir. Pooling boyutu ne kadar büyükse, çıktı boyutu o kadar küçülmektedir.

iii) **Konvolüsyon ve Aktivasyon Fonksiyonu (Rectified Linear Unit - ReLU) etkileşimi:** Konvolüsyon işleminin ardından aktivasyon fonksiyonu kullanılmaktadır. En yaygın kullanılan aktivasyon fonksiyonu ReLU'dur:

$$f(x) = \max(0, x) \quad (7.29)$$

Bu, negatif değerleri sıfıra eşitleyerek yalnızca pozitif değerleri korumaktadır. Bu durum eğer konvolüsyon işlemi sonucunda negatif değerler varsa, ReLU bunları sıfıra çevirmektedir. Ya da o bölgede aktivasyonun olmaması anlamına gelmektedir. Bu yüzden model, yalnızca önemli olan özellikleri öğrenerek etkilemektedir.

Yani, Konvolüsyon işlemi, ReLU'ya girdi sağlamaktadır. ReLU, bu girdiyi işleyerek negatif değerleri kaldırır ve özellikleri daha belirgin hale getirmektedir.

iv) **Çıktı boyutu ve parametre sayısı (ağın kompleksliği)**

CNN mimarisinde her katmanda belirli sayıda parametre bulunmaktadır. Parametreler şunlara bağlıdır:

1. Filtre büyüklüğü (Kernel Size)
2. Giriş kanalı sayısı (Input Channels)
3. Çıkış kanalı sayısı (Output Channels)

Toplam parametre sayısı şu formülle hesaplanmaktadır:

$$\text{ParameterSayısı} = (\text{FilterSize} \times \text{InputChannels} + 1) \times \text{OutputChannels} \quad (7.30)$$

Bu formülün CNN ağına etkisi: Eğer daha fazla filtre (output channel) eklenirse, parametre sayısı artmaktadır. Eğer daha büyük filtreler kullanılırsa, modelin karmaşıklığı artmaktadır. Eğer çok fazla parametre varsa, model aşırı öğrenmeye yatkın hale gelmektedir.

### Bağlantı:

- Konvolüsyon işlemi çıktı ürettiği için bu çıktı parametre hesaplamalarını doğrudan etkilemektedir.
  - Daha küçük çıktılar (max-pooling gibi işlemlerle) oluşturulursa, parametre sayısı azalmaktadır.
  - Daha büyük model daha fazla parametre gerektirir, bu da daha uzun eğitim süresi anlamına gelmektedir
- v) **Olasılık ve konvolüsyonun ilişkisi:** Konvolüsyon sadece görüntü işleme için değil, olasılık dağılımlarının hesaplanmasında da kullanılmaktadır. Bağımsız iki rastgele değişkenin toplamının olasılık yoğunluk fonksiyonu şu şekilde hesaplanmaktadır:

$$P_Z^{(t)} = \int_{-\infty}^{\infty} P_x^{(\tau)} P_y^{(t-\tau)} d\tau \quad (7.31)$$

CNN'lerde, olasılık yoğunluk fonksiyonu, modelin sınıflandırma çıktılarıyla ilgilidir. Softmax gibi son katman, modelin belirli sınıfa ait olasılığını hesaplamak için kullanılmaktadır.

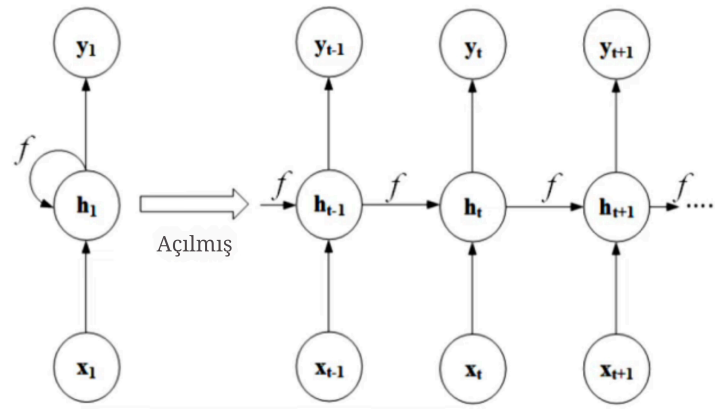
- vi) **Son katman (softmax) ve eğitim süreci ile etkileşim:** Konvolüsyon, aktivasyon fonksiyonları, pooling ve diğer işlemler sonucunda, çıktı tamamen bağlı (fully connected) katmanlara iletilmektedir. Son olarak, softmax veya sigmoid fonksiyonu kullanılarak modelin tahminleri hesaplanmaktadır. Örneğin, softmax fonksiyonu:

$$\sigma(z_i) = \frac{e^{(z_i)}}{\sum_j e^{(z_j)}} \quad (7.32)$$

CNN'nin ürettiği son özellik haritası (feature map), tam bağlı katmana ( $W$ ) bağlanarak sınıflandırma yapılmaktadır. Softmax, modelin sınıf olasılıklarını hesaplar ve en yüksek olasılığa sahip sınıf tahmin edilmektedir. Eğer daha fazla konvolüsyon katmanı veya daha büyük filtreler kullanılmışsa, daha fazla özellik çıkarılacağından softmax işlemi daha fazla bilgi ile çalışılmaktadır.

## 7.6 Tekrarlayan Sinir Ağları (Recurrent Neural Networks - RNN)

RNN, sıralı veya zaman serisi verilerini analiz etmek için geliştirilmiş yapay sinir ağı modelidir (Goodfellow et al., 2016). Geleneksel ileri beslemeli sinir ağlarından (Feedforward Neural Networks - FNN) farklı olarak, RNN'ler girdiler arasındaki zaman bağımlılıklarını öğrenerek ardışık verileri işleyerek tahmin yapabilmektedir. Bu özellikleri sayesinde, doğal dil işleme (NLP), konuşma tanıma, duygu analizi, görüntü altyazılama ve finansal zaman serisi tahmini gibi pek çok alanda başarıyla uygulanmaktadır (Graves, 2012; Pascanu et al., 2013).



Şekil 7.7: RNN çalışma prensibi

RNN'lerin en önemli özelliği, geçmiş girdileri belleğinde tutabilmesi ve bunları mevcut çıktılarının belirlenmesinde kullanabilmesidir (Hochreiter and Schmidhuber, 1997). Şekil 7.7'de gösterildiği üzere bunu sağlamak için RNN, her zaman adımında gizli durum (hidden state) oluşturmaktadır. Bu durum, önceki zaman adımlarından gelen bilgiyi mevcut girdilerle birleştirerek güncellemektedir (Elman, 1990; Sutton, 2018). Böylece, ağın geçmiş verileri hatırlaması ve sıralı veriler arasındaki ilişkileri anlamlandırması mümkün olmaktadır. Geleneksel sinir ağlarında girişler bağımsız olarak değerlendirilirken, RNN'ler girdiler arasında bağlam ilişkisi kurarak daha anlamlı tahminler yapar.

RNN'lerin çalışma prensibi, önceki adımlardan gelen bilgileri modelin ağırlıklarının zaman içinde güncellenmesiyle öğrenmeye dayanmaktadır (Rumelhart et al., 1986). Model, ileri besleme (forward propagation) adımında girdiyi işlemektedir. Çıktı üretirken de, zaman içinde geri yayılım (Backpropagation Through Time - BPTT) algoritması ile hataları hesaplayarak ağırlıkları güncellenmektedir (Werbos, 1990). Ancak, RNN'lerin eğitimi sırasında gradyan kaybolma (vanishing gradient) veya gradyan patlaması (exploding gradient) gibi sorunlarla karşılaşabilmektedir (Pascanu et al., 2013). Gradyan kaybolma

problemi, uzun sekanslardaki bağımlılıkların öğrenilmesini zorlaştırırken, gradyan patlaması ise modelin dengesiz öğrenmesine neden olabilmektedir. Bu sorunları çözmek için Uzun Kısa Süreli Bellek (Long Short-Term Memory - LSTM) ve Kapılı Tekrarlayan Birim (Gated Recurrent Unit - GRU) gibi gelişmiş RNN mimarileri geliştirilmiştir (Hochreiter and Schmidhuber, 1997). LSTM'ler, giriş, çıkış ve unutmaya kapıları ile bilgiyi uzun vadeli olarak saklayabilirken, GRU'lar daha hafif ve hesaplama açısından daha verimli yapı sunmaktadır (Chung et al., 2014).

RNN'lerin farklı mimarileri arasında Hopfield Ağları, Elman Ağları ve Jordan Ağları da bulunmaktadır. Hopfield ağı, örüntü tanıma ve hafıza depolama gibi görevler için geliştirilmiş özel ağ yapısıdır. Elman ve Jordan ağları ise bağlam katmanı (context layer) kullanarak geçmiş bilgiyi saklayan ve tekrar kullanabilen basit tekrarlayan ağ yapılarıdır (Hochreiter and Schmidhuber, 1997).

### **RNN algoritması**

RNN, girdileri zaman içinde işleyerek sekans bağımlılıklarını öğrenmek için tasarlanmıştır. Bu modelde, her zaman adımında önceki gizli durum (hidden state) kullanılarak yeni gizli durum hesaplanmaktadır.

- i) **Gizli Durum (Hidden State) hesaplaması:** Her giriş için gizli durum şu şekilde hesaplanmaktadır (Aishwarya.27, 2024):

$$h_t = \sigma(U * X_t + W * h_{(t-1)} + B) \quad (7.33)$$

- $h_t$ : Mevcut zaman adımındaki gizli durum.
- $h_{t-1}$ : Bir önceki zaman adımındaki gizli durum.
- $X_t$ : Mevcut giriş verisi.
- $U, W$  : Ağırlık matrisleri.
- $B$  : Bias terimi.
- $\sigma$ : Aktivasyon fonksiyonu.

- ii) **Çıktı hesaplaması:**

RNN modelinde, her gizli durumdan çıktı elde edilmektedir. Bu işlem aşağıdaki formülle ifade edilmektedir (Aishwarya.27, 2024):

$$Y_t = O(V * h_t + C) \quad (7.34)$$

- $Y_t$ : Modelin ürettiği çıktı.
  - $h_t$ : Mevcut gizli durum.
  - $V$ : Ağırlık matrisi.
  - $C$ : Bias terimi.
  - $O$ : Aktivasyon fonksiyonu.
- iii) **Genel fonksiyon (Overall Function)**: Tüm işlemleri tek fonksiyon ile ifade edilirse (Aishwarya.27, 2024):

$$Y = f(X, h, W, U, V, B, C) \quad (7.35)$$

Bu fonksiyon, tüm zaman adımları boyunca RNN modelinin çalışma prensibini göstermektedir. RNN'nin eğitim sürecinde güncellenen parametreleri tablo 7.35'te paylaşılmaktadır (Aishwarya.27, 2024).

**Tablo 7.3:** RNN değer gösterim tablosu

Ağırlık Matrisleri	Bias Terimleri
W: Önceki gizli durumu yeni gizli duruma taşıyan ağırlık matrisi.	B: Gizli durum için bias.
U: Giriş verisinin gizli duruma etkisini belirleyen ağırlık matrisi.	C: Çıktı için bias.
V: Gizli durumdan çıktıya dönüşüm için kullanılan ağırlık matrisi.	

### 7.7 Kapılı Tekrarlayan Birim (Gated Recurrent Unit - GRU)

RNN türü olarak, özellikle uzun dizilerle çalışırken yüksek performans göstermektedir (Dey and Salem, 2017). LSTM mimarisine benzer şekilde, zaman serileri gibi sıralı verilerin analizinde etkin rol oynayan GRU, derin öğrenme süreçlerinde sıkça karşılaşılan gradyan patlaması veya sönümlenmesi (vanishing gradient) sorunlarını minimize etmek amacıyla tasarlanmıştır. Mimari açıdan LSTM'e kıyasla daha sade bir yapıya sahip olan GRU, bünyesinde "sıfırlama" (reset) ve "güncelleme" (update) adı verilen iki temel kapı mekanizması barındırır. Sıfırlama kapısı, önceki hücreden gelen hafıza bilgisinin ne kadarının göz ardı edileceğini belirlerken; güncelleme kapısı, geçmiş bilgi birikimi ile mevcut girdi arasındaki dengenin kurulmasını sağlar. Bu kapı mekanizmaları sayesinde GRU hem bellek yönetimi hem de hesaplama maliyeti açısından verimli bir çalışma prensibi

sunar. Söz konusu yapısal avantajlar, GRU'yu bilhassa konuşma tanıma ve zaman serisi tahminleme gibi görevler için oldukça elverişli hale getirmektedir. Ayrıca daha az sayıda parametre içermesi nedeniyle eğitim süreci genellikle daha hızlı tamamlanmakta, bu da onu aşırı karmaşık olmayan modellemelerde öncelikli bir tercih yapmaktadır. Sonuç olarak GRU, LSTM'in sağladığı performans avantajlarının pek çoğunu daha yalın bir mimari ile sunarak geniş bir uygulama alanı için cazip bir seçenek oluşturmaktadır (Sajjad vd., 2020).

### GRU algoritması

- i) **Sıfırlama kapısı (reset gate -  $r_t$ ):** Bu kapı, önceki gizli durumdan ( $h_{t-1}$ ) gelen bilginin ne kadarının unutulacağına (sıfırlanacağına) karar vermektedir. Eğer  $r_t$  0'a yakınsa, önceki bilgiler ihmal edilir ve sadece o anki yeni girdi  $x_t$  ile işlem yapılmaktadır.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (7.36)$$

- $\sigma$ : fonksiyonu sonucu 0 ile 1 arasına sıkıştırmaktadır.

- ii) **Güncelleme kapısı (update gate -  $z_t$ ):** Bu kapı, geçmiş bilgilerin ( $h_{t-1}$ ), ne kadarının geleceğe taşınacağını belirlemektedir. LSTM'deki "unutma" ve "giriş" kapılarının birleşimi gibi çalışır.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (7.37)$$

Sonuç 1'e yakınsa geçmiş bilgi korunur, 0'a yakınsa yeni bilgiye öncelik verilir.

- iii) **Aday gizli durum (candidate hidden state -  $\tilde{h}_t$ ):** Burada, sıfırlama kapısı ( $r_t$ ) devreye girer. Geçmiş bilgi ( $h_{t-1}$ ), sıfırlama kapısı ile çarpılarak ( $z_t * h_t$ ) filtrelenir ve yeni girdi ( $x_t$ ) ile birleştirilerek  $\tanh$  fonksiyonundan geçirilir. Bu, o anki adım için oluşturulan "ham" bilgidir.

$$\tilde{h}_t = \tanh(r_t \cdot [r_t * h_{t-1}, x_t] + b_h) \quad (7.38)$$

- iv) **Final gizli durum (hidden state -  $h_t$ ):** Son aşamada, güncelleme kapısı ( $z_t$ ) kullanılarak "Eski bilgiyi mi tutayım, yoksa yeni hesapladığım aday bilgiyi mi kullanayım?" sorusunu sorarak bir karar verilir:

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (7.39)$$

- $(1 - z_t) * h_{t-1}$ : Geçmişin korunma oranı.
- $z_t * \tilde{h}_t$ : Yeni bilginin eklenme oranı.

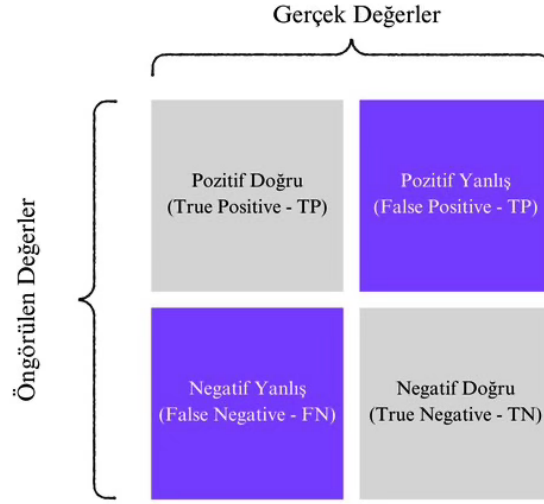
## 7.8 Makine Öğrenmesi Modellerinin Performansları

Bu bölümde, oluşturulan makine öğrenmesi ve derin öğrenme modellerinin test edilme süreci detaylandırılmaktadır. Model test aşaması, geliştirilen modellerin performanslarını değerlendirmek, doğruluklarını ölçmek ve tahmin başarımlarını analiz etmek amacıyla gerçekleştirilmektedir. Fiyat balonlarının tespiti, finansal tahmin problemlerinde model doğruluğunu değerlendirmek için çeşitli metrikler kullanılmıştır. Modelin test edilmesi süreci, test veri kümesi kullanılarak yapılan tahminler ve bu tahminlerin gerçek değerlerle karşılaştırılması esasına dayanmaktadır.

Test aşamasında, her model için önceden ayrılan test veri kümesi kullanılmıştır. Makine öğrenmesi modellerinde, test kümesinde elde edilen tahminler doğrudan sınıflandırma sonuçlarına dönüştürülmektedir. Derin öğrenme modellerinde ise, doğruluk ve hata metrikleri belirli eşik değerine göre değerlendirilmiştir. Modelin başarı düzeyini ölçmek için Accuracy (Doğruluk), F1-Score, Receiver Operating Characteristic – Area Under Curve (ROC-AUC), Precision-Recall (Hassasiyet-Duyarlılık Eğrisi), Brier Skoru ve SHAP analizleri kullanılmıştır. Test aşamasında uygulanan yöntemler aşağıda detaylandırılmıştır.

### 7.8.1 Karmaşıklık matrisi (confusion matrix)

Karmaşıklık matrisi, sınıflandırma modelinin yaptığı tahminlerin dağılımını gösteren tablodur. Modelin doğru ve yanlış tahminlerini Şekil 7.8’de gösterildiği gibi dört kategoriye ayırarak analiz edilmesini sağlamaktadır (Salmi et al., 2023):



**Şekil 7.8:** Karmaşıklık matrisi değer gösterim tablosu

Bu matris sayesinde, yanlış pozitif (FP) ve yanlış negatif (FN) tahminler analiz edilerek modelin hangi tür hatalar yaptığı belirlenmektedir. Örneğin, yüksek FP oranı, modelin yanlış alarm verme eğiliminde olduğunu gösterirken, yüksek FN oranı, modelin balon olaylarını gözden kaçırdığını göstermektedir (Stehman, 1997).

Karmaşıklık matrisi ayrıca dengesiz veri kümelerinde modelin başarılarını daha iyi değerlendirmeye yardımcı olmaktadır. ROC ve Precision-Recall eğrilerinin temelini oluşturmaktadır (Hand and Till, 2001).

### 7.8.2 Doğruluk (Accuracy)

Doğruluk metriği, modelin yaptığı doğru tahminlerin, toplam tahminlere oranını ifade etmektedir. Matematiksel olarak şu şekilde tanımlanmaktadır:

$$\text{Doğruluk} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (7.40)$$

Doğruluk metriği doğrudan karmaşıklık matrisinden (confusion matrix) türetilen temel ölçütlerden biridir (Stehman, 1997). Doğruluk metriği genel başarıyı ölçmek için yaygın olarak kullanılsa da, dengesiz veri kümelerinde yanıltıcı olabilmektedir (Fawcett, 2006).

### 7.8.3 Hassasiyet (Precision)

Hassasiyet, modelin pozitif olarak tahmin ettiği sınıfların ne kadarının gerçekten pozitif olduğunu ölçer:

$$\text{Hassasiyet} = \frac{TP}{(TP + FP)} \quad (7.41)$$

Hassasiyet metriği, özellikle yanlış pozitif tahminlerin maliyetli olduğu durumlarda önemlidir (Sokolova and Lapalme, 2009).

#### 7.8.4 Duyarlılık (Recall)

Duyarlılık, modelin gerçekte pozitif olan verileri ne kadar iyi tespit ettiğini gösterir:

$$\text{Duyarlılık} = \frac{TP}{(TP + FN)} \quad (7.42)$$

Bu metrik, özellikle gözden kaçırılmaması gereken durumlar için önemlidir. Örneğin, fiyat balonlarının yanlış negatif tahmin edilmesi, yatırımcılar için ciddi ekonomik kayıplara yol açabilir (Powers, 2011).

#### 7.8.5 F1-Skoru

F1-Skoru, hassasiyet ve duyarlılık arasında denge kurarak modelin genel performansını değerlendiren ölçüttür:

$$F1 = 2 * \frac{(\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (7.43)$$

Bu metrik, sınıflar arasında dengesizlik olduğunda daha güvenilir performans değerlendirmesi sunar (Van Rijsbergen, 1979).

#### 7.8.8 ROC-AUC ve Precision-Recall Eğrileri

Makine öğrenmesi modellerinde sınıflandırma performansını ölçmek için kullanılan en yaygın yöntemlerden biri ROC (Receiver Operating Characteristic) eğrisi ve AUC (Area Under the Curve) skorudur. Bunlara ek olarak, Precision-Recall eğrisi özellikle dengesiz veri kümelerinde modelin başarısını analiz etmek için önemli araçtır.

- i) **ROC Eğrisi ve AUC Skoru:** ROC eğrisi, yanlış pozitif oranı (FPR) ile doğru pozitif oranı (TPR) arasındaki ilişkiyi gösteren grafikdir:

$$FPR = \frac{FP}{(FP + TN)} \quad (7.44)$$

$$TPR = \frac{TP}{(TP + FN)} \quad (7.45)$$

Bu eğrinin altındaki alan (AUC), modelin sınıflandırma performansının genel göstergesi olarak değerlendirilir. AUC skoru 1'e ne kadar yakınsa, modelin sınıflandırma başarısı o kadar yüksektir (Fawcett, 2006).

- ii) **Precision-Recall Eğrisi:** Precision-Recall eğrisi, özellikle dengesiz veri kümelerinde, modelin gerçek pozitifleri ne kadar doğru tahmin ettiğini ve yanlış pozitifleri nasıl sınırladığını ölçmek için daha doğru analiz sunmaktadır. Eğrinin altındaki alan AP (Average Precision) şeklinde ifade edilmektedir. Modelin balonları doğru tahmin etme kabiliyetini ölçmek için kullanılmaktadır (Saito and Rehmsmeier, 2015). Şu şekilde hesaplanmaktadır:

$$AP = \sum_n (R_n - R_{n-1})P_n \quad (7.46)$$

### 7.8.9 Shapley katkılarına dayalı açıklayıcı yöntem (SHapley Additive Explanations - SHAP)

SHAP analizi, makine öğrenmesi modellerinin tahminlerini açıklamak için kullanılan en güçlü yöntemlerden biridir. Modelin her tahmininde hangi değişkenlerin ne kadar etkili olduğunu belirlemeye yardımcı olmaktadır.

SHAP değerleri, oyun teorisinde kullanılan Shapley değerlerinden türetilmiştir ve modelin tahminlerine hangi girdilerin ne kadar katkıda bulunduğunu hesaplar. SHAP değeri şu şekilde ifade edilmektedir (Lundberg and Lee, 2017):

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|! (|N| - |S| - 1)!}{|N|!} [v(S \cup \{j\}) - v(S)] \quad (7.47)$$

- $j$  :  $j$  'inci değişkenin tahmine katkı miktarıdır.
- $v(S)$ :  $S$  alt kümesiyle yapılan tahmin sonucudur.
- $N$  : Toplam değişken sayısıdır.

SHAP analizi sayesinde modelin hangi deęişkenlere dayalı karar verdięi ve hangi faktörlerin fiyat balonlarının oluşumunu en çok etkiledięi anlaşılabilir (Lundberg and Lee, 2017).

## **7.9 Model Doğrulama ve Seçimi**

Makine öğrenmesi ve derin öğrenme modellerinin başarısını deęerlendirmek ve en iyi performansı sağlayan modeli belirlemek için doğrulama ve seçim süreçleri kritik rol oynamaktadır. Model doğrulama, geliştirilen modellerin genelleme yeteneęini ölçmek için kullanılan yöntemleri kapsarken, model seçimi ise farklı algoritmalar arasında karşılaştırma yaparak en uygun modelin belirlenmesini sağlamaktadır (Hastie et al., 2009).

Bu çalışmada, çapraz doğrulama teknikleri, öznitelik seçimi (feature selection) ve model seçim kriterleri dikkate alınarak en iyi model belirleme süreci gerçekleştirilmiştir.

### **7.9.1 Çapraz doğrulama yöntemleri**

Makine öğrenmesi modellerinin genelleme yeteneęini test etmek için yaygın olarak kullanılan yöntemlerden biri çapraz doğrulama (cross-validation) tekniğidir. Çapraz doğrulama, modeli farklı veri alt kümeleri üzerinde eğiterek ve test ederek aşırı öğrenmeyi önlemeyi amaçlamaktadır (Bishop, 2006). Bundan dolayı modelin aşırı öğrenme gösterip göstermedięi belirlenmiş ve modelin yeni veriler üzerindeki başarısı daha sağlıklı şekilde ölçülmüştür.

Bu çalışmada, k-katlı çapraz doğrulama (k-fold cross-validation) yöntemi kullanılmıştır. K-katlı çapraz doğrulama sürecinde, veri seti eşit büyüklükte k parçaya bölünmekte, her iterasyonda parça test seti olarak ayrılırken geri kalan k-1 parça modelin eğitimi için kullanılmaktadır. Bu süreç k kez tekrarlanarak modelin performansı daha sağlam şekilde ölçülmektedir (Kohavi, 1995).

Çalışmada, makine öğrenmesi modelleri için 5-katlı çapraz doğrulama uygulanmıştır. Derin öğrenme modelleri için ise, eğitim ve doğrulama veri setleri zaman serisi yapısını koruyacak şekilde sıralı olarak bölünmüş ve modelin gelecekteki veriler üzerinde nasıl performans gösterdięi incelenmiştir (Bergmeir and Benítez, 2012).

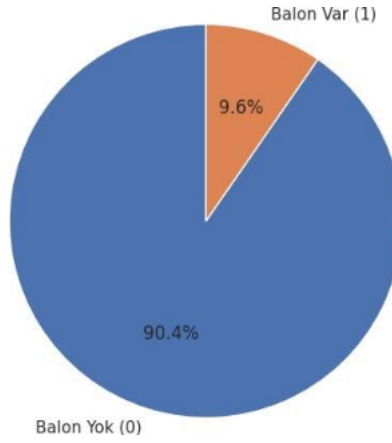
### 7.9.2 Öznitelik Seçimi (Feature Selection)

Öznitelik seçimi, modelin tahmin performansını artırmak ve hesaplama maliyetini azaltmak için gereksiz veya düşük katkı sağlayan değişkenleri elimine etme sürecidir (Guyon and Elisseeff, 2003). Öznitelik mühendisliği sürecinde, veride bulunan tüm değişkenlerin model için eşit derecede önemli olmadığı gözlemlenmiştir. Bazı değişkenlerin modelin öğrenme sürecine katkısı düşük olabilirken, bazıları modelin kararlarını doğrudan etkileyebilir.

Bu çalışmada, SHAP analizi kullanılarak özniteliklerin model üzerindeki etkileri analiz edilmiştir (Lundberg and Lee, 2017). SHAP analizi sayesinde, modelin verdiği kararları hangi özniteliklerin ne kadar etkilediği tespit edilmiş ve en önemli değişkenler belirlenmiştir. Öznitelik seçimi süreci sonunda, modelin açıklanabilirliği artırılmış, eğitim süresi azaltılmış ve gereksiz özniteliklerden arındırılmış modelleme süreci gerçekleştirilmiştir.

## 8. ANALİZ VE BULGULAR

Bu bölümde, GSADF temelli balon tespiti yaklaşımının çıktıları ayrıntılı biçimde incelenmiş, uygulanan makine öğrenmesi ve derin öğrenme modellerinin performansları hem nicel hem de görsel olarak değerlendirilmiştir. İlk olarak, veri setinin yapısal özellikleri ve dengesizlik düzeyi analiz edilmiştir. Ardından sınıflandırma sonuçları, SHAP değerleri ve doğruluk metrikleri temelinde ayrıntılı model değerlendirmeleri yapılmıştır. Son olarak, elde edilen bulgular istatistiksel testlerle yorumlanarak, modellerin güçlü ve zayıf yönleri tartışılmıştır.



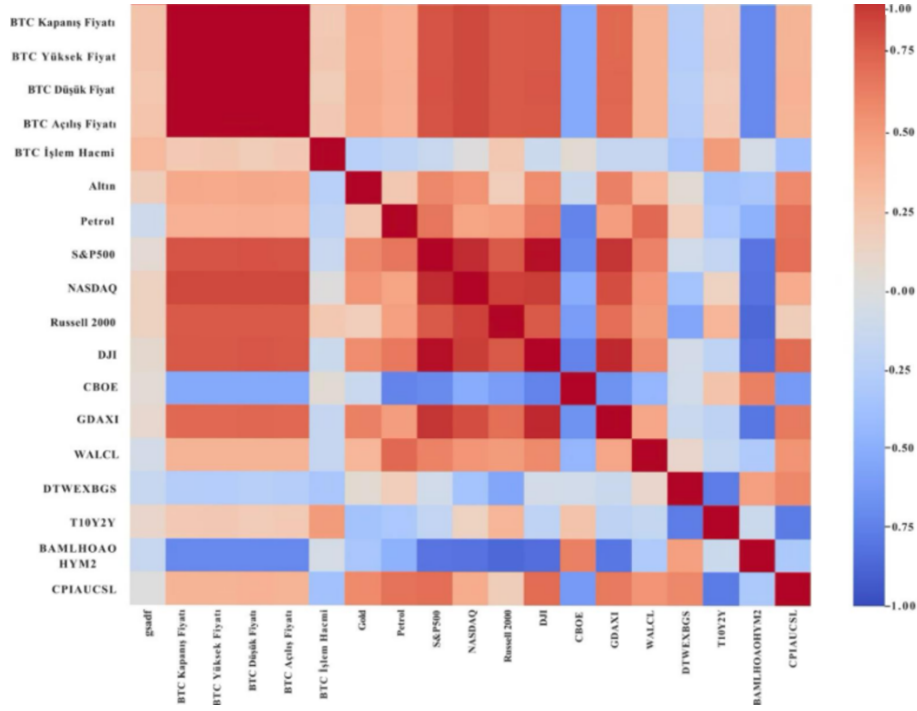
**Şekil 8.1:** Eviews sonuç dağılım grafiği

Şekil 8.1’de, GSADF testi ile elde edilen hedef değişkenin sınıf dağılımı pasta grafiği olarak gösterilmektedir. Grafik, gözlemlerin yaklaşık %90.4’inin “balon yok” (0), yalnızca %9.6’inin “balon var” (1) sınıfına ait olduğunu ortaya koymaktadır. Bu yapı, pozitif sınıfın oldukça nadir olduğu belirgin sınıf dengesizliğine işaret etmektedir. Dolayısıyla, çalışmada yalnızca doğruluk oranına odaklanmak yanıltıcı olacağından dolayı F1 skoru ve PR-AUC gibi dengesiz veri setlerine duyarlı ölçütler tercih edilmiş ve modellerde sınıf ağırlıkları kullanılarak azınlık sınıfına duyarlılık artırılmaya çalışılmıştır.



**Şekil 8.2:** Kripto varlık piyasalarında GSADF testi ile balon tespit sonuçları

Şekil 8.2’de zaman serisine uygulanan GSADF testinin çıktısı yer almaktadır. Grafikteki kahverengi çizgi GSADF değerini, kırmızı kesikli çizgi ise kritik eşiği (1.95) göstermektedir. GSADF bu eşiğin üzerine çıktığında, serinin “normal” seyrinden çıkarak kısa bir süre içinde aşırı hızlandığı ve balon benzeri bir davranış sergileyebileceği anlaşılmaktadır. Bu dönemler pembe alanlarla “balon sinyali” olarak işaretlenmiştir. Grafikte özellikle 2021 başında belirgin ve güçlü bir sinyal görülürken, 2024 başında daha kısa süreli ama dikkat çeken eşik aşımı bulunmaktadır. Eşik altında kalan dönemler ise serinin daha dengeli/normal bir yapıda ilerlediğini göstermektedir. Şekil 8.2 balon dönemlerinin hangi zaman aralıklarında yoğunlaştığını netleştirerek, çalışmanın ilerleyen adımlarında yapılacak balon–normal etiketleme ve modelleme süreçlerine temel oluşturmaktadır.



**Şekil 8.3:** Korelasyon ısı haritası

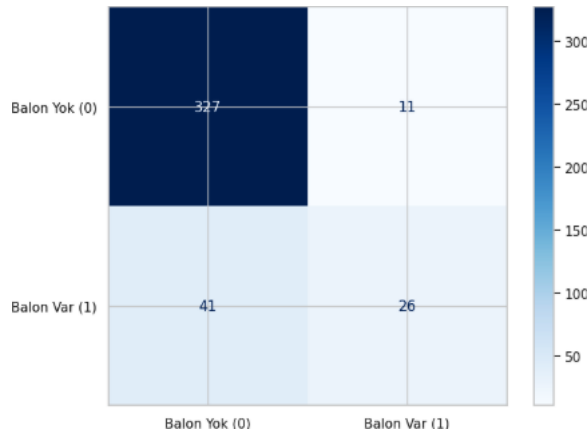
Şekil 8.3'te verilen korelasyon ısı haritası, çalışmadaki değişkenlerin birbirleriyle nasıl etkileşim içinde olduğunu net şekilde ortaya koymaktadır. Haritaya ilk bakışta, Bitcoin'in kendi fiyat kalemleri (Açılış, Kapanış, En Yüksek, En Düşük) arasında beklendiği üzere kusursuza yakın pozitif ilişki olduğu görülmektedir. Ancak bu durum, söz konusu verilerin aslında aynı bilgiyi tekrarladığını ve modelde teknik karmaşaya (çoklu doğrusal bağlantı) yol açmamak için dikkatli olunması gerektiğini göstermektedir. Ayrıca GSADF balon göstergesinin fiyatla paralel hareket etmesi, balonların genellikle piyasanın coşkulu olduğu yükseliş trendlerinde şiştiğini doğrulamaktadır.

En çarpıcı bulgulardan biri Bitcoin'in geleneksel piyasalarla olan şeklidir. Harita, Bitcoin'in S&P500 ve NASDAQ gibi küresel borsa endeksleriyle şaşkıncı derecede uyumlu hareket ettiğini göstermektedir. Bu durum, incelenen dönemde Bitcoin'in piyasadan bağımsız "güvenli liman" olmaktan ziyade, küresel risk iştahına duyarlı varlık gibi davrandığını kanıtlar niteliktedir. Nitekim piyasadaki korkuyu ölçen CBOE endeksi ile görülen ters yönlü (negatif) ilişki de bunu desteklemektedir. Piyasalarda belirsizlik ve korku arttığında yatırımcılar riskli gördükleri kripto varlıklardan kaçmaktadır.

Buna karşın, enflasyon ve dolar endeksi gibi makroekonomik verilerin fiyatlarla ilişkisi daha zayıf ve karmaşıktır. Ancak bu "düşük korelasyon" zayıflık değil, aksine bu değişkenlerin

fiyattan bağımsız, modele "farklı pencereden" ek bilgi sağladığının işaretidir. Benzer şekilde işlem hacminin de fiyattan ayrıışan yapısı, hacim verisinin kendine has, özgün sinyal taşıdığını düşündürmektedir.

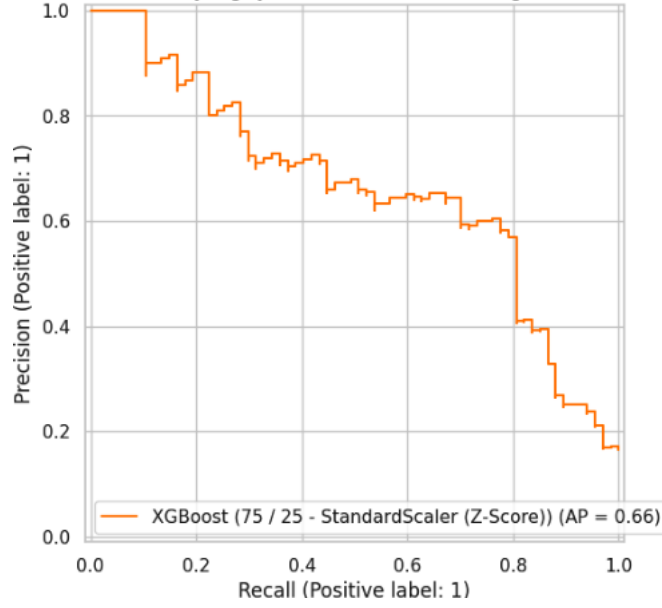
Şekil 8.3'te, ağaç tabanlı yöntemler arasında en iyi performansı gösteren XGBoost modelinin karmaşıklık matrisi sunulmaktadır (75/25 zaman temelli bölünme ve z-skor kullanımı ile). Gerçek sınıfı "balon yok (0)" olan 327 gözlem doğru sınıflandırılırken, 11 gözlem yanlışlıkla "balon var (1)" olarak etiketlenmiştir. Buna karşılık, gerçek sınıfı "balon var (1)" olan gözlemlerin 26'sı doğru, 41'i ise kaçırılarak "balon yok (0)" şeklinde tahmin edilmiştir.



**Şekil 8.4:** En iyi ağaç modeli karmaşıklık matrisi (XGBoost)

Bu sonuçlar, XGBoost'un negatif sınıfı çok başarılı biçimde ayırt edebildiğini, ancak balon dönemlerini temsil eden pozitif sınıfta önemli sayıda yanlış negatif ürettiğini göstermektedir. Kriz veya balon dönemlerini kaçırmamanın maliyetinin yüksek olduğu finansal uygulamalarda bu durum, modelin daha yüksek hatırlama (recall) için iyileştirilmesi gerektiğine işaret etmektedir.

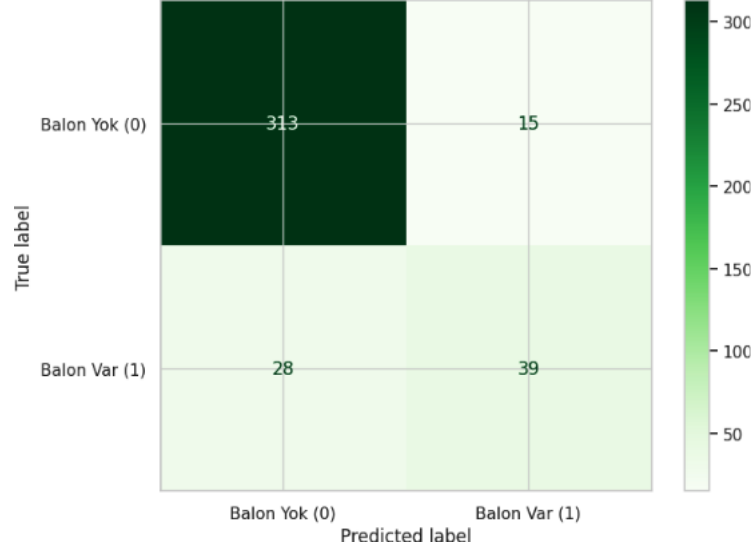
Şekil 8.4'te, XGBoost modeli için elde edilen hassasiyet – duyarlılık eğrisi yer almakta ve eğri altında kalan alan yaklaşık 0.66 (AP = 0.66) olarak görülmektedir. Düşük duyarlılık değerlerinde hassasiyetin 0.9 seviyelerine kadar çıktığı, duyarlılık arttıkça hassasiyetin kademeli olarak azaldığı izlenmektedir.



**Şekil 8.5:** En iyi ağaç modeli hassasiyet – duyarlılık eğrisi (XGBoost)

Şekil 8.5’teki eğri, modelin daha katı eşik değerleri ile tahmin yaptığıında “balon var” dediği dönemlerin büyük kısmının gerçekten balon olduğunu göstermektedir. Ancak bu durumda sakıncalı biçimde birçok balon dönemini hiç işaretlemediğini göstermektedir. Eşik gevşetildikçe daha fazla balon dönemi yakalanmakta fakat yanlış pozitifler artarak hassasiyet düşmektedir. Bu davranış, sınıf dengesizliğinin yoğun olduğu veri setlerinde ağaç tabanlı modellerin dikkatli eşik ayarı gerektirdiğini göstermektedir.

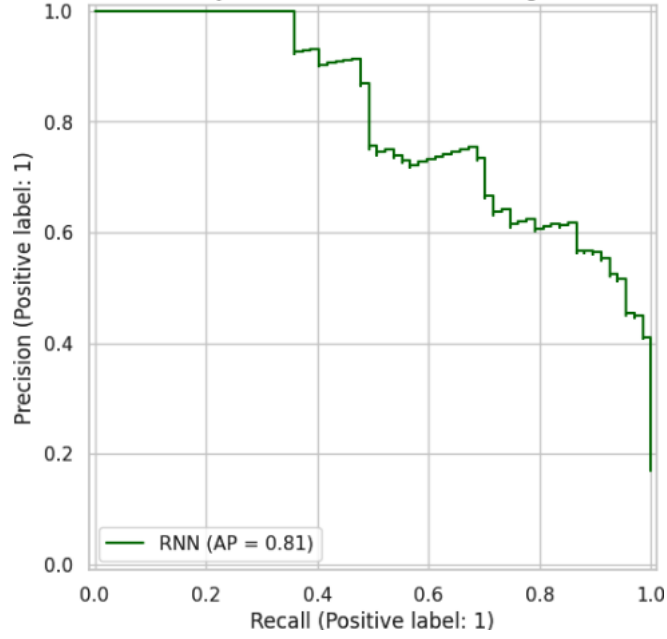
Şekil 8.6’da derin öğrenme modelleri içerisinde en iyi sonucu veren RNN modeline ait karmaşıklık matrisi gösterilmektedir. Gerçek sınıfı “balon yok (0)” olan 313 gözlem doğru sınıflandırılırken, 15 gözlem yanlışlıkla “balon var (1)” şeklinde tahmin edilmiştir. Pozitif sınıf açısından bakıldığında, “balon var (1)” olan gözlemlerin 39’u doğru, 28’i ise “balon yok (0)” olarak yanlış sınıflandırılmıştır.



**Şekil 8.6:** En iyi DL modeli karmaşıklık matrisi (RNN)

XGBoost ile karşılaştırıldığında, RNN modelinin pozitif sınıftaki doğru tahmin sayısını (TP) artırdığı, yanlış negatif sayısını ise belirgin biçimde azalttığı görülmektedir (XGBoost: 26 TP / 41 FN; RNN: 39 TP / 28 FN). Diğer ifadeyle, RNN modeli hem daha yüksek duyarlılık hem de benzer hatta biraz daha yüksek hassasiyet sağlayarak, balon dönemlerini yakalama konusunda ağaç tabanlı modele göre üstün performans sergilemektedir. Bu bulgu, zaman bağımlılığını doğrudan modelleyebilen derin öğrenme yöntemlerinin kripto fiyat balonlarının tespitinde avantaj sağladığını desteklemektedir.

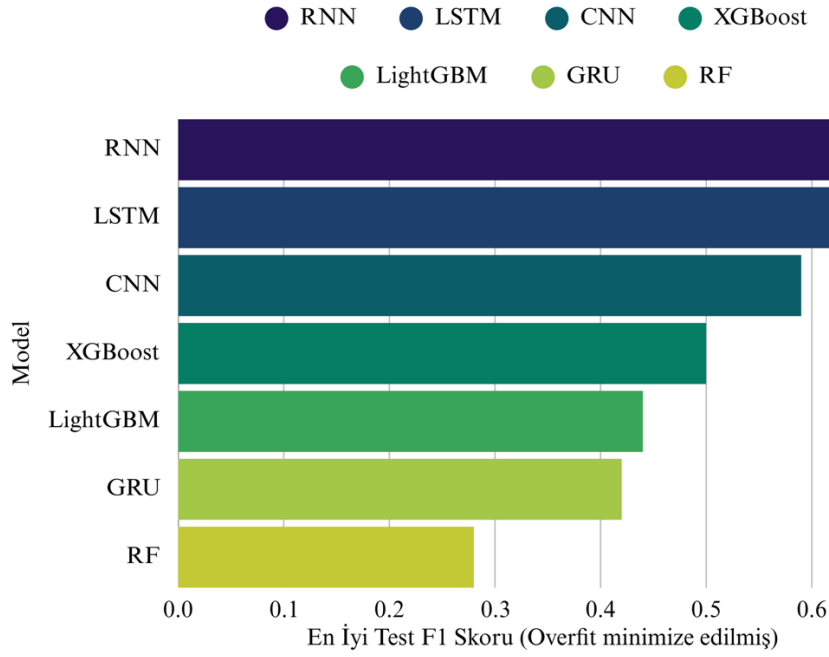
Şekil 8.7’de, RNN modeli için çizilen hassasiyet–duyarlılık eğrisi görülmekte ve eğri altındaki alan 0.81 ( $AP = 0.81$ ) olarak hesaplanmaktadır. Eğri, düşük duyarlılık seviyelerinde hassasiyetin 1’e çok yakın seyrettiğini gözlemlenmektedir. Duyarlılık 0.4-0.7 aralığında iken hassasiyetin hâlâ yaklaşık 0.7 düzeyinde kaldığını göstermektedir.



**Şekil 8.7:** En iyi DL modeli hassasiyet - duyarlılık eğrisi (RNN)

XGBoost'un hassasiyet - duyarlılık eğrisi ile karşılaştırıldığında (bkz. Şekil 8.4), RNN modelinin hem daha büyük alan ürettiği hem de geniş duyarlılık aralığında daha yüksek hassasiyet sağladığı görülmektedir. Bu sonuç, RNN'in "balon" sınıfını daha istikrarlı şekilde ayırt edebilmektedir. Ayrıca bu durum dengesiz veri yapısına rağmen balon dönemleri tespitinde daha güvenilir sinyal ürettiğini göstermektedir.

Şekil 8.7'de, çalışmada kullanılan tüm modellerin en yüksek Test F1 skorları yatay bar grafiği ile karşılaştırmalı olarak sunulmuştur. Grafikte, RNN modelinin en yüksek F1 skoruna ulaştığı, onu çok yakın performansla LSTM modelinin izlediği görülmektedir. CNN modeli de derin öğrenme grubunda tatmin edici performans sergileyerek, ağaç tabanlı modellere (XGBoost, LightGBM, RandomForest) göre daha yüksek F1 skorları üretmiştir.



**Şekil 8.8:** Modellerin en iyi test F1 skorlarına göre karşılaştırılması

Ağaç temelli yöntemler arasında XGBoost, LightGBM ve RF'e göre daha iyi sonuç vermekle birlikte, genel sıralamada derin öğrenme modellerinin gerisinde kalmaktadır. Şekil 8.8'de görülen bu karşılaştırma, kripto varlık fiyat balonlarının tahmininde zaman serisi dinamiklerini doğrudan modelleyebilen derin öğrenme mimarilerinin (özellikle RNN ve LSTM) klasik ağaç tabanlı yöntemlere göre daha yüksek sınıflandırma başarısı sağladığını nicel olarak ortaya koymaktadır.

**Tablo 8.1:** Modellerin farklı bölümleri ve ölçekleyicilere göre performans karşılaştırması

Senaryo	Ölçekleyici	Model	Test F1	Train F1	Overfit Farkı	Test PR-AUC	Test Precision	Test Recall
70 / 30	Min-Max	RNN	0,70588	0,58585	-0,12002	0,79780	0,69565	0,71641
70 / 30	Min-Max	CNN	0,68181	0,56493	-0,11688	0,78138	0,69230	0,67164
75 / 25	Min-Max	CNN	0,67200	0,47910	-0,19289	0,77030	0,72413	0,62686
75 / 25	Min-Max	LSTM	0,66165	0,53538	-0,12626	0,71755	0,66666	0,65671
70 / 30	Min-Max	LSTM	0,65185	0,49710	-0,15474	0,75795	0,64705	0,65671
80 / 20	Min-Max	RNN	0,60344	0,64393	0,04049	0,72422	0,71428	0,52238
80 / 20	Z-Skor	RNN	0,56000	0,45945	-0,10054	0,67838	0,84848	0,41791
70 / 30	Min-Max	XGBoost	0,55445	1	0,44554	0,76523	0,82352	0,41791
70 / 30	Z-Skor	XGBoost	0,54545	1	0,45454	0,75380	0,84375	0,40298
80 / 20	Min-Max	CNN	0,54368	0,54421	0,00052	0,75528	0,77777	0,41791

**Tablo 8.1 (devam)**

75 / 25	Min-Max	RNN	0,53846	0,48048	-0,05798	0,74554	0,75675	0,41791
80 / 20	Z-Skor	XGBoost	0,51428	0,91578	0,40150	0,70121	0,71052	0,40298
80 / 20	Min-Max	XGBoost	0,51428	0,91578	0,40150	0,70138	0,71052	0,40298
80 / 20	Z-Skor	CNN	0,5	0,49068	-0,00931	0,61993	0,49275	0,50746
75 / 25	Z-Skor	XGBoost	0,5	0,91578	0,41578	0,66088	0,70270	0,38805
75 / 25	Min-Max	XGBoost	0,5	0,91578	0,41578	0,66088	0,70270	0,38805
80 / 20	Z-Skor	LSTM	0,49090	0,5	0,00909	0,63600	0,62790	0,40298
70 / 30	Min-Max	LightGBM	0,48351	1	0,51648	0,64308	0,91666	0,32835
80 / 20	Min-Max	LSTM	0,47706	0,52117	0,044108	0,67670	0,61904	0,38805
75 / 25	Z-Skor	RNN	0,473118	0,54952	0,07640	0,73304	0,84615	0,32835
70 / 30	Z-Skor	CNN	0,45098	0,62548	0,17450	0,63372	0,65714	0,34328
75 / 25	Z-Skor	LightGBM	0,44680	1	0,55319	0,63119	0,77777	0,31343
70 / 30	Z-Skor	RNN	0,43478	0,56375	0,12897	0,58846	0,52083	0,37313
70 / 30	Z-Skor	LightGBM	0,43181	1	0,56818	0,64383	0,90476	0,28358
80 / 20	Z-Skor	GRU	0,425532	0,757991	0,332459	0,575674	0,740741	0,298507
70 / 30	Min-Max	GRU	0,424870	0,412698	-0,012172	0,586525	0,325397	0,611940
75 / 25	Z-Skor	CNN	0,40000	0,59712	0,19712	0,58710	0,60606	0,29850
75 / 25	Min-Max	LightGBM	0,3956	0,99428	0,59868	0,64101	0,75	0,26865
80 / 20	Z-Skor	LightGBM	0,38297	1	0,61702	0,63282	0,66666	0,26865
70 / 30	Z-Skor	LSTM	0,38000	0,64591	0,26591	0,56671	0,57575	0,28358
80 / 20	Min-Max	GRU	0,379845	0,402844	0,022999	0,566915	0,256545	0,731343
80 / 20	Min-Max	LightGBM	0,36956	1	0,63043	0,61523	0,68000	0,25373
75 / 25	Min-Max	GRU	0,363636	0,343434	-0,020202	0,572402	0,240385	0,746269
75 / 25	Z-Skor	LSTM	0,35714	0,65151	0,29437	0,55878	0,88235	0,22388
75 / 25	Min-Max	RF	0,33707	0,90526	0,56818	0,54082	0,68181	0,22388
75 / 25	Min-Max	RF	0,33707	0,90526	0,56818	0,54082	0,68181	0,22388
70 / 30	Z-Skor	RF	0,33707	0,91978	0,58270	0,56822	0,68181	0,22388
70 / 30	Min-Max	RF	0,33707	0,91978	0,58270	0,56966	0,68181	0,22388
70 / 30	Z-Skor	GRU	0,314607	0,790698	0,476091	0,493516	0,636364	0,208955
75 / 25	Z-Skor	GRU	0,309524	0,798122	0,488598	0,549960	0,764706	0,194030
75 / 25	Z-Skor	RF	0,27906	0,90526	0,62619	0,53551	0,63157	0,17910
80 / 20	Z-Skor	RF	0,19753	0,91978	0,72225	0,54470	0,57142	0,11940
80 / 20	Min-Max	RF	0,19753	0,91978	0,72225	0,54701	0,57142	0,11940

Tablo 8.1, aynı veri üzerinde 70/30, 75/25, 80/20 veri bölme oranları, Min-Max ve Z-Skor ölçekleme yöntemleri ve farklı modellerin (RNN, CNN, LSTM, GRU, RF) nasıl sonuç verdiğini bütüncül olarak görmemizi sağlamaktadır. Tablo 8.1’de özellikle Test F1 metriğine odaklanmak anlamlıdır. Çünkü F1 skoru, “model ne kadar doğru tahmin yaptı?” sorusunu tek başına hassasiyet ya da duyarlılığa bakmadan, ikisini birlikte dengeleyerek değerlendirmektedir. Tabloya göre en iyi sonuç 70/30 bölünme Min-Max ölçeklendirici + RNN kombinasyonunda elde edilmiştir (Test F1  $\approx$  0.706). Bu satırın öne çıkmasının nedeni sadece test başarısının yüksek olması değildir. Aynı zamanda eğitim ve test sonuçları arasında aşırı kopma olmamasıyla modelin genelleme yapabildiğini göstermesidir. Genel

eğilime bakıldığında, derin öğrenme tabanlı modellerde (RNN/CNN/LSTM/GRU) Min-Max ölçeklemenin daha tutarlı sonuçlar ürettiği görülmektedir. Üst sıraların önemli kısmı Min-Max kullanılan senaryolardan oluşmaktadır.

Buna karşılık, bazı Z-Skor senaryolarında özellikle derin öğrenme modellerinin Train F1 değerinin 1.00'a kadar çıkmasına rağmen Test F1'in düşük kalması, modelin eğitim verisini "fazla iyi" öğrenip testte aynı başarıyı gösteremediğini düşündürür. Bu durum pratikte çoğu zaman aşırı öğrenme olarak yorumlanmaktadır. Model eğitim verisini ezberler, fakat gerçek hayatta performans düşer. GRU modeli bazı denemelerde duyarlılık tarafında iyi görünebilse de hassasiyet ile duyarlılık dengesi her zaman korunamadığı için toplam başarıyı gösteren Test F1 açısından genelde orta-alt seviyelerde kalmıştır. RF gibi ağaç tabanlı yöntemler ise bazı metriklerde fena görünmese de zaman serisinin ardışık yapısını derin öğrenme kadar iyi yakalayamadığından Tablo 8.1'de çoğunlukla alt sıralarda yer almıştır.

## 9. SONUÇ VE ÖNERİLER

Bu tez çalışması, finansal piyasalarda volatilité ve spekülasyon ile özdeşleşen Bitcoin'in fiyat hareketlerindeki balon oluşumlarının tespitini ve öngörüsünü hedeflemiştir. Araştırma metodolojisi, Phillips, Shi ve Yu (2015) tarafından geliştirilen GSADF testi ile belirlenen balon dönemlerinin, denetimli öğrenme algoritmaları için birer hedef değişken olarak kullanılması prensibine dayanmaktadır. Çalışma 1 Ocak 2020 – 5 Haziran 2024 dönemine ait günlük veriler üzerinden gerçekleştirilmiştir. RF, XGBoost, LightGBM, LSTM, CNN, RNN ve GRU modellerinin tahmin performansları karşılaştırmalı olarak analiz edilmiştir. Elde edilen ampirik bulgular ve literatür bağlamında geliştirilen öneriler aşağıda sunulmaktadır. Ayrıca ekonometrik analiz sonuçları, kripto varlık piyasalarının irrasyonel fiyat davranışlarına ve spekülatif balon oluşumlarına yatkınlığını istatistiksel olarak doğrulamaktadır. GSADF testi ile tespit edilen çoklu ve patlayıcı balon dönemleri Cheung, Roca ve Su (2015) ile Corbet, Lucey ve Yarovaya (2018) çalışmalarındaki "kripto piyasalarındaki döngüsel ve patlayıcı fiyat hareketleri" bulgusuyla paralellik göstermektedir. Bu çalışmanın literatüre özgün katkısı, ekonometrik test yöntemi olan GSADF'nin makine öğrenmesi algoritmalarına entegre edilerek, balon tespitinin sınıflandırma problemi olarak başarılı şekilde modellenmesidir.

Modelleme aşamasındaki en belirgin bulgu, Tekrarlayan Sinir Ağları (RNN) modelinin sergilediği üstün performanstır. RNN modeli, %70,5 Test F1 skoru ve 0,797 PR-AUC değeri ile diğer tüm modelleri geride bırakmıştır. Bu sonuç, McNally (2018) ve Di Persio ve Honchar (2017) tarafından da vurgulandığı üzere, finansal zaman serilerinin doğrusal olmayan, ardışık ve stokastik yapısının modellenmesinde derin öğrenme mimarilerinin, özellikle de geçmiş bilgiyi koruma yeteneğine sahip "hücre durumu" (cell state) yapısıyla RNN ve türevlerinin, geleneksel yöntemlere göre daha başarılı olduğunu teyit etmektedir. Buna karşılık, ağaç tabanlı modeller (RF, XGBoost, LightGBM) eğitim veri setinde yüksek başarı göstermelerine rağmen test veri setinde performans düşüşü yaşamışlardır. Bu durum söz konusu modellerin zaman serisindeki yapısal kırılmaları ve volatilité şoklarını genelleştirmekte zorlandığını ve aşırı öğrenme eğiliminde olduğunu göstermektedir. RNN modelinin özellikle ATR ve RSI gibi teknik göstergelerle balonları başarılı şekilde öngörebilmesi, Fry ve Cheah (2016) tarafından belirtilen "spekülatif şoklara açık piyasa yapısı" tezini desteklemekte ve teknik analizin makine öğrenmesi ile optimize edilerek öngörü gücünün artırılabilirliğini ortaya koymaktadır.

Bu bağlamda çalışma, elde edilen bulgular ışığında üç temel alanda somut katkı sağlamaktadır:

- I. Geliştirilen RNN modeli, yatırımcılar için spekülasyon balon dönemlerinde yanlış negatif sinyalleri minimize ederek sermaye kaybını önleyen dinamik risk yönetim aracı sunmaktadır.
- II. Düzenleyici kurumlar için piyasa manipülasyonlarını ve aşırı oynaklığı tespit edebilen, aşırı öğrenme riski düşük, gerçek zamanlı piyasa gözetim mekanizması önermektedir.
- III. Akademik literatür açısından finansal zaman serilerindeki yapısal kırılmaların ve sınıf dengesizliğinin yönetiminde, ağaç tabanlı modeller yerine ardışık bağımlılıkları öğrenen RNN tabanlı mimarilerin genelleştirme yeteneğinin üstünlüğünü ampirik olarak kanıtlamaktadır.

Ancak bu çalışmanın bazı kısıtları bulunmaktadır. Çalışma, veri seti olarak günlük frekanstaki verileri kullanmış olup, gün içi yüksek frekanslı (saatlik veya dakikalık) volatilitenin oluşturduğu mikro balonları kapsam dışı bırakmıştır. Ayrıca analiz, baskın kripto varlık olan Bitcoin ile sınırlandırılmıştır. Altcoin piyasalarının farklı dinamikleri çalışmaya dahil edilmemiştir.

Bu kısıtlar ve çalışmanın sağladığı katkılar göz önüne alındığında, gelecek çalışmalar için şu öneriler sunulabilmektedir. İlk olarak yatırımcılar için geliştirme olarak modelin sunduğu dinamik risk yönetimi avantajı, gelecekteki çalışmalarda otomatik alım-satım botlarına entegre edilerek, yüksek frekanslı verilerle (HFT) test edilebilir ve portföy optimizasyonundaki etkisi ölçülebilir. İkinci olarak düzenleyici kurumlar için genişletme de önerilen gözetim mekanizması, sadece Bitcoin özelinde değil, manipülasyona daha açık olan DeFi tokenleri ve NFT piyasalarına da uyarlanarak düzenleyici kurumların daha geniş piyasa hakimiyeti kurmasına katkı sağlanabilir. Son olarak akademik literatür için derinleşme, RNN mimarisinin kanıtlanan başarısı üzerine inşa edilerek, gelecek araştırmalarda Transformer tabanlı modellerin veya Hibrit (CNN-LSTM) yapıların, sosyal medya duyarlılık verileri (sentiment analysis) ile zenginleştirilerek kullanılması, tahmin başarımını daha da artırabilir.

## 10. KAYNAKLAR

- Abreu, D. and Brunnermeier, K. M.** (2002). Synchronization risk and delayed arbitrage. *Journal of Financial Economics* 66 341–360
- Aghalibayli, N.** (2019). Bitcoin as a cryptocurrency and its relationship with gold, crude oil and Euro exchange rate (Master's thesis, Marmara Universitesi (Turkey)).
- Aishwarya.27.** (2024). Introduction to recurrent neural networks. GeeksforGeeks. <https://www.geeksforgeeks.org/introduction-to-recurrent-neural-network/> (Eriřim Tarihi:12 řubat 2025).
- Akgül, A., řahin, E. E., ve řenol, F. Y.** (2022). Blockchain-based cryptocurrency price prediction with chaos theory, onchain analysis, sentiment analysis and fundamental-technical analysis. *Chaos Theory and Applications*, 4(3), 157-168.
- Akıncı, M., Akıncı, G. Y., ve Yılmaz, Ö.** (2014). Lale ılgınlığı'ndan Mortgage Krizi'ne spekülâtif balonlar. *Tarih Okulu Dergisi*, 7(9), 719-749.
- Akkaya, M.** (2022). Kripto para işlemlerini yönlendiren ekonomik ve finansal faktörler: Bitcoin fiyat oluşumu. *TESAM Akademi Dergisi*, 9(1), 209-226.
- Akkuş, H. T.** (2021). Housing price bubbles and factors affecting the formation of bubbles: The Turkish case. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (46), 271-292.
- Akkuş, H. T., Gursoy, S., Dogan, M., ve Demir, A. B.** (2022). Metaverse and metaverse cryptocurrencies (meta coins): bubbles or future?. *Journal of Economics Finance and Accounting*, 9(1), 22-29.
- Akshisaxena.** (2024). CatBoost in Machine Learning. <https://www.geeksforgeeks.org/catboost-ml/> (Eriřim Tarihi: 8 Ocak 2025)
- Aleksejeva, V., Lavrinenko, O., Betlej, A., and Danileviča, A.** (2021). Analysis Of disparities in the use of information and communication technology (Ict) in the eu countries. *Entrepreneurship And Sustainability Issues* 9 (2), 332-345
- Ali, M., Nelson, J., Shea, R., and Freedman, M. J.** (2016). Blockstack: A global naming and storage system secured by blockchains. In 2016 USENIX annual technical conference (USENIX ATC 16) (pp. 181-194).
- Anavatan, A. ve Kayacan, Y. E.** (2018). BİST 100 Endeksinde Balon Etkisinin İncelenmesi. *Avrasya Sosyal ve Ekonomik Arařtırmalar Dergisi* 5 (8) 124- 131.
- Ante, L.** (2022). The non-fungible token (NFT) market and its relationship with Bitcoin and Ethereum. *FinTech*, 1(3), 216-224.

## **KAYNAKLAR DİZİNİ (devam)**

- Baker, M., and Wurgler, J. (2007).** Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2), 129-151.
- Balı, S., ve Büyüksalvarcı, A. (2011).** 1630'dan 2010'a finansal krizler tarihi: Balonlar, panikler, buhranlar ve küresel finansal krizi. Çatı Kitapları.
- Baur, D. G., Hong, K., and Lee, A. D. (2018).** Bitcoin: Medium of exchange or speculative assets?. *Journal of International Financial Markets, Institutions and Money*, 54, 177-189.
- Beck, R., Müller-Bloch, C., and King, J. L. (2018).** Governance in the blockchain economy: A framework and research agenda. *Journal of the association for information systems*, 19(10), 1.
- Bergmeir, C., and Benítez, J. M. (2012).** On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 102-120. <https://doi.org/10.1016/j.ins.2011.12.028>
- Bergstra, J., and Bengio, Y. (2012).** Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(1), 281-305.
- Bishop, C. M. (2006).** *Pattern recognition and machine learning*. Springer.
- Blanchard, O. J., and Watson, M. W. (1982).** Bubbles, rational expectations and financial markets.
- Blockchain Türkiye Platformu. (2019).** Dünyada Blokzinciri Regülasyonları ve Uygulama Örnekleri Karşılaştırma Raporu. Hukuk, Düzenlemeler ve Kamu İlişkileri Çalışma Grubu.
- Bouoiyour, J., and Selmi, R. (2015).** What does Bitcoin look like?. *Annals of Economics & Finance*, 16(2).
- Bouri, E., Gupta, R., and Roubaud, D. (2019).** Herding behaviour in cryptocurrencies. *Finance Research Letters*, 29, 216–221. DOI: 10.1016/j.frl.2018.07.008
- Bouri, E., Molnár, P., Azzi, G., Roubaud, D., and Hagfors, L. I. (2017).** On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier? *Finance Research Letters*, 20, 192–198. <https://doi.org/10.1016/j.frl.2016.09.025>
- Breiman, L. (2001).** Random forests. *Machine learning*, 45, 5-32.
- Buluş, A., ve Kabaklarlı, E. (2010).** 1929 Ekonomik buhranı ile son dönem global krizinin karşılaştırılması.
- Caetano, R. (2015).** *Learning Bitcoin*. Packt Publishing Ltd.

## **KAYNAKLAR DİZİNİ (devam)**

- Canbaz, M. F., ve Berkun, G.** (2021). Kripto Para Teknolojisi ve İslami Finans Açısından Meşruiyeti. *Uluslararası Bankacılık, Ekonomi ve Yönetim Araştırmalar Dergisi*, 4(1), 103.
- Carlson, D.** (2024). Financial regulators: The market police. *Encyclopedia Britannica*. <https://www.britannica.com/money/financial-market-regulators> (Erişim Tarihi: 26 Ekim 2024)
- Caspi, I.** (2017). Rtdaf: Testing for bubbles with EViews. *Journal of Statistical Software*, 81(Code Snippet 1). <https://doi.org/10.18637/jss.v081.c01>
- Cheah, E. T., and Fry, J.** (2015). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics letters*, 130, 32-36.
- Chen, C. Y.-H., and Hafner, C. M.** (2019). Sentiment-induced bubbles in the cryptocurrency market. *Journal of Risk and Financial Management*, 12(2), 53. DOI: 10.3390/jrfm12020053
- Chen, T., and Guestrin, C.** (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794).
- Chen, Y., and Bellavitis, C.** (2020). Blockchain disruption and decentralized finance: The rise of decentralized business models. *Journal of Business Venturing Insights*, 13, e00151.
- Cheng, R., Wu, N., Chen, S., and Han, B.** (2022). Will metaverse be nextg internet? vision, hype, and reality. *IEEE network*, 36(5), 197-204.
- Cheung, A., Roca, E., and Su, J. J.** (2015). Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices. *Applied Economics*, 47(23), 2348-2358.
- Chohan, U. W.** (2021a). Decentralized finance (DeFi): An emergent alternative financial architecture. *Critical Blockchain Research Initiative (CBRI) Working Papers*.
- Chohan, U. W.** (2021b). A History of Dogecoin. *Discussion Paper Series: Notes on the 21st Century*. 1-11.
- Chohan, U. W.** (2021c). Non-Fungible Tokens (NFTs): Early Thoughts and a Research Agenda. *Critical Blockchain Research Initiative (CBRI) Working Papers*.
- Christin, N.** (2012). *Traveling the Silk Road: A Measurement Analysis of A Large Anonymous Online Marketplace*. Carnegie Mellon University Pittsburgh, 3.

## KAYNAKLAR DİZİNİ (devam)

- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y.** (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.1412.3555>
- Corbet, S., Lucey, B., and Yarovaya, L.** (2018). Datestamping the Bitcoin and Ethereum bubbles. *Finance Research Letters*, 26, 81-88.
- Corbet, S., Lucey, B., Urquhart, A., and Yarovaya, L.** (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, 62, 182-199.
- Cui, C., He, M., Di, F., Lu, Y., Dai, Y., and Lv, F.** (2020). Research on power load forecasting method based on LSTM model. In 2020 IEEE 5th information technology and mechatronics engineering conference (ITOEC) (pp. 1657-1660). IEEE.
- Çarkacıoğlu, A.** (2016). Bitcoin üzerine bir değerlendirme. *Muhasebe ve Finansman Dergisi*, 69, 115–126. <https://doi.org/10.25095/mufad.396605>
- Dale, R.** (2014). *The first crash: lessons from the South Sea Bubble*.
- De Filippi, P., and Wright, A.** (2018). *Blockchain and the law: The rule of code*. Harvard University Press.
- Demir, F., Karabıyık, A., ve Ayhan Küçük, E. E.** (2008). ABD Mortgage Krizi, BDDK Çalışma Tebliği Sayı 3.
- Dey, R., and Salem, F. M.** (2017). Gate-variants of gated recurrent unit (GRU) neural networks. In 2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS) (pp. 1597-1600). IEEE.
- Dubey, A.** (2018). Feature selection using random forest: The wisdom of crowds. *Towards Data Science*. <https://towardsdatascience.com/feature-selection-using-random-forest-26d7b747597f> (Erişim Tarihi: 25.12.2024).
- Eğilmez, M.** (2023). Amerikan Bankacılık Sisteminde Yaşanan Sorunlar ve Türkiye'ye Olası Etkileri. *Kendime Yazılar*. <https://www.mahfiegilmez.com/2023/03/amerikan-bankaclk-sistemindeyasanan.html> (Erişim Tarihi: 3 Eylül 2024)
- Elman, J. L.** (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211.
- Eren, A., ve Saraçoğlu, M.** (2017). Kapitalizmin Tarihinde Finansal Krizler: Tarih Tekerrür mü Ediyor?. *Ömer Halisdemir Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 10(1), 87-97.
- Fawcett, T.** (2006). "An introduction to ROC analysis." *Pattern Recognition Letters*, 27(8), 861-874.

## **KAYNAKLAR DİZİNİ (devam)**

- Fernández, A., Garcia, S., Herrera, F., and Chawla, N. V.** (2018). “SMOTE for Learning from Imbalanced Data.” *IEEE Transactions on Neural Networks and Learning Systems*, 30(9), 2804-2816.
- Fry, J., and Cheah, E. T.** (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, 47, 343-352.
- Galbraith, J. K.** (2009). *The great crash 1929*. Houghton Mifflin Harcourt.
- Garber, P. M.** (2001). *Famous first bubbles: The fundamentals of early manias*. mit Press.
- Garcia, D., Tessone, C. J., Mavrodiev, P., and Perony, N.** (2014). The digital traces of bubbles: feedback cycles between socio-economic signals in the Bitcoin economy. arXiv preprint arXiv:1408.1494.
- Géron, A.** (2019). *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.
- Gers, F. A., Schmidhuber, J., and Cummins, F.** (2000). Learning to Forget: Continual Prediction with LSTM. *Neural Computation*, 12(10), 2451–2471.
- The Guardian.** (2022). FTX collapse: How one of crypto’s biggest exchanges fell apart. <https://www.theguardian.com> (Erişim Tarihi: 25 Ocak 2025)
- Goodfellow, I., Bengio, Y., and Courville, A.** (2016). *Deep learning*. MIT Press.
- Graves, A., and Schmidhuber, J.** (2005). Framewise phoneme classification with bidirectional LSTM networks. In *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005*. (Vol. 4, pp. 2047-2052). IEEE.
- Graves, A., Mohamed, A. R., and Hinton, G.** (2013). Speech recognition with deep recurrent neural networks. In *2013 IEEE international conference on acoustics, speech and signal processing* (pp. 6645-6649). Ieee.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... and Chen, T.** (2018). Recent Advances in Convolutional Neural Networks. *Pattern Recognition*, 77, 354–377.
- Guyon, I., and Elisseeff, A.** (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3, 1157–1182.
- Güleç, T. C., ve Aktaş, H.** (2019). Kripto para piyasasında spekülâtif fiyat balonlarının analizi. *Muhasebe ve Finansman Dergisi*, (84), 149-164.
- Güven, V., ve Şahinöz, E.** (2018). *Blokzincir-kripto paralar - Bitcoin Satoshi dünyayı değiştiriyor*. Kronik Kitap.
- Haffar, A., and Le Fur, E.** (2021). Structural vector error correction modelling of Bitcoin price. *The Quarterly Review of Economics and Finance*, 80, 170-178.

## **KAYNAKLAR DİZİNİ (devam)**

- Han, J., Kamber, M., and Pei, J.** (2011). *Data Mining: Concepts and Techniques*. Elsevier.
- Hand, D. J., and Till, R. J.** (2001). "A simple generalisation of the area under the ROC curve for multiple class classification problems." *Machine Learning*, 45(2), 171-186.
- Hassan, S., and De Filippi, P.** (2021). Decentralized autonomous organization. *Internet Policy Review*, 10(2).
- Hastie, T., Tibshirani, R., and Friedman, J.** (2009). *The Elements of Statistical Learning*. Springer.
- He, K., Zhang, X., Ren, S., and Sun, J.** (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778.
- Hochreiter, S., and Schmidhuber, J.** (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780.
- Homm, U., and Breitung, J.** (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1), 198-231.
- Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q.** (2017). Densely Connected Convolutional Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4700–4708.
- Graves, A.** (2012). Supervised sequence labelling with recurrent neural networks. Springer. <https://doi.org/10.1007/978-3-642-24797-2>
- IBM.** What are convolutional neural networks? <https://www.ibm.com/think/topics/convolutional-neural-networks> (Erişim Tarihi: 25 Ocak 2025)
- İnci, A. C., ve Alper, D.** (2018). The predictability of cryptocurrency returns with technical indicators. *Journal of Risk and Financial Management*, 11(4), 66. <https://doi.org/10.3390/jrfm11040066>
- İnci, S., ve Alper, İ.** (2018). *Bitcoin devrimi*. Ankara: Elma.
- James, G., Witten, D., Hastie, T., and Tibshirani, R.** (2013). *An Introduction to Statistical Learning*. Springer.
- Jin, R., and Agrawal, G.** (2003). Communication and memory efficient parallel decision tree construction. In *Proceedings of the 2003 SIAM international conference on data mining* (pp. 119-129). Society for Industrial and Applied Mathematics.

## **KAYNAKLAR DİZİNİ (devam)**

- Judmayer, A., Stifter, N., Krombholz, K., and Weippl, E.** (2022). Blocks and chains: introduction to bitcoin, cryptocurrencies, and their consensus mechanisms. Springer Nature.
- Kavlakoglu, E., and Russi, E.** (2024). What is XGBoost?. IBM. Retrieved from [https://www.ibm.com/think/topics/xgboost?mhsrc=ibmsearch\\_a&mhq=XGBOOST](https://www.ibm.com/think/topics/xgboost?mhsrc=ibmsearch_a&mhq=XGBOOST) (Erişim Tarihi: 7 Aralık 2024)
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... and Liu, T. Y.** (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3146-3154.
- Kim, J.** (2021). Advertising in the metaverse: Research agenda. *Journal of Interactive Advertising*, 21(3), 141-144.
- Kindleberger, C. P., Manias, P., and Crashes, A.** (2005). *History of financial crises*. London: Macmillan.
- Kohavi, R.** (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 1137-1145.
- Kristoufek, L.** (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific reports*, 3(1), 3415.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E.** (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25, 1097–1105.
- Kuhn, M., and Johnson, K.** (2013). *Applied Predictive Modeling*. Springer.
- Landgraf, C.** (2016). Testing for speculative bubbles using recursive right-tailed unit root tests. *Economics Letters*, 146, 13–17. <https://doi.org/10.1016/j.econlet.2016.06.013>
- LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P.** (1998). Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Li, P., Wu, Q., and Burges, C.** (2007). Mcrank: Learning to rank using multiple classification and gradient boosting. *Advances in neural information processing systems*, 20.
- Li, Y., Wang, Z., Wang, H., Wu, M., and Xie, L.** (2021). Identifying price bubble periods in the Bitcoin market-based on GSADF model. *Quality & Quantity*, 1-16.

## **KAYNAKLAR DİZİNİ (devam)**

- LightGBM Development Team.** (2024). LightGBM Release 4.5.0. Microsoft Corporation. Retrieved from <https://lightgbm.readthedocs.io/en/stable/FAQ.html> (Erişim Tarihi: 8 Aralık 2024)
- Liu, Y., Tsyvinski, A., and Wu, X.** (2023). Common risk factors in cryptocurrency markets. *Journal of Finance*, 78(2), 1133–1172. <https://doi.org/10.1111/jofi.13119>
- Ljumović, M., and Klar, M.** (2015). Estimating expected error rates of random forest classifiers: A comparison of cross-validation and bootstrap. 4th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 212-215.
- Lundberg, S. M., and Lee, S. I.** (2017). “A unified approach to interpreting model predictions.” *Advances in Neural Information Processing Systems*, 4765-4774.
- Mackay, C.** (2003). *Extraordinary popular delusions*. Courier Corporation.
- Malhotra, A., and Maloo, M.** (2014). Bitcoin—is it a bubble? Evidence from unit root tests. *Evidence from Unit Root Tests*.
- Mete, S., Koy, A., ve Ersoy, H.** (2019). Kriptoparalarda fiyat balonu incelemesi. *BDDK Bankacılık ve Finansal Piyasalar Dergisi*, 13(1), 105-120.
- Mian, A., and Sufi, A.** (2009). The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis. *The Quarterly journal of economics*, 124(4), 1449-1496.
- Mystakidis, S.** (2022). Metaverse. *Encyclopedia*, 2(1), 486-497.
- Nadini, M., Alessandretti, L., Di Giacinto, F., Martino, M., Aiello, L. M., and Baronchelli, A.** (2021). Mapping the NFT revolution: market trends, trade networks, and visual features. *Scientific reports*, 11(1), 20902.
- Nakamoto, S.** (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*.
- Nguyen, Q. K.** (2016). Blockchain-a financial technology for future sustainable development. In 2016 3rd International conference on green technology and sustainable development (GTSD) (pp. 51-54). IEEE.
- O’Shea, K., and Nash, R.** (2015). *An Introduction to Convolutional Neural Networks*. arXiv.
- Ofek, E., and Richardson, M.** (2003). Dotcom mania: The rise and fall of internet stock prices. *the Journal of Finance*, 58(3), 1113-1137.
- Okina, K., Shirakawa, M., and Shiratsuka, S.** (2001). The asset price bubble and monetary policy: Japan’s experience in the late 1980s and the lessons. *Monetary and Economic Studies (special edition)*, 19(2), 395-450.

## **KAYNAKLAR DİZİNİ (devam)**

- Oran, A.** (2011). Balonlari Daha iyi Tanimaya Çalışmak: Balon Tanimlari, Modelleri ve Lale Çilginliđi Örneđi. Dokuz Eylül Üniversitesi İktisadi İdari Bilimler Fakóltesi Dergisi, 26(1), 151-161.
- Öncü, A.** (2021). Dogecoin Fiyat Dalgalanmaları ve Elon Musk Etkisi. Kripto Para Analizleri.
- Özbayođlu, A. M., Gudelek, M. U., ve Sezer, O. B.** (2020). Deep learning for financial applications: A survey. Applied Soft Computing, 93, 106384. DOI: 10.1016/j.asoc.2020.106384
- Panagiotidis, T., Stengos, T., and Vravosinos, O.** (2018). On the determinants of bitcoin returns: A LASSO approach. Finance Research Letters, 27, 235-240.
- Panko, R. R.** (2008). IT employment prospects: beyond the dotcom bubble. European Journal of Information Systems, 17(3), 182-197.
- Park, S., and Yang, J. S.** (2024). Machine learning models based on bubble analysis for Bitcoin market crash prediction. Engineering Applications of Artificial Intelligence, 135, 108857.
- Parr, T., Turgutlu, K., Csiszar, C., and Howard, J.** (2018). Beware default random forest importances. explained. ai.
- parthshukla211.** (2023). GradientBoosting vs AdaBoost vs XGBoost vs CatBoost vs LightGBM. <https://www.geeksforgeeks.org/gradientboosting-vs-adaboost-vs-xgboost-vs-catboost-vs-lightgbm/> (Eriřim Tarihi: 8 Ocak 2025)
- Pascanu, R., Mikolov, T., and Bengio, Y.** (2013). On the difficulty of training recurrent neural networks. Proceedings of the 30th International Conference on Machine Learning (ICML-13), 1310-1318.
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al.** (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.
- Phillips, P. C., Shi, S., and Yu, J.** (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. International economic review, 56(4), 1043-1078.
- Phillips, P. C., Wu, Y., and Yu, J.** (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?. International economic review, 52(1), 201-226.
- Powers, D. M.** (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. Journal of Machine Learning Technologies, 2(1), 37-63. <https://doi.org/10.48550/arXiv.2010.16061>

## **KAYNAKLAR DİZİNİ (devam)**

- Priya, C. R., Rahman, T., Talla Narayana Gowd, D. G. K. P., Subramanian, N., and Saxena, A.** (2023). Critical Investigation of Cryptocurrency Data and Analysis: A Comprehensive Overview. *Journal of Informatics Education and Research*, 3(2).
- Ranka, S., and Singh, V.** (1998). CLOUDS: A decision tree classifier for large datasets. In *Proceedings of the 4th knowledge discovery and data mining conference (Vol. 2, No. 8, pp. 2-8)*.
- Reinhart, C. M., and Rogoff, K. S.** (2009). *This time is different: Eight centuries of financial folly*. Princeton University Press.
- Rodriguez, J. J., Kuncheva, L. I., and Alonso, C. J.** (2006). Rotation forest: A new classifier ensemble method. *IEEE transactions on pattern analysis and machine intelligence*, 28(10), 1619-1630.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J.** (1986). Learning representations by back-propagating errors. *nature*, 323(6088), 533-536.
- Sagheer, A., and Kotb, M.** (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, 323, 203-213.
- Saito, T., and Rehmsmeier, M.** (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Sajjad, M., Khan, Z. A., Ullah, A., Hussain, T., Ullah, W., Lee, M. Y., and Baik, S. W.** (2020). A novel CNN-GRU-based hybrid approach for short-term residential load forecasting. *Ieee Access*, 8, 143759-143768.
- Salmi, Z. J., Khodakarami, M. I., and Behnamfar, F.** (2023). Development of seismic fragility curves for RC/MR frames using machine learning methods. *Asian Journal of Civil Engineering*, 24(3), 823-836.
- Santoni, G. J.** (1987). The great bull markets 1924-29 and 1982-87: Speculative bubbles or economic fundamentals. *Federal Reserve Bank of St. Louis Review*, 69(9), 16-29.
- Sawhney, A., Kour, P., and Sawhney, R.** (2022). CryptoBubbles: Predicting Bitcoin bubbles with hybrid econometric and machine learning models. *Expert Systems with Applications*, 201, 117001. <https://doi.org/10.1016/j.eswa.2022.117001>
- Schubert, E. S.** (1988). Innovations, debts, and bubbles: International integration of financial markets in Western Europe, 1688–1720. *The journal of economic history*, 48(2), 299-306.

## **KAYNAKLAR DİZİNİ (devam)**

- Segendorf, B.** (2014). What is bitcoin. *Sveriges Riksbank Economic Review*, 2014, 2-71.
- Serada, A.** (2020). Cryptomarkets gamified: What can we learn by playing cryptokitties?. In *Proceedings of DiGRA 2020 Conference: Play Everywhere*.
- Shah, D., Isah, H., and Zulkernine, F.** (2019). Stock market analysis: A review and taxonomy of prediction techniques. *International Journal of Financial Studies*, 7(2), 26. DOI: 10.3390/ijfs7020026
- Shiller, R. J.** (2000). Measuring bubble expectations and investor confidence. *The Journal of Psychology and Financial Markets*, 1(1), 49-60.
- Shiller, R. J.** (2015). *Irrational exuberance: Revised and expanded third edition*.
- Shu, M., and Zhu, W.** (2020). Real-time prediction of Bitcoin bubble crashes. *Physica A: Statistical Mechanics and its Applications*, 548, 124477.
- Shu, M., Song, R., and Zhu, W.** (2021). The 2021 bitcoin bubbles and crashes—detection and classification. *Stats*, 4(4), 950-970.
- Simonyan, K., and Zisserman, A.** (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv.
- Smart, John, Cascio, J. and Paffendorf, J.** (2007). “Pathwaysto the 3D Web: A Cross-Industry Public Foresight Project.”Metaverse Roadmap. <https://metaverseroadmap.org/> (Erişim Tarihi: 21 Eylül 2024)
- Sokolova, M., and Lapalme, G.** (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Sompolinsky, Y., and Zohar, A.** (2013). Accelerating bitcoin's transaction processing. fast money grows on trees, not chains. *Cryptology ePrint Archive*.
- Songur, M., ve Ordu, S.** (2023). Bitcoin Haberlerinin Bitcoin Fiyat ve Getirisi Üzerine Etkisi. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (25), 220-234.
- Stanford University.** (2024-2025). Convolutional Neural Networks for Visual Recognition (CS231n) <https://explorecourses.stanford.edu/search?academicYear=20242025&filter-coursestatus-Active=on&q=CS+231N%3A+Deep+Learning+for+Computer+Vision&view=catalog> (Erişim Tarihi: 25 Ocak 2025)
- Stehman, S. V.** (1997). Selecting and interpreting measures of thematic classification accuracy. *Remote Sensing of Environment*, 62(1), 77-89. [https://doi.org/10.1016/S0034-4257\(97\)00083-7](https://doi.org/10.1016/S0034-4257(97)00083-7)

## **KAYNAKLAR DİZİNİ (devam)**

- Stephenson, N.** (1994). Snow crash. Bantam Books.
- Sun, S., Cao, Z., Zhu, H. and Zhao, J.** (2019). A Survey of Optimization Methods from a Machine Learning Perspective. arXiv preprint arXiv:1906.06821.
- Sutskever, I., Vinyals, O., and Le, Q. V.** (2014). Sequence to Sequence Learning with Neural Networks. Neural Information Processing Systems (NIPS).
- Sutton, R. S., and Barto, A. G.** (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... and Rabinovich, A.** (2015). Going Deeper with Convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9.
- Taha, A. A., and Malebary, S. J.** (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. IEEE access, 8, 25579-25587.
- Teigland, R., and Puschmann, T.** (2013). Breaking Out of the Bank in Europe: Exploring Collective Emergent Institutional Entrepreneurship, through Bitcoin. Breaking out of the Bank, 15-16.
- TensorFlow.** (2022). CNNs in Deep Learning. <https://www.tensorflow.org/tutorials/images/cnn?hl=tr> ( Erişim Tarihi: 25 Ocak 2025)
- Tsay, R. S.** (2005). Analysis of financial time series. John wiley & sons.
- Van Rijsbergen, C.** (1979). Information retrieval: theory and practice. In Proceedings of the joint IBM/University of Newcastle upon tyne seminar on data base systems (Vol. 79, pp. 1-14). Oxford, UK: Butterworth-Heinemann.
- Van Wijk, D.** (2013). What can be expected from the BitCoin. Erasmus Universiteit Rotterdam, 18, 1-21.
- Wang, S., Ding, W., Li, J., Yuan, Y., Ouyang, L., and Wang, F. Y.** (2019). Decentralized autonomous organizations: Concept, model, and applications. IEEE Transactions on Computational Social Systems, 6(5), 870-878.
- Wei, D.** (2017). Price bubbles in bitcoin: Evidence, causes and implications. Journal of Shanghai University of Finance and Economics, 19(02), 50-62.
- Wei, Y., Li, Y., and Wang, Z.** (2022). Multiple price bubbles in global major emission trading schemes: Evidence from European Union, New Zealand, South Korea and China. Energy Economics, 113, 106232.

## **KAYNAKLAR DİZİNİ (devam)**

- Werbos, P. J.** (1990). Backpropagation through time: What it does and how to do it. Proceedings of the IEEE, 78(10), 1550-1560.
- XGBoost Developers.** (2024). XGBoost: Release 2.1.1. Retrieved from <https://xgboost.readthedocs.io/en/stable/> (Erişim Tarihi: 7 Aralık 2024)
- Yavuz, E.** (2021). 1929 Dünya Ekonomik Buhranının Türkiye'ye ve diğer ülkelere olan etkisinin istatistik analizi. Erzincan Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 14(2), 377–390.
- Yermack, D.** (2013). Is Bitcoin a real currency? An economic appraisal. NBER Working Paper No. 19747. <https://doi.org/10.3386/w19747>
- Yıldırım, H.** (2021). Kripto Para Çılgınlığı Finansal Krizlerden Kripto Paralara. İstanbul: Himalaya Yayınları.
- Zeren, F., ve Esen, S.** (2018). Geleceğin para birimi ya da sadece bir balon: bitcoin. Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 21(39), 433-448.
- Zetsche, D. A., Arner, D. W., and Buckley, R. P.** (2020). Decentralized finance. Journal of Financial Regulation, 6(2), 172-203.
- Zhang, H., Jin, R., and Zhou, Z. H.** (2019). Understanding data-driven feature selection in light gradient boosting machine (LightGBM). Journal of Machine Learning Research, 20, 1-33.