

TC
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



KRİPTO PARA FİYAT ANALİZİ İÇİN MAKİNE ÖĞRENME YAKLAŞIMLARI

İrem Sevda İNCE

YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

DANIŞMAN
Dr. Öğr. Üyesi Yusuf ÇELİK

TUNCELİ – 2025

TC
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ



KRİPTO PARA FİYAT ANALİZİ İÇİN MAKİNE ÖĞRENME YAKLAŞIMLARI

İrem Sevda İNCE

YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

DANIŞMAN
Dr. Öğr. Üyesi Yusuf ÇELİK

TUNCELİ – 2025

TC
MUNZUR ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

KRİPTO PARA FİYAT ANALİZİ İÇİN MAKİNE ÖĞRENME YAKLAŞIMLARI

İrem Sevda İNCE
YÜKSEK LİSANS TEZİ
HESAPLAMALI BİLİMLER VE MÜHENDİSLİK ANABİLİM DALI

Bu tez 28/05/2025 tarihinde aşağıdaki jüri üyeleri tarafından **oybirliği** ile kabul edilmiştir.

İmza

İmza

İmza

Doç. Dr. Vedat TÜMEN
(Bitlis Eren Üniversitesi)
Başkan

Dr. Öğr. Üyesi Yusuf ÇELİK
(Munzur Üniversitesi)
Danışman

Dr. Öğr. Üyesi Ümit CAN
Munzur Üniversitesi)
Üye

Bu tez, Enstitümüz Hesaplamalı Bilimler ve Mühendislik Anabilim Dalı'nda hazırlanmıştır.

Enstitü Müdürü
Doç. Dr. Nagihan KARAASLAN AYHAN

NOT: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı “Fikir ve Sanat Eserleri Kanunu”ndaki hükümlere tabidir.

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalarında bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi ve hiçbir şekilde “intihal içermediğini” beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçları kabul ettiğimi bildiririm.

İmza

İrem Sevda İNCE

Danışman

Dr. Öğr. Üyesi Yusuf ÇELİK

TEŞEKKÜR

Bu çalışmanın her aşamasında akademik bir titizlik ve özveriyle rehberlik eden, bilgi ve deneyiminden istifade etme fırsatı bulduğum değerli danışmanım Dr. Öğr. Üyesi Yusuf ÇELİK'e en içten teşekkürlerimi sunarım.

İrem Sevda İNCE
Tunceli - 2025



İÇİNDEKİLER

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ.....	I
TEŞEKKÜR	II
İÇİNDEKİLER	III
ŞEKİLLER LİSTESİ	V
TABLolar LİSTESİ	VII
KISALTMALAR LİSTESİ.....	VIII
ÖZET	IX
ABSTRACT.....	X
1. GİRİŞ	1
1.1. Amaç ve Kapsam	1
1.2. Literatür Değerlendirmesi.....	2
1.3. Tezin Organizasyonu ve Katkıları	4
2.FİNANSAL VERİLER VE DERİN ÖĞRENME TEMELLERİ	6
2.1. Finansal zaman serilerinin temel özellikleri	6
2.2. Kripto para piyasalarının yapısı ve dinamikleri.....	6
2.3. Teknik analiz göstergelerinin teorik temeli	8
2.3.1 Trend tabanlı göstergeler	9
2.3.2 Momentum tabanlı göstergeler	10
2.3.3 Volatilite göstergeleri	11
2.3.4 Diğer göstergeler.....	12
3.MAKİNE ÖĞRENMESİ	14
3.1 Makine öğrenmesine genel bakış.....	14
3.1.1 Denetimli öğrenme	16
3.1.2 Denetimsiz öğrenme	16
3.1.3 Pekiştirmeli öğrenme	17
3.2. Yapay Sinir Ağları	17
3.2.1. Yapay nöronlar ve sinirsel temelleri	17
3.2.2 Yapay sinir sğlarının katmanlı yapısı	20
3.2.3. Derin sinir ağları ve derin öğrenme yaklaşımı.....	21
3.2.4 Derin öğrenme paradigmaları	22
3.3 Tekrarlı Sinir Ağları.....	24
3.3.1 Uzun kısa süreli bellek.....	25
3.3.2. Geçitli tekrarlayan birim	28

3.3.3. Çift yönlü uzun kısa süreli bellek	30
4. YÖNTEM VE MODEL TASARIMI	32
4.1 Veri Kümesi ve Özellikler	32
4.2. Yöntem Performans Değerlendirilmesi	36
4.2.1. Zaman serisi tahmin performans metrikleri.....	37
4.3 Model Tasarımı.....	39
4.3.1 Model mimari karşılaştırması	40
4.3.2. Farklı kayıp fonksiyonlarının model performansına etkisi.....	43
5.TEKNİK GÖSTERGELERİN ENTEGRE EDİLDİĞİ MODELLER.....	47
5.1 Saatlik Çözünürlükte Teknik Göstergelerle Zenginleştirilmiş LSTM Modeli	48
5.1.1. Modelin eğitim süreci ve doğrulama performansı	49
5.1.2. Test veri seti üzerindeki performans karşılaştırmaları.....	51
5.2. Dakikalık Zaman Serileri İçin Teknik Göstergelerle Desteklenmiş LSTM Tabanlı Tahmin Modeli	54
5.2.1. Dakikalık girişle gerçekleştirilen saatlik tahminler	54
5.2.2. 15 dakikalık periyotlarda yapılan kısa vadeli tahminler	55
5.2.3. Test Veri Seti Üzerindeki Performans ve Tahmin-Gözlem Karşılaştırmaları.....	57
6. SONUÇLAR.....	61
7. KAYNAKLAR	62
ÖZGEÇMİŞ	Hata! Yer işareti tanımlanmamış.

ŞEKİLLER LİSTESİ

Sayfa

Şekil 2.1. Blokzincir sisteminde bir işlemin doğrulanma ve ödüllendirme süreci.....	8
Şekil 3.1. Makine öğrenmesinin temel türleri ve alt dallarına şematik gösterimi.....	15
Şekil 3.2. Sık kullanılan bazı aktivasyon fonksiyonları. (a) sigmoid aktivasyon fonksiyonu (b) tanh aktivasyon fonksiyonu (c) ReLU aktivasyon fonksiyonu	19
Şekil 3.3. Yapay nöronun temel hesaplama süreci.....	20
Şekil 3.4. Üç katmanlı (bir gizli katmanlı) bir yapay sinir ağının şematik gösterimi	21
Şekil 3.5. LSTM hücresinin iç yapısı	26
Şekil 4.1. BTC/USDT işlem çiftine ait 1 dakikalık kapanış fiyatlarının zaman serisi (2017- 2025)	34
Şekil 4.2. BTC/USDT işlem çiftine ait 1 dakikalık işlem hacimlerinin zaman serisi (2017- 2025)	34
Şekil 4.3. ETH/USDT işlem çiftine ait 1 dakikalık kapanış fiyatlarının zaman serisi (2017- 2025)	34
Şekil 4.4. ETH/USDT işlem çiftine ait 1 dakikalık işlem hacimlerinin zaman serisi (2017- 2025)	34
Şekil 4.5. Bağımsız test dönemine ait BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerinin 1 dakikalık kapanış fiyatlarının zaman serileri (27 Mart – 12 Nisan 2025)	35
Şekil 4.6. BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerine ait saatlik kapanış fiyatlarının zaman serileri	37
Şekil 4.7. Modelleme sürecinin genel iş akışını gösteren diyagram	40
Şekil 4.8. Farklı mimariler ve batch size değerlerine göre epoch bazlı doğrulama kaybı değişimi.....	41
Şekil 5.1. Modelleme sürecine ait veri akış diyagramı	47
Şekil 5.2. Modelin eğitim ve doğrulama kayıplarını gösteren öğrenme eğrileri (a) (BTC) ve (b) (ETH) göstermektedir.....	49
Şekil 5.3. Modelin BTC/USDT işlem çifti için bağımsız test veri seti üzerindeki performansı: (a) Tüm test süresine ait gerçek ve tahmin fiyatlarının zaman serisi üzerindeki karşılaştırması. (b) ilk 200 saatlik örneğe ait detaylı karşılaştırma. (c) ilk 5 saatlik fiyat hareketlerine ait yakınlaştırılmış görünüm..	51
Şekil 5.4. Modelin ETH/USDT işlem çifti için bağımsız test veri seti üzerindeki performansı:(a) Tüm test süresine ait gerçek ve tahmin fiyatlarının zaman serisi üzerindeki karşılaştırması. (b) İlk 200 saatlik örneğe ait detaylı karşılaştırma. (c) İlk 5 saatlik fiyat hareketlerine ait yakınlaştırılmış görünüm..	52
Şekil 5.5. 15 dakikalık tahmin modeli için BTC/USDT ve ETH/USDT işlem çiftlerinde model eğitimi sırasında gözlemlenen doğrulama ve eğitim kayıplarının epoch bazında değişimi.(a) BTC modeli, (b) ETH modeli. Her iki modelde de erken	

durdurma mekanizması etkin şekilde çalışarak eğitim ve doğrulama kayıplarının kısa sürede dengelendiği gözlemlenmiştir..	57
Şekil 5.6. BTC için test seti üzerindeki tahmin-gözlem karşılaştırmaları.....	58
Şekil 5.7. ETH için test seti üzerindeki tahmin-gözlem karşılaştırmaları.....	59



TABLÖLAR LİSTESİ

	Sayfa
Tablo 3.1. Makine öğrenmesi yöntemlerinin larşılaştırılması	15
Tablo 4.1. BTC/USDT işlem çiftine ait örnek veri satırları	32
Tablo 4.2. ETH/USDT işlem çiftine ait örnek veri satırları	33
Tablo 4.3. Farklı derin öğrenme modelleri ve batch size değerlerine göre fiyat tahmin performans metrikleri	42
Tablo 4.4. Farklı kayıp fonksiyonlarıyla eğitilen derin öğrenme modellerinin performans karşılaştırması (LSTM, GRU ve BiLSTM)	45
Tablo 5.1. Teknik göstergelerle zenginleştirilmiş LSTM modelinin BTC ve ETH doğrulama veri seti üzerindeki performans metrikleri.....	50
Tablo 5.2. Test verisi üzerinde LSTM modelinin performans metrikleri.....	51
Tablo 5.3. Model performans sonuçları.....	55
Tablo 5.4. Teknik gösterge çözünürlüğüne göre doğrulama performans karşılaştırması ..	56
Tablo 5.5. Dakikalık tahmin modeli için BTC/USDT ve ETH/USDT işlem çiftlerine ait doğrulama seti üzerindeki performans metrikleri	57
Tablo 5.6. Model testseti performans metrikleri.....	60

KISALTMALAR LİSTESİ

BiLSTM	: Bidirectional Long Short-Term Memory
BTC	: Bitcoin
EMA	: Exponential Moving Average
ETH	: Ethereum
GRU	: Gated Recurrent Unit
LSTM	: Long Short-Term Memory
MACD	: Moving Average Convergence Divergence
MAE	: Mean Absolute Error
MSE	: Mean Squared Error
RNN	: Recurrent Neural Network
RMSE	: Root Mean Squared Error
ROC	: Rate of Change
RSI	: Relative Strength Index
SMA	: Simple Moving Average

ÖZET

Bu çalışma, finansal zaman serilerinin kısa vadeli tahmini amacıyla çeşitli veri yapıları ve model konfigürasyonlarının kapsamlı bir değerlendirmesini sunmakta; teknik göstergelerle zenginleştirilmiş Long Short-Term Memory (LSTM) tabanlı bir derin öğrenme modeli önermektedir. Model, saatlik ve dakikalık çözünürlükte yapılandırılmış veri setleri kullanılarak, fiyat-hacim bilgileri ile teknik analiz göstergelerinin ayrı girişler olarak işlendiği farklı senaryolar altında test edilmiştir. Ayrıca LSTM, GRU ve BiLSTM mimarileri; MSE, MAE, Huber ve büyük hataları daha fazla cezalandıracak şekilde tasarlanmış özel bir ağırlıklı kayıp fonksiyonu ile değerlendirilmiştir. Deneysel sonuçlar, teknik göstergelerle desteklenen LSTM modelinin hem fiyat hareketlerinin yönünü hem de büyüklüğünü doğru şekilde tahmin etmede yüksek performans sergilediğini göstermektedir. Bu çalışma, derin öğrenme tabanlı finansal tahmin bağlamında farklı veri türleri ve modelleme stratejilerini birlikte ele alan bütüncül çerçevesiyle literatüre anlamlı bir katkı sunmaktadır.

Anahtar kelimeler: Finansal zaman serileri, derin öğrenme, kripto para fiyat tahmini

ABSTRACT

Machine Learning Approaches for Cryptocurrency Price Analysis

This study offers a thorough examination of various data formats and model setups for the short-term prediction of financial time series, presenting a deep learning framework based on Long Short-Term Memory (LSTM) enriched with technical indicators. The proposed model was tested on datasets prepared at both hourly and minute-level resolutions, where price-volume features and technical indicators were fed into the network through distinct input paths. Moreover, different model architectures namely LSTM, GRU, and BiLSTM were evaluated in combination with multiple loss functions including MSE, MAE, Huber, and a specially designed weighted loss function that applies stronger penalties to larger errors. The experimental findings indicate that the LSTM model, when supported by technical indicators, achieves high accuracy in predicting both the trend direction and magnitude of price movements. This research makes a valuable contribution to the field by adopting a holistic approach that integrates diverse data types and modeling strategies within the scope of deep learning-based financial forecasting.

Keywords: Financial time series, deep learning, cryptocurrency price forecasting

1. GİRİŞ

1.1. Amaç ve Kapsam

Son yıllarda artan veri hacmi, donanım kapasitesindeki gelişmeler ve açık kaynak yazılım araçlarının yaygınlaşması, makine öğrenmesi ve özellikle derin öğrenme alanında önemli bir dönüşüm yaratmıştır. Geçmişte sınırlı işlem gücü, algoritmaların eğitim sürecindeki istikrarsızlıklar ve özellikle kaybolan gradyan gibi yapısal sorunlar nedeniyle yaygın kullanılamayan derin öğrenme modelleri, GPU teknolojilerinin gelişimi ve iyileştirilmiş öğrenme algoritmaları sayesinde daha erişilebilir ve verimli hale gelmiştir (LeCun ve ark., 2015; Bengio ve ark., 1994). Özellikle uzun bağımlılıkların öğrenilmesi sorununa çözüm olarak önerilen LSTM yapısı, bu ilerlemenin en kritik adımlarından biri olarak kabul edilmektedir (Hochreiter ve Schmidhuber, 1997). Bu teknolojik ve yöntemsel gelişmeler yalnızca görüntü işleme veya doğal dil işleme gibi alanlarda değil; aynı zamanda karmaşık zaman serisi verilerinin modellenmesinde de derin öğrenmeyi ön plana çıkarmıştır. Finansal piyasa verileri gibi yüksek oynaklık içeren ve doğrusal olmayan yapılar, geleneksel istatistiksel yöntemlerin sınırlılıklarını ortaya koymakta; bu bağlamda zaman bağımlılıklarını öğrenebilen yapay sinir ağı mimarileri, özellikle de tekrarlayan sinir ağları (RNN) etkili alternatifler sunmaktadır (Yu ve ark., 2019).

Bu tez çalışmasının temel amacı, kripto para piyasalarında işlem gören Bitcoin (BTC) fiyatlarını, teknik göstergelerle zenginleştirilmiş bir LSTM (Long Short-Term Memory) modeli kullanarak yüksek doğrulukla tahmin etmektir. Bu amaç doğrultusunda, Binance platformundan elde edilen dakikalık ve saatlik fiyat verileri kullanılmış, farklı çözünürlüklerdeki verilerle çalışan modeller karşılaştırmalı olarak değerlendirilmiştir. Ayrıca, RSI, MACD, SMA, Bollinger Bands gibi yaygın kullanılan teknik analiz göstergeleri modele eklenerek, tahmin performansına olan etkileri araştırılmıştır.

Çalışmanın kapsamı, yalnızca geçmiş fiyat ve işlem hacmi verilerini değil, aynı zamanda teknik göstergeleri de dikkate alarak çok boyutlu bir veri yapısı oluşturmayı içermektedir. Bu kapsamda LSTM mimarisi, sıralı verilerdeki zaman bağımlılıklarını öğrenme yeteneği nedeniyle tercih edilmiş; model, farklı pencere boyutları ve zaman çözünürlükleri ile test edilmiştir. Ayrıca, farklı kayıp fonksiyonları ve model mimarileri karşılaştırılarak en yüksek performansı sağlayan yapı belirlenmiştir.

Bu tez, teknik göstergelerle desteklenmiş derin öğrenme modellerinin kripto para fiyat tahminindeki etkinliğini göstermeyi ve literatüre yöntemsel katkı sağlamayı hedeflemektedir.

1.2. Literatür Değerlendirmesi

Finansal zaman serileri; doğrusal olmayan yapı, yüksek oynaklık, ani fiyat değişimleri ve dışsal etkilere kaynaklanan bozulmalar gibi karmaşık özellikler barındırmaktadır. Bu durum, fiyat tahmini gibi problemlerde geleneksel istatistiksel yöntemlerin başarısını sınırlamaktadır. Özellikle ARIMA ve GARCH gibi modeller, belirli varsayımlara dayalı oldukları için dinamik ve doğrusal olmayan piyasa yapısına tam uyum sağlayamamaktadır (Rundo ve ark., 2019). Bununla birlikte, bu modeller zamanla yerini, daha esnek öğrenme yeteneklerine sahip makine öğrenmesi ve derin öğrenme tabanlı yaklaşımlara bırakmıştır.

Makine öğrenmesi yöntemleri, finansal piyasalarda veri odaklı tahmin sistemlerinin gelişmesini sağlamış; zaman serilerindeki karmaşık örüntüleri geçmiş verilere dayanarak öğrenebilme kapasiteleriyle öne çıkmıştır (Masini ve ark.,2019). Başlangıçta destek vektör makineleri, rastgele ormanlar ve gradyan artırma gibi yöntemler uygulansa da, bu algoritmaların ardışık bağımlılıkları yeterince işleyememesi, özellikle dalgalı piyasa verilerinde doğruluk sorunlarına yol açmıştır. Bu nedenle zaman içi örüntüleri daha başarılı şekilde öğrenebilen derin öğrenme mimarilerine yönelim artmıştır. Rundo ve ark. (2019) ,LSTM temelli yapıların klasik makine öğrenmesi algoritmalarına kıyasla daha düşük hata oranları ve daha yüksek yön tahmini doğruluğu sağladığını rapor etmektedir. Benzer şekilde, Çelik (2024) çalışmasında, LSTM ve GRU modelleri BIST100 endeksinin genel eğilimlerini başarılı şekilde öğrenmiş, ancak ani piyasa dalgalanmaları karşısında sınırlı duyarlılık sergilediği belirtilmiştir.

Bu bağlamda, özellikle ardışık verilerle çalışan derin öğrenme yapıları ön plana çıkmıştır. Uzun Kısa Süreli Bellek (LSTM) modelleri, Hochreiter ve Schmidhuber (1997) tarafından geliştirilmiş olup, kaybolan gradyan problemine çözüm getirmesi ve uzun vadeli bağımlılıkları öğrenebilme yeteneği sayesinde finansal zaman serisi tahmininde çığır açmıştır. LeCun ve ark. (2015) ile Yu ve ark. (2019) tarafından yapılan çalışmalarda, LSTM ve türevi modellerin, klasik RNN ve diğer makine öğrenmesi yöntemlerine kıyasla daha yüksek tahmin doğruluğu sağladığı rapor edilmiştir. Mudassir ve ark. (2020), ANN, SANN,

SVM ve LSTM modellerini Bitcoin fiyat tahmininde karşılaştırmış ve LSTM'nin hem kısa hem uzun vadeli tahminlerde açık ara üstün performans gösterdiğini ortaya koymuştur.

Gerek hisse senedi piyasalarında gerekse kripto paralarda uygulanan bu modeller, yatırımcı davranışlarını yansıtan teknik göstergelerle desteklendiğinde daha başarılı sonuçlar vermektedir. Bu nedenle, finansal verilerin doğrusal olmayan yapısını daha iyi yakalayabilmek amacıyla, teknik göstergelerle zenginleştirilmiş hibrit modeller literatürde yaygın olarak kullanılmaktadır. RSI, MACD, SMA ve Bollinger Bands gibi teknik analiz göstergeleri, yatırımcı psikolojisi ve piyasa momentumuna dair anlamlı sinyaller taşıdıkları için tahmin modellerinin başarımını artırmaktadır (Neely ve ark., 2014; Achelis, 2001). Chen ve ark. (2019), dikkat mekanizması ve Empirical Mode Decomposition yöntemini LSTM mimarisıyla birleştirerek, klasik modellere kıyasla daha düşük tahmin hatası elde etmiş ve hibrit yapının avantajlarını ortaya koymuştur. Özüpek ve ark. (2024) tarafından geliştirilen EMD-TI-LSTM modeli ise hem zaman serisi bileşenleri hem de teknik göstergeleri doğrudan giriş olarak kullanarak, tahmin doğruluğunu anlamlı biçimde artırmıştır. Benzer şekilde, Agrawal ve ark. (2022), RSI ve MACD gibi teknik göstergeleri doğrudan model girdisi olarak kullandıkları Enhanced Deep Learning Algorithm mimarisi ile, LSTM tabanlı yapıların doğruluk seviyesini önemli ölçüde yükseltmiştir.

Kripto para piyasaları, yüksek oynaklık, merkeziyetsizlik ve sınırlı regülasyon gibi özgün yapıları nedeniyle, finansal tahmin modelleri açısından ayrı bir araştırma alanı olarak değerlendirilmektedir. Alessandretti ve ark. (2018), Bitcoin ve benzeri dijital varlıkların klasik para birimlerinden farklı olarak sosyal medya, duyarlılık analizi ve zincir üstü verilerden de etkilenebildiğini ortaya koymuştur. Kalra ve ark. (2024) tarafından geliştirilen H.BLSTM adlı hibrit model, çift yönlü LSTM ve artımlı öğrenme yaklaşımlarını birleştirerek, teknik göstergeler ve hacim temelli özelliklerle desteklenen çok değişkenli bir yapı sunmuş ve klasik yöntemlere kıyasla düşük hata oranı ve yüksek adaptasyon başarısı sağlamıştır. Rundo (2019) ise, döviz piyasalarında uygulanan modelinde, derin LSTM mimarisini pekiştirmeli öğrenme katmanı ile entegre ederek, kısa ve orta vadeli eğilim tahminlerinde yüksek doğruluk ve otomatik işlem stratejileri açısından yüksek getiriler elde edildiğini raporlamıştır.

Tüm bu çalışmalar göstermektedir ki; teknik göstergelerle desteklenmiş LSTM tabanlı derin öğrenme modelleri hem zaman serisi korelasyonlarını hem de yatırımcı davranışlarını öğrenebilme açısından klasik yöntemlere göre anlamlı avantajlar sunmaktadır.

Ancak literatürde, çok çözünürlüklü veri (örneğin saatlik ve dakikalık) ile test edilmiş, yön tahmini başarımı ve finansal getiri açısından değerlendirilmiş yapay zekâ sistemlerinin sayısı halen sınırlıdır. Bu tez çalışması, farklı zaman çözünürlüklerinde LSTM modelini teknik göstergelerle birleştirerek hem doğruluk hem de stratejik kârlılık açısından kapsamlı bir analiz sunmayı hedeflemektedir.

1.3. Tezin Organizasyonu ve Katkıları

Bu tez çalışması, kripto para piyasalarında derin öğrenme temelli fiyat tahmini modellerinin doğruluk ve stratejik uygulanabilirlik açısından performansını değerlendirmeyi amaçlamaktadır. Bu kapsamda, özellikle LSTM (Long Short-Term Memory) mimarisi kullanılarak Bitcoin (BTC) fiyatlarının saatlik ve dakikalık çözünürlükte tahmini gerçekleştirilmiş; farklı model yapılarına, kayıp fonksiyonlarına ve pencere boyutlarına dayalı deneyler yürütülmüştür. Ayrıca, RSI, MACD, SMA ve Bollinger Bandı gibi yaygın teknik analiz göstergeleri modele entegre edilerek, bu göstergelerin tahmin performansına olan katkıları analiz edilmiştir.

Tez dört ana bölümden oluşmaktadır;

- Birinci bölümde, kripto para piyasalarının yapısı, zaman serisi modelleme gereksinimleri ve çalışmanın amacı ile kapsamı ortaya konulmuştur.
- İkinci bölümde, kullanılan veri seti ve teknik göstergeler detaylandırılmış; seçilen model mimarileri ve kullanılan yöntemler açıklanmıştır.
- Üçüncü bölümde, deneysel aşamada test edilen modellerin performans karşılaştırmaları yapılmış hem istatistiksel hem de finansal metrikler üzerinden analizler sunulmuştur.
- Son bölüm olan dördüncü bölümde ise genel değerlendirmeler yapılmış, sonuçlar bütünsel biçimde özetlenmiş ve ileride yapılabilecek çalışmalara yönelik öneriler sunulmuştur.

Bu çalışmanın literatüre katkısı üç temel başlık altında özetlenebilir:

Zaman çözünürlüğü farkındalığı: Hem saatlik hem de dakikalık veriyle çalışan, çok çözünürlüklü LSTM tabanlı modeller geliştirilerek, çözünürlüğün tahmin doğruluğu ve stratejik kârlılık üzerindeki etkisi kapsamlı biçimde karşılaştırılmıştır.

Farklı hiper parametrelerinin ve çeşitli kayıp fonksiyonlarının model performansı üzerindeki etkisi sistematik olarak incelenmiş; bu bağlamda optimal yapıların belirlenmesine yönelik bir yol haritası sunulmuştur.

Kripto para piyasalarında sıklıkla kullanılan teknik göstergeler, derin öğrenme modeline ayrı bir giriş olarak entegre edilmiş ve bu entegrasyonun model doğruluğuna katkısı detaylı biçimde gösterilmiştir. Özellikle dakikalık çözünürlükte teknik göstergelerle zenginleştirilmiş LSTM yapısı, yüksek performanslı bir çözüm olarak önerilmiştir.

Bu yönleriyle çalışma, derin öğrenme uygulamalarının kripto para piyasalarında fiyat tahminini nasıl optimize edilebileceğine dair yeni bir perspektif sağlamaktadır.



2. FİNANSAL VERİLER VE DERİN ÖĞRENME TEMELLERİ

Bu bölümde, finansal piyasalarda sıklıkla karşılaşılan zaman serilerinin temel yapısal özellikleri ele alınmakta; ardından, kripto para piyasalarının kendine özgü dinamikleri ve bu piyasalarda kullanılan teknik analiz göstergelerinin teorik temelleri açıklanmaktadır. Bölümün devamında ise, makine öğrenmesi ve derin öğrenme tabanlı yaklaşımlar tanıtılarak, özellikle tekrarlayan sinir ağı mimarilerine dayalı tahmin modellerinin finansal veriler üzerindeki uygulamaları detaylı biçimde incelenmektedir.

Finansal veriler, ekonomik faaliyetlerin ölçülmesinde ve piyasa hareketlerinin analizinde kritik rol oynayan zaman serisi biçiminde kaydedilir. Bu veriler arasında özellikle fiyat, hacim ve getiri gibi değişkenler ön plandadır. Söz konusu değişkenler zaman içinde gözlemlendiğinde, belirli istatistiksel örüntüler ve yapısal özellikler sergiler. Bu nedenle, finansal piyasalara yönelik veri analizi süreçlerinde, öncelikle bu zaman serilerinin temel istatistiksel ve dinamik niteliklerinin anlaşılması gerekmektedir.

2.1. Finansal zaman serilerinin temel özellikleri

Finansal zaman serileri, ardışık zaman noktalarında gözlemlenen fiyat, hacim veya getiri gibi değişkenlerin dinamik yapısını temsil eden veri kümeleridir. Bu tür seriler genellikle yüksek volatilité, doğrusal olmayanlık, mevsimsellik ve trend gibi karmaşık istatistiksel özellikler gösterir. Ekonomik göstergeler, politik gelişmeler veya piyasa duyarlılığı gibi dışsal faktörlere duyarlılıkları nedeniyle, sabit varyans ve durağanlık varsayımları çoğu zaman geçerli olmaz. Ayrıca, otokorelasyon gibi zaman bağımlılığı içeren yapılar, bu verilerin öngörülebilirliğini zorlaştırırken, kısa vadeli dalgalanmaların ve uzun vadeli eğilimlerin aynı anda modellenmesini gerekli kılar. Bu nedenle finansal zaman serilerinin analizi, klasik istatistiksel yöntemlerin ötesinde, esnek ve veri odaklı modelleme yaklaşımlarını gerekli kılmaktadır (Rundo ve ark.,2019).

2.2. Kripto para piyasalarının yapısı ve dinamikleri

Kripto para piyasaları, geleneksel finansal sistemlere alternatif olarak gelişen, merkeziyetsizlik ilkesi üzerine inşa edilmiş dijital varlık ekosistemidir. Bu piyasaların kökeni, ilk merkeziyetsiz dijital para birimi olan Bitcoin'in 2008 yılında Satoshi Nakamoto tarafından tanıtılmasına dayanmaktadır. Ancak bu yapı, yalnızca bir para birimi değil; aynı

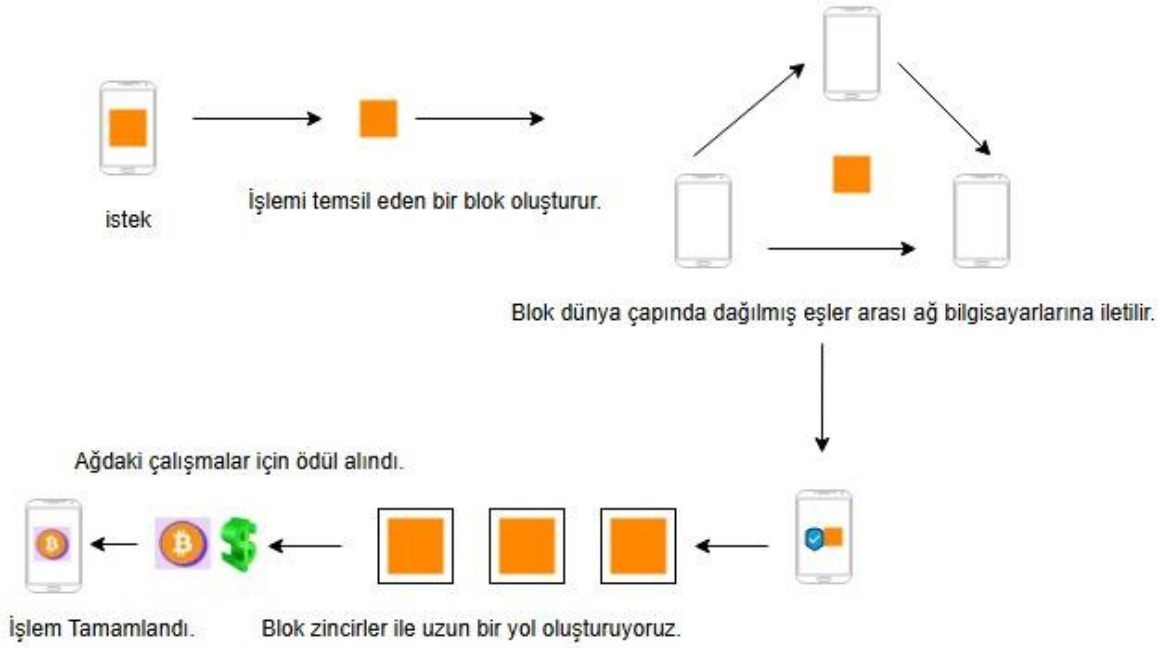
zamanda yeni nesil bir teknolojik altyapı olan blokzincir sistemini de içermektedir (Corbet ve ark., 2019).

Kripto varlıklar; dijital paralar, token'lar, stablecoin'ler ve merkeziyetsiz finans ürünlerini kapsayan geniş bir yelpazeyi ifade eder. Bu varlıkların temelini oluşturan blockchain teknolojisi, işlem verilerinin güvenli, şeffaf ve değiştirilemez şekilde tutulmasını sağlar. Blokzincir, işlem taleplerini dijital bloklara dönüştürerek dağıtık ağ üzerindeki tüm düğümlere iletir. Her yeni blok, bir önceki bloğun kriptografik özeti ile zincire eklenir ve doğrulama süreci sonrası bu kayıt kalıcı hale gelir. Bu yapı sayesinde işlemler, merkezi olmayan bir sistemde güvenle tamamlanır.

Kripto para birimlerinin değeri çoğu zaman spekülasyon etmenlere bağlı olarak şekillenmektedir. Hızla dalgalanan fiyatlar, yatırımcılar için hem yüksek getiri hem de yüksek risk anlamına gelmektedir. Bununla birlikte, büyük kripto paraların (örneğin Bitcoin ve Ethereum) zamanla geleneksel riskli varlıklarla daha korelasyonlu hareket ettiği gözlemlenmiştir (Oerth ve Dowsett, 2023). Ayrıca, atomik yerleşim gibi yeni kavramlar, varlık transferlerinin eş zamanlı olarak gerçekleştirilmesini sağlayarak piyasalarda dönüştürücü potansiyeller sunmaktadır.

Öte yandan, blockchain teknolojisinin sadece finans değil; sağlık, dağıtım, enerji ve kamu yönetimi gibi birçok sektörde uygulama alanı bulması, bu sistemlerin çok yönlülüğünü ortaya koymaktadır. Teknolojik altyapının temel avantajları arasında güvenlik, ölçeklenebilirlik, programlanabilirlik ve merkeziyetsizlik yer almaktadır. Ancak blok boyutu, işlem hızı ve enerji tüketimi gibi bazı teknik sınırlılıklar da bu teknolojinin gelişiminde önemli rol oynamaktadır (Rajasekaran ve ark., 2022).

Aşağıda, blokzincir işlemlerinin temel aşamalarını gösteren şematik bir diyagram Şekil 2.1'de sunulmuştur. Bu görsel, bir işlemin nasıl başlatıldığını, doğrulandığını, zincire nasıl eklendiğini ve son kullanıcıya nasıl yansıdığını adım adım göstermektedir:



Şekil 2.1. Blokzincir sisteminde bir işlemin doğrulanma ve ödüllendirilme süreci

Bu yapısal süreç, blokzincir teknolojisinin güvenlik, şeffaflık ve merkeziyetsizlik ilkeleri doğrultusunda nasıl işlediğini somut biçimde ortaya koymaktadır.

2.3. Teknik analiz göstergelerinin teorik temeli

Teknik analiz göstergeleri, finansal varlıkların geçmiş fiyat ve hacim verilerini temel alarak gelecekteki fiyat yönelimlerini tahmin etmeye yönelik araçlardır. Bu göstergeler, piyasadaki eğilimleri tanımlamak, alım-satım sinyalleri üretmek, volatilitiyi değerlendirmek ve işlem zamanlaması yapmak gibi birçok işlevi yerine getirir. Özellikle trendin yönünü belirlemek amacıyla kullanılan hareketli ortalamalar, fiyatın ivmesini izlemeye yarayan momentum göstergeleri ve dalgalanmayı ölçen volatilité temelli göstergeler, yatırımcıların karar alma süreçlerinde önemli rol oynar. Teknik göstergelerin temel varsayımı, piyasa fiyatlarının geçmişteki davranışlarının gelecekte de benzer şekilde tekrar edeceğidir (Murphy, 1999). Bu çerçevede, teknik analiz yaklaşımları, özellikle kısa ve orta vadeli tahminler için finansal piyasalarda yaygın biçimde kullanılmaktadır.

2.3.1 Trend tabanlı göstergeler

Basit Hareketli Ortalama (SMA), belirli bir zaman penceresi boyunca fiyatların aritmetik ortalamasını alarak genel trend yönünü belirlemeye yardımcı olan temel bir teknik analiz göstergesidir. Genellikle kısa ve orta vadeli eğilimleri izlemek için tercih edilir (Murphy, 1999).

SMA'nın hesaplanma yöntemi Denklem (2.1)'de verilmiştir:

$$SMA_{t,n} = \frac{1}{n} \sum_{i=0}^{n-1} P_{t-i} \quad (2.1)$$

Burada;

$SMA_{t,n}$: t anındaki n dönemlik basit hareketli ortalama

P_{t-i} : $t - i$ zamanındaki kapanış fiyatı,

n: Ortalama alınan dönem sayısıdır.

Üssel Hareketli Ortalama (EMA), SMA'nın aksine daha yeni fiyatlara daha fazla ağırlık vererek trend değişimlerine daha duyarlı bir gösterge sunar. Bu sayede fiyatlardaki ani değişimlere daha hızlı tepki verir. Bu çalışmada, 72 dönemlik EMA(72) kullanılmıştır.

EMA'nın hesaplanma yöntemi Denklem (2.2)'de verilmiştir:

$$EMA_t = \alpha \cdot P_t + (1-\alpha)EMA_{t-1} \quad , \quad \alpha = \frac{2}{n+1} \quad (2.2)$$

Burada;

EMA_t : t anındaki EMA değeri

EMA_{t-1} : t-1 anındaki EMA değeri

α : Üssel ağırlık katsayısı

P_t : Güncel kapanış fiyatı

n: EMA için seçilen dönem sayısı

MACD, kısa ve uzun vadeli EMA'lar (EMA_{12} ve EMA_{26}) arasındaki farkı hesaplayarak momentum ve trend dönüşlerini analiz eder (Appel, 2005).

Bu çalışmada, MACD hesaplamasında EMA_{12} , EMA_{26} ve sinyal çizgisi için EMA_9 kullanılmıştır.

Hesaplanma yöntemi Denklem (2.3a) ve (2.3b)'de verilmiştir:

$$MACD_t = EMA_{12}P(t) - EMA_{26}P(t) \quad (2.3a)$$

$$Signal_t = EMA_9(MACD_t) \quad (2.3b)$$

Burada;

$MACD_t$: t anındaki MACD değeri

EMA_{12}, EMA_{26} : 12 ve 26 dönemlik üssel hareketli ortalamalar

$P(t)$: t anındaki kapanış fiyatı

$Signal_t$: MACD'nin dokuz dönemlik üssel hareketli ortalamasıdır.

2.3.2 Momentum tabanlı göstergeler

Momentum göstergeleri, fiyat değişimlerinin yönünü ve gücünü ölçen teknik analiz araçlarıdır. Bu çalışmada üç temel momentum göstergesi kullanılmıştır: RSI, Momentum ve ROC.

RSI, belirli bir dönem boyunca fiyatların artış ve azalış oranlarını karşılaştırarak aşırı alım ya da aşırı satım bölgelerini belirler (Wilder, 1978). Bu çalışmada, 14 dönemlik RSI(14) hesaplanmıştır. RSI hesaplanma yöntemi Denklem (2.4)'de verilmiştir:

$$RSI_t = 100 - \left(\frac{100}{1 + RS} \right), \quad RS = \frac{AvgGain_n}{AvgLoss_n} \quad (2.4)$$

Burada;

RSI_t : Göreli Güç Endeks

$AvgGain_n$: Son n dönemdeki ortalama fiyat artışı

$AvgLoss_n$: Son n dönemdeki ortalama fiyat kaybı

Momentum göstergesi, güncel fiyat ile belirli bir dönem önceki fiyat arasındaki farkı olarak fiyat ivmesini hesaplar (Neely ve ark., 2014). Bu çalışmada, 24 dönemlik Momentum(24) kullanılmıştır. Momentum hesaplanma yöntemi Denklem (2.5)'te verilmiştir:

$$Momentum_t = P_t - P_{t-n} \quad (2.5)$$

Burada;

$Momentum_t$: Momentum değeri

P_t : Güncel kapanış fiyatı

P_{t-n} : n dönem önceki kapanış fiyatı

Rate of Change (ROC), fiyatın geçmişe göre yüzde değişimini ölçer (Achelis, 2001).

Bu çalışmada, 24 dönemlik ROC(24) hesaplanmıştır. ROC hesaplanma yöntemi Denklem (2.6)'da verilmiştir:

$$ROC_t = 100 \times \left(\frac{P_t - P_{t-n}}{P_{t-n}} \right) \quad (2.6)$$

Burada;

ROC_t : t zamandaki ROC değeri,

P_t : Güncel kapanış fiyatı,

P_{t-n} : n dönem önceki kapanış fiyatı göstermektedir.

2.3.3 Volatilite göstergeleri

Volatilite göstergeleri, piyasanın dalgalanma derecesini ölçmek amacıyla kullanılır. Bu çalışmada Bollinger Band Genişliği ve Standart Sapma kullanılmıştır.

Bollinger Bantları, fiyatın bir ortalama etrafında standart sapma bazlı oluşturulan bantlar içinde ne kadar dalgalandığını ölçer (Bollinger, 2002). Bu çalışmada, bantlar 2 standart sapma ($\pm 2\sigma$) ve 20 dönemlik SMA ile hesaplanmıştır.

Bollinger Bantları hesaplanma yöntemi Denklem (2.7)'de verilmiştir:

$$Width_t = \frac{BB_{upper} - BB_{lower}}{SMA_n}, \quad BB = SMA_n \pm k \cdot \sigma_n \quad (2.7)$$

Burada;

$Width_t$: Bant genişliği

SMA_n : Orta bant (n dönemlik SMA)

σ_n : n dönemlik standart sapma

k: Bant çarpanı genellikle 2 olarak alınır

Standart sapma, fiyatların ortalamadan ne kadar saptığını ölçerek volatilitiyi temsil eder. Bu çalışmada, 24 dönemlik STD(24) hesaplanmıştır.

Standart sapma hesaplanma yöntemi Denklem (2.8)'de verilmiştir:

$$\sigma_t = \sqrt{\frac{1}{n} \sum_{i=0}^{n-1} (P_{t-i} - \bar{P})^2} \quad (2.8)$$

Burada;

σ_t : Standart sapma

P_{t-i} : Geçmiş fiyatlar

$\bar{P}$: Ortalama fiyat

n: Dönem sayısını göstermektedir.

2.3.4 Diğer göstergeler

Z-score, bir fiyatın ortalamadan kaç standart sapma uzaklıkta olduğunu göstererek norm dışı davranışları tespit etmeye yarar. Bu çalışmada, kapanış fiyatının z-normalize edilmiş değeri kullanılmıştır. Z-score, hesaplanma yöntemi Denklem (2.9)'de verilmiştir:

$$Z_t = \frac{P_t - \mu_n}{\sigma_n} \quad (2.9)$$

Z_t : Z-normalize edilmiş değer

P_t : Güncel kapanış fiyatı

μ_n : Ortalama fiyat

σ_n : Standart sapma

DayOfWeek değişkeni, her veri noktasının haftanın hangi gününe ait olduğunu belirtir. Bu çalışmada, zaman serisindeki işlem günlerinin haftalık döngüsü 0 = Pazartesi, 6 = Pazar olacak şekilde kodlanmıştır.



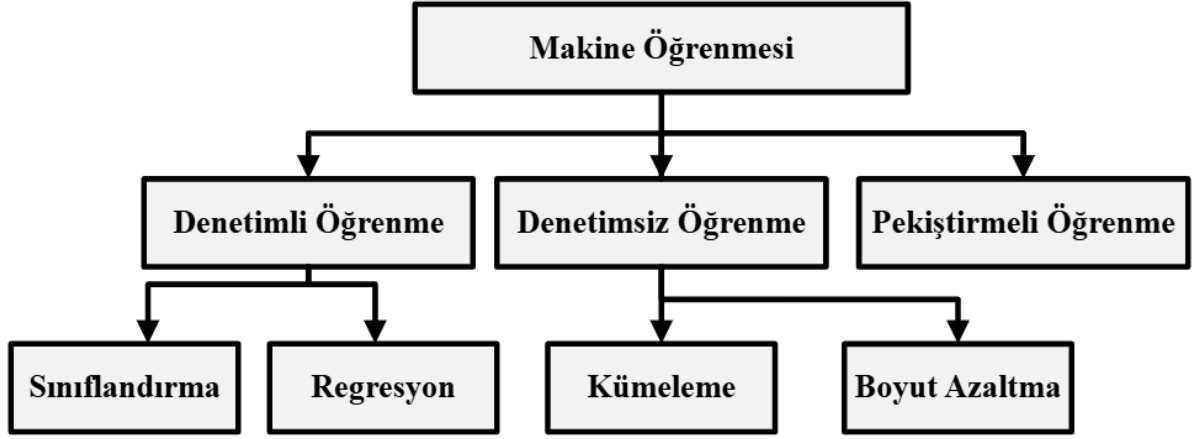
3. MAKİNE ÖĞRENMESİ

Makine öğrenmesi, veriden anlamlı desenler çıkarmayı ve bu desenleri kullanarak gelecekteki olayları öngörmeyi amaçlayan bir yapay zekâ alanıdır. Finansal zaman serileri gibi karmaşık, gürültülü ve öngörülemez veri yapılarının modellenmesinde, klasik istatistiksel yöntemlerin ötesine geçilmesini sağlayan güçlü araçlar sunar. Özellikle büyük veri setleriyle çalışırken, makine öğrenmesi algoritmaları; öğrenme, genelleme ve uyarlama yetenekleri sayesinde veri içindeki gizli ilişkileri ortaya koyabilir (Alpaydın, 2020).

3.1 Makine öğrenmesine genel bakış

Makine öğrenmesi, bilgisayar sistemlerinin açıkça programlanmaksızın veri üzerinden öğrenmesini sağlayan bir yapay zekâ alt alanıdır. Bu öğrenme, bir bilgisayarın geçmiş verilere dayanarak gelecekteki örüntüleri veya olayları tahmin etmesini sağlar. Günümüzde, makine öğrenmesi finansal tahminlerden sağlık uygulamalarına, doğal dil işleme sistemlerinden otonom araçlara kadar birçok alanda yaygın olarak kullanılmaktadır.

Makine öğrenmesi, genel olarak üç ana kategoriye ayrılmaktadır: denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme. Her bir kategori, kendi içinde farklı problem türlerini ve çözüm yaklaşımlarını barındırır. Denetimli öğrenme, sınıflandırma ve regresyon gibi etiketli veriyle çalışan görevleri kapsarken; denetimsiz öğrenme, kümeleme ve boyut indirgeme gibi veri içindeki yapıları keşfetmeye yönelik teknikleri içerir. Pekiştirmeli öğrenme ise ödül temelli bir yaklaşımla, ajanların çevreyle etkileşerek optimum karar stratejileri geliştirmesini hedefler. Bu üçlü ayrım, makine öğrenmesinin temel yapı taşlarını oluşturur ve uygulama alanlarının doğasına göre yöntem seçiminde yol gösterici olur (Mitchell, 1997; Sutton ve Barto, 2018). Bu yapının genel görünümü Şekil 3.1’de diyagram şeklinde sunulmuştur.



Şekil 3.1. Makine öğrenmesinin temel türleri ve alt dallarının şematik gösterimi

Makine öğrenmesi yöntemleri, kullanılan veri türüne, hedeflenen çıktıya ve öğrenme yaklaşımına göre farklı gruplara ayrılmaktadır. Bu ayrım, yöntemin temel amaçlarını, giriş-çıkış veri yapısını ve örnek algoritmalarını anlamada önemli bir temel oluşturur. Aşağıda, makine öğrenmesinin üç ana türü olan denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme; karşılaştırmalı biçimde Tablo 3.1’de sunulmuştur. Bu tablo, hem kuramsal sınıflamayı hem de uygulama düzeyindeki farklılıkları sade bir biçimde özetlemektedir.

Tablo 3.1. Makine öğrenmesi yöntemlerinin karşılaştırılması

Öğrenme Türü	Giriş Verisi	Çıktı Verisi	Amaç	Örnek Algoritmalar
Denetimli Öğrenme	Etiketli	Etiketli	Girdileri çıktılarına eşlemek	Karar Ağaçları, SVM, Lojistik Regresyon
Denetimsiz Öğrenme	Etiketsiz	Etiketsiz	Veri yapısını keşfetmek	K-Means, PCA, DBSCAN
Pekiştirmeli Öğrenme	Etiketsiz (Ödül/Ceza)	Etiketsiz (Ödül/Ceza)	Optimal politika öğrenimi	Q-Learning, SARSA, DDPG

3.1.1 Denetimli öğrenme

Denetimli öğrenme, algoritmanın eğitim sırasında hem giriş hem de doğru çıkış bilgisine erişebildiği bir öğrenme yaklaşımıdır. Bu yöntemde model, eğitim verisi üzerinde giriş-çıkış eşleştirmelerini öğrenerek genelleştirme yapar. Temel hedef, daha önce görülmemiş veriler için doğru tahminlerde bulunabilmektir (Mitchell, 1997).

- **Sınıflandırma:** sınıflandırma problemlerinde hedef değişken kategoriktir. Örneğin, bir e-postanın gereksiz olup olmadığını tahmin etmek, bir hastalık teşhisini koymak veya bir görseldeki nesneyi tanımak sınıflandırma problemine örnek verilebilir. Bu tür problemlerde genellikle karar ağaçları, destek vektör makineleri, yapay sinir ağları ve lojistik regresyon gibi yöntemler kullanılmaktadır (Bishop, 2006). Derin öğrenme mimarileri, özellikle görüntü ve metin tabanlı sınıflandırma problemlerinde yüksek doğruluk oranlarına ulaşarak bu alanda önemli ilerlemeler sağlamıştır.
- **Regresyon:** regresyon, çıktının sürekli (sayısal) olduğu problemleri kapsar. Örneğin, bir evin satış fiyatını, bir ürünün gelecekteki talebini ya da borsa endekslerinin yönünü tahmin etmek gibi. Regresyon modelleri genellikle doğrusal ya da doğrusal olmayan ilişkiler üzerinden kurulur. LSTM gibi zaman serisi temelli yapılar da regresyon problemlerinde yaygın biçimde kullanılmaktadır (Bishop, 2006).

3.1.2 Denetimsiz öğrenme

Denetimsiz öğrenme, eğitim verilerinde herhangi bir çıktı (etiket) bilgisinin olmadığı durumlarda kullanılır. Bu tür algoritmalar, verinin içsel yapısını, benzerliklerini ve örüntülerini keşfetmeye çalışır. En yaygın kullanılan teknikler kümeleme ve boyut indirgeme yöntemleridir (Mitchell, 1997).

Kümeleme yöntemleri, benzer gözlemleri gruplandırarak popülasyon içinde yapısal alt kümeler elde etmeyi sağlar. K-means gibi algoritmalar, müşteri segmentasyonu, pazar analizi ve biyoinformatik uygulamalarda yaygın olarak kullanılır.

Boyut azaltma teknikleri ise çok boyutlu veri setlerinde, bilgi kaybını en aza indirerek daha az sayıda ama anlamlı değişkenle veri temsilini sağlar. PCA (Principal Component Analysis) ve t-SNE gibi yöntemler, özellikle görselleştirme ve ön işleme aşamalarında önemlidir.

3.1.3 Pekiştirmeli öğrenme

Pekiştirmeli öğrenme, bir ajanın çevresiyle etkileşime girerek en iyi kararları öğrenmesini sağlayan bir algoritma ailesidir. Ajan, bir ortamda eylem gerçekleştirdikçe ödül veya ceza alır. Zaman içinde ödülleri maksimize eden bir strateji (politika) geliştirmesi beklenir (Sutton ve Barto, 2018).

Bu öğrenme yöntemi, klasik denetimli veya denetimsiz modellerden farklı olarak, sıralı karar problemlerine odaklanır. Örneğin, satranç ya da Go gibi oyunlarda en iyi hamleyi seçmek, bir robotun engellerden kaçınarak hedefe ulaşması veya bir sistemin kaynakları verimli şekilde tahsis etmesi bu kapsamda değerlendirilebilir.

3.2. Yapay Sinir Ağları

Yapay zekâ alanında yaşanan gelişmeler, öğrenme algoritmalarının daha karmaşık görevleri çözebilecek şekilde evrilmesine olanak sağlamıştır. Bu evrim sürecinde derin öğrenme, çok katmanlı yapay sinir ağları kullanılarak gerçekleştirilen yeni bir yaklaşım olarak öne çıkmıştır. Derin öğrenme yöntemleri, büyük veri setlerinden özellikleri otomatik olarak çıkarabilme kapasitesiyle klasik makine öğrenmesi yöntemlerinden ayrılır (LeCun ve ark. 2015).

3.2.1. Yapay nöronlar ve sinirsel temelleri

Yapay sinir ağları, bilgi işleme açısından insan beyninden esinlenerek geliştirilmiş, birbirine bağlı yapay nöronlardan oluşan hesaplama sistemleridir. Bu yapay sistemlerin temel yapı taşı olan yapay nöron, biyolojik nöronların sinyalleri alıp işleyerek diğer hücrelere iletme yeteneğinden ilham alır. Biyolojik bir nöronda dendritler aracılığıyla alınan sinyaller gövde (soma) tarafından işlenerek akson boyunca diğer nöronlara iletilirken; yapay nöron, dışarıdan gelen sayısal girdileri alır, bu girdileri belirli ağırlıklarla çarpar ve toplar, ardından bir aktivasyon fonksiyonu yardımıyla çıktıya dönüştürür (Aggarwal, 2018).

Bir yapay nöronun temel hesaplama süreci şu adımlardan oluşur:

Girdiler: $(x_0, x_1, x_2 \dots x_n)$

Ağırlıklar: $(w_0, w_1, w_2 \dots w_n)$

Net girdi:

$$\sum_{i=1}^n x_i * w_i + b \quad (3.1)$$

Aktivasyon fonksiyonu: $Y = f(X)$

Burada b terimi bias (sapma) olarak adlandırılır ve modelin esnekliğini artırmak için kullanılır. Aktivasyon fonksiyonu f doğrusal olmayan bir özellik kazandırarak sinir ağına öğrenme yeteneği kazandırır. Sıklıkla kullanılan aktivasyon fonksiyonları arasında sigmoid, tanh ve ReLU bulunmaktadır.

Sigmoid Fonksiyonu, çıkışı 0 ile 1 arasında sınırlayan doğrusal olmayan bir aktivasyon fonksiyonudur ve özellikle ikili sınıflandırma problemlerinde yaygın olarak kullanılır.

$$Sigmoid(ag) = \frac{1}{1 + e^{-ag}} \quad (3.2)$$

Tanh (Hiperbolik Tanjant) Fonksiyonu, giriş değerlerini -1 ile 1 arasında normalize ederek daha güçlü bir merkezileştirme sağlar.

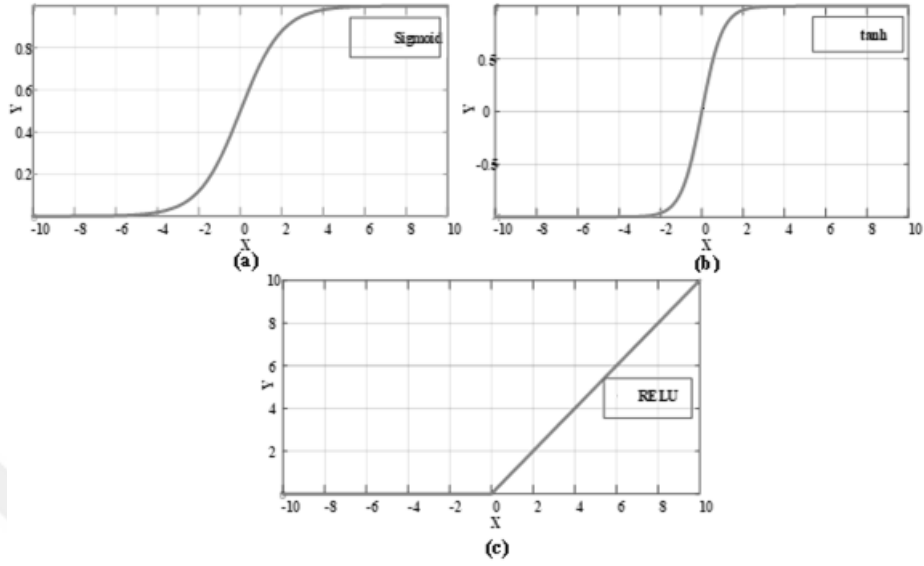
$$tanh(ag) = \frac{e^{ag} - e^{-ag}}{e^{ag} + e^{-ag}} \quad (3.3)$$

ReLU (Rectified Linear Unit) Fonksiyonu, negatif girdileri sıfıra eşitleyerek yalnızca pozitif değerleri geçirir ve bu sayede derin ağlarda gradyan sönümlene problemi azaltır.

$$ReLU(x) = \begin{cases} x & \text{Eğer } x > 0 \\ 0 & \text{Diğer Durumlar} \end{cases} \quad (3.4)$$

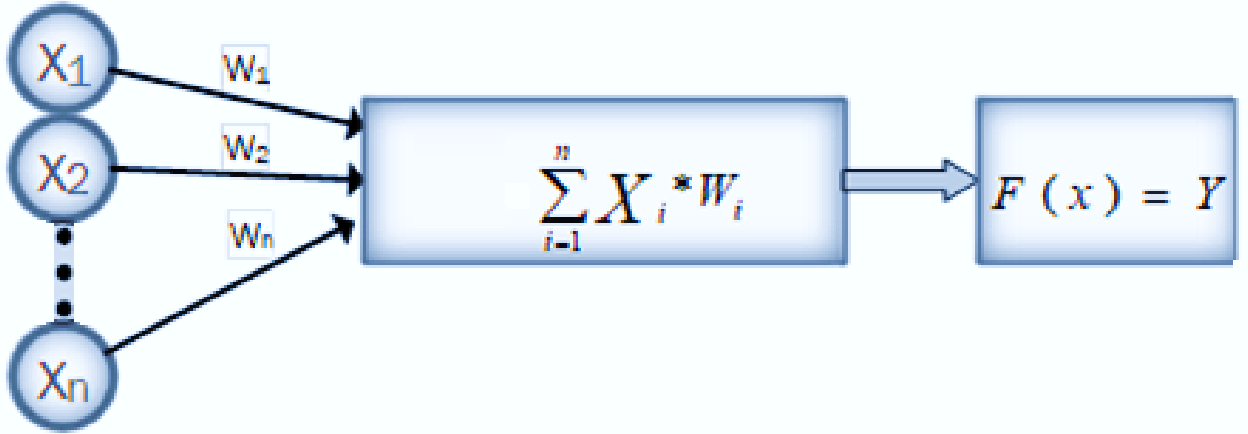
Bu aktivasyon fonksiyonlarının doğrusal olmayan davranışlarını daha iyi kavrayabilmek amacıyla, her bir fonksiyonun karakteristik eğrisi Şekil 3.2 görsel olarak sunulmuştur. Grafiklerden de görülebileceği üzere, Sigmoid fonksiyonu 0 ile 1 arasında sürekli artan bir eğri çizerken, tanh fonksiyonu -1 ile 1 arasında simetrik bir yapı sergiler.

ReLU ise yalnızca pozitif değerleri geçirip negatif değerleri sıfıra eşitleyerek parçalı doğrusal bir yapıya sahiptir



Şekil 3.2. Sık kullanılan bazı aktivasyon fonksiyonları. (a) sigmoid aktivasyon fonksiyonu (b) tanh aktivasyon fonksiyonu (c) ReLU aktivasyon fonksiyonu

Yapay sinir ağlarının temelini, 1943'te McCulloch ve Pitts tarafından tanımlanan ilk yapay nöron modeli oluşturmuştur. Bu modelde, girdilerin ağırlıklı toplamı bir eşiği aştığında çıktı "1", aksi halde "0" olarak belirleniyordu. Günümüzdeki derin sinir ağı mimarileri ise bu basit modelden türeyen, çok daha esnek ve karmaşık yapılara sahiptir (LeCun ve ark., 2015). Modern nöronlar, çeşitli aktivasyon fonksiyonları sayesinde sürekli değerli çıktılar üretebilirken, derin ağlar çok sayıda katmanla karmaşık veri ilişkilerini modelleyebilir hale gelmiştir. McCulloch-Pitts nöronu, modern derin öğrenmenin temelini atan ilk önemli adımdır. Yapay nöronun temel yapısı şekil 3.3 te gösterilmiştir.



Şekil 3.3. Yapay nöronun temel hesaplama süreci

Şekil 3.3 yapay bir nöronun temel bileşenlerini gösteren şematik temsili göstermektedir. Nöron, giriş değerlerini ağırlıklarla çarpar, toplar ve aktivasyon fonksiyonu aracılığıyla çıktıya dönüştürür. Bu yapı, sinir ağlarının öğrenme yeteneğini oluşturan temel yapı taşını temsil eder.

3.2.2 Yapay sinir ağlarının katmanlı yapısı

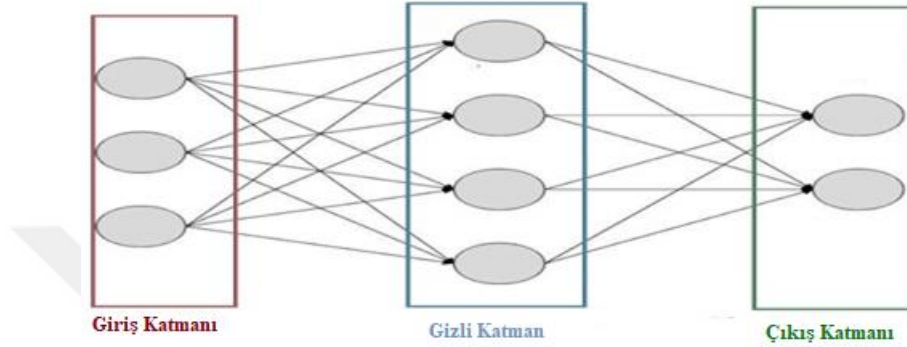
Yapay sinir ağları (YSA), girişten çıkışa doğru veri akışının katmanlar arasında gerçekleştiği çok katmanlı yapılardır. Bu yapılar genellikle üç temel katmandan oluşur: giriş katmanı, bir veya daha fazla gizli katman ve çıkış katmanı. Giriş katmanı, dış dünyadan alınan verileri nöronlara iletir. Gizli katmanlar, bu verileri işleyerek anlamlı temsillere dönüştürür. Çıkış katmanı ise problemin doğasına göre sınıflandırma ya da regresyon çıktısı üretir.

Her katmandaki nöronlar, bir önceki katmandaki nöronlara ağırlıklı bağlantılar ile bağlanır. Bu ağırlıklar, öğrenme sürecinde güncellenerek modelin veriye uyumu sağlanır. Ayrıca her nörona genellikle bir sapma terimi de eklenir. Nöronun toplam girdi değeri, belirli bir aktivasyon fonksiyonu aracılığıyla işlenir ve doğrusal olmayan çıktı üretilir.

Bu yapı, modelin karmaşık ilişkileri öğrenmesini sağlar. Gizli katman sayısı arttıkça, ağın temsil kapasitesi genişler ve daha karmaşık örüntüler öğrenilebilir. Ancak bu durum aynı zamanda aşırı öğrenme riskini artırır. Bu nedenle katman sayısı, aktivasyon fonksiyonu

seçimi ve diğer mimari parametreler, uygulamanın gereksinimlerine göre dikkatle belirlenmelidir (Aggarwal, 2018).

Şekil 3.4, bir giriş, bir gizli ve bir çıkış katmanından oluşan üç katmanlı yapay sinir ağı şematik olarak sunulmuştur.



Şekil 3.4. Üç katmanlı (bir gizli katmanlı) bir yapay sinir ağı şematik gösterimi

3.2.3. Derin sinir ağları ve derin öğrenme yaklaşımı

Yapay sinir ağlarının evrimi, daha katmanlı ve karmaşık ağ yapılarının ortaya çıkmasına öncülük etmiştir. Çok sayıda gizli katmana sahip bu mimariler, derin sinir ağları olarak adlandırılırken, bu yapılarla gerçekleştirilen öğrenme süreçleri derin öğrenme olarak adlandırılır. Derin öğrenme, yüksek boyutlu ve karmaşık verilerden otomatik olarak anlamlı özellikler çıkarma yeteneği sayesinde geleneksel makine öğrenmesi yaklaşımlarından ayrışır.

Derin sinir ağları, ardışık her katmanda girdiden daha soyut ve yüksek seviyeli özelliklerin öğrenilmesini sağlayarak, özellikle görüntü işleme, doğal dil anlama ve konuşma tanıma gibi alanlarda çığır açan başarılar elde etmiştir. Bu çok katmanlı yapı, verideki karmaşık ilişkilerin basitten başlayıp giderek artan bir soyutlama düzeyinde modellenmesine olanak tanır.

Derin öğrenme modellerinin başarısı, genellikle büyük ölçekli veri setlerinin varlığına, yüksek performanslı işlem gücüne ve probleme uygun ağ mimarilerinin seçimine sıkı sıkıya bağlıdır. Özellikle grafik işlem birimleri (GPU'lar) ve tensör işlem birimleri

(TPU'lar) gibi özel donanımların yaygınlaşması, derin ağların eğitimini önemli ölçüde hızlandırmış ve verimli hale getirmiştir. Ek olarak, aşırı öğrenmeyi (overfitting) engellemeye yönelik dropout, toplu normalleştirme (batch normalization) gibi düzenleme teknikleri ve gelişmiş optimizasyon algoritmaları (örneğin Adam, RMSProp) derin ağların daha istikrarlı ve genelleme yeteneği yüksek öğrenme süreçleri geçirmesine katkıda bulunur.

Özetle, derin sinir ağları, yapay sinir ağlarının çok katmanlı bir biçimi olup, ham veriden otomatik olarak özellik öğrenme kapasitesine sahip güçlü modelleme araçlarıdır. Bu modeller, insan müdahalesine gerek kalmadan hiyerarşik veri temsilleri oluşturabilmeleri sayesinde, birçok uygulama alanında önemli ilerlemeler ve yenilikçi sonuçlar ortaya koymuştur (LeCun ve ark.,2015).

3.2.4 Derin öğrenme paradigmaları

Derin öğrenme, farklı veri türleri ve problem yapıları için optimize edilmiş çeşitli sinir ağı mimarilerini kapsamaktadır. Özellikle uzamsal örüntülerin öğrenilmesi, düşük boyutlu temsillerin çıkarılması ve sıralı verilerin modellenmesi gibi görevler için evrişimli sinir ağları, otoenkoderler ve tekrarlayan yapılar gibi özel paradigmlar geliştirilmiştir (LeCun ve ark.,2015).

3.2.4.1 Evrişimli sinir ağları

Evrişimli sinir ağları (Convolutional Neural Networks – CNN'ler), derin öğrenmenin en güçlü ve yaygın mimarilerinden biridir. Özellikle görsel veri analizi alanında devrim niteliğinde ilerlemelere öncülük etmiştir. Görüntü ve video verilerinin karmaşık yapısını anlama ve bu verilerden anlamlı özellikler çıkarma konusundaki üstün yetenekleri sayesinde, CNN'ler, geleneksel görüntü işleme tekniklerinin ve manuel özellik mühendisliğinin sınırlamalarını aşarak birçok uygulamada çığır açmıştır. Görüntü sınıflandırma, nesne tespiti, görüntü segmentasyonu, yüz tanıma, video analizi, tıbbi görüntüleme ve hatta doğal dil işlemenin bazı alanları gibi geniş bir yelpazede yüksek başarılar elde etmişlerdir.

CNN'lerin temelini, biyolojik görsel korteksin çalışma prensiplerinden ilham alan özgün bir katmanlama yapısı oluşturur. Bu yapıda, temel olarak üç ana katman türü birbirini takip eder: evrişim katmanları, havuzlama katmanları ve tam bağlantılı katmanlar. Bu katmanların her biri, ağı görsel veriyi işleme ve anlaması sürecinde kritik bir rol oynar.

Evrişim Katmanları: CNN'lerin kalbini oluşturan bu katmanlar, giriş verisindeki yerel örüntüleri otomatik olarak algılamak için özel olarak tasarlanmış filtreler (veya çekirdekler) kullanır. Bu filtreler, giriş görüntüsü üzerinde kayarak her bir yerel bölgedeki belirli özellikleri (kenarlar, köşeler, dokular gibi) tespit eder. Her bir filtre, farklı türde bir özelliği yakalamak üzere eğitilir. Birden fazla filtre aynı anda uygulanarak, giriş verisinin zengin bir özellik haritası elde edilir. Bu sayede CNN'ler, manuel olarak tanımlanması zor olabilecek karmaşık görsel özellikleri doğrudan veriden öğrenebilirler.

Havuzlama Katmanları: Evrişim katmanlarından elde edilen özellik haritalarının ardından genellikle havuzlama katmanları gelir. Bu katmanların temel amacı, özellik haritalarının boyutunu azaltmak ve böylece hesaplama maliyetini düşürmek, aynı zamanda öğrenilen özelliklerin konumlarına karşı daha değişmez hale gelmelerini sağlamaktır. En yaygın havuzlama türleri arasında maksimum havuzlama ve ortalama havuzlama bulunur. Maksimum havuzlama, belirli bir bölgedeki en belirgin özelliği korurken, ortalama havuzlama bölgedeki tüm özelliklerin ortalamasını alır. Havuzlama katmanları, ağın daha genel ve soyut özelliklere odaklanmasına yardımcı olur.

Tam Bağlantılı Katmanlar: CNN mimarisinin sonuna doğru genellikle bir veya daha fazla tam bağlantılı katman bulunur. Bu katmanlar, önceki evrişim ve havuzlama katmanları tarafından öğrenilen yüksek seviyeli özellikleri alarak, nihai görev olan sınıflandırma veya regresyon gibi işlemleri gerçekleştirmek için kullanılır. Tam bağlantılı bir katmanda, bir katmandaki her bir nöron, bir sonraki katmandaki tüm nöronlara bağlıdır. Bu katmanlar, öğrenilen tüm özellikleri bir araya getirerek bir karar verme mekanizması oluşturur. Örneğin, bir görüntü sınıflandırma görevinde, son tam bağlantılı katman, farklı sınıflara ait olasılıkları çıktı olarak verebilir.

CNN'lerin bu katmanlı yapısı, verideki karmaşık örüntülerin basitten karmaşığa doğru hiyerarşik bir şekilde modellenmesine olanak tanır. İlk evrişim katmanları daha düşük seviyeli özellikleri (kenarlar, renk geçişleri gibi) öğrenirken, derinlere inildikçe daha soyut ve nesneye özgü özellikler (göz, tekerlek, kedi kulağı gibi) öğrenilir. Bu hiyerarşik özellik öğrenme yeteneği, CNN'leri görsel veri analizinde bu kadar güçlü kılan temel faktörlerden biridir. Ayrıca, ağırlık paylaşımı ve yerel bağlantılar gibi mekanizmalar sayesinde, CNN'ler parametre sayısını azaltır ve daha verimli bir öğrenme süreci sunar (Krizhevsky ve ark., 2012).

3.2.4.2 Otoenkoderler

Otoenkoderler, denetimsiz öğrenme paradigmaları içerisinde yer alan ve girdinin düşük boyutlu bir gizli uzayda temsilini öğrenmeyi ve bu temsilden orijinal girdiyi yeniden yapılandırmayı hedefleyen sinir ağı mimarileridir. Temel bir otoenkoder yapısı, girdi verisini sıkıştırılmış bir gizli temsile dönüştüren bir kodlayıcı ve bu gizli temsili kullanarak girdi verisini tekrar oluşturmaya çalışan bir çözücü olmak üzere iki temel bileşenden meydana gelir.

Otoenkoderler, girdi verisinin anlamlı ve sıkıştırılmış özelliklerini öğrenme kabiliyetleri sayesinde, boyut indirgeme, gürültülü veriden arındırma, anomali tespiti ve özellik çıkarımı gibi çeşitli makine öğrenmesi görevlerinde etkin bir şekilde kullanılmaktadır. Ayrıca, varyasyonel otoenkoderler gibi olasılıksal türevleri, latent uzayda öğrenilen dağılımlardan yeni veri örnekleri üretebilme potansiyeline sahip olup, üretken modelleme alanında önemli bir yer edinmiştir (Berahmand ve ark. ,2024).

3.2.4.3 Diğer paradigmlar

Derin öğrenme alanında, yukarıda bahsedilenlerin dışında da çeşitli mimariler bulunmaktadır. Örneğin, Yinelenen Sinir Ağları (Recurrent Neural Networks - RNN), sıralı verilerin işlenmesinde kullanılırken, Uzun Kısa Süreli Bellek (Long Short-Term Memory - LSTM) ve Kapı Kontrollü Tekrarlayan Ünite (Gated Recurrent Unit - GRU) gibi RNN türevleri, uzun vadeli bağımlılıkların öğrenilmesinde daha etkilidir. Bu mimariler, özellikle zaman serisi verileri, doğal dil işleme ve konuşma tanıma gibi uygulamalarda yaygın olarak kullanılmaktadır.

3.3 Tekrarlı Sinir Ağları

Zaman serileri doğal dil ifadeleri ve ses sinyalleri gibi ardışık verilerdir. Zaman serileri önceki girdilerin sonraki çıktılar üzerinde etkili olduğu örüntüler içerir. Bu tür veri yapılarında, geçmiş bilgiyi taşıyabilmek ve sonraki adımlarda değerlendirebilmek, öğrenme açısından kritik öneme sahiptir. Bu ihtiyaca yönelik olarak geliştirilen Tekrarlayan Sinir Ağları (Recurrent Neural Networks – RNN), klasik yapay sinir ağlarından farklı olarak zaman adımlarını birbirine bağlayan bir yapı sunar. Elman (1990) tarafından önerilen bu mimari, her adımda hem mevcut girdi x_t hem de önceki gizli durumu h_{t-1} işleyerek yeni bir çıktı üretir. Böylece model, zaman içerisindeki ardışık bağımlılıkları öğrenebilir.

Ancak RNN'ler, özellikle uzun dizilerde öğrenme sırasında karşılaşılan vanishing gradient (kaybolan gradyan) problemi nedeniyle uzun vadeli bağımlılıkları öğrenmekte zorlanır. Bu durum, geçmişteki önemli bilgilerin öğrenme sürecinde etkisizleşmesine neden olur.

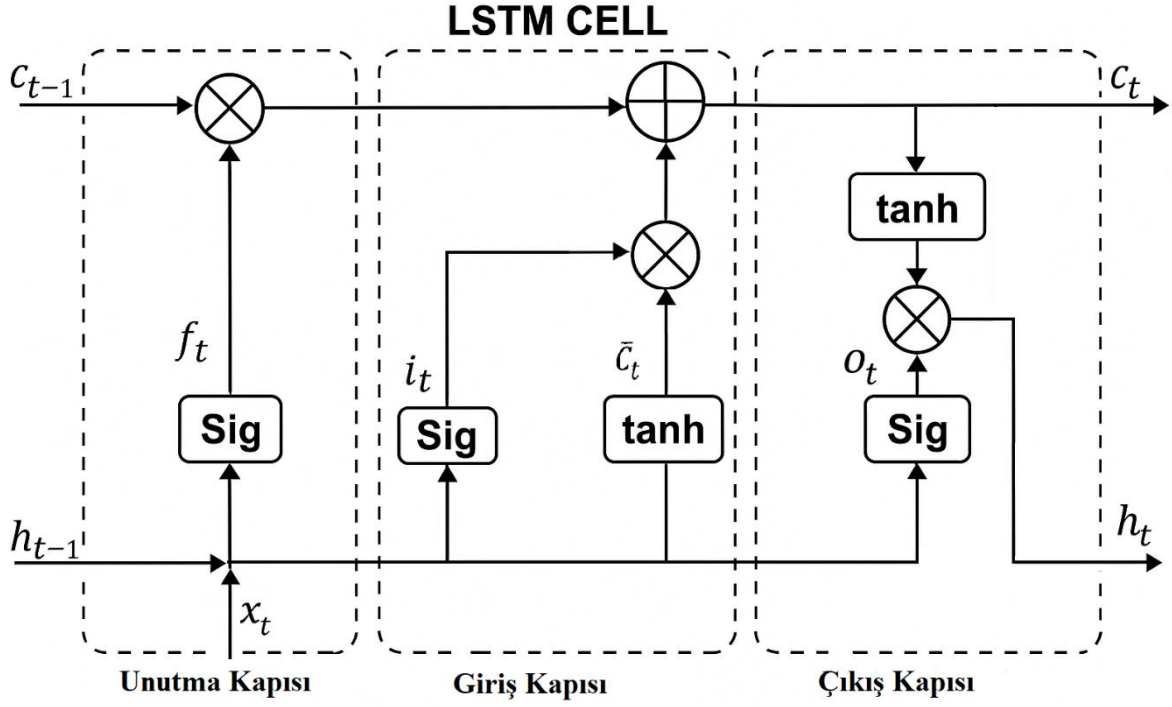
Bu sınırlamaların üstesinden gelmek amacıyla, kapı tabanlı kontrol mekanizmalarıyla geliştirilen LSTM yapısı geliştirilmiştir.

3.3.1 Uzun kısa süreli bellek

Uzun Kısa Süreli Bellek (Long Short-Term Memory – LSTM) mimarisi, Hochreiter & Schmidhuber (1997) tarafından, klasik tekrarlayan sinir ağlarının uzun vadeli bağımlılıkları öğrenme konusundaki yetersizliklerini aşmak üzere önerilmiştir. Bu yapı, bilgi aktarımını düzenleyen özel kapı mekanizmaları aracılığıyla hücre durumunu kontrol ederek, zaman boyunca önemli bilgilerin taşınmasını mümkün kılar. LSTM hücresinde yer alan unutma kapısı, giriş kapısı ve çıkış kapısı bilginin ne zaman korunacağını, güncelleneceğini ve çıktı olarak aktarılacağını belirler (Lim ve Zohren, 2021).

LSTM'ler, standart RNN'lere kıyasla çok daha istikrarlı öğrenme sunar ve günümüzde dil modelleme, duygu analizi, karmaşık zaman serisi tahmini gibi birçok alanlarda yoğun olarak kullanılmaktadır (Hua ve ark., 2018; Wang ve Raj, 2017)

Bir LSTM ünitesi genellikle üç ana kapı içerir: unut kapısı, giriş kapısı ve çıkış kapısı. Bu kapılar, zaman boyunca önemli bilgilerin işlenmesini ve korunmasını sağlayarak LSTM'nin sıralı veri üzerindeki öğrenme performansını artırır (Gers ve ark., 2000).



Şekil 3.5. LSTM hücresinin iç yapısı.

LSTM hücresi, Şekil 1’de gösterildiği gibi, zaman adımları boyunca bilgi akışını düzenlemek amacıyla kapı mekanizmalarıyla yapılandırılmıştır. Girdi x_t ve önceki gizli durum h_{t-1} , her biri belirli bir işlevi olan kapılardan geçirilerek işlenir. Öncelikle, unutma kapısı önceki hücre durumunun ne kadarının korunacağını belirler. Ardından, giriş kapısı ve aday hücre durumu, yeni bilginin hücreye ne ölçüde ekleneceğini tanımlar. Bu iki etkileşim sonucunda güncellenen hücre durumu, çıkış kapısı ile kontrol edilerek, çıkışı oluşturur. Bu mimari, hem önceki zaman adımlarındaki bağlamı korumaya olanak tanır hem de girdi sinyalleriyle güncel bilgileri birleştirerek dizesel örüntülerin daha doğru modellenmesini sağlar.

Bu işlem akışı, her kapının özel işleviyle denetlenir.

Giriş kapısı, mevcut girdiden hangi bilgilerin hücre durumuna eklenmesi gerektiğini belirler. Denklem 3.1a ve 3.2b’de giriş kapısına ait matematiksel gösterimler sunulmuştur.

$$i_t = \sigma(x_t U^i + h_{t-1} W^i) \quad (3.1a)$$

$$\tilde{c}_t = \tanh(x_t U^g + h_{t-1} W^g) \quad (3.1b)$$

İlk olarak, giriş kapısı için bir aktivasyon vektörü i_t hesaplanır. Bu aktivasyon, mevcut girdi ve önceki hücre çıktısının birleşimiyle belirlenen bir ağırlık matrisi ile çarpılır ve sonra bir sigmoid fonksiyonu (σ) ile 0 ile 1 arasında bir değere dönüştürülür. U ile W , sırasıyla girdi ve önceki çıktı için giriş kapısına ait ağırlık matrisleridir. $\tilde{C}$ mevcut girdinin hücre durumuna potansiyel katkısını temsil eder ve $\tanh$ fonksiyonu sayesinde bu değerler -1 ile 1 arasında normalleştirilir.

Unut kapısı, geçmiş hücre durumu C_{t-1} 'den hangi bilgilerin unutulacağını belirler. σ sigmoid fonksiyonudur ve 0 ile 1 arasında değerler alır. x_t mevcut girdidir. h_{t-1} ise önceki hücre çıktısıdır. U ve W ağırlık matrisleridir. Unut kapısına ait hesaplamalar Denklem 3.2'de verilmiştir.

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \quad (3.2)$$

Çıkış kapısı, sonraki katmana ve hücrenin kendisine gönderilecek çıktıyı belirleyen kapıdır. Denklem 3.3a ve 3.3b'de çıkış kapısına ait matematiksel gösterimler sunulmuştur.

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \quad (3.3a)$$

$$h_t = \tanh(C_t) * o_t \quad (3.3b)$$

Bu denklemlerde, o_t çıkış kapısının aktivasyonu, sigmoid fonksiyonu σ ile hesaplanır. Bu, kapının ne kadar "açık" olduğunu belirler ve böylece çıktının ne kadarının geçeceğini düzenler. Denklemden yer alan h_t elde edilen hücre çıktısıdır ve bu değer sonraki katmana iletilir. Hücre durumu, uzun süreli hafızayı temsil eder ve geçmiş durum ile mevcut girdinin etkisine göre güncellenir. Denklem 3.4'de hücre durumunun nasıl güncellendiğini gösteren hesaplamalar verilmiştir.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (3.4)$$

Bu denklemde yer alan f_t unutma kapısının çıktısıdır ve geçmiş hücre durumunun hangi kısmının korunacağını belirler. i_t giriş kapısının çıktısıdır ve mevcut girdinin hücre durumuna hangi kısmının eklenmesi gerektiğini belirler. $\tilde{C}_t$ mevcut girdinin hücre durumuna potansiyel katkısını temsil eden ve $\tanh$ fonksiyonu ile normalleştirilen değerdir (Yu ve ark., 2019).

3.3.2. Geçitli tekrarlayan birim

Geçitli Tekrarlayan Birim (Gated Recurrent Unit – GRU), Chung ve arkadaşları tarafından 2014 yılında önerilmiştir (Chung ve ark., 2014). Bu yapı, LSTM mimarisinde yer alan kapı mekanizmalarını sadeleştirerek, yalnızca güncelleme ve sıfırlama işlevlerine dayalı bir yapı sunar. GRU, her bir tekrarlayan birimin farklı zaman ölçeklerine uyum sağlayarak, bağımlılıkları adaptif biçimde modellemeyi amaçlar. LSTM'ye benzer şekilde, GRU da birim içerisindeki bilgi akışını denetleyen kapı mekanizmalarına sahiptir; ancak bu yapıda ayrı bir hücre durumu bulunmaz. Daha sade ve parametre açısından verimli mimarisiyle GRU, sıralı verilerle çalışan zaman serisi tahmini gibi uygulamalarda yaygın olarak kullanılmaktadır.

$$z_t = \sigma(x_t U^z + h_{t-1} W^z) \quad (3.5)$$

Denklem 3.5'te görüldüğü üzere güncelleme kapısı z_t , önceki bilginin ne ölçüde korunacağını ve yeni bilginin ne ölçüde ekleneceğini denetler. z_t değeri yüksek olduğunda, model büyük ölçüde önceki gizli durumu korur ve mevcut girdiden gelen bilgi daha sınırlı şekilde dahil edilir. Buna karşılık, z_t değeri düşük olduğunda yeni bilgi hücreye daha fazla entegre edilirken, geçmiş bilgi büyük ölçüde güncellenir. Bu mekanizma sayesinde GRU, her bir zaman adımında hangi bilgilerin saklanacağına veya güncelleneceğine dinamik olarak karar verebilir ve böylece karmaşık ardışık veri örüntülerini etkili biçimde öğrenebilir.

Sıfırlama kapısı, GRU mimarisinde geçmiş bilginin güncel işlem sırasında ne ölçüde dikkate alınacağını belirler. Eğer bu kapının aktivasyonu düşükse, önceki gizli durumun katkısı azaltılır ve model daha çok mevcut girdiye odaklanır. Buna karşılık, aktivasyonun yüksek olduğu durumlarda, geçmiş bilgi daha aktif biçimde korunur ve yeni girdilerle birlikte işlenir. Bu mekanizma, modelin geçmiş ve güncel bilgiyi birlikte değerlendirme esnekliğini artırır ve zaman içindeki bağlamı daha etkili şekilde modellemesini sağlar.

$$r_t = \sigma(x_t U^r + h_{t-1} W^r) \quad (3.6)$$

Denklem 3.6’da gösterilen r_t , sıfırlama kapısının aktivasyonunu ifade eder ve 0 ile 1 arasında bir değer alır. σ , girişleri bu aralığa dönüştüren sigmoid aktivasyon fonksiyonudur. x_t , mevcut zaman adımıdaki giriş vektörünü, h_{t-1} ise bir önceki zaman adımıdaki gizli durumu temsil eder. U ve W , sırasıyla giriş ve önceki gizli duruma uygulanan ağırlık matrisleridir; bu matrisler modelin eğitimi sırasında optimize edilerek öğrenilir.

$$\tilde{h}_t = \tanh(x_t U^h + (r_t * h_{t-1}) W^h) \quad (3.7)$$

GRU yapısında, yeni gizli durumun adayı $\tilde{h}_t$, mevcut giriş x_t ile önceki gizli durum h_{t-1} ’nin sıfırlama kapısı (r_t) tarafından modüle edilmiş birleşimiyle hesaplanır. Bu işlem, modelin geçmiş bilgiyi ne ölçüde dikkate alacağını kontrol eden sıfırlama kapısının etkisi altında gerçekleştirilir. Bu hesaplamalar denklem 3.7’de verilmiştir. Aday gizli durum, güncellenmiş gizli duruma katkıda bulunacak yeni bilgiyi temsil eder ve mevcut giriş ile geçmiş durum arasındaki ilişkinin ne ölçüde korunacağını belirler.

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (3.8)$$

Denklem 3.8’de gösterilen mevcut zaman adımıdaki yeni gizli durum h_t , güncelleme kapısının z_t arafından belirlenen oranlarda, önceki gizli durum ile aday gizli durumun birleşiminden oluşur. Güncelleme kapısının aktivasyonu z_t , yeni bilginin ne ölçüde hücreye dahil edileceğini belirlerken, $(1 - z_t)$ çarpanı, önceki gizli bilginin ne oranda korunacağını ifade eder. Bu işlem, $\tilde{h}_t$ ile gösterilen yeni gizli durum adayı ile önceki gizli durumun birleştirilmesi yoluyla gerçekleştirilir. Böylece GRU hücresi, zaman serisi verisinde bilginin hangi oranda saklanıp hangi oranda güncelleneceğine dinamik olarak karar verir. LSTM mimarisinde, hücre durumunun ne kadarının çıktı olarak aktarılacağını kontrol eden ayrı bir çıkış kapısı bulunur. Bu sayede model, hücre içinde tuttuğu bilginin ne kadarını dış dünyaya yansıtacağını belirleyebilir. GRU’da ise böyle bir çıkış kapısı bulunmaz. Bunun yerine, güncellenen gizli durum doğrudan çıktı olarak kullanılır. Bu yapısal fark nedeniyle GRU, her adımda içsel durum bilgisinin tamamını çıktı olarak verir; yani çıktı üzerinde ek bir filtreleme yapılmaz (Cho ve ark., 2014; Yin ve ark., 2017; Yu ve ark., 2019).

3.3.3. Çift yönlü uzun kısa süreli bellek

Klasik LSTM yalnızca geçmiş bilgileri kullanırken, BiLSTM aynı zamanda gelecekteki bağlamı da dikkate alarak her bir zaman adımında daha zengin bir temsil oluşturur. BiLSTM, klasik LSTM mimarisinin zaman dizilerinde hem geçmiş hem de gelecek bilgiyi işleyebilmesi amacıyla çift yönlü olarak geliştirilmiş halidir. Bu yapı, ilk kez Graves ve Schmidhuber (2005) tarafından önerilmiş ve medikal sinyal işleme (Imani, 2023), ses tanıma (Efanov ve ark., 2022) ve finansal zaman serileri (Lu ve ark., 2021) gibi geniş bir yelpazede birçok alana uygulanmıştır. BiLSTM, giriş zaman dizisini hem ileri ($t_0 \rightarrow t_n$) hem de geri ($t_n \rightarrow t_0$) yönde çalışarak, her zaman adımı için çift yönlü bağlamsal öğrenme gerçekleştirir.

BiLSTM mimarisinde, ileri ve geri yönlü olmak üzere iki ayrı LSTM katmanı yer alır. Böylece, yalnızca geçmiş bilgilere dayanan klasik LSTM yapısının sınırlamaları aşılmış olur.

İleri yöndeki LSTM hücresi Denklem 3.9 gibi tanımlanır:

$$h_{\text{forward}_t} = \text{LSTM}_f(x_t, h_{\text{forward}_{t-1}}) \quad (3.9)$$

İleri yönlü LSTM hücresi, t anındaki giriş vektörü x_t ile bir önceki zaman adımındaki gizli durumu $h_{\text{forward}_{t-1}}$ kullanarak güncel gizli durumu h_{forward_t} üretir. Bu yapı, klasik LSTM hücresinin ileri yönlü akışını temsil eder. Zamansal bağımlılıklar geçmişten günümüze doğru yakalanır.

Geri yöndeki LSTM hücresi Denklem 3.10'daki gibi tanımlanır:

$$h_{\text{backward}_t} = \text{LSTM}_b(x_t, h_{\text{backward}_{t+1}}) \quad (3.10)$$

Geri yönlü LSTM hücresi, yine aynı giriş vektörü x_t 'yi kullanır fakat bu kez bir sonraki zaman adımındaki gizli durum olan $h_{\text{backward}_{t+1}}$ ile işlem yapar. Böylece model gelecekteki bağlam bilgisini de dikkate alır. Bu yaklaşım, özellikle iki yönlü bağımlılıklar içeren zaman serilerinde oldukça etkilidir.

$$h_t = [h_{\text{forward}_t} ; h_{\text{backward}_t}] \quad (3.11)$$

Son aşamada, Denklem 3.11 gösterildiği üzere her iki yönde hesaplanan gizli durum vektörleri birleştirilerek, BiLSTM katmanının nihai çıktısı elde edilir. Bu mimari, klasik LSTM'nin tek yönlü bağlam işleme sınırlamasını aşarak, her zaman adımında daha zengin ve bağlamsal olarak güçlü temsiller oluşturur. Bu yapı, klasik LSTM'nin tek yönlü kısıtını aşarak daha güçlü özellik temsilleri üretir.



4. YÖNTEM VE MODEL TASARIMI

Bu bölümde, çalışmada kullanılan finansal veri seti, teknik gösterge hesaplamaları, derin öğrenme tabanlı model mimarisi, eğitim süreci ve değerlendirme kriterleri detaylı olarak açıklanmaktadır. Ayrıca, tahmin sonuçlarının analizinde kullanılan stratejik simülasyon adımlarına da yer verilmektedir.

4.1 Veri Kümesi ve Özellikler

Bu çalışmada kullanılan veri setleri, Binance kripto para borsasının sunduğu halka açık API servisleri aracılığıyla oluşturulmuştur (Binance, 2025). BTC/USDT ve ETH/USDT işlem çiftlerine ait 1 dakikalık çözünürlükte zaman serisi verileri, Binance'in resmi REST API servisi üzerinden çekilmiştir. Tüm zaman damgaları UTC formatında kaydedilmiştir.

Binance'in geçmişe dönük veri sağlamaya başladığı en erken tarih olan 17 Ağustos 2017 07:00:00 UTC itibarıyla başlatılan veri çekimi, 26 Mart 2025 23:59:00 UTC zamanına kadar olan tüm dakikalık verileri kapsamaktadır. Bu şekilde hem BTC/USDT hem de ETH/USDT veri setleri oluşturulmuştur. Her iki veri seti de tam olarak 3.992.708 satır içermekte ve 7 yıl 7 ay 10 günlük bir dönemi temsil etmektedir.

Veri setlerindeki her bir satır; ilgili dakikaya ait açılış zamanı (Open Time), kapanış zamanı (Close Time), açılış fiyatı (Open), en yüksek fiyat (High), en düşük fiyat (Low), kapanış fiyatı (Close) ve o dakikada gerçekleşen işlem hacmi (Volume) olmak üzere yedi temel sütun içermektedir. Bu yapı, fiyat dinamiklerinin yanı sıra piyasa aktivitesini de yüksek frekanslı çözünürlükle analiz etmeye olanak tanımaktadır.

Aşağıda, BTC/USDT ve ETH/USDT işlem çiftlerine ait ilk beş dakikalık zaman aralığını gösteren örnek veri satırları yer almaktadır. BTC/USDT işlem çiftine ait örnek veri görünümü Tablo 3.1'de sunulmuştur:

Tablo 4.1. BTC/USDT işlem çiftine ait örnek veri satırları

Open Time	Close Time	Open	High	Low	Close	Volume
2017-08-17 07:00:00	2017-08-17 07:00:59.999	4261.48	4261.48	4261.48	4261.48	1.775183
2017-08-17 07:01:00	2017-08-17 07:01:59.999	4261.48	4261.48	4261.48	4261.48	0.000000

2017-08-17 07:02:00	2017-08-17 07:02:59.999	4280.56	4280.56	4280.56	4280.56	0.261074
2017-08-17 07:03:00	2017-08-17 07:03:59.999	4261.48	4261.48	4261.48	4261.48	0.012008
2017-08-17 07:04:00	2017-08-17 07:04:59.999	4261.48	4261.48	4261.48	4261.48	0.140796

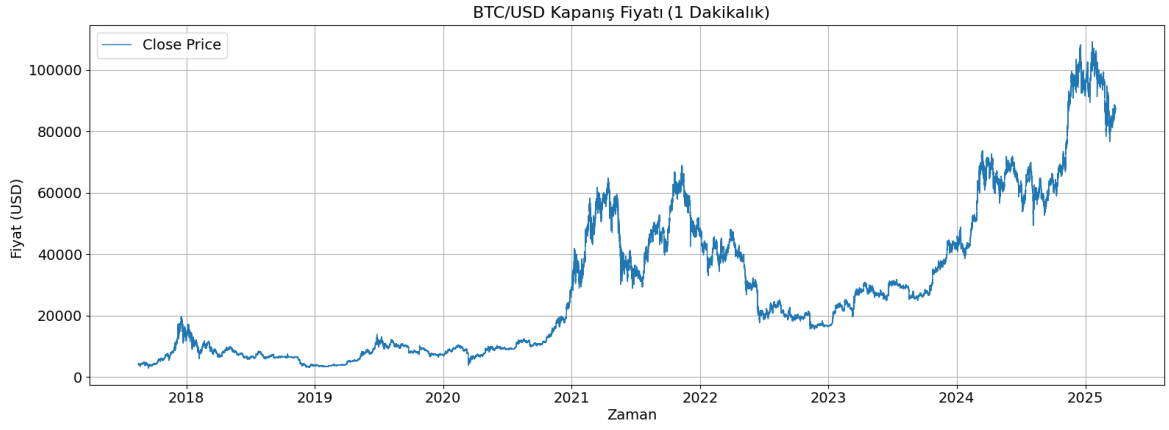
Benzer yapı ETH/USDT işlem çifti için de geçerlidir. Aşağıda ETH veri setine ait örnek satırlar sunulmuştur:

Tablo 4.2. ETH/USDT işlem çiftine ait örnek veri satırları

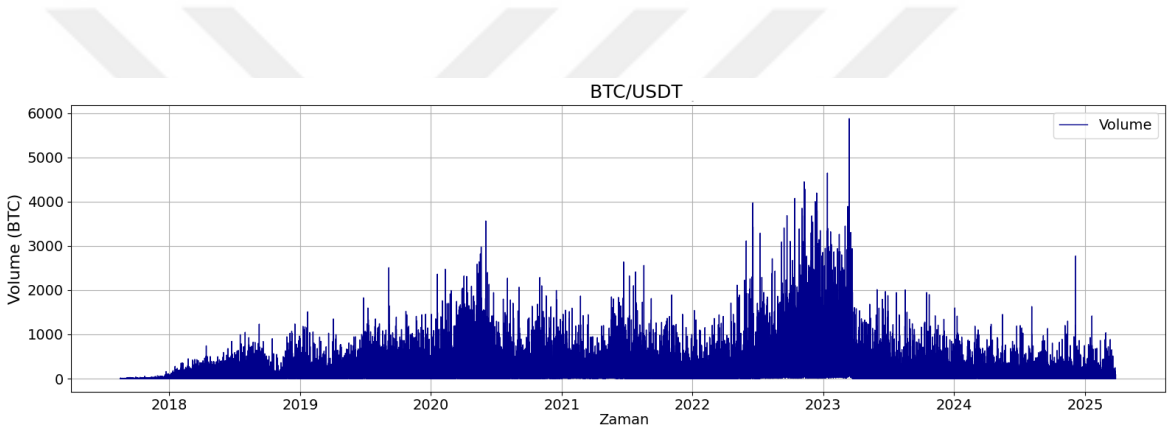
Open Time	Close Time	Open	High	Low	Close	Volume
2017-08-17 07:00:00	2017-08-17 07:00:59.999	301.13	301.13	301.13	301.13	0.42643
2017-08-17 07:01:00	2017-08-17 07:01:59.999	301.13	301.13	301.13	301.13	2.75787
2017-08-17 07:02:00	2017-08-17 07:02:59.999	300.00	300.00	300.00	300.00	0.09930
2017-08-17 07:03:00	2017-08-17 07:03:59.999	300.00	300.00	300.00	300.00	0.31389
2017-08-17 07:04:00	2017-08-17 07:04:59.999	301.13	301.13	301.13	301.13	0.23202

Her iki veri seti, zaman serisi analizleri, kısa vadeli fiyat tahmini modelleri, hacim analizleri ve teknik gösterge üretimi için sağlam ve yüksek çözünürlüklü bir kaynak teşkil etmektedir. Ayrıca, veri setleri senkronize edilerek karşılaştırmalı analizlerin tutarlılığı güvence altına alınmıştır.

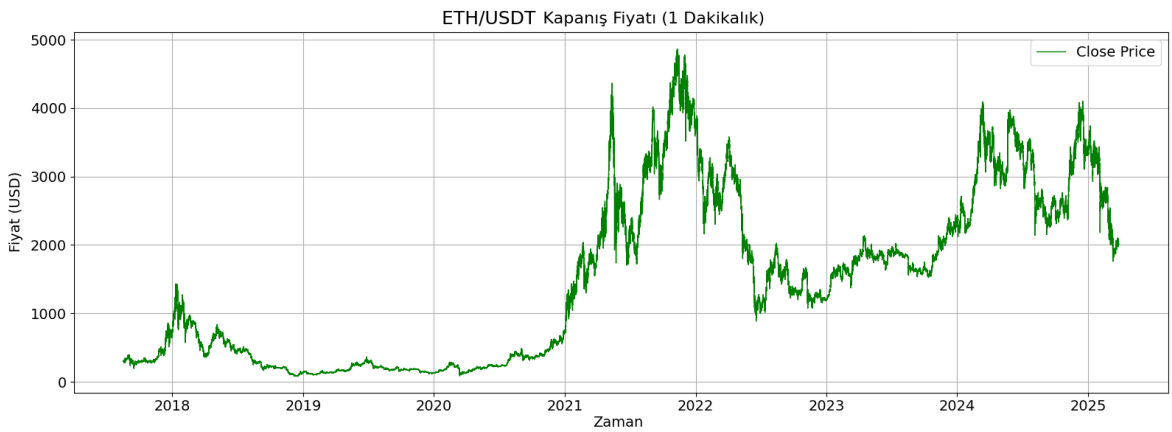
BTC/USDT ve ETH/USDT işlem çiftlerine ait 1 dakikalık çözünürlükteki kapanış fiyatı ve işlem hacmi verileri Şekil 4.1, Şekil 4.2, Şekil 4.3 ve Şekil 4.4'te sunulmuştur. Her iki varlık için 2017–2025 yılları arasında fiyat seviyelerinde yüksek volatilité gözlemlenirken, işlem hacimlerinde özellikle 2021–2022 döneminde belirgin bir yoğunluk yaşanmış, sonraki yıllarda hacimlerin görece azaldığı gözlemlenmiştir. Elde edilen bu dakikalık zaman serileri, kripto para piyasasının hem uzun vadeli eğilimlerini hem de kısa vadeli volatilité dinamiklerini modellemek açısından yüksek çözünürlüklü ve zaman uyumlu bir kaynak sunmaktadır.



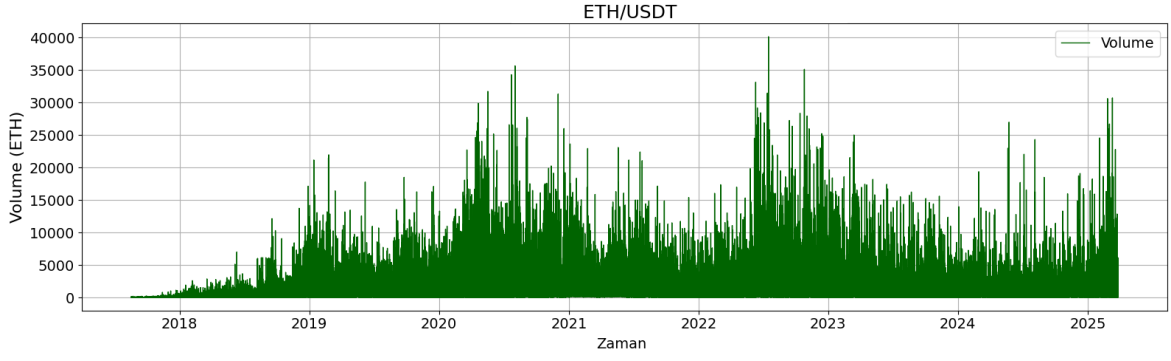
Şekil 4.1. BTC/USDT işlem çiftine ait 1 dakikalık kapanış fiyatlarının zaman serisi (2017–2025)



Şekil 4.2. BTC/USDT işlem çiftine ait 1 dakikalık işlem hacimlerinin zaman serisi (2017–2025)

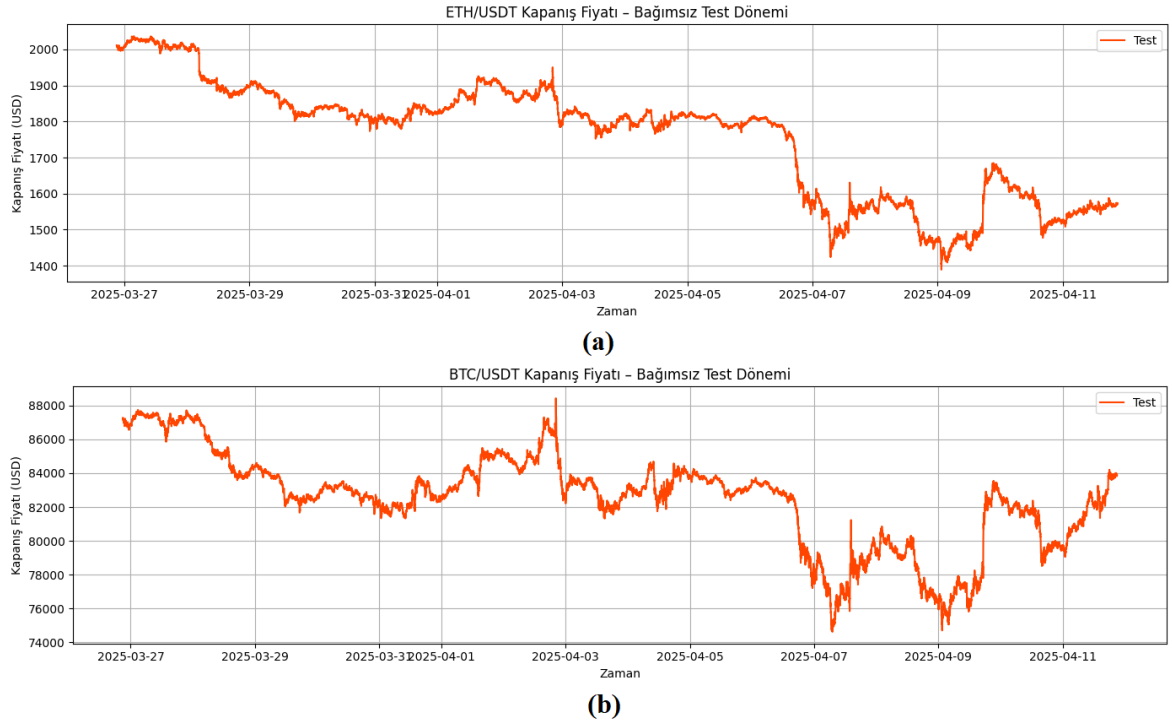


Şekil 4.3. ETH/USDT işlem çiftine ait 1 dakikalık kapanış fiyatlarının zaman serisi (2017–2025)



Şekil 4.4. ETH/USDT işlem çiftine ait 1 dakikalık işlem hacimlerinin zaman serisi (2017–2025)

Test amacıyla, eğitim sürecine dahil edilmeyen ayrı bir bağımsız test veri seti oluşturulmuştur. Bu test veri seti, 27 Mart 2025 00:00 ile 12 Nisan 2025 00:00 tarihleri arasındaki BTC/USDT ve ETH/USDT işlem çiftlerine ait dakikalık fiyat ve hacim verilerini içermektedir. Toplamda yaklaşık 23.000 dakikalık gözlem içeren bu veri bölümü, yalnızca model performansını değerlendirmek amacıyla kullanılmıştır.



Şekil 4.5. Bağımsız test dönemine ait BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerinin 1 dakikalık kapanış fiyatlarının zaman serileri (27 Mart – 12 Nisan 2025)

Şekil 4.5. 27 Mart – 12 Nisan 2025 tarihleri arasındaki bağımsız test dönemine ait BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerinin 1 dakikalık kapanış fiyatlarının zaman

serilerini göstermektedir. Eğitim sürecine dahil edilmeyen bu veri seti, yalnızca modelin genel performansını değerlendirmek amacıyla kullanılmıştır.

4.2. Yöntem Performans Değerlendirilmesi

Bu çalışmada, kripto para piyasalarındaki kısa vadeli fiyat hareketlerini tahmin etmeye yönelik olarak, farklı zaman çözünürlüklerinde (saatlik ve dakikalık) derin öğrenme tabanlı zaman serisi modelleri geliştirilmiştir. Modelleme süreci iki ana aşamadan oluşmaktadır. İlk aşamada, veri hacmini daha yönetilebilir kılmak ve analiz altyapısını sistematik biçimde kurmak amacıyla, BTC/USDT ve ETH/USDT işlem çiftlerine ait dakikalık veriler 1 saatlik zaman dilimlerine indirgenmiştir.

Bu dönüşüm işlemiyle, her bir saatlik zaman aralığı için:

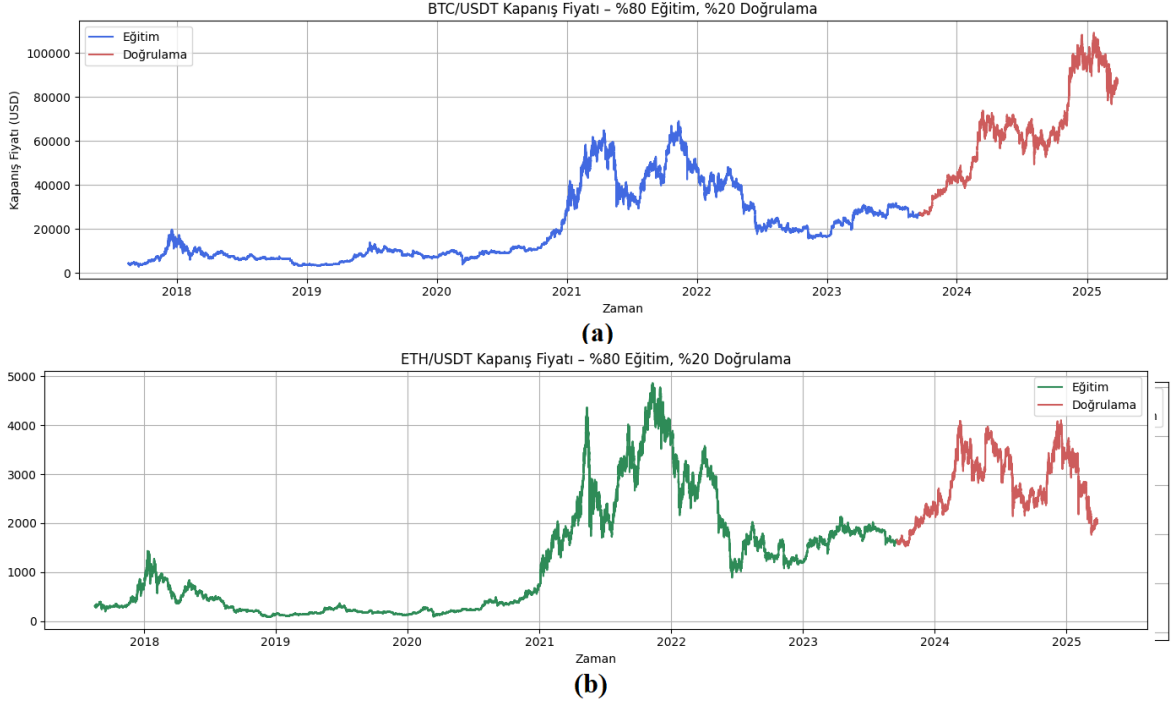
Kapanış fiyatı (Close): ilgili saatin son işlem fiyatı,

İşlem hacmi (Volume): o saate ait toplam işlem hacmi olarak hesaplanmıştır.

Elde edilen saatlik veri seti, hem BTC hem de ETH için toplam 66.561 saatlik gözlem içermekte olup, senkronize ve eksiksiz yapısıyla modelleme yaklaşımının ilk prototipleri için sağlam bir temel sunmuştur.

İkinci aşamada ise, daha yüksek çözünürlükteki öngörüler elde edebilmek amacıyla dakikalık verilerle çalışan daha gelişmiş modeller geliştirilmiştir. Bu modellerde yalnızca kapanış fiyatı ve hacim bilgileri değil, aynı zamanda teknik göstergelerle zenginleştirilmiş veri temsilleri de giriş olarak sunulmuştur.

Tüm deneysel çalışmalarda, veri setleri %80 eğitim ve %20 doğrulama olacak şekilde ayrılmıştır. Bu oran hem saatlik hem de dakikalık çözünürlükteki modellemelerde tutarlı biçimde uygulanmış; geliştirilen modellerin hem öğrenme kapasitesi hem de genelleme başarımı güvenilir biçimde değerlendirilmiştir.



Şekil 4.6. BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerine ait saatlik kapanış fiyatlarının zaman serileri

Modelleme sürecinde kullanılan eğitim ve doğrulama veri bölümlerine ait görsel zaman serisi yapısı, Şekil 4.6’da sunulmuştur. Bu grafiklerde BTC/USDT (a) ve ETH/USDT (b) işlem çiftlerinin saatlik kapanış fiyatları gösterilmekte; eğitim bölümleri sırasıyla mavi ve yeşil tonlarla, doğrulama bölümleri ise donuk kırmızı renkle vurgulanmaktadır. Bu ayırım, modelin yalnızca eğitim verisi üzerinde değil, aynı zamanda doğrulama seti üzerinden de ne derece genelleyebildiğini değerlendirmek amacıyla yapılmıştır.

4.2.1. Zaman serisi tahmin performans metrikleri

Zaman serisi tahmin modellerinin başarımı, çeşitli hata metrikleri ile değerlendirilir. Bu çalışmada kullanılan başlıca regresyon metrikleri şunlardır: Ortalama Mutlak Hata (MAE), Kök Ortalama Kare Hata (RMSE), Ortalama Mutlak Yüzde Hata (MAPE) ve Belirleme Katsayısı (R^2). Bu metriklerin her biri model başarısını farklı yönlerden ele alarak tahmin performansına dair kapsamlı bir değerlendirme sunar.

Ortalama Mutlak Hata (MAE): tahmin edilen değerler ile gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasıdır. MAE’ye ait hesaplama denklem (4.1)’de verilmiştir.

$$MAE = (1/n) \sum |y_i - \hat{y}_i| \quad (4.1)$$

Ortalama Kare Hata (MSE): MSE, tahmin hatalarının karelerinin ortalamasıdır. Büyük hataları daha fazla cezalandırdığı için, modelin büyük hatalara duyarlılığını ölçmede etkilidir. MSE'ye ait hesaplama denklem (4.2)'de verilmiştir.

$$MSE = (1/n) \sum (y_i - \hat{y}_i)^2 \quad (4.2)$$

Kök Ortalama Kare Hata (RMSE): RMSE, MSE'nin karekökü alınarak elde edilir ve hata ölçüsünü orijinal veri birimiyle aynı düzeye getirir. Bu, hataların büyüklüğünü daha anlamlı bir şekilde yorumlamayı sağlar. RMSE'ye ait hesaplama denklem (4.3)'de verilmiştir.

$$RMSE = \sqrt{(1/n) \sum (y_i - \hat{y}_i)^2} \quad (4.3)$$

Ortalama Mutlak Yüzde Hata (MAPE): MAPE, tahmin hatalarının gerçek değerlere oranla yüzde cinsinden ortalamasını hesaplar. Farklı ölçeklerdeki veriler arasında karşılaştırma yapmayı kolaylaştırır. MAPE'ye ait hesaplama denklem (4.4)'te verilmiştir.

$$MAPE = (100/n) \sum |(y_i - \hat{y}_i) / y_i| \quad (4.4)$$

Belirleme Katsayısı (R²): Modelin, verideki değişimleri ne kadar iyi açıkladığını gösteren bir ölçüdür. 1'e yaklaştıkça modelin daha başarılı olduğu anlaşılır. 0'a yakınsa model zayıftır. Negatif değerler, modelin sabit ortalama tahmininden bile kötü olduğunu gösterir. Belirleme Katsayısı'na ait hesaplama denklem (4.5)'te verilmiştir.

$$r^2 = 1 - \frac{\sum(y_i - \hat{y})}{\sum(y_i - \bar{y})} \quad (4.5)$$

Açıklamalar:

y_i : Gerçek değer

$\hat{y}$: Model tarafından tahmin edilen değer

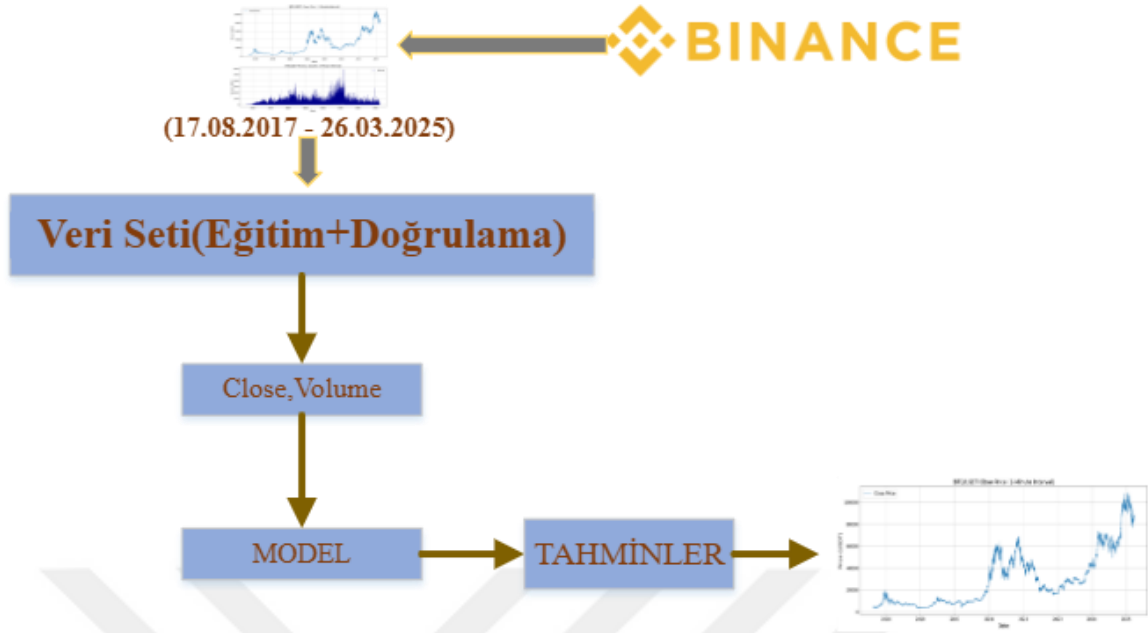
$\bar{y}$: Gerçek değerlerin aritmetik ortalaması

n : Gözlem sayısı

Zaman serisi tahmin modellerinin performansını değerlendirirken, tek bir metriğe dayanmak yerine birden fazla metriği birlikte kullanmak daha kapsamlı bir analiz sağlar. Her bir metrik, modelin farklı yönlerini değerlendirir ve bu sayede modelin güçlü ve zayıf yönleri daha iyi anlaşılabilir (Chicco ve ark., 2021).

4.3 Model Tasarımı

Bu çalışmada, saatlik çözünürlükteki BTC/USDT verileri kullanılarak üç farklı derin öğrenme mimarisi (LSTM, GRU ve BiLSTM) eğitilmiştir. Modellerin amacı, geçmiş 24 saate ait kapanış fiyatı ve işlem hacmi bilgisini kullanarak 25. saatin kapanış fiyatını tahmin etmektir. Yönteme ait gösterim Şekil 4.7 de gösterilmektedir.



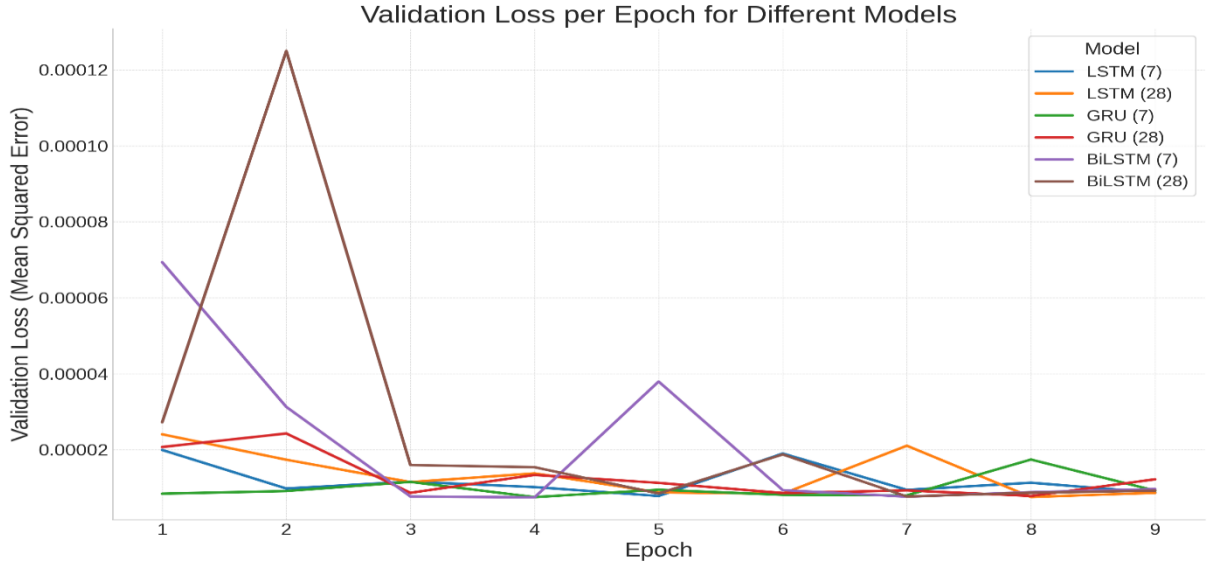
Şekil 4.7. Modelleme sürecinin genel iş akışını gösteren diyagram

Bu çalışmada derin öğrenme tabanlı kripto para fiyat tahmini için üç farklı mimari (LSTM, GRU ve BiLSTM) test edilmiştir. Modelleme süreci aşamalı olarak yapılandırılmış; ilk adımda yalnızca kapanış fiyatı ve işlem hacmi içeren temel modeller değerlendirilmiş, ardından farklı batch boyutları ve dört ayrı kayıp fonksiyonu (MSE, MAE, Huber ve Weighted Loss) ile deneyler gerçekleştirilmiştir. Bu deneysel süreç, modelin genel doğruluğunu ve genelleme kapasitesini optimize etmeye yönelik kapsamlı bir mimari karşılaştırma sunmuştur.

4.3.1 Model mimari karşılaştırması

Her bir mimari, iki farklı mini-batch boyutu (batch_size=7 ve batch_size=28) ile eğitilmiş; böylece hem mimari seçiminin hem de batch boyutunun performansa etkisi gözlemlenmiştir. Model eğitiminde mean squared error (MSE) kayıp fonksiyonu kullanılmış, maksimum epoch sayısı 100 olarak belirlenmiştir. Aşırı öğrenmeyi önlemek amacıyla erken durdurma mekanizması uygulanmıştır.

Eğitim sırasında modellerin doğrulama veri seti üzerindeki performansları sürekli takip edilmiştir. Şekil 7’de, farklı mimari ve batch boyutlarının ilk 9 epoch boyunca doğrulama kaybındaki değişimleri sunulmuştur.



Şekil 4.8. Farklı mimariler ve batch size değerlerine göre epoch bazlı doğrulama kaybı değişimi

Şekil 4.8’de görüldüğü üzere:

- **BiLSTM:** BiLSTM modelleri, yüksek doğrulama başarımları ile dikkat çekmiş; özellikle batch_size=28 ile daha kararlı ve dengeli bir öğrenme süreci gerçekleştirmiştir.
- **GRU:** GRU modelleri, doğrulama kaybını erken epoch’larda hızlı şekilde azaltarak etkili bir öğrenme performansı sergilemiş; bu mimaride batch_size=28 görece daha başarılı sonuçlar vermiştir.
- **LSTM:** LSTM modelleri, doğrulama kaybı açısından genel olarak dengeli bir öğrenme süreci sergilemiş; özellikle batch_size=28 ile daha düşük kayıp değerlerine ulaşmıştır. Ancak, öğrenme hızı bakımından diğer mimarilere göre daha yavaş bir gelişim göstermiştir.

Tablo 4.7’de, farklı derin öğrenme mimarilerinin (LSTM, GRU, BiLSTM) ve farklı batch boyutlarının (7 ve 28) performans karşılaştırmaları hem normalize edilmiş fiyatlar hem de gerçek BTC fiyatı üzerinden hesaplanan metrikler ile sunulmuştur. Her bir model için RMSE, MAE ve R^2 değerleri ayrı ayrı raporlanmış, böylece mimari ve batch boyutuna bağlı performans değişimleri karşılaştırmalı olarak değerlendirilmiştir.

Tablo 4.3. Farklı derin öğrenme modelleri ve batch size değerlerine göre fiyat tahmin performans metrikleri

Model	Batch Size	RMSE (USD)	MAE (USD)	RMSE (USD)	MAE (USD)	R ²
LSTM	7	968.70	703.59	0.0092	0.0067	0.9967
LSTM	28	652.88	468.49	0.0062	0.0044	0.9985
GRU	7	977.19	723.13	0.0093	0.0069	0.9967
GRU	28	576.24	409.82	0.0055	0.0039	0.9988
BiLSTM	7	640.41	459.05	0.0061	0.0044	0.9986
BiLSTM	28	498.20	334.50	0.0047	0.0032	0.9991

Genel olarak, modeller yüksek doğruluk seviyelerine ulaşmış olup, R² değerleri tüm senaryolarda %99 seviyesinin üzerinde gerçekleşmiştir. Bu, modellerin fiyat değişimlerini büyük ölçüde açıklayabildiğini göstermektedir.

- BiLSTM modeli, batch_size=28 ile en iyi genel performansı göstermiştir. Bu model 498.20 USD RMSE ve %99.91 R² skoruyla, hem hata oranını minimuma indirmiş hem de modelin tahminleri ile gerçek değerler yüksek düzeyde örtüşmektedir.
- GRU modelleri, batch_size=28 ile daha düşük RMSE ve MAE değerleri sunarak başarılı bir performans sergilemiş; buna karşın batch_size=7 ile elde edilen sonuçlar görece daha yüksek hata oranlarına işaret etmektedir.
- LSTM mimarileri, özellikle batch_size=28 ile iyi sonuçlar üretmiş olsa da diğer mimarilere kıyasla daha yüksek hata seviyelerine sahiptir.

Modelleme sürecinde farklı batch size değerlerinin model başarımı üzerindeki etkileri de incelenmiştir. batch_size=28 kullanımı, özellikle parametre güncelleme frekansını azaltarak daha dengeli öğrenme eğrileri üretmiş ve hem doğrulama kaybı hem de doğrulama metrikleri açısından daha tutarlı sonuçlar sağlamıştır. 28 değeri, veri setinin haftalık (4 haftalık pencere) döngüsel yapısıyla da uyumlu olduğu için zaman serisi yapısına daha anlamlı bir biçimde entegre olmuştur. Bu bağlamda, modelin genel genelleme yeteneğini artırdığı gözlemlenmiş ve sonraki tüm deneylerde batch_size=28 tercih edilmiştir.

4.3.2. Farklı kayıp fonksiyonlarının model performansına etkisi

Uygulamanın önceki aşamalarında, zaman serisi tahmini görevine yönelik olarak LSTM, GRU ve BiLSTM gibi farklı derin öğrenme mimarileri, öncelikle Mean Squared Error (MSE) kayıp fonksiyonu ile değerlendirilmiştir. Bu ilk değerlendirme, modellerin genel öğrenme kapasiteleri ve kısa vadeli fiyat hareketlerini yakalama potansiyelleri hakkında temel bir referans sunmuştur.

Bu bölümde ise yalnızca mimari yapıların değil; aynı zamanda kayıp fonksiyonlarının da model performansı üzerindeki etkisi sistematik biçimde analiz edilmiştir. Her bir model aşağıda açıklanan dört farklı kayıp fonksiyonu ile eğitilmiş ve test edilmiştir:

- Ortalama Kare Hatası (Mean Squared Error - MSE), MSE, tahmin edilen değerler ile gerçek değerler arasındaki farkların karelerinin ortalamasını alarak hesaplanır. Bu yöntem, büyük hataları daha fazla cezalandırdığı için, modelin büyük sapmaları minimize etmesine yardımcı olur. MSE'ye ait hesaplama denklem (4.6)'te verilmiştir.

$$MSE = \frac{1}{n} \sum (y_i - \hat{y}_i)^2 \quad (4.6)$$

- Ortalama Mutlak Hata (Mean Absolute Error - MAE), MAE, tahmin edilen değerler ile gerçek değerler arasındaki mutlak farkların ortalamasını alarak hesaplanır. Bu yöntem, tüm hataları eşit şekilde cezalandırır ve özellikle uç değerlerin etkisini azaltmada etkilidir. MAE'ye ait hesaplama denklem (4.7)'te verilmiştir.

$$MAE = \frac{1}{n} \sum |y_i - \hat{y}_i| \quad (4.7)$$

- Huber Kayıp Fonksiyonu (Huber Loss), MSE ve MAE'nin avantajlarını birleştirerek, küçük hatalar için kare kaybı, büyük hatalar için ise mutlak kayıp kullanır. Bu, modelin hem duyarlılığını hem de sağlamlığını artırır. Delta (δ) parametresi, hangi hata büyüklüğünden itibaren geçiş yapılacağını belirler (Chollet, 2021). Huber kayıp fonksiyonuna ait hesaplama denklem (4.8)'de verilmiştir.

$$\text{HuberLoss} = \begin{cases} \frac{1}{2}(y_i - \hat{y}_i)^2 & \text{if } |y_i - \hat{y}_i| \leq \delta \\ \delta \left(|y_i - \hat{y}_i| - \frac{1}{2}\delta \right) & \text{otherwise} \end{cases} \quad (4.8)$$

- Ağırlıklı Kayıp Fonksiyonu Regresyon problemlerinde büyük hata örneklerine odaklanmak amacıyla literatürde çeşitli dinamik ağırlıklı loss fonksiyonları geliştirilmiştir (Rudy ve Sapsis, 2021). Bu çalışmada ise, finansal tahmin bağlamına özel olarak, doğrudan tahmin hatasının büyüklüğüne bağlı bir ağırlıklandırma yaklaşımı benimsenmiştir. Önerilen yöntemde, klasik MSE fonksiyonu, hata büyüklüğüne göre lineer olarak artan bir katsayı ile çarpılarak büyük sapmaların daha güçlü şekilde cezalandırılması sağlanmıştır. Bu yapı sayesinde, özellikle yüksek volatiliteye sahip dönemlerdeki fiyat hareketleri modele daha etkili biçimde aktarılmakta ve tahmin doğruluğu kritik seviyelerde artırılmaktadır. Ağırlıklı kayıp fonksiyonuna ait hesaplama denklem (4.9)'da verilmiştir.

$$\text{Weighted Loss} = \frac{1}{n} \sum [(y_i - \hat{y}_i)^2 \cdot (1 + \lambda |y_i - \hat{y}_i|)] \quad (4.9)$$

Burada kullanılan denklemlerde:

y_i : Gerçek (gözlemlenen) değerleri,

$\hat{y}_i$: Model tarafından tahmin edilen değerleri,

n : Gözlem sayısını,

$\lambda=0.1$, büyük hataların klasik MSE'ye göre daha fazla cezalandırılmasını sağlamak amacıyla belirlenmiştir.

δ : Huber kaybı için hata duyarlılık eşikini, ifade etmektedir.

Bu dört farklı kayıp fonksiyonu, modelin hata büyüklüğüne karşı farklı tepkilerini değerlendirme imkânı sunmuştur. Özellikle Weighted Loss, büyük hataları daha fazla cezalandırarak, klasik MSE'ye kıyasla modelin öğrenme sürecini daha etkili biçimde yönlendirmiştir.

Tüm deneylerde batch_size=28 sabit tutularak, eğitim koşulları standartlaştırılmıştır. Modeller aynı veri kümesi üzerinde, aynı hiperparametrelerle eğitilmiş ve performanslar

hem normalize edilmiş metrikler (ölçekten bağımsız hata oranları) hem de gerçek BTC fiyatları (USD cinsinden RMSE, MAE, R^2) üzerinden karşılaştırmalı olarak analiz edilmiştir.

Tablo 4.4. Farklı kayıp fonksiyonlarıyla eğitilen derin öğrenme modellerinin performans karşılaştırması (LSTM, GRU ve BiLSTM)

Model	Loss Function	RMSE (USD)	MAE (USD)	RMSE (Norm)	MAE (Norm)	R^2
LSTM	MSE	652.88	468.49	0.006194	0.004445	0.998521
	Weighted Loss	470.13	304.33	0.004460	0.002887	0.999233
	MAE	593.81	421.38	0.005634	0.003998	0.998777
	Huber	600.99	424.64	0.005702	0.004029	0.998747
GRU	MSE	576.24	409.82	0.005467	0.003888	0.998848
	Weighted Loss	619.16	442.81	0.005874	0.004201	0.998670
	MAE	595.36	425.61	0.005649	0.004038	0.998770
	Huber	816.11	599.66	0.007743	0.005689	0.997689
BiLSTM	MSE	498.20	334.50	0.004727	0.003174	0.999139
	Weighted Loss	492.45	327.32	0.004672	0.003105	0.999159
	MAE	501.14	322.84	0.004755	0.003063	0.999129
	Huber	589.62	393.50	0.005594	0.003733	0.998794

Farklı kayıp fonksiyonlarının LSTM, GRU ve BiLSTM mimarileri üzerindeki etkilerini karşılaştıran analiz sonuçlarına göre, en başarılı sonuçlar LSTM modeli ile Weighted Loss kullanılarak elde edilmiştir. Bu kombinasyon, hem normalize hem de gerçek fiyatlar üzerinden hesaplanan tüm metriklerde (RMSE, MAE ve R^2) en düşük hata ve en yüksek açıklayıcılık değerlerine ulaşmıştır. Özellikle yön bilgisine duyarlı kayıpların, LSTM'nin zaman serisindeki trend yönünü öğrenme kabiliyetiyle yüksek uyum gösterdiği görülmektedir.

GRU modeli ise en iyi performansı klasik Mean Squared Error (MSE) fonksiyonu ile sergilemiştir. MAE, Huber ve Weighted Loss fonksiyonları ile başarı düzeyi belirgin biçimde düşmüş, özellikle Huber Loss kullanımı model doğruluğunu anlamlı şekilde azaltmıştır. Bu durum, GRU'nun yön odaklı veya robust kayıplara karşı daha az esnek olabileceğini düşündürmektedir.

BiLSTM modeli genel olarak tüm kayıp fonksiyonları altında istikrarlı ve yüksek doğrulukta sonuçlar üretmiştir. Özellikle Weighted Loss ve MAE kullanımı, BiLSTM mimarisinin düşük hata ve yön tahmini başarımında üstün bir denge sunduğunu göstermiştir. Ancak genel kıyaslama yapıldığında, LSTM + Weighted Loss kombinasyonu küçük farklarla daha yüksek genel başarı sağlamıştır.

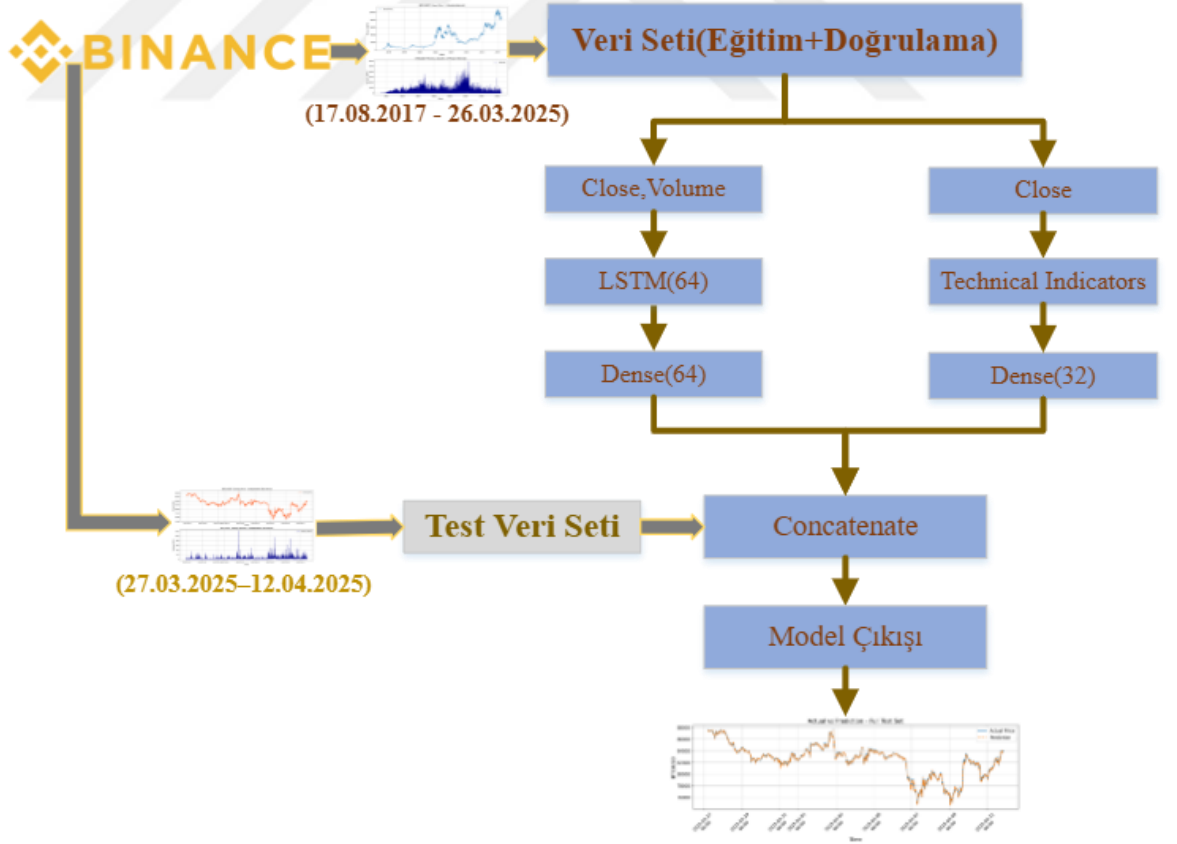
Bu değerlendirmeler doğrultusunda, uygulamanın ilerleyen aşamalarında LSTM mimarisi ve Weighted Loss kayıp fonksiyonu tercih edilmiştir. Ayrıca, batch_size parametresi olarak 28 değeri kullanılmış ve bu yapı ile modelin hem öğrenme stabilitesi hem de genelleme başarımı optimize edilmiştir.



5. TEKNİK GÖSTERGELERİN ENTEGRE EDİLDİĞİ MODELLER

Modelleme sürecinin bu aşamasında, LSTM tabanlı yapının tahmin gücünü artırmak amacıyla giriş verileri teknik göstergelerle zenginleştirilmiş ve modelin daha geniş bir bağlamda öğrenme yapması sağlanmıştır. Bu kapsamda, kapanış fiyatı ve işlem hacmi gibi temel zaman serisi verilerine ek olarak, piyasa eğilimlerini, momentumu ve volatilitiyi temsil eden çeşitli teknik analiz göstergeleri de modele dahil edilmiştir.

Bu teknik göstergeler, model yapısına doğrudan entegre edilerek farklı bilgi kaynaklarının birleşik biçimde değerlendirilmesine olanak tanımıştır. Model girişlerinde yalnızca zaman serisi özellikleri değil, aynı zamanda teknik analiz göstergeleri de kullanılmış; böylece fiyat dinamikleri ile piyasa sinyalleri birlikte işlenmiştir. Zaman serisine ait kapanış fiyatı ve hacim bilgileri LSTM katmanında, teknik göstergeler ise ayrı bir yoğun katmanda değerlendirilmiş ve iki farklı bilgi akışı Concatenate katmanı ile birleştirilmiştir. Geliştirilen bu model yapısı, farklı zaman çözünürlüklerinde ve pencere uzunluklarında test edilerek esnek bir yapı ortaya koymuştur.



Şekil 5.1. Modelleme sürecine ait veri akış diyagramı

Bu çalışmada önerilen derin öğrenme modeli, kripto para zaman serisinin tahminine yönelik bütünlük bir yapı sunmaktadır. Şekil 5.1.'de modelleme sürecine ait genel veri akışı, eğitim ve test ayrımı ile birlikte tahmin-değerlendirme adımları ayrıntılı biçimde görselleştirilmiştir. Saatlik çözünürlükteki BTC/USDT verileri, 17.08.2017–26.03.2025 tarih aralığında eğitim ve doğrulama amacıyla kullanılmış; 27.03.2025–12.04.2025 arası ise test süreci için ayrılmıştır. Model, giriş olarak geçmiş 24 saate ait kapanış fiyatı ve işlem hacmini alırken, aynı zamanda bu verilere dayalı olarak hesaplanan teknik analiz göstergeleriyle de zenginleştirilmiştir.

Girdi verileri iki ayrı akışta işlenmiştir. Birinci akışta kapanış fiyatı ve hacim bilgisi LSTM(64) ve ardından Dense(64) katmanında modellenmiştir. İkinci akışta ise aynı kapanış fiyatından elde edilen teknik göstergeler (trend, momentum ve volatilité göstergeleri) Dense(32) katmanı ile işlenmiştir. Bu iki bilgi akışı, Concatenate katmanında birleştirilerek modelin tahmin çıktısı üretilmiştir. Model, daha önce hiç görmediği test verileri ile çalıştırılmış ve tahmin edilen değerler gerçek BTC kapanış fiyatlarıyla karşılaştırılmıştır.

Model performansı, RMSE, MAE ve R^2 gibi yaygın regresyon metrikleriyle değerlendirilmiştir. Bu süreç, modelin genel tahmin gücünü ve farklı bilgi kaynaklarını bütünlük olarak sunduğu katkıyı sayısal olarak ortaya koymak amacıyla yapılandırılmıştır.

5.1 Saatlik Çözünürlükte Teknik Göstergelerle Zenginleştirilmiş LSTM Modeli

Bu aşamada, önceki bölümde en başarılı performansı sunan LSTM + Weighted Loss + batch_size=28 kombinasyonu temel alınarak, teknik göstergelerle zenginleştirilmiş bir model geliştirilmiştir. Modelin girdisi, saatlik çözünürlükte yapılandırılmış 24 saatlik bir pencereye ait kapanış fiyatı ve işlem hacmi verilerinden oluşmakta; bu giriş 24x2 boyutunda zaman serisi verisi şeklinde LSTM katmanına sunulmuştur. Buna ek olarak, her örnek için hesaplanan 11 farklı teknik gösterge sabit bir öznitelik vektörü olarak modele ikinci bir giriş şeklinde dahil edilmiştir. Teknik göstergeler zaman bağımlılığı taşımayan sabit bir yapı ile Dense(32) katmanında işlenmiş, zaman serisi verisinden gelen LSTM çıkışı ile Concatenate katmanında birleştirilmiştir. Modelin tahmin hedefi, giriş penceresini takip eden 25. saatin kapanış fiyatı olup, saatlik çözünürlükte tek adımlı bir fiyat tahmini görevi gerçekleştirilmiştir. Eğitim sürecinde %80 eğitim – %20 doğrulama oranı korunmuş, batch_size=28 (4 hafta) ile dört haftalık örneklem temsili sağlanmış ve erken durdurma

mekanizması uygulanmıştır. Bu yapı, teknik göstergelerden gelen bağlamsal bilgiyi LSTM tabanlı zaman serisi modellemesiyle bütünleştirerek, saatlik bazda kısa vadeli fiyat tahminlerinin doğruluğunu artırmayı hedeflemiştir.

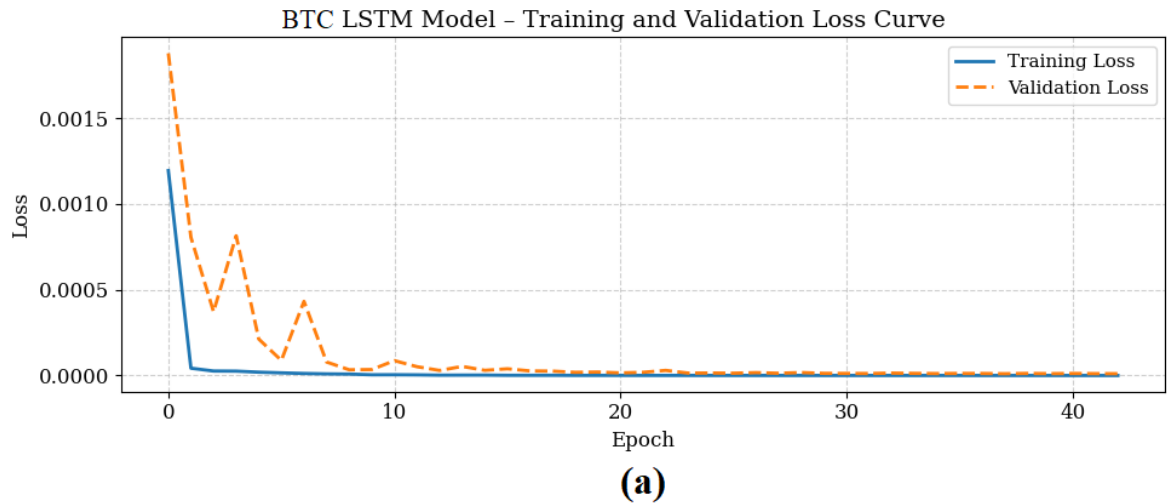
5.1.1. Modelin eğitim süreci ve doğrulama performansı

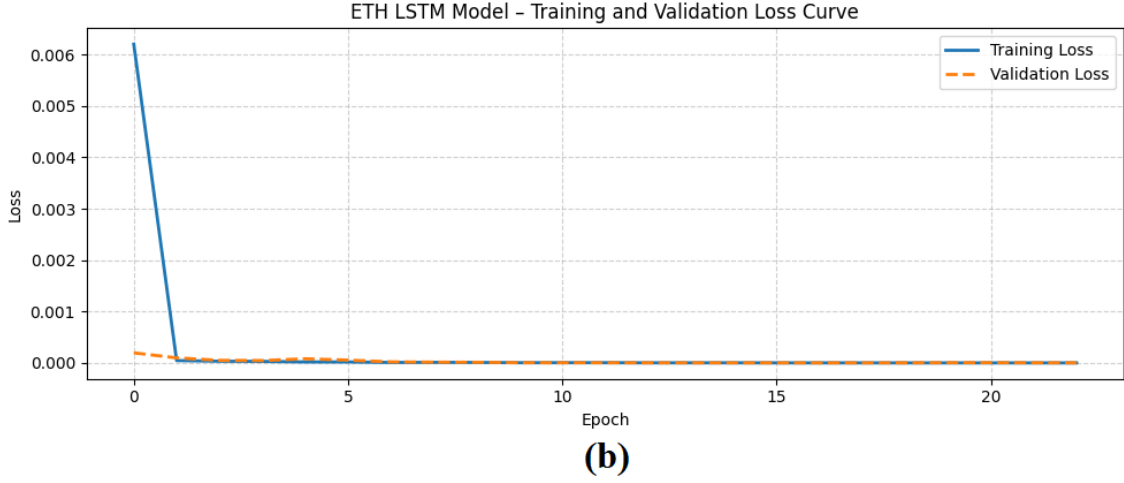
Eğitim süreci boyunca hem Bitcoin (BTC) hem de Ethereum (ETH) için doğrulama kaybında istikrarlı bir azalma gözlemlenmiştir. Her iki modelde de erken durdurma mekanizması etkin biçimde çalışmış, böylece modelin aşırı öğrenmeye eğilimi engellenmiştir.

- BTC modeli yaklaşık 43. epoch civarında minimum doğrulama kaybına ulaşarak öğrenme sürecini başarıyla tamamlamıştır.
- ETH modeli ise daha erken bir noktada hızlı şekilde doğrulama kaybını düşürerek stabilize olmuştur.

Modelin eğitim ve doğrulama kayıplarını gösteren öğrenme eğrileri Şekil 5.2 (a) (BTC) ve Şekil 5.2 (b) (ETH) karşılaştırmalı olarak sunulmuştur. Bu grafikler, modelin hem stabil hem de genellenebilir bir öğrenme stratejisi izlediğini göstermektedir.

. BTC'ye ait kayıp eğrileri öğrenme sürecindeki istikrarı ve modelin aşırı öğrenmeden kaçındığını gösterirken, ETH'ye ait kayıp eğrileri modelin farklı veri dağılımları üzerinde de benzer bir öğrenme davranışı sergilediğini ortaya koymaktadır. Her iki durumda da erken durdurma mekanizmasının etkili çalıştığı ve aşırı uyumun önüne geçildiği görülmektedir.





Şekil 5.2. Modelin eğitim ve doğrulama kayıplarını gösteren öğrenme eğrileri (a) (BTC) ve (b) (ETH) göstermektedir.

Modelin eğitim süreci boyunca elde ettiği doğrulama sonuçları, hem BTC hem de ETH için oldukça yüksek doğruluk düzeyleriyle tamamlanmıştır. Tablo 5.1’de, eğitim sırasında doğrulama veri seti üzerinden elde edilen performans metrikleri hem normalize edilmiş değerler hem de gerçek USD fiyatları cinsinden özetlenmiştir.

Tablo 5.1 Teknik göstergelerle zenginleştirilmiş LSTM modelinin BTC ve ETH doğrulama veri seti üzerindeki performans metrikleri

Metrik	BTC (Normalize)	BTC (USD)	ETH (Normalize)	ETH (USD)
MSE	0.000005	23,069.58	0.000001	18.69
RMSE	0.002311	151.89	0.000907	4.32
MAE	0.001121	73.69	0.000632	3.01
R ²	0.999951		0.999956	

Modelin doğrulama seti üzerinde gösterdiği yüksek başarı, öğrenme sürecinin etkinliğini ve kullanılan mimarinin tahmin görevine uygunluğunu ortaya koymaktadır. BTC için RMSE değeri 151.89 USD, ETH için ise yalnızca 4.32 USD olarak ölçülmüştür. Bu sonuçlar, modelin hem yüksek fiyat dalgalanmalarına sahip hem de daha düşük volatilitte gösteren varlıklarda etkili biçimde çalışabildiğini ortaya koymaktadır.

5.1.2. Test veri seti üzerindeki performans karşılaştırmaları

Modelin doğrulama seti üzerindeki yüksek başarımının ardından, eğitim sürecinden tamamen bağımsız olan test veri seti üzerinde de performans değerlendirmesi gerçekleştirilmiştir. Bu test seti, modelin daha önce hiç görmediği saatlik verileri içermekte olup, modelin eğitim dışında kalan veriler üzerindeki başarısını değerlendirmek amacıyla kullanılmıştır.

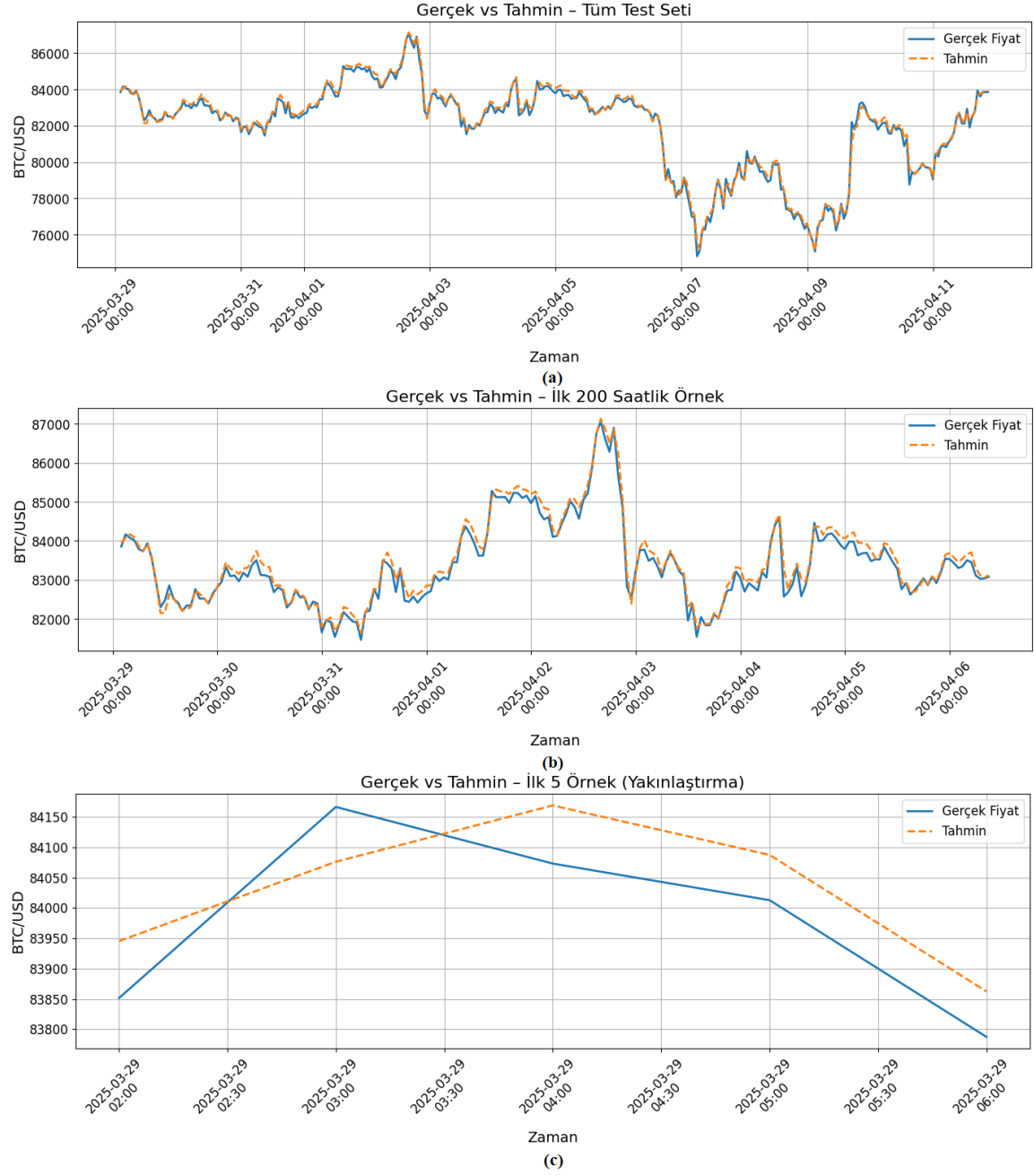
Tablo 5.2.'de BTC ve ETH için test seti üzerinden elde edilen performans metrikleri hem normalize edilmiş değerler hem de gerçek USD fiyatları cinsinden sunulmuştur. BTC için RMSE değeri 205.29 USD olarak hesaplanırken, ETH için bu değer yalnızca 3.41 USD düzeyinde gerçekleşmiştir. MAPE değerlerinin %0.20'nin altında kalması, modelin tahminlerinin göreceli hata açısından oldukça düşük olduğunu ve yüksek hassasiyet sunduğunu ortaya koymaktadır.

Tablo 5.2. Test verisi üzerinde LSTM modelinin performans metrikleri

Metrik	BTC (Normalize)	BTC (USD)	ETH (Normalize)	ETH (USD)
RMSE	0.003124	205.29	0.000716	3.41
MAE	0.002331	153.20	0.000480	2.29
MAPE	0.19%	0.19%	0.14%	0.14%
R ²	0.992987	0.992987	0.999431	0.999431

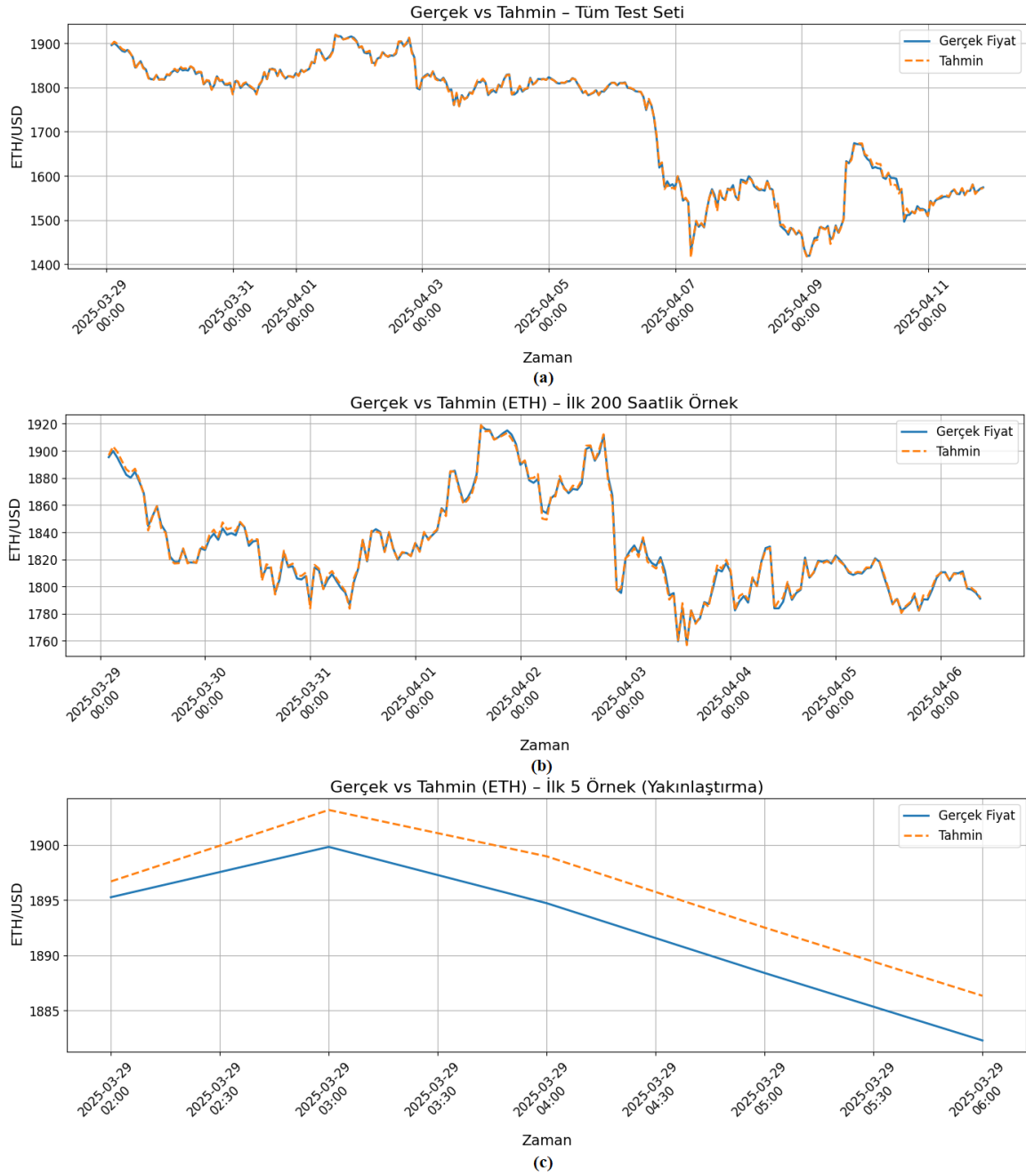
Test verisi üzerinde yapılan görsel analizler de bu metriklerle tutarlılık göstermektedir. Tablo 5.2'te BTC için modelin tahmin ettiği fiyatlar ile gerçek fiyatlar arasındaki uyum zaman serisi üzerinde sunulmuştur. Şekil 5.3-a, tüm test setini kapsarken; Şekil 5.3-b, ilk 200 saate ait daha detaylı bir karşılaştırma sunmaktadır. Şekil 5.3-c ise ilk 5 saate ait yakınlştırılmış bir görünümle modelin kısa vadeli dalgalanmalara karşı ne derece duyarlı olduğunu göstermektedir.

Benzer şekilde, Şekil 3.4'te modelin ETH fiyat tahminleri görsel olarak değerlendirilmiştir. Her iki görsel dizisi birlikte incelendiğinde, modelin hem yüksek hacimli (BTC) hem de nispeten daha düşük volatiliteye sahip (ETH) piyasalarda etkili şekilde çalıştığı görülmektedir.



Şekil 5.3. Modelin BTC/USDT işlem çifti için bağımsız test veri seti üzerindeki performansı: (a) Tüm test süresine ait gerçek ve tahmin fiyatlarının zaman serisi üzerindeki karşılaştırması. (b) İlk 200 saatlik örneğe ait detaylı karşılaştırma. (c) İlk 5 saatlik fiyat hareketlerine ait yakınlaştırılmış görünüm

Grafikler, modelin kısa ve orta vadeli fiyat eğilimlerini yüksek doğrulukla tahmin edebildiğini ve kısa vadeli dalgalanmalara karşı duyarlı olduğunu ortaya koymaktadır.



Şekil 5.4. Modelin ETH/USDT işlem çifti için bağımsız test veri seti üzerindeki performansı:(a) Tüm test süresine ait gerçek ve tahmin fiyatlarının zaman serisi üzerindeki karşılaştırması. (b) İlk 200 saatlik örneğe ait detaylı karşılaştırma. (c) İlk 5 saatlik fiyat hareketlerine ait yakınlaştırılmış görünüm

ETH fiyatları üzerinde de modelin hem genel trendleri hem de kısa vadeli mikro hareketleri isabetli biçimde yakalayabildiği gözlemlenmektedir.

5.2. Dakikalık Zaman Serileri İçin Teknik Göstergelerle Desteklenmiş LSTM Tabanlı Tahmin Modeli

Bu bölüm, önceki kısımlarda açıklanan saatlik çözünürlükteki modelleme yaklaşımlarının devamı niteliğinde olup, aynı veri işleme prensiplerinin dakikalık verilerle genellenmesini konu almaktadır.

Aşağıda iki farklı senaryo altında dakikalık zaman serisi verileriyle geliştirilen tahmin modelleri sunulmuştur:

5.2.1. Dakikalık girişle gerçekleştirilen saatlik tahminler

Dakikalık zaman serisinin yüksek çözünürlüğü ve değişkenliği nedeniyle, bu bölümde yalnızca temel LSTM yapısı değil, aynı zamanda daha yüksek gizli birim sayısına sahip konfigürasyon (LSTM(128)) ve yerel zaman pencerelerinden özellik çıkarımı yapabilen Conv1D tabanlı modeller de değerlendirilmiştir. Bu yapıların kullanımı, yüksek frekanslı veride daha karmaşık örüntülerin yakalanabilmesi ve modelin genel öğrenme kapasitesinin artırılabilmesi amacıyla tercih edilmiştir. Dakikalık veri yoğunluğu nedeniyle modellerin hem performans hem de hesaplama verimliliği açısından optimize edilebilmesi hedeflenmiş, teknik göstergeler ise saatlik çözünürlükte hesaplanarak ek girdi olarak modele sunulmuştur. Model girişi saatlik veri girişi ile benzer özellikler göstermektedir. Zaman serisi girişi dakika bazlı bir günlük veriyi içermektedir. Teknik özellikler saatlik modeller ile aynı şekilde giriş olarak verilmiştir.

- Zaman serisi girişi: 1440x2 boyutunda dakikalık fiyat ve hacim verisi.
- Teknik göstergeler: Her saat başında hesaplanan 11 adet teknik gösterge: DayOfWeek, Close_Change, SMA_30, RSI, MACD, EMA_72, Momentum_24, ROC_24, STD_24, Bollinger_Width, Zscore_Close. Bu göstergeler, saatlik olarak hesaplanmış ve saat başlarında ilgili zaman serisi penceresiyle eşlenmiştir.

Her iki giriş ayrı alt yapılarda işlenmiş, ardından Concatenate katmanında birleştirilerek regresyon çıktısına yönlendirilmiştir.

Bu deneysel çalışmada üç farklı model mimarisi test edilmiştir:

- LSTM(64): Tek katmanlı, 64 birimli LSTM + Dense yapı. En sade ve en iyi performansı veren modeldir.
- LSTM(128) + Dropout: Model kapasitesi artırılmış, %20 dropout uygulanmıştır. Overfitting riski incelenmiştir.
- Conv1D + LSTM: Girişe Conv1D katmanı eklenerek kısa vadeli desenlerin çıkarımı amaçlanmış, ardından LSTM ile zamansal bağlam işlenmiştir.

Tüm modellerde çıkış hedefi aynıdır: geçmiş 1440 dakikalık girişten, 60 dakika sonrası BTC kapanış fiyatını tahmin etmek. Kayıp fonksiyonu olarak yön duyarlılığı sağlayan Weighted Loss fonksiyonu kullanılmıştır.

Tablo 5.3. Model performans sonuçları

Mimari Yapı	RMSE (USD)	MAE (USD)	R ²
LSTM(64)	137.57	76.87	0.9999
LSTM(128)	192.69	112.44	0.9999
Conv1D(64) → LSTM(64)	539.12	279.19	0.9993

Tablo 5.3'te gösterildiği üzere LSTM(64), hem en düşük hata değerlerine ulaşmış hem de en yüksek R² skorunu sağlamıştır. Bu durum, veri seti için en uygun mimarinin temel LSTM yapısı olduğunu göstermektedir. LSTM(128) kapasite artışı aşırı öğrenme riskiyle birlikte başarıyı sınırlamıştır. Conv1D(64) +LSTM(64), Conv1D ile kısa vadeli desen çıkarımı amaçlanmış olsa da, fiyat tahmini bağlamında istenilen başarıya ulaşamamıştır.

Bu deneysel çalışma sonucunda, dakikalık çözünürlükte dahi temel LSTM mimarisinin güçlü genelleme yeteneğini koruduğu ve özellikle fiyat seviyesi tahmininde uzun vadeli bağlam bilgisinin kritik rol oynadığı anlaşılmıştır. Model karmaşıklığından çok, mimari ve veri yapısı uyumunun başarıyı belirlediği sonucuna varılmıştır.

5.2.2. 15 dakikalık periyotlarda yapılan kısa vadeli tahminler

Bu bölümde, daha yüksek frekanslı piyasa öngörülleri yapabilmek amacıyla, her 15 dakikalık zaman diliminde güncellenen kısa vadeli tahmin modelleri geliştirilmiştir. Model, geçmişe dönük 360 dakikalık (6 saatlik) kapanış fiyatı ve işlem hacmi bilgisini giriş olarak almakta; tahmin hedefi olarak ise bu pencereyi takip eden 15. dakikanın kapanış fiyatını

öngörmektedir. Model mimarisi, bir önceki başlıkta kullanılan LSTM tabanlı yapıyla aynıdır.

Bu yapı kapsamında iki farklı senaryo değerlendirilmiştir. İlk senaryoda, teknik göstergeler saatlik çözünürlükte hesaplanarak her 15 dakikalık tahmin için uygun saatlik gösterge eşlenmiş; ikinci senaryoda ise aynı göstergeler bu kez 15 dakikalık çözünürlükte yeniden oluşturulmuştur. Her iki senaryoda da teknik göstergeler zaman serisi içermeyen, sabit öznitelik vektörleri olarak modele ayrı bir giriş şeklinde verilmiştir. Ana zaman serisi girişi olarak yalnızca 15 dakikalık kapanış fiyatı ve hacim verileri kullanılmış, tahmin süreci dakikalık çözünürlükte ve tek adımlı olarak gerçekleştirilmiştir.

Tablo 5.4'te her iki senaryonun doğrulama verisi üzerindeki performans karşılaştırması sunulmuştur:

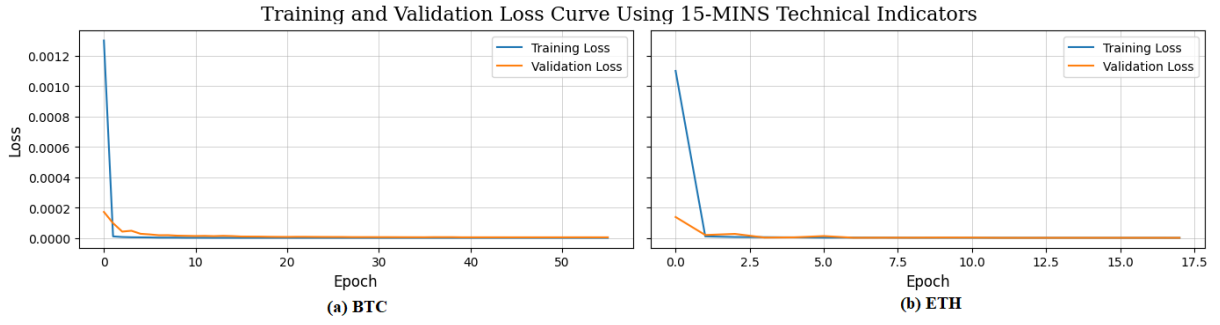
Tablo 5.4. Teknik gösterge çözünürlüğüne göre doğrulama performans karşılaştırması (BTC/USDT)

Teknik Gösterge Çözünürlüğü	RMSE (USD)	MAE (USD)
Saatlik	185.90	123.39
15 Dakikalık	88.16	34.52

Tablo 5.4, teknik göstergelerin farklı çözünürlüklerde hesaplanmasının model doğruluk metriklerine etkisini göstermektedir. Sonuçlar, tahmin çözünürlüğü ile aynı frekansta hesaplanan (15 dakikalık) teknik göstergelerin, model performansını anlamlı düzeyde artırdığını ortaya koymaktadır. Elde edilen hata değerleri, saatlik çözünürlükte hesaplanan göstergelere kıyasla yaklaşık yarı yarıya azalmıştır.

Bu bulgular doğrultusunda, ETH modeli de dâhil olmak üzere sonraki tüm test ve değerlendirme aşamalarında yalnızca 15 dakikalık çözünürlükte hesaplanmış teknik göstergeler kullanılmıştır.

Bu yapı ile oluşturulan modellerin eğitim süreci boyunca sergilediği öğrenme davranışı, doğrulama başarımını da destekler niteliktedir. Epoch bazında kaydedilen kayıp değerlerinin seyri, Şekil 5.5. (a) (BTC) ve Şekil 5.5. (b) (ETH)'de sunulmuştur. Her iki modelde de eğitim ve doğrulama kayıpları kısa sürede dengeye ulaşmış, erken durdurma mekanizması etkin şekilde çalışarak aşırı öğrenme önlenmiştir.



Şekil 5.5. 15 dakikalık tahmin modeli için BTC/USDT ve ETH/USDT işlem çiftlerinde model eğitimi sırasında gözlemlenen doğrulama ve eğitim kayıplarının epoch bazında değişimi.(a) BTC modeli, (b) ETH modeli. Her iki modelde de erken durdurma mekanizması etkin şekilde çalışarak eğitim ve doğrulama kayıplarının kısa sürede dengelendiği gözlemlenmiştir.

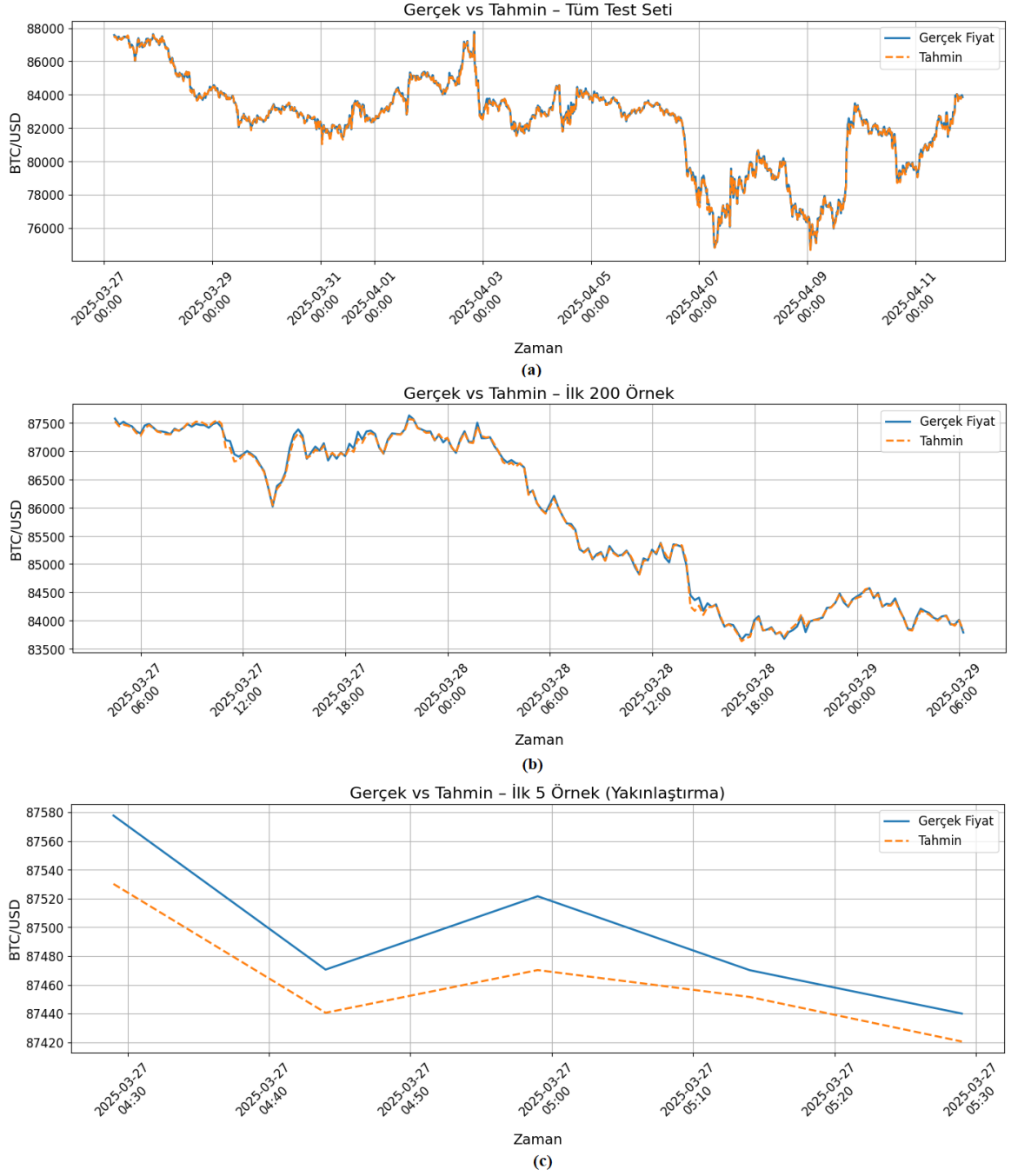
Doğrulama verisi üzerindeki performans sonuçları oldukça tatmin edicidir. Özellikle hem BTC hem de ETH işlem çiftlerinde, gerçek fiyatlar ile tahmin edilen fiyatlar arasındaki fark oldukça düşük seviyelerde seyretmiştir. Bu duruma ilişkin performans metrikleri Tablo 5.5’te özetlenmiştir.

Tablo 5.5. Dakikalık tahmin modeli için BTC/USDT ve ETH/USDT işlem çiftlerine ait doğrulama seti üzerindeki performans metrikleri

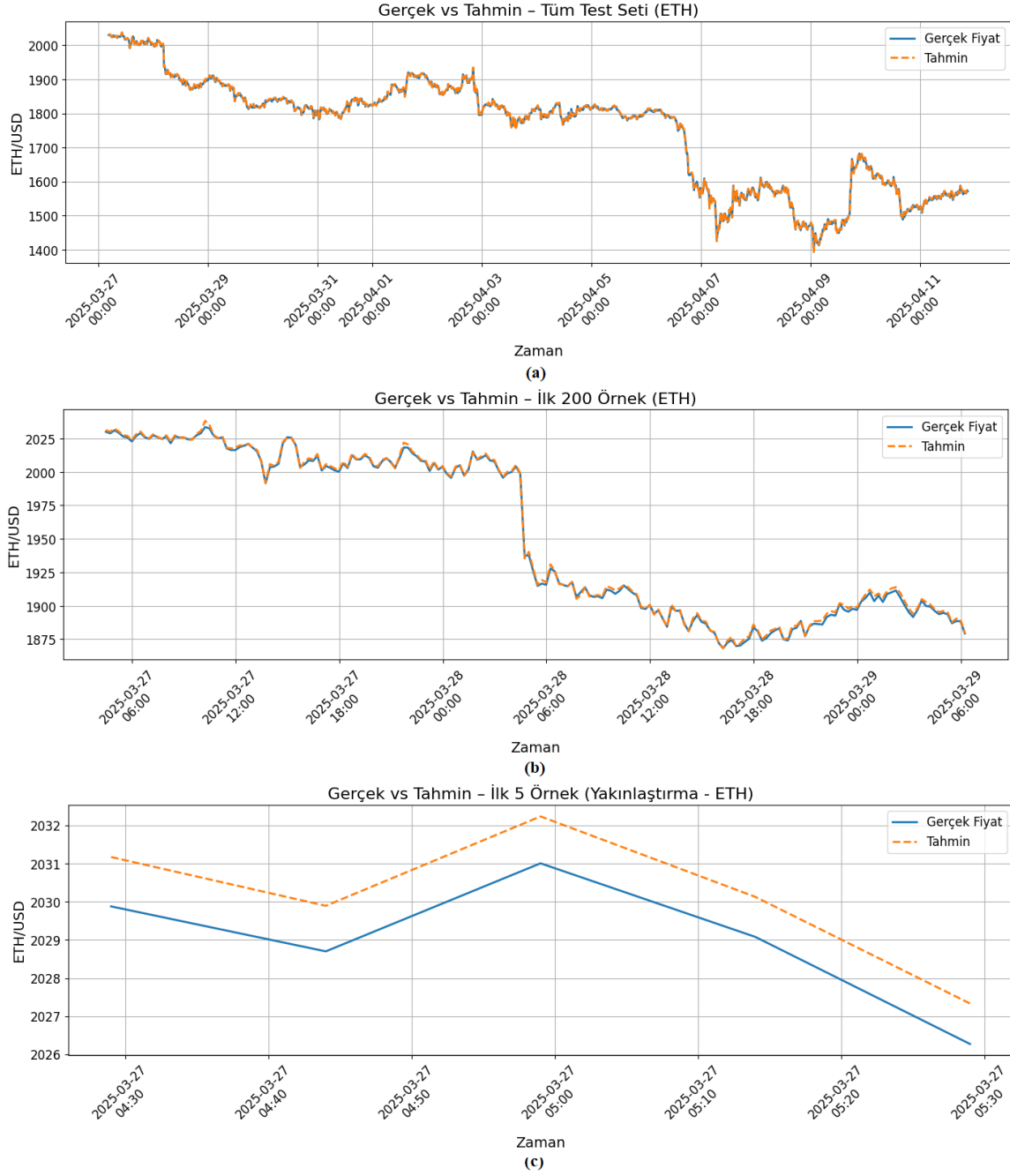
İşlem Çifti	RMSE	MAE	R ²
BTC/USDT	88.16	34.52	0.9995
ETH/USDT	1.83	1.28	0.9999

5.2.3. Test Veri Seti Üzerindeki Performans ve Tahmin-Gözlem Karşılaştırmaları

Eğitim sürecinde hiç kullanılmamış olan bağımsız test veri seti üzerinde yapılan değerlendirme sonuçları, modellerin genel geçerliliği açısından güçlü bulgular sunmaktadır. BTC ve ETH için test seti üzerindeki tahmin-gözlem karşılaştırmaları sırasıyla Şekil 5.6 ve Şekil 5.7’de gösterilmiştir. Şekillerin (a) bölümlerinde tüm test verisi boyunca elde edilen tahmin sonuçları; (b) bölümlerinde ilk 200 test örneği; (c) bölümlerinde ise ilk 5 örnek yakınlaştırılmış biçimde sunulmuştur. Bu görseller, modelin yalnızca trendi değil, kısa vadeli dalgalanmaları da yüksek doğrulukla yakalayabildiğini göstermektedir.



Şekil 5.6. BTC için test seti üzerindeki tahmin-gözlem karşılaştırmaları



Şekil 5.7. ETH için test seti üzerindeki tahmin-gözlem karşılaştırmaları

Test seti performans metrikleri Tablo 5.6’da sunulmuştur. Her iki işlem çifti için elde edilen RMSE ve MAE değerleri oldukça düşüktür; özellikle BTC/USDT işlem çifti için 67.96 USD RMSE ve 44.08 USD MAE, ETH/USDT için ise sırasıyla 2.01 USD ve 1.55 USD olarak hesaplanmıştır. Bu düşük hata seviyeleri, modelin fiyat tahminlerinin mutlak anlamda gerçek değerlere oldukça yakın olduğunu göstermektedir.

Tablo 5.6. Model test seti performans metrikleri

İşlem Çifti	RMSE	MAE	MAPE	R ²
BTC/USDT	67.96	44.08	0.05%	0.9994
ETH/USDT	2.01	1.55	0.09%	0.9998

Finansal zaman serilerinde MAPE metriği, özellikle düşük fiyatlı varlıklarda görece hata oranlarının abartılı hale gelmesi nedeniyle sınırlı kullanılabilirliğe sahip olabilir. Ancak BTC ve ETH gibi yüksek değerli varlıklar için bu durum geçerli değildir. Bu çalışmada MAPE'nin %0.05 ve %0.09 gibi son derece düşük değerlerde gerçekleşmesi hem mutlak hem de görece hata düzeylerinin oldukça düşük olduğunu göstermekte ve modelin pratik uygulanabilirliğini desteklemektedir.

6. SONUÇLAR

Bu çalışma, yüksek frekanslı kripto para piyasalarında kısa vadeli fiyat tahmini gerçekleştirmek üzere yapılandırılmış LSTM tabanlı derin öğrenme modelleri aracılığıyla, veri çözünürlüğü, mimari yapı ve kayıp fonksiyonu tercihleri gibi çok sayıda bileşenin sistematik olarak analiz edildiği bütüncül bir yaklaşım sunmaktadır. Modelleme süreci boyunca Bitcoin ve Ethereum işlem çiftlerine ait hem saatlik hem de dakikalık veriler kullanılmış; her iki zaman ölçeğinde de LSTM mimarisinin istikrarlı ve yüksek doğrulukta tahminler sunduğu gözlemlenmiştir. Özellikle kapanış fiyatı ve hacim gibi temel zaman serisi değişkenlerine ek olarak teknik analiz göstergelerinin modele entegre edilmesi, modelin öğrenme kapasitesini belirgin ölçüde artırmıştır.

Farklı zaman çözünürlüklerinde hesaplanan teknik göstergeler karşılaştırıldığında, tahmin hedefiyle aynı frekansta üretilen göstergelerin daha başarılı sonuçlar verdiği tespit edilmiştir. Bu bulgu, zaman uyumlu veri-gösterge eşleşmesinin model başarımı açısından kritik öneme sahip olduğunu göstermektedir. Diğer yandan, farklı kayıp fonksiyonları ile yapılan deneylerde, özellikle hata büyüklüğüne duyarlı özel bir kayıp fonksiyonu Weighted Loss fonksiyonunun LSTM mimarisi ile birlikte en yüksek tahmin performansını sağladığı belirlenmiştir.

Genel olarak çalışma, fiyat seviyesinin yanı sıra yön bilgisini de içeren bütünsel bir tahmin yaklaşımı ortaya koymakta; veri çözünürlüğü, mimari yapı, teknik gösterge kullanımı ve kayıp fonksiyonu seçimlerinin birlikte optimize edilmesinin derin öğrenme temelli finansal modelleme sistemlerinde başarıyı belirleyen temel unsurlar olduğunu ortaya koymaktadır. Elde edilen bulgular, benzer piyasa koşullarında uygulanabilecek esnek ve genellenebilir tahmin sistemlerinin tasarımı açısından önemli bir referans niteliğindedir.

7. KAYNAKLAR

- Achelis, S. B.** (2001). **Technical Analysis from A to Z** (2. baskı). New York: McGraw-Hill.
- Aggarwal, C. C.** (2018). *Neural networks and deep learning*. Cham: Springer.
- Agrawal, M., Shukla, P. K., Nair, R., Nayyar, A., Masud, M.** (2022). Stock prediction based on technical indicators using deep learning model. *Computers, Materials & Continua*, 70(1), 287–304. <https://doi.org/10.32604/cmc.2022.014637>
- Alpaydin, E.** (2020). *Introduction to machine learning*. MIT press.
- Appel, G.** (2005). *Technical analysis: Power tools for active investors*. Financial Times Prentice Hall.
- Alessandretti, L., ElBahrawy, A., Aiello, L. M., & Baronchelli, A.** (2018). Anticipating cryptocurrency prices using machine learning. *Complexity*, 2018, Article ID 8983590. <https://doi.org/10.1155/2018/8983590>
- Bengio, Y., Simard, P., Frasconi, P.** (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166.
- Berahmand, K., Daneshfar, F., Salehi, E. S., Li, Y., Xu, Y.** (2024). Autoencoders and their applications in machine learning: A survey. *Artificial Intelligence Review*, 57(2), 28. <https://doi.org/10.1007/s10462-023-10662-6>
- Binance.** (2025). Public REST API Documentation. <https://binance-docs.github.io/apidocs/>
- Bishop, C. M.** (2006). *Pattern Recognition and Machine Learning*. Springer.
- Bollinger, J.** (2002). *Bollinger on Bollinger Bands*. McGraw-Hill.
- Chen, L., Chi, Y., Guan, Y., Fan, J.** (2019). A hybrid attention-based EMD-LSTM model for financial time series prediction. *2nd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 113–118. <https://doi.org/10.1109/ICAIBD.2019.8837038>
- Chicco, D., Warrens, M. J., Jurman, G.** (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, e623.
- Cho, K., van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. ve Bengio, Y.** (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint, *arXiv:1409.1259*. <https://arxiv.org/abs/1409.1259>
- Chollet, F.** (2021). *Deep Learning with Python* (2. baskı). Manning Publications.

- Chung, J., Gulcehre, C., Cho, K., Bengio, Y.** (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint*, arXiv:1412.3555. <https://arxiv.org/abs/1412.3555>
- Corbet, S., Lucey, B., Urquhart, A., Yarovaya, L.** (2019). Cryptocurrencies as a financial asset: A systematic analysis. *International Review of Financial Analysis*, 62, 182–199. <https://doi.org/10.1016/j.irfa.2018.09.003>
- Çelik, Y.** (2024). Bellek tabanlı LSTM ve GRU makine öğrenmesi algoritmaları kullanarak BIST100 endeks tahmini. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 36(2), 553–561.
- Efanov, D., Aleksandrov, P., Karapetyants, N.** (2022). The BiLSTM-based synthesized speech recognition. *Procedia Computer Science*, 213, 415–421.
- Elman, J. L.** (1990). Finding structure in time. *Cognitive Science*, 14(2), 179–211.
- Gers, F. A., Schmidhuber, J., Cummins, F.** (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, 12(10), 2451–2471. <https://doi.org/10.1162/089976600300015015>
- Graves, A. ve Schmidhuber, J.** (2005). Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks*, 18(5–6), 602–610. <https://doi.org/10.1016/j.neunet.2005.06.042>
- Hochreiter, S., Schmidhuber, J.** (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Imani, M.** (2023). Alzheimer’s diseases diagnosis using fusion of high informative BiLSTM and CNN features of EEG signal. *Biomedical Signal Processing and Control*, 86, 105298.
- Kalra, R., Singh, T., Mishra, S., Satakshi, Kumar, N., Kim, T., Kumar, M.** (2024). An efficient hybrid approach for forecasting real-time stock market indices. *Journal of King Saud University - Computer and Information Sciences*, 36(8), 102180. <https://doi.org/10.1016/j.jksuci.2024.102180>
- LeCun, Y., Bengio, Y., Hinton, G.** (2015). Deep learning. *Nature*, 521(7553), 436–444.
- Lim, B. ve Zohren, S.** (2021). Time-series forecasting with deep learning: A survey. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194), 20200209. <https://doi.org/10.1098/rsta.2020.0209>
- Lu, W., Zhang, Y., Ma, K.** (2021). A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications* 33(3):1-13.

- Krizhevsky, A., Sutskever, I., Hinton, G. E.** (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- Masini, R., Medeiros, M. C., Mendes, E. F.** (2021). Machine learning advances for time series forecasting. *Journal of Economic Surveys*, 35(3), 887–918. <https://doi.org/10.1111/joes.12371>
- Mitchell, T. M.** (1997). Machine Learning. McGraw-Hill.
- Mudassir, M., Bennbaia, S., Unal, D., Hammoudeh, M.** (2020). Time-series forecasting of Bitcoin prices using high-dimensional features: a machine learning approach. *Neural Computing and Applications*, 32, 17299–17317. <https://doi.org/10.1007/s00521-020-04814-w>
- Murphy, J. J.** (1999). Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications (2. baskı). New York: New York Institute of Finance.
- Neely, C. J., Rapach, D. E., Tu, J., ve Zhou, G.** (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, 60(7), 1772–1791.
- Oerth, A. ve Dowsett, D.** (2023). Exploring cryptocurrencies: The history, the hype, and the future. *The Journal of Investing*, 32(3), 19–48. <https://doi.org/10.3905/joi.2023.1.258>
- Ozupek, O., Yilmaz, R., Ghasemkhani, B., Birant, D., Kut, R. A.** (2024). A Novel Hybrid Model (EMD-TI-LSTM) for Enhanced Financial Forecasting with Machine Learning. *Mathematics*, 12(17), 2794.
- Rajasekaran, A. S., Azees, M., Al-Turjman, F.** (2022). A comprehensive survey on blockchain technology. *Sustainable Energy Technologies and Assessments*, 52, 102039. <https://doi.org/10.1016/j.seta.2022.102039>
- Rudy, S. H., Sapsis, T. P.** (2023). Output-weighted and relative entropy loss functions for deep learning precursors of extreme events. *Physica D: Nonlinear Phenomena*, 443, 133570. <https://doi.org/10.1016/j.physd.2022.133570>
- Rundo, F.** (2019). Deep LSTM with reinforcement learning layer for financial trend prediction in FX high-frequency trading systems. *Applied Sciences*, 9(20), 4460. <https://doi.org/10.3390/app9204460>
- Rundo, F., Trubiani, C., Tamburrano, G.** (2019). Machine Learning for Quantitative Finance Applications: A Survey. *Applied Sciences*, 9(24), 5574. <https://doi.org/10.3390/app9245574>
- Sutton, R. S., Barto, A. G.** (2018). Reinforcement Learning: An Introduction (2, Baskı). MIT Press.

- Wang, H., Raj, B.** (2017). On the origin of deep learning. *arXiv preprint*, arXiv:1702.07800
- Wilder, J. W.** (1978). *New Concepts in Technical Trading Systems*. Greensboro, NC: Trend Research.
- Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., Zhang, H.** (2018). Deep learning with long short-term memory for time series prediction. *arXiv preprint*, arXiv:1810.10161. <https://arxiv.org/abs/1810.10161>
- Yin, W., Kann, K., Yu, M., Schütze, H.** (2017). Comparative study of CNN and RNN for natural language processing. *arXiv preprint*, arXiv:1702.01923. <https://arxiv.org/abs/1702.01923>
- Yu, Y., Si, X., Hu, C., Zhang, J.** (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 31(7), 1235–1270.

