

HİTİT SOSYAL BİLİMLER DERGİSİ

Hitit Journal of Social Sciences

e-ISSN: 2757-7949

Cilt | Volume: 18 • Sayı | Number: 2

Ağustos | August 2025

Mevsimsel Trendler ve Piyasa Volatilitésinin Entegrasyonu: Kripto Para Tahmini için Hibrit SARIMA-XGBoost Modeli

Integrating Seasonal Trends and Market Volatility: A Hybrid SARIMA-XGBoost
Model for Cryptocurrency Forecasting

Melikşah AYDIN

Corresponding Author | Sorumlu Yazar

Arş. Gör. | Res. Asst.

Niğde Ömer Halisdemir Üniversitesi, İktisadi ve İdari Bilimler
Fakültesi, Finans ve Bankacılık Bölümü, Niğde Türkiye
Niğde Ömer Halisdemir University, Faculty of Economics and
Administrative Sciences, Department of Finance and Banking,
Niğde, Türkiye

maydin@ohu.edu.tr

<https://orcid.org/0000-0001-7711-1874>

<https://ror.org/03ejnre35>

Eyyüp Ensari ŞAHİN

Doç. Dr. | Assoc. Dr.

Niğde Ömer Halisdemir Üniversitesi, İktisadi ve İdari Bilimler
Fakültesi, Finans ve Bankacılık Bölümü, Niğde, Türkiye
Niğde Ömer Halisdemir University, Faculty of Economics and
Administrative Sciences, Department of Finance and Banking,
Niğde, Türkiye

ensarisahin@ohu.edu.tr

<https://orcid.org/0000-0003-2110-7571>

<https://ror.org/03ejnre35>

Makale Bilgisi | Article Information

Makale Türü | Article Type: Araştırma Makalesi | Research Article

Geliş Tarihi | Received: 05.01.2025

Kabul Tarihi | Accepted: 21.07.2025

Yayın Tarihi | Published: 31.08.2025

Atıf | Cite As

Aydın, M., & Şahin, E. E. (2025). Mevsimsel trendler ve piyasa volatilitésinin entegrasyonu: kripto para tahmini için hibrit SARIMA-XGBoost modeli. *Hitit Sosyal Bilimler Dergisi*, 18(2), 581-602. <https://doi.org/10.17218/hititsbd.1613531>

Yazar Katkıları: %50-%50

Değerlendirme: Bu makalenin ön incelemesi iki iç hakem (editörler-yayın kurulu üyeleri) içerik incelemesi ise iki dış hakem tarafından çift taraflı kör hakemlik modeliyle incelendi. Benzerlik taraması yapılarak (Turnitin) intihal içermediği teyit edildi.

Etik Beyan: Bu çalışmanın hazırlanma sürecinde bilimsel ve etik ilkelere uyulduğu ve yararlanılan tüm çalışmaların kaynakçada belirtildiği beyan olunur.

Etik Bildirim: husbeditor@hitit.edu.tr
<https://dergipark.org.tr/tr/pub/hititsbd>

Çıkar Çatışması: Çıkar çatışması beyan edilmemiştir.

Finansman: Bu araştırmayı desteklemek için dış fon kullanılmamıştır.

Telif Hakkı & Lisans: Yazarlar dergide yayınlanan çalışmalarının telif hakkına sahiptirler ve çalışmaları CC BY-NC 4.0 lisansı altında yayımlanmaktadır.

Author Contributions: 50%-50%

Review: Single anonymized-Two Internal (Editorial board members) and Double anonymized - Two External Double-blind Peer Review. It was confirmed that it did not contain plagiarism by similarity scanning (Turnitin).

Ethical Statement: It is declared that scientific and ethical principles have been followed while conducting and writing this study and that all the sources used have been properly cited.

Complaints: husbeditor@hitit.edu.tr -
<https://dergipark.org.tr/tr/pub/hititsbd>

Conflicts of Interest: The author(s) has no conflict of interest to declare.

Grant Support: The author(s) acknowledge that they received no external funding to support this research.

Copyright & License: Authors publishing with the journal retain the copyright to their work licensed under the CC BY-NC 4.0.

Integrating Seasonal Trends and Market Volatility: A Hybrid SARIMA-XGBoost Model for Cryptocurrency Forecasting

Abstract

This paper explores the application of a hybrid SARIMA-XGBoost model for forecasting cryptocurrency prices within the context of financial time series analysis. Due to their dynamic nature, high volatility, seasonal patterns, and asymmetric market behavior, cryptocurrencies pose significant challenges for traditional forecasting methods. The study combines the strength of the SARIMA model in capturing seasonal and trend components with the XGBoost algorithm's ability to model nonlinear relationships. This hybrid approach aims to better represent the complex and non-stationary nature of cryptocurrency markets, thereby improving forecasting accuracy. Cryptocurrencies, such as Bitcoin, exhibit volatile price movements influenced by factors like periodic "halving" effects and volatility clustering. Additionally, investor behaviors, herd psychology, and social media sentiment contribute to the difficulty of price forecasting. In this study, residuals from the SARIMA model are used as inputs for the XGBoost algorithm to capture nonlinear patterns, effectively integrating both linear and nonlinear elements in a single model. The paper analyzes the model's performance using BTC (Bitcoin), XRP (Ripple), and ETH (Ethereum) price data. Metrics such as the Akaike Information Criterion (AIC), Mean Squared Error (MSE), and Mean Absolute Error (MAE) are employed to evaluate the model's accuracy. Results indicate that the SARIMA-XGBoost hybrid model outperforms standalone SARIMA and XGBoost models, achieving higher accuracy, especially during periods of market volatility. The hybrid model successfully captures the dynamic nature of cryptocurrency markets, producing lower error rates compared to other models. The findings demonstrate the effectiveness of hybrid modeling approaches in forecasting cryptocurrency prices and highlight the potential of optimized hybrid models in investment decision-making and risk management. However, the study also notes that the model's performance could be further enhanced during short-term fluctuations and highly volatile market conditions.

Keywords: SARIMA, XGBoost, Cryptocurrency Forecasting, Financial Time Series, Volatility, Hybrid Modeling

Mevsimsel Trendler ve Piyasa Volatilitesinin Entegrasyonu: Kripto Para Tahmini için Hibrit SARIMA-XGBoost Modeli

Öz

Bu makale, finansal zaman serisi analizi bağlamında, kripto para birimlerinin fiyat tahmini için hibrit bir model olan SARIMA-XGBoost yöntemini ele almaktadır. Kripto para birimlerinin dinamik yapısı, yüksek volatilité, mevsimsel desenler ve asimetrik piyasa davranışları nedeniyle geleneksel yöntemlerle analiz edilmesi oldukça zordur. Makalede, SARIMA modelinin mevsimsel ve trend bileşenlerini ele almadaki gücü ile XGBoost algoritmasının doğrusal olmayan ilişkileri modelleme yeteneği birleştirilmiştir. Bu hibrit model, kripto para piyasasının karmaşık ve durağan olmayan yapısını daha iyi temsil ederek tahmin performansını artırmayı amaçlamaktadır. Kripto varlıklar, Bitcoin gibi önde gelen para birimlerinin dönemsel "yarılanma" etkileri ve volatilité kümeleşmesi gibi faktörler nedeniyle dalgalı fiyat hareketleri sergilemektedir. Aynı zamanda, yatırımcı davranışları, sürü psikolojisi ve sosyal medya etkisi gibi değişkenler de fiyat tahminlerini zorlaştırmaktadır. Bu çalışmada, SARIMA modelinin artık değerleri, doğrusal olmayan desenlerin modellenmesi için XGBoost'a girdi olarak verilmiştir. Böylece hem doğrusal hem de doğrusal olmayan desenler aynı modelde bütünleştirilmiştir. Makale, BTC (Bitcoin), XRP (Ripple) ve ETH (Ethereum) fiyat verileri üzerinde modelin performansını analiz etmektedir. Model performansı, Akaike Bilgi Kriteri (AIC), Ortalama Kare Hata (MSE) ve Ortalama Mutlak Hata (MAE) ölçütleriyle değerlendirilmiştir. Sonuçlar, SARIMA-XGBoost modelinin bağımsız SARIMA ve XGBoost modellerine kıyasla daha yüksek doğruluk sağladığını göstermektedir. Özellikle volatil dönemlerde hibrit model, kripto para piyasasının dinamiklerini daha iyi yakalayarak daha küçük hata oranları üretmektedir. Bu çalışmanın bulguları, kripto para birimlerinin gelecekteki fiyatlarının tahmin edilmesi için hibrit modellemelerin etkili bir yaklaşım olduğunu kanıtlamaktadır. Aynı zamanda, optimize edilmiş hibrit modellerin yatırım kararları ve risk yönetimi süreçlerinde kullanılabilirliği vurgulanmaktadır. Ancak, kısa vadeli dalgalanmalar ve aşırı volatilité dönemlerinde modelin performansında iyileştirmeler yapılabileceği belirtilmektedir.

Anahtar Kelimeler: SARIMA, XGBoost, Kripto Para Tahmini, Finansal Zaman Serisi, Volatilité, Hibrit Modelleme

Giriş

Finansal zaman serisi verilerinin tahmini, seriler arasındaki ilişkinin belirlenmesi, verilerin karmaşık yapısı, trendler ve mevsimsel etkiler nedeniyle oldukça zordur (Alhnaity ve Abbod, 2020). Özellikle kripto varlıklar yapısı gereği riskli varlıklardır. Kripto varlıkların henüz erken aşama evresi dediğimiz kabullenme aşaması bu karmaşıklığı önemli ölçüde arttırmaktadır. Kripto paralar, her 4 yıllık süreçte Bitcoinin yarılanma (Halving) sistemi nedeniyle fiyatlarında aşırı dalgalanmalar yaşamaktadır (Takaishi, 2021; Nikolova ve diğerleri, 2020). Halvingin bir sonucu olarak kripto para zaman serileri, fiyatların zaman içinde yukarı veya aşağı yönlü hareket sergilediği trendler içerebildiği farklı çalışmalarda bulunmuştur (Gao ve diğerleri, 2022). Ayrıca, bazı kripto paralar periyodik dalgalanmalar gösteren mevsimsel desenler sergileyebilmektedir (Gao ve diğerleri, 2022).

Literatürde yapılan pek çok çalışmada kripto para fiyatlarının zaman serilerinde genellikle volatilité kümeleşmesi sergilediği; bu durumun yüksek volatilité dönemlerinin yine yüksek volatilité ile, düşük volatilité dönemlerinin ise düşük volatilité ile takip edilmesi şeklinde gözlemlendiği farklı pek çok çalışmada vurgulanmıştır (Liu ve diğerleri, 2021; Fleischer ve diğerleri, 2022). Yine çalışmalar incelendiğinde kripto varlıkların zaman serilerinin durağan olmadığı pek çok çalışmada bulgu olarak verilmiştir. Bu durum ortalama ve varyans gibi istatistiksel özelliklerin zaman içinde değiştiğini göstermekte ve durağan olmayan yapının ele alınmasının, kripto para fiyatlarının doğru şekilde modellenmesinin ve tahmini için kritik öneme sahip olduğunu sonucunu doğurmuştur ((Maleki ve diğerleri 2020; Zheng ve diğerleri,, 2023; Memic ve diğerleri, 2021; Phillips ve Gorse, 2018).

Kripto varlık zaman serilerinde bir diğer önemli konu ise asimetrik volatilitedir. Kripto para zaman serileri sıklıkla asimetrik volatilité sergilemektedir. Asimetrik volatilitenin etkisi ise negatif şokların volatilité üzerinde pozitif şoklara kıyasla daha büyük bir etkiye sahip olduğunu bize göstermektedir (Takaishi, 2021; Gupta ve Chaudhary, 2022). Asimetrik volatilitenin modellenmesi, kripto para piyasalarının dinamiklerinin doğru şekilde ölçülmesi için oldukça önemlidir (Papadimitriou ve diğerleri, 2022). Kripto varlıkların fraktal ve çoklu bağlantı sorunu da literatürde sıklıkla tartışılmıştır. Geçmiş olayların gelecekteki olaylar üzerinde kalıcı bir etkisi olduğu uzun vadeli bağımlılık sergilemektedir (Takaishi, 2018; Derbentsev ve diğerleri, 2019). Ayrıca, kripto varlıklar fraktal özellikler sergileyerek farklı zaman ölçeklerinde karmaşık ölçeklenme davranışları ve öz-benzerlik göstermektedir (Takaishi, 2021; Takaishi, 2018). Son olarak kripto varlık zaman serileri yatırımcıların sürü davranışı etkisinde içerisinde barındırmaktadır. Kripto para piyasaları, yatırımcıların diğerlerinin eylemlerini takip etme eğiliminde olduğu sürü davranışından etkilenebilir ve bu da korelasyonlu ticaret desenlerine yol açabilir (Bouri ve diğerleri, 2019; Haykır ve Yağlı, 2022). Ayrıca, sosyal medya ve diğer çevrimiçi kaynaklarda yansıtılan yatırımcı duyarlılığı, kripto para fiyatları üzerinde önemli bir etkiye sahip olabilmektedir (Karalevicius ve diğerleri, 2018; Feizian, 2023).

Tüm bu karmaşık yapı finansal zaman serilerinin tespit edilmesi ve çözümlenebilmesi amacıyla literatürde farklı pek çok çalışmada farklı yöntemlerle araştırılmıştır. Son dönemlerde yapılan çalışmalarda ise geleneksel yöntemlerin veri karmaşıklığını tespit etmede yetersiz kaldığını bu karmaşık yapının hibrit modellerle çalışması gerekliliği üzerine durulmaktadır (Li ve diğerleri,, 2020; Alhnaity ve Abbod, 2020; Yu ve diğerleri, 2020; Wu ve diğerleri, 2020). Literatürde çalışmalar incelendiğinde önemli ölçüde önerilmekte olan hibrit yaklaşım ise Mevsimsel Otoregresif Entegre Hareketli Ortalama (SARIMA) modeli ile eXtreme Gradient Boosting (XGBoost) algoritmasını birleşimi olan modeldir. (Hoa ve diğerleri, 2023). SARIMA,

akademik çalışmalarda sıklıkla kullanılan mevsimsel ve trend desenlerini tanımlayabilen güvenilir bir doğrusal tahmin modeli olarak pek çok çalışmada bulunmaktadır (Li ve diğerleri, 2020; Alhnaity ve Abbod, 2020). XGBoost modeli ise, doğrusal olmayan ve karmaşık ilişkileri modellemede üstün performans göstermesi ile öne çıkmaktadır. Hibrit model öncelikle SARIMA modelini kullanarak verilerdeki doğrusallığı ve mevsimsel etkileri belirlemektedir. SARIMA modelinin çıktıları (artık değerler) ise XGBoost modeline girdi olarak kullanılmaktadır. Girdiler sonucunda doğrusal olmayan desenler modellenmeye devam etmektedir (Hoa ve diğerleri, 2023). Sistemin işleyişi çalışmanın uygulama bölümünde detaylı açıklanacak olup özetle değerlendirildiğinde yapılan çalışmalarda Hibrit modellerin iyi performans sergilediği görülmektedir. SARIMA-XGBoost modeli, durağan olmayan ve karmaşık desenler gibi zorlukları etkili bir şekilde ele alarak tahmin doğruluğunu artırdığına yönelik çalışmalar bu durumu desteklemektedir (Li ve diğerleri, 2020; Alhnaity ve Abbod, 2020; Sako ve diğerleri, 2022).

Sonuç olarak, SARIMA-XGBoost hibrit modeli, klasik doğrusal modellerin güçlü yönlerini ileri düzey makine öğrenimi algoritmalarıyla birleştiren güçlü bir finansal zaman serisi tahmin aracı olarak öne çıkmaktadır. Bağımsız modeller ve diğer hibrit yöntemlerle karşılaştırıldığında üstün performansı, bu modeli finansal tahmin alanında araştırmacılar ve uygulayıcılar için vazgeçilmez bir araç haline getirmektedir. Çalışmada BTC, XRP ve ETH fiyat verileri Hibrit model ile test edilmiştir.

1. Literatür İncelemesi

Kripto para piyasalarının tahminine yönelik literatürde, finansal zaman serileri üzerine geliştirilen hibrit modeller giderek daha fazla ilgi çekmektedir. Alessandretti ve diğerleri (2018), yaptıkları çalışmada, makine öğrenimi tekniklerinin kripto para piyasasına uygulanmasını incelemiş ve özellikle Bitcoin, Ripple ve Ethereum fiyatlarının tahmininde bu yöntemlerin etkinliğini vurgulamışlardır. Çalışmada, ileri düzey algoritmaların kısa vadeli piyasa hareketlerini tahmin etmede geleneksel yöntemlere kıyasla üstün performans sergileyebileceği sonucuna ulaşılmıştır. Bu bulgu, kripto para piyasalarının dinamik doğası göz önüne alındığında, veri odaklı yaklaşımların potansiyelini ortaya koymaktadır.

Benzer şekilde, Valencia ve diğerleri (2019), Bitcoin, Ethereum ve Ripple odaklı çalışmalarında, duygu analizi ve SARIMA-XGBoost modeli dahil olmak üzere makine öğrenimi tekniklerinin kripto para piyasasındaki fiyat hareketlerini tahmin etmede uygulanabilirliğini ele almışlardır. Çalışmada, hibrit modellerin piyasa fiyatlarını tahmin etmedeki başarısı vurgulanmış ve özellikle duygu analizinin eklenmesiyle tahmin performansının arttığı ifade edilmiştir.

Yu ve diğerleri (2020), finansal zaman serisi tahmininde ARIMA ve Extreme Learning Machine (ELM) entegrasyonunu içeren bir hibrit model önererek, doğrusal ve doğrusal olmayan desenleri aynı anda ele alma kapasitesini incelemişlerdir. Bu çalışma, hibrit SARIMA-XGBoost modellerinin geliştirilmesi açısından önemli bir temel sunmaktadır. Aynı yıl, Li ve diğerleri (2020), geleneksel ARIMA modelinin derin öğrenme teknikleri ile desteklenmesinin önemini vurgulamış ve doğrusal ile doğrusal olmayan modelleme yaklaşımlarının birleştirilmesinin finansal zaman serisi tahminindeki rolüne dikkat çekmişlerdir. Her iki çalışma da hibrit modellerin tahmin doğruluğunu artırmadaki potansiyelini ortaya koymaktadır.

Patel ve diğerleri (2020) ise Bitcoin, Ethereum ve Litecoin gibi kripto paraların yüksek volatilité ve düzensiz davranışları ile ilgili zorlukları ele alan bir stokastik sinir ağı modeli

geliştirmiştir. Bu model, piyasa volatilitisini simüle etmek amacıyla rastgelelik içermekte ve deterministik modellere kıyasla daha yüksek tahmin doğruluğu sağlamaktadır. Çalışma, kripto para fiyatlarının tahmininde hibrit yaklaşımların önemini destekleyen bulgular sunmaktadır.

Rampini ve Cecconi (2021), gayrimenkul fiyatlarının tahmini için makine öğrenimi tekniklerini, özellikle XGBoost algoritmasını uygulamış ve XGBoost'un karmaşık, doğrusal olmayan problemleri ele alma konusundaki etkinliğini vurgulamışlardır. Kim ve diğerleri (2021) ise Seul'deki arazi fiyatlarının tahmini için XGBoost ve rastgele orman (random forest) algoritmalarını incelemiş ve bu yöntemlerin geleneksel istatistiksel yöntemlere göre üstün olduğunu belirtmişlerdir. Her iki çalışma, farklı sektörlerde uygulanmış olmasına rağmen, XGBoost'un karmaşık desenleri yakalama kapasitesinin, Bitcoin, Ripple ve Ethereum gibi kripto para fiyatlarının tahmininde hibrit SARIMA-XGBoost modeli geliştirilmesine de katkı sağlayabileceğini göstermektedir.

Jabeur ve diğerleri (2021), COVID-19 pandemisi döneminde petrol fiyatlarının tahmini için gelişmiş makine öğrenimi tekniklerini, özellikle XGBoost algoritmasını incelemişlerdir. Çalışma, hibrit modellerin finansal zaman serilerindeki eğilimleri yakalama ve tahmin doğruluğunu artırma kapasitesini göstermiştir. Benzer şekilde, Iqbal ve diğerleri (2021), Bitcoin odaklı çalışmalarında, ARIMA ve XGBoost tekniklerinin kripto para fiyatlarının zaman serisi tahmini için uygulanabilirliğini değerlendirmiş ve bu piyasaların benzersiz volatilité özelliklerine dikkat çekerek hibrit SARIMA-XGBoost modelinin geliştirilmesine ilişkin önemli bilgiler sunmuştur.

Bitto ve diğerleri (2022), Bitcoin, Ethereum, Litecoin ve Tether gibi kripto para fiyatlarını tahmin etmek için makine öğrenimi yaklaşımlarını, özellikle SARIMA-XGBoost modelini uygulamış ve bu modelin yüksek doğrulukla tahmin yapma kapasitesini vurgulamışlardır. Çalışmada ayrıca, AR modelinin bu kripto paraları öngörmede sağladığı başarı ön plana çıkarılmıştır. Bu bulgular, SARIMA-XGBoost modelinin kripto para piyasası için güçlü bir araç olabileceğini desteklemektedir.

Huang (2023), COVID-19 pandemisinin Ethereum fiyatları üzerindeki etkisini inceleyerek, COVID-19 vakaları ile Ethereum getirileri arasındaki ilişkiyi analiz etmek için vektör otoregresyon (VAR) ve ARMA-GARCHX modellerini kullanmıştır. Çalışma, kripto paraların dinamiklerini anlamada çeşitli ekonometrik ve makine öğrenimi yaklaşımlarının entegrasyonuna duyulan ihtiyacı vurgulamaktadır. Song (2023) ise makine öğrenimi tekniklerinin, özellikle XGBoost ve LightGBM'in, Bitcoin ve Ethereum gibi kripto paraların kapanış fiyatlarını tahmin etmekteki etkinliğini incelemiş ve bu yöntemlerin uygun yatırım fırsatlarını belirlemedeki başarısını ortaya koymuştur. Shen ve Wang (2023) tarafından yapılan çalışmada ise Bitcoin, Ethereum ve Dogecoin gibi büyük kripto paraların değerlendirilmesi ve tahmini için XGBoost, LightGBM, GARCH ve ARIMA gibi farklı modellerin karşılaştırılması ele alınmış ve bu yöntemlerin tahmin doğruluğu üzerinde etkileri analiz edilmiştir.

Bu kapsamda, literatürdeki çalışmalar, finansal zaman serilerinin tahmininde hibrit modellerin, özellikle SARIMA-XGBoost entegrasyonunun hem doğrusal hem de doğrusal olmayan desenleri yakalama kapasitesini ve tahmin doğruluğunu artırmadaki potansiyelini ortaya koymaktadır. Bitcoin, Ripple ve Ethereum gibi kripto paraların fiyat tahmininde bu modelin uygulanabilirliği, literatürdeki bulgularla tutarlı olarak, güçlü bir araştırma alanı sunmaktadır.

2. Veri Seti ve Yöntem

Veri seti Bitcoin (BTC), Ethereum (ETH) ve Ripple (XRP) kripto paralarından oluşmaktadır ve yahoofinance.com üzerinden 24.06.2019 – 24.06.2024 dönemini kapsayan günlük fiyatlar ABD doları cinsinden alınmıştır. Veri setinin birim kök içerip-içermediğini kontrol etmek amacıyla, bu çalışmada literatürde sıklıkla kullanılan ADF (Augmented Dickey-Fuller) ve KPSS (Kwiatkowski-Phillips-Schmidt-Shin) testleri kullanılmıştır. Her iki testin birlikte kullanılması, test sonuçlarının birbiriyle uyumlu ve birbirini destekler nitelikte olup olmadığını değerlendirmeyi amaçlamaktadır.

Ayrıca bu çalışmada, finansal zaman serilerinin tahmininde hem doğrusal hem de doğrusal olmayan desenleri yakalayabilmek amacıyla hibrit bir SARIMA-XGBoost modeli tercih edilmiştir. SARIMA modeli, özellikle mevsimsellik ve trend gibi belirgin zaman serisi bileşenlerini modellemede istatistiksel olarak güçlü bir yöntem olması nedeniyle tercih edilmiştir. XGBoost algoritmasının seçimi ise, yüksek doğruluk, hız ve overfitting'e karşı sunduğu düzenleştirci (regularization) mekanizmaları gibi avantajlara dayanmaktadır. XGBoost, gradient boosting yöntemine dayanarak, özellikle karmaşık ve doğrusal olmayan ilişkilerin bulunduğu veri setlerinde güçlü bir tahmin performansı göstermektedir. Alternatif olarak değerlendirilebilecek LSTM ve Random Forest gibi algoritmalar, kendi avantajlarına sahip olmakla birlikte, LSTM'nin derin öğrenme tabanlı yapısı nedeniyle daha büyük veri setlerine ve uzun eğitim sürelerine ihtiyaç duyması, ayrıca hiperparametre optimizasyonunun daha karmaşık olması gibi sınırlılıkları bulunmaktadır. Random Forest ise bireysel ağaçlar arasındaki bağımsız yapısı nedeniyle boosting yöntemlerinin sağladığı hata düzeltme mekanizmalarını sunmamaktadır. XGBoost'un parametrik esnekliği, paralel hesaplama desteği ve hiperparametre ayarlarıyla model karmaşıklığı üzerinde sağladığı kontrol, çalışmanın kapsamı ve veri setinin boyutu dikkate alındığında daha uygun bir seçim olarak değerlendirilmiştir. Dolayısıyla, XGBoost'un bu bağlamda modelin genel doğruluğunu ve genellenebilirliğini artırmada daha etkili olacağı öngörülerek, hibrit SARIMA-XGBoost modeli uygulanmıştır.

Çalışmada kullanılan hibrit SARIMA-XGBoost modeli, finansal zaman serilerinde hem doğrusal hem de doğrusal olmayan desenleri yakalamayı hedefleyen iki aşamalı bir modelleme sürecine dayanmaktadır. İlk aşamada, zaman serisinin doğrusal ve mevsimsel bileşenlerini modellemek için SARIMA (Seasonal Autoregressive Integrated Moving Average) modeli uygulanmıştır. SARIMA modeli, veri setindeki trend ve mevsimsel örüntüleri tahmin ettikten sonra, bu tahminlerden elde edilen artık değerler (residuals) hesaplanmıştır. Artık değerler, SARIMA modelinin açıklayamadığı, veri setindeki doğrusal olmayan ve karmaşık yapıların bir göstergesi niteliğindedir. İkinci aşamada ise, bu artık değerler bir bağımlı değişken olarak XGBoost algoritmasına girdi olarak verilmiştir. XGBoost, bu artık değerler ile oluşturulan veri setini kullanarak, SARIMA modelinin yakalayamadığı doğrusal olmayan ilişkileri ve karmaşık desenleri öğrenmeyi hedeflemiştir. Sonuç olarak, nihai tahminler, SARIMA modelinden elde edilen tahminler ile XGBoost'un artık değerler üzerine yaptığı tahminlerin toplanması yoluyla elde edilmiştir. Bu süreç, istatistiksel ve makine öğrenimi yöntemlerinin güçlü yönlerini birleştirerek, finansal zaman serisi tahmininde daha yüksek doğruluk ve genellenebilirlik sağlamayı amaçlamaktadır. Böylece hibrit model, hem SARIMA'nın istatistiksel

açıklayıcılığından hem de XGBoost'un doğrusal olmayan ilişkileri öğrenme yeteneğinden yararlanarak bir tahmin yaklaşımı sunmaktadır.

2.1. ADF Birim Kök Testi

ADF testi, Dickey-Fuller testinin geliştirilmiş bir versiyonudur ve otokorelasyon sorunlarını çözmek için gecikme terimlerini denkleme eklemektedir. Modelin genel denklemi şu şekildedir:

$$\Delta Y_t = \alpha + \beta_t + \gamma Y_{t-1} + \sum_{i=1}^p \phi_i \Delta Y_{t-i} + \epsilon_t \quad (1)$$

burada;

ΔY_t : Değişkenin birinci farkı,

α : Sabit terim,

β_t : Trend terimi,

γ : Birim kökün varlığını test eden katsayı,

ϕ_i : Gecikme katsayıları,

ϵ_t : Hata terimi

Test istatistiği kritik değerden küçükse (genellikle %1, %5 veya %10 anlamlılık seviyesinde), H_0 reddedilir ve zaman serisinin durağan olduğu sonucuna varılmaktadır. Aksi halde, zaman serisinin birim köke sahip olduğu (durağan olmadığı) kabul edilmektedir.

H_0 : Zaman serisi birim köke sahiptir (durağan değildir).

H_1 : Zaman serisi birim köke sahip değildir (durağandır).

2.2. KPSS Birim Kök Testi

KPSS KPSS testi, zaman serisinin durağanlık varsayımını test etmektedir. ADF testinin tersine, KPSS testi durağanlığın kabul edildiği bir null hipotez ile çalışmaktadır. KPSS testi, zaman serisinin sabit bir bileşene ve rastgele bir yürüyüş (random walk) terimine ayrıldığını varsaymaktadır. Modelin genel denklemi şu şekildedir:

$$Y_t = \mu + \beta t + u_t + \epsilon_t \quad (2)$$

burada;

μ : Sabit terim,

βt : Trend bileşeni,

u_t : Rastgele yürüyüş bileşeni,

ϵ_t : Hata terimi

Test istatistiği kritik değerden küçükse, H_0 kabul edilir ve zaman serisinin durağan olduğu sonucuna varılmaktadır. Test istatistiği kritik değerden büyükse, H_0 reddedilir ve zaman serisinin durağan olmadığı kabul edilmektedir.

H_0 : Zaman serisi birim köke sahip değildir (durağandır).

H_1 : Zaman serisi birim köke sahiptir (durağan değildir).

2.3. SARIMA Modeli

SARIMA (Seasonal Autoregressive Integrated Moving Average) modeli, zaman serilerindeki hem mevsimsel (seasonal) hem de mevsimsel olmayan (non-seasonal) bileşenleri modellemek için kullanılan bir tahmin yöntemidir. ARIMA modelinin mevsimselliği ele almak için genişletilmiş bir versiyonudur. Özellikle, mevsimsel desenlerin bulunduğu ve verinin

durağan hale getirilmesi gerektiği durumlarda kullanılmaktadır (Box ve diğerleri, 2015; Hyndman ve Athanasopoulos, 2018).

SARIMA modeli, üç temel bileşen üzerine kuruludur:

i. ARIMA Bileşenleri:

- AR (Autoregressive): Zaman serisinin önceki değerlerinin doğrusal kombinasyonu.
- I (Integrated): Zaman serisinin durağan hale getirilmesi için fark alma işlemi.
- MA (Moving Average): Hata terimlerinin geçmiş değerlerinin doğrusal kombinasyonu.

ii. Mevsimsel (Seasonal) Bileşenler:

- Mevsimsel otoregresif, mevsimsel fark alma ve mevsimsel hareketli ortalama terimleri.

iii. Genel Model:

$$SARIMA(p, d, q)(P, D, Q)_s \quad (3)$$

burada;

p,d,q: ARIMA modelinin otoregresif (p), fark alma (d) ve hareketli ortalama (q) dereceleri.

P,D,Q: Mevsimsel otoregresif (P), mevsimsel fark alma (D) ve mevsimsel hareketli ortalama (Q) dereceleri,

s: Mevsimsellik periyodu

SARIMA modeli şu şekilde ifade edilir:

$$\phi_p(B^s)(1 - B^s)^D(1 - B)^d y_t = \theta_q(B^s)\theta_p(B)\varepsilon_t \quad (4)$$

burada;

B: Gecikme operatörü, ($By_t = y_{t-1}$).

$(1 - B)^d$: Birinci fark alma operatörü.

$(1 - B^s)^D$: Mevsimsel fark alma operatörü (ss, mevsimsel periyodu ifade eder).

$\phi_p(B^s)$: Mevsimsel otoregresif (AR) terimlerin polinomu.

$\theta_q(B^s)$: Mevsimsel hareketli ortalama (MA) terimlerin polinomu.

$\theta_p(B)$: Hareketli ortalama (MA) parametrelerini temsil eden polinomdur.

y_t : Zaman serisinin gözlemlenen değeri.

ε_t : Beyaz gürültü (hata terimi).

2.4. XGBoost Modeli

XGBoost (Extreme Gradient Boosting), gradyan artırma (gradient boosting) algoritmasının yüksek performanslı ve optimize edilmiş bir uygulamasıdır. Bu model, büyük veri setlerinde doğruluk ve hız sağlamak amacıyla geliştirilmiştir. XGBoost, özellikle makine öğrenimi yarışmalarında ve çeşitli uygulamalarda sıkça kullanılan bir yöntemdir. XGBoost, karar ağaçlarının gradyan artırma yöntemiyle sıralı olarak eğitildiği bir modeldir. Modelin ana özellikleri şunlardır:

- *Regülerizasyon:* L1 ve L2 regülerizasyonu sayesinde aşırı öğrenme (overfitting) azaltılmaktadır.
- *Paralel İşleme:* Model, paralel işlemeye izin vererek eğitim süresini kısaltılmaktadır.
- *Hata Azaltma:* Gradyan artırma yöntemi, önceki modellerin hatalarını minimize ederek tahmin doğruluğunu artırmaktadır.
- *Kayıp Fonksiyonları:* XGBoost, kullanıcı tanımlı kayıp fonksiyonlarını destekleyerek esneklik sağlamaktadır.

Model, doğrusal olmayan ilişkilerin bulunduğu karmaşık veri setlerinde etkili sonuçlar üretir ve genellikle tahmin, sınıflandırma ve zaman serisi analizinde kullanılmaktadır (Chen ve Guestrin, 2016).

XGBoost, her bir iterasyonda kayıp fonksiyonunu optimize etmek için aşağıdaki matematiksel yapıyı kullanmaktadır:

(a) Tahmin Fonksiyonu

Bir veri kümesindeki n adet veri için y_i gerçek değerlerinin tahmini şu şekilde ifade edilmektedir:

$$\hat{y}_i = \sum_{\{k=1\}}^K \ell(y_i, \{\hat{y}\}_i) + \sum_{\{k=1\}}^K \Omega(f_k) \quad (5)$$

burada;

$\hat{y}_i$: i -inci gözlemin tahmin edilen değeri.

K : Toplam karar ağacı sayısı.

f_k : k -inci karar ağacı tarafından i -inci gözlem için üretilen tahmin.

F : Tüm karar ağaçlarının olası uzayı

(b) Kayıp Fonksiyonu

Modelin toplam hatasını ölçmek için optimize edilen kayıp fonksiyonu:

$$Obj = \sum_{i=1}^n L(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (6)$$

burada;

$L(y_i, \hat{y}_i)$: y_i gerçek değeri ile $\hat{y}_i$ tahmini arasındaki hata (genellikle kare hata veya log-loss).

$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2$: Model karmaşıklığını kontrol eden regülerizasyon terimi.

T : Karar ağacındaki yaprak sayısı.

$\|w\|$: Yaprak ağırlıkları.

γ ve λ : Regülerizasyon parametreleri.

(c) Gradyan ve Hessian (Optimizasyon)

Her iterasyonda kayıp fonksiyonunu optimize etmek için ikinci dereceden Taylor genişlemesi kullanılır:

$$L(y_i, \hat{y}_i^{(t)}) \approx L(y_i, \hat{y}_i^{(t-1)}) + g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i) \quad (7)$$

burada;

$g_i = \frac{\partial L(y_i, \hat{y}_i)}{\partial y_i}$: Gradyan (kayıp fonksiyonunun birinci türevi).

$g_i = \frac{\partial^2 L(y_i, \hat{y}_i)}{\partial y_i^2}$: Hessian (kayıp fonksiyonunun ikinci türevi).

$f_t(x_i)$: Yeni ağacın i-inci gözlem üzerindeki tahmini.

3. Analiz ve Bulgular

Tablo 1 üzerinde BTC, XRP ve ETH değişkenleri için ADF ve KPSS birim kök testi sonuçları gösterilmiştir. Tablo 1 üzerindeki p-değerleri incelendiğinde bütün değişkenlerin her iki test sonucuna göre seviyede durağan olduğu görülmektedir.

Tablo 1. ADF ve KPSS Test Sonuçları

	Test türü	Test istatistiği	p-değeri	Kritik Değerler		
				%1	%5	%10
BTC	ADF	-45.388	0.000	-3.434	-2.863	-2.568
	KPSS	0.105	0.100	0.347	0.463	0.574
XRP	ADF	-2.980	0.030	-3.4339	2.8631	2.567
	KPSS	1.121	0.100	0.739	0.463	0.574
ETH	ADF	-43.795	0.000	-2.434	0.347	0.463
	KPSS	0.102	0.100	0.347	0.563	0.574

Tablo 1’de, ADF test istatistikleri kritik değerlerden küçük çıkarken, KPSS sonuçlarının büyük ölçüde kabul edilebilir sınırlar içinde kalması, serilerin durağanlık varsayımını desteklemektedir. Bu bulgular, sonraki modelleme aşamalarında serilerin güvenle kullanılabileceğini ortaya koymaktadır.

Tablo 2. SARIMA Modeli Sonuçları

Parametre	Değer
Model	SARIMAX(1, 1, 1)x(1, 1, 1, 12)
Bağımlı Değişken	Price
Gözlem Sayısı	1462
Log-Likelihood	-15643.522
Akaike Bilgi Kriteri (AIC)	31297.043
Model Tanımı Tablosu	Değer
ARIMA (1, 1, 1)	Otoregresif (AR), fark (I), hareketli ortalama (MA) modelleme yapılmış.
Mevsimsel Bileşen	(1, 1, 1, 12): Mevsimsel otoregresif, fark ve hareketli ortalama bileşenleri 12 periyot için belirlenmiş.

Tablo 2’de, SARIMAX (1, 1, 1) x (1, 1, 1, 12) modeli, hem mevsimsel hem de mevsimsel olmayan bileşenleri ele alan bir yapıya sahiptir. ARIMA kısmı, otoregresif (1), fark alma (1) ve hareketli ortalama (1) terimlerini içermektedir. Mevsimsel bileşen, 12 periyodu (örneğin, aylık mevsimsellik) dikkate almaktadır. Model, bağımlı değişken olarak fiyat kullanılarak yapılandırılmıştır. 1462 gözlem, modelin tahmin performansını desteklemek için yeterli bir veri setine sahip olduğunu göstermektedir. Log-Likelihood:-15643.522 değeri, modelin verilere uyumunu ifade etmektedir. Daha yüksek (daha az negatif) log-likelihood değeri genelde daha iyi bir uyuma işaret etmektedir. AIC değeri olan 31297.043, modelin

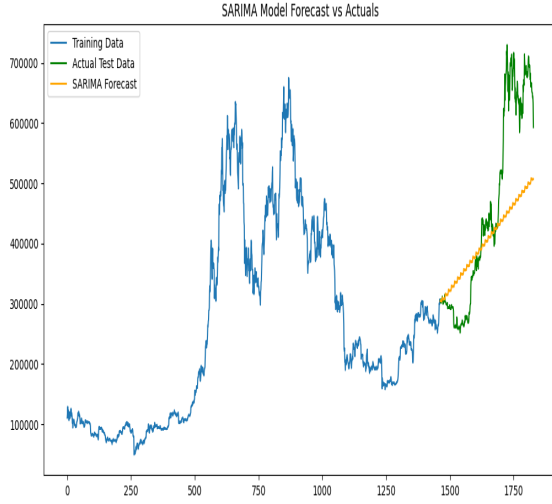
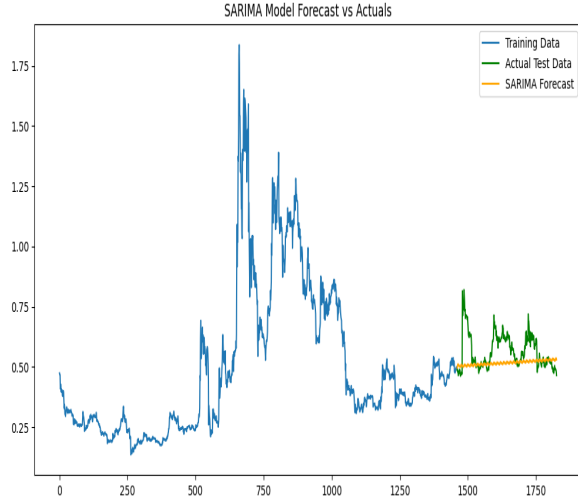
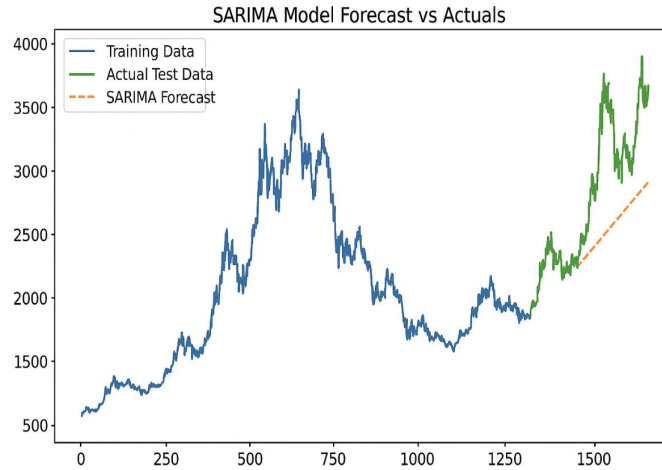
karmaşıklığı ve veri uyumunu bir arada değerlendirmektedir. Daha düşük AIC değerleri, daha iyi bir model seçimine işaret etmektedir. ARIMA (1, 1, 1), ARIMA modeli, kısa dönemli desenleri ve verinin doğrusal özelliklerini yakalamak için yapılandırılmıştır.

Tablo 3. SARIMA Modeli Performansı

	Index	Gerçek Fiyat	SARIMA Tahmini	Fark
BTC	1462	304.850	304,341.75	508,25
	1463	302.870	304,353.03	-1,483.03
	1464	307.110	305,833.82	1,276.18
	1465	300.940	305,356.32	-4,416.32
	1466	304.710	303,857.03	852.97
XRP	1462	0.4866	0.496276	-0.009676
	1463	0.4902	0.501659	-0.011459
	1464	0.4791	0.510309	-0.031209
	1465	0.4838	0.511358	-0.027558
	1466	0.4647	0.510789	-0.046089
ETH	1462	3048500	3043283	5217
	1463	3028700	3043337	-14637
	1464	3071100	3058064	13036
	1465	3009400	3053422	-44022
	1466	3047100	3.037817e	9283

Not: Modelin AIC değeri 31,297.043'tür, bu da modelin veriyle uyumlu olduğu göstermektedir.

Tablo 3'de, SARIMA modeli, Bitcoin (BTC), Ripple (XRP) ve Ethereum (ETH) fiyatlarını tahmin etmek için uygulanmış ve model performansını değerlendirmek amacıyla tahmin sonuçları gerçek değerlerle karşılaştırılmıştır. Ayrıca modelin Akaike Bilgi Kriteri (AIC) değeri hesaplanmıştır. Bu göstergeler modelin etkinliğini ve iyileştirilmesi gereken yönlerini değerlendirmek için çalışmada verilmiştir. Bitcoin için SARIMA modeli, genelde gerçek değerlere yakın tahminler üretmiştir. Ancak, bazı durumlarda (örneğin, 1465 indeksindeki -4,416.32'lik fark) tahmin ile gerçek fiyat arasında belirgin farklılıklar görülmektedir. Bu durum, modelin bazı volatilité dönemlerinde performans kaybına uğradığını göstermektedir. Ripple için model, küçük hatalarla tutarlı tahminler üretmiştir. Fark değerleri genellikle 0.05'in altındadır. Bu, modelin Ripple'ın fiyat hareketlerini tahmin etmekte oldukça başarılı olduğunu göstermektedir. Ethereum için model, bazı indekslerde (örneğin, 1465'te -44,022.00 fark) daha büyük hatalar yapmıştır. Bu, Ethereum fiyatlarının volatilitésinin model tarafından tam olarak yakalanamadığını göstermektedir. Model, Ripple (XRP) fiyatlarını daha tutarlı ve küçük hatalarla tahmin etmiştir. Ayrıca Bitcoin tahminlerinde çoğunlukla kabul edilebilir hata seviyeleri gözlenmiştir. Ethereum tahminlerinde, özellikle fiyat hareketlerinin yüksek olduğu dönemlerde daha büyük hatalar yapılmıştır. Bu, modelin volatilitéyi yakalamada iyileştirilmeye ihtiyaç duyabileceğini göstermektedir.

Grafik 1. SARIMA Modeli BTC Grafiği**Grafik 2. SARIMA Modeli XRP Grafiği****Grafik 3. SARIMA Modeli ETH Grafiği**

Grafik 1’de, BTC için, SARIMA modeli, eğitim verilerinde (mavi çizgi) genel trendi ve zaman serisinin dalgalı yapısını yakalayabilmiştir. Bu, modelin hem otoregresif hem de mevsimsel bileşenleri etkili bir şekilde öğrenme yeteneğini göstermektedir. Geçmiş verilere güçlü bir uyum sağlandığı açıkça görülmektedir. Test verileri (yeşil çizgi), modelin tahminlerinin (turuncu kesikli çizgi) doğruluğunu değerlendirmek için kullanılmıştır. Tahminler genel olarak gerçek fiyatlarla uyumlu görünmektedir. Model, uzun dönemli artış trendini başarıyla tahmin edebilmiştir. Ancak, kısa vadeli volatilité ve ani fiyat değişikliklerinde tahminler, gerçek fiyatlardan sapmalar göstermektedir. Model hem genel trendi hem de mevsimsel yapıyı başarılı bir şekilde modellemiştir. Uzun vadeli tahminlerde doğru yönü işaret etmektedir. SARIMA modelinin mevsimsel bileşenleri etkili bir şekilde ele alması, dönemsel değişimleri tahmin etmede avantaj sağladığı grafikten görülmektedir.

Grafik 2’de, XRP için, SARIMA modeli Eğitim verileri, modelin geçmiş zaman serisi verilerine uyum sağlamak için kullandığı veri setini temsil etmektedir. XRP’nin eğitim verilerinde önemli trendler ve dalgalanmalar gözlemlenmiştir. Grafik boyunca zaman serisinin genel yapısında hem yükseliş hem de düşüş eğilimleri dikkat çekmektedir. Özellikle eğitim dönemindeki bu dalgalanmalar modelin öğrenme kapasitesini test etmiştir. Model, bu veriler üzerinde eğitim aldığı için, genel bir uyum sağladığı ve bu geçmiş desenleri etkili bir şekilde öğrendiği söylenebilir. Test dönemi boyunca gerçek gözlemler yeşil çizgi ile gösterilmektedir. XRP fiyatları test dönemi boyunca daha dar bir bant içinde hareket etmiş ve düşük bir volatilité sergilemiştir. Fiyatların genelde sabit bir seviyede olduğu görülmektedir. Gerçek test verilerinde genel olarak yatay bir hareket gözlenmiştir. Bu, XRP fiyatlarının bu dönemde göreceli bir istikrar sergilediğini göstermektedir. SARIMA modeli, XRP fiyatlarının genel trendini test verileri boyunca doğru bir şekilde takip etmiştir. Ancak, modelin kısa vadeli dalgalanmaları yeterince yakalayamadığı ve bu nedenle bazı sapmaların meydana geldiği gözlemlenmiştir. Model, genel trendi ve yatay hareketi iyi bir şekilde takip etmiş olsa da, kısa vadeli fiyat değişimlerini yakalamakta yetersiz kalmış olabilir. Özellikle volatilité dönemlerinde, modelin tahmin doğruluğu bir miktar düşüş göstermiştir. SARIMA modeli, XRP fiyatlarının genel trendlerini ve test verilerindeki sabit hareketleri iyi bir şekilde tahmin etmiş olsa da volatil dönemlerdeki hassasiyeti artırmak için iyileştirmelere ihtiyaç duyulmaktadır. Bu model, mevsimsel ve trend bileşenlerini yakalamakta etkili bir araç olarak değerlendirilebilir.

Grafik 3’de, ETH için, SARIMA modeli eğitim verilerinde Ethereum fiyatlarının zamanla önemli dalgalanmalar gösterdiği görülmektedir. Özellikle 500. ve 700. indeks civarında fiyatların hızlı bir şekilde arttığı ve zirve yaptığı dikkat çekmektedir. Bu dönemlerde volatilitenin yüksek olduğu anlaşılmaktadır. Eğitim verilerinin genel trendi, zamanla yukarı yönlü bir hareket sergilemektedir. Fiyatlar, belirli dönemlerde düzeltme hareketleriyle geçici düşüşler yaşasa da genel eğilim artış yönündedir. SARIMA modeli, bu geçmiş verilere uyum sağlayarak hem trendleri hem de dönemsel dalgalanmaları başarılı bir şekilde öğrenmiştir. Eğitim verilerindeki fiyat hareketleri, modelin eğitimi için güçlü bir temel sağlamıştır. Test verilerinde Ethereum fiyatlarının belirgin bir artış trendi içinde olduğu gözlemlenmektedir. Fiyatlar 1500. indeks civarından itibaren hızlı bir yükseliş göstermiş ve zamanla yüksek volatilité ile dalgalı bir yapı sergilemiştir. Test dönemi boyunca fiyatların ani yükselişler ve düşüşler yaşadığı fark edilmektedir. Bu, Ethereum piyasasının genellikle yüksek volatilitéye sahip bir varlık olduğunu doğrulamaktadır. Gerçek test verilerinde fiyatların genel olarak artış eğiliminde olması, Ethereum’un bu dönemde değer kazandığını göstermektedir. SARIMA modeli, genel trendi doğru bir şekilde yakalamış ve test verileri boyunca fiyatların artış eğilimini yansıtmıştır. Ancak modelin, volatil dönemlerde fiyat dalgalanmalarını yeterince yakalayamadığı görülmektedir. Özellikle ani yükselişler ve düşüşlerde tahminlerin gerçek fiyatların gerisinde kaldığı anlaşılmaktadır. Model, uzun vadeli trendi başarılı bir şekilde tahmin etmiş ve fiyatların genel artış yönünü takip edebilmiştir. Ancak, kısa vadeli dalgalanmalar ve volatilité model tarafından tam olarak temsil edilememiştir.

SARIMA modeli sonuçları topluca değerlendirildiğinde kısa vadeli dalgalanmalar ve volatilité SARIMA tarafından tespit edilmesi noktasında önemli eksikleri içerdiği görülmüştür. Daha kesin sonuçlar için hibrit model SARIMA-XGBoost analizi yapılmıştır.

Tablo 4. SARIMA- XGBoost Hibrit Modeli Performans Sonuçları

	Tarih	Fiyat	Günlük Fiyat Değişim	Lag 1	Lag 2	SARIMA Tahmin
BTC	2019-06-28	123490	11490	112000	128760	112552.96
	2019-06-29	119210	-4280	123490	112000	123091.32
	2019-06-30	107450	11760	119210	123490	119370.72
	2019-07-01	106390	-1060	107450	119210	107811.96
	2019-07-02	108660	2270	106390	107450	106400.59
	2023-06-29	304710	3770	300940	307110	303857.03
	2023-06-30	305000	290	304710	300940	304755.42
	2023-07-01	306070	1070	305000	304710	303463.04
	2023-07-02	306630	560	06070	305000	305404.08
	2023-07-03	311760	5130	306630	306070	310091.27
SARIMA MSE: 13,830,743,516.54; XGBoost MSE: 1,602,024,980.89; SARIMA MAE: 89,194.40; XGBoost MAE: 28,094.09						
XRP	2019-06-28	0.426	0.016	0.410	0.409	0.467
	2019-06-29	0.426	0,0004	0.426	0.426	0.409
	2019-06-30	0.397	-0.029	0,426	0.426	0.426
	2019-07-01	0.406	0.009	0.397	0.397	0.426
	2019-07-02	0.398	-0.008	0.406	0.406	0.397
	2023-06-28	0.464	-0.019	0.510	0.483	0.479
	2023-06-29	0.474	0.009	0.504	0.464	0.483
	2023-06-30	0.472	-0.001	0.504	0.474	0.464
	2023-07-01	0.473	0.0004	0.506	0.472	0.474
	2023-07-02	0.484	0.011	0.503	0.473	0.472
ARIMA MSE: 0.0078; XGBoost MSE: 0.0014; SARIMA MAE: 0.0627; XGBoost MAE: 0.0237						
ETH	2019-06-28	0.426	0.016	0.409	0.467	0.410
	2019-06-29	0.426	0.0004	0.426	0.409	0.426
	2019-06-30	0.397	-0.029	0.426	0.426	0.426
	2019-07-01	0.406	0.009	0.397	0.426	0.397
	2019-07-02	0.398	-0.008	0.406	0.397	0.406
	2023-06-28	0.464	-0.019	0.483	0.479	0.510
	2023-06-29	0.474	0,009	0.464	0.483	0.504
	2023-06-30	0.472	-0.001	0.474	0.464	0.504
	2023-07-01	0.472	-0.001	0.474	0.464	0.504
	2023-07-02	0.473	0.0004	0.472	0.474	0.506
SARIMA MSE: 1,368,829,203,493.34; XGBoost MSE: 183,298,721,903.92; SARIMA MAE: 888,078.97; XGBoost MAE: 298,072.48						

Tablo 4'te verilen sonuçlar değerlendirildiğinde BTC, XRP ve ETH için SARIMA ve XGBoost tabanlı hibrit modellerin performans karşılaştırmasını içermektedir. Burada, MSE (Mean Squared Error) ve MAE (Mean Absolute Error) değerleri her bir modelin tahmin performansını göstermektedir. Sonuçları genel hatlarıyla BTC için; SARIMA MSE: 13.830.743.516,54 ve MAE: 89.194,4. Bu, SARIMA modelinin BTC fiyatları üzerinde oldukça yüksek hata paylarına sahip olduğunu göstermektedir. Hibrit XGBoost MSE 1.602.024.980,90 ve MAE: 28.094,10. XGBoost tabanlı hibrit model, SARIMA modeline göre daha düşük MSE ve MAE değerlerine sahip, bu da modelin daha iyi performans gösterdiğini işaret etmektedir. XRP için; SARIMA MSE: 0,0078 ve MAE: 0,0627. SARIMA modelinin XRP üzerinde yüksek doğruluk sağladığı ve düşük hata verdiği görülüyor. Hibrit XGBoost MSE: 0,0014 ve MAE: 0,0237. XGBoost tabanlı hibrit modelin performansı, SARIMA modeline göre çok daha iyi ve düşük hata paylarına sahip olduğu görülmüştür. XRP için, Hibrit XGBoost modelinin daha düşük hata oranlarına sahip olduğu ve bu modelin tahminleri geliştirmek adına daha uygun olduğu anlaşılmaktadır. ETH için yorumlar ise şu şekildedir. SARIMA MSE: 1.368.829.203.493,34 ve MAE: 888.078,97. SARIMA modelinin ETH için büyük hatalar verdiği ve yüksek MSE ile MAE değerlerine sahip olduğu görülmektedir. Hibrit XGBoost MSE: 183.298.721.903,93 ve MAE: 298.072,49. XGBoost tabanlı hibrit model de ETH için daha düşük MSE ve MAE değerleri sağlıyor, ancak hâlâ önemli hata payları mevcuttur. ETH için de, Hibrit XGBoost modelinin SARIMA'dan daha iyi performans gösterdiği anlaşılmaktadır, ancak modelin optimizasyonu daha da faydalı olacağı düşünülerek tüm kripto varlıklar optimizasyonu gerçekleştirilmiştir.

XGBoost gibi modeller, Grid Search, Random Search veya Bayesian Optimization gibi tekniklerle hiperparametre araması yaparak en iyi parametreleri bulmaktadır. Bu arama sürecinde, model belirli parametreler üzerinde çeşitli denemeler yapar ve her denemenin performansını değerlendirir. Bu parametreler genellikle cross-validation veya validation set üzerinde en düşük hata oranını veren değerlerdir. Hiperparametre optimizasyonu, modelin performansını artırmak için önemlidir çünkü bu değerler modelin öğrenme sürecini doğrudan etkilemektedir.

XGBoost modeli için hiperparametre optimizasyon sürecinde, modelin genel performansını artırmak amacıyla belirli hiperparametrelerin en uygun değerlerinin sistematik bir şekilde belirlenmesi hedeflenmiştir. Bu çalışmada hiperparametre optimizasyonu için Grid Search yöntemi tercih edilmiştir. Grid Search, önceden belirlenen bir parametre aralığında her olası kombinasyonu deneyerek en iyi sonuç veren parametre setini bulmayı amaçlayan bir arama yöntemidir. Optimizasyon sürecinde değerlendirilen temel hiperparametreler şunlardır: max_depth (ağacın maksimum derinliği), learning_rate (öğrenme oranı), n_estimators (oluşturulacak ağaç sayısı), subsample (alt örnekleme oranı) ve colsample_bytree (ağaç başına sütun örnekleme oranı). Her parametre kombinasyonu, k-katlı çapraz doğrulama (cross-validation) yöntemi kullanılarak doğrulama seti üzerinde test edilmiş ve en düşük hata oranını veren parametre seti optimal kombinasyon olarak belirlenmiştir. Bu süreç, modelin hem aşırı uyum (overfitting) hem de düşük uyum (underfitting) risklerini minimize ederek, genel tahmin performansını iyileştirmeyi amaçlamaktadır. Aşağıda yer alan Tablo 5'te, SARIMA- XGBoost Hibrit Modeli Parametreleri verilmiştir.

Tablo 5. SARIMA- XGBoost Hibrit Modeli Parametreleri

Model	BTC	XRP	ETH
Ağaç başına sütun örnekleme oranı	0.80	0.80	0.80
Öğrenme oranı (Eta)	0.20	0.20	0.20
Maksimum derinlik	8	7	5
Ağaç sayısı	200	200	200
Alt örnekleme oranı	0.80	0.80	0.80

Tablo 5'te verilen sonuçlar incelendiğinde, Ağaç başına sütun örnekleme oranı parametresi, her karar ağacını oluştururken kullanılan veri setinin ne kadarının kullanılacağını belirtmektedir. Örneğin, 0.8 değeri, her ağaç için verilerin %80'inin kullanılacağı anlamına gelmektedir. Öğrenme oranı (Eta), modelin her iterasyonda (adımda) ne kadar ilerlemesi gerektiğini belirlemektedir. Düşük bir öğrenme oranı (örneğin 0.2), modelin küçük adımlarla öğrenmesini sağlamaktadır, bu da genelleme gücünü artırmaktadır. Maksimum derinlik, her bir karar ağacının maksimum derinliğini belirlemektedir. Derin ağaçlar daha karmaşık ilişkiler öğrenebilirken, sıf ağaçlar daha basit ilişkileri modelleyebilir ve overfitting (aşırı öğrenme) riskini azaltmaktadır. Ağaç sayısı, Modelde kullanılacak toplam ağaç sayısını belirtmektedir. Daha fazla ağaç kullanmak modelin daha fazla öğrenmesine olanak tanımaktadır. Alt örnekleme oranı, modelin her iterasyonda eğitim verilerinin rastgele seçilen bir alt kümesini kullanıp kullanmayacağını belirlemektedir. 0.8 değeri, her iterasyonda verilerin %80'inin seçileceği anlamına gelmektedir. Bu, aşırı öğrenmeyi (overfitting) engellemeye yardımcı olmaktadır.

Tablo 6. SARIMA- XGBoost Hibrit Modeli Optimizasyon Sonuçları

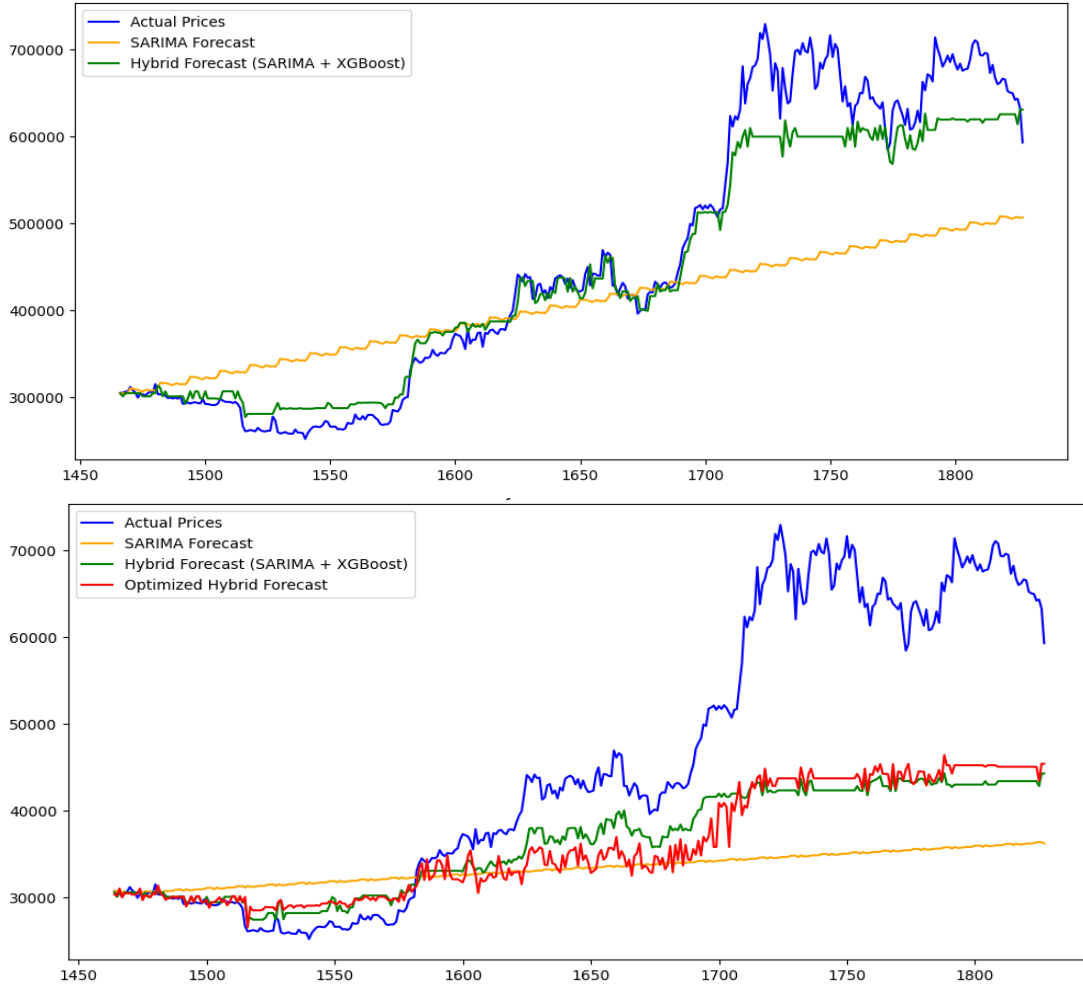
	Model	MSE	MAE
BTC	SARIMA (Set 1)	354,606,457.517	14,155.061
	İlk Hibrit XGBoost (Set 1)	195,534,867.609	9,906.886
	Optimize Hibrit XGBoost (Set 1)	188,831,157.640	10,349.065
	SARIMA (Tekrar)	354,606,457.517	14,155.061
	ARIMA	1.361,313,954,728.494	883,356.170
	İlk Hibrit XGBoost (Set 2)	206,742.679.245.070	332,217.382
	Optimize Hibrit XGBoost (Set 2)	191,403,725,932.141	321,731.079
XRP	SARIMA (Set 1)	354,606,457.517	14,155.061
	İlk Hibrit XGBoost (Set 1)	195,534,867.609	9,906.886
	Optimize Hibrit XGBoost (Set 1)	188,831,157.640	10,349.065
	ARIMA (Set 2)	354,606,457.517	14,155.061
	İlk Hibrit XGBoost (Set 2)	195,534,867.609	9,906.886
	Optimize Hibrit XGBoost (Set 2)	188,831,157.640	10,349.065
ETH	SARIMA (Set 1)	1,361,313,954,728.494	883,356.170
	İlk Hibrit XGBoost (Set 1)	191,403,725,932.141	321,731.079
	Optimize Hibrit XGBoost (Set 1)	321,731.079	321,731.079
	ARIMA (Set 2)	1,361,313,954,728.494	191,403,725,932.141
	İlk Hibrit XGBoost (Set 2)	206,742,679,245.070	332,217.382
	Optimize Hibrit XGBoost (Set 2)	191,403,725,932.141	321,731.079

Tablo 6'da BTC, XRP ve ETH için SARIMA, ARIMA ve HibritXGBoost modellerini MSE (Mean Squared Error) ve MAE (Mean Absolute Error) değerleri karşılaştırılmaktadır. Modellerin performansları optimize edilmiş ve optimize edilmemiş sürümleriyle birlikte sunulmuştur. Her bir varlık için bu sonuçları detaylı yorumları şu şekildedir;

SARIMA modeli BTC için daha yüksek hata (MSE ve MAE) verirken, İlk Hibrit XGBoost modeli önemli bir iyileşme göstermektedir. Optimize Hibrit XGBoost modeli ise biraz daha kötü sonuçlar verse de, hibrit XGBoost modeline göre hala daha iyi performans göstermektedir. ARIMA modelinin çok yüksek MSE ve MAE değerleri, bu modelin BTC için oldukça kötü bir tahmin yaptığına işaret etmektedir. Set 2'deki Hibrit XGBoost modelleri, Set 1'e göre daha yüksek MSE ve MAE değerleri göstermekte, bu da modelin optimizasyonunun BTC için ilk sette tamamlanmış olduğu anlamına gelmektedir.

XRP için sonuçlar, BTC ile benzer şekilde, İlk Hibrit XGBoost modelinin daha iyi sonuçlar verdiğini göstermektedir. Optimize Hibrit XGBoost modelinin MSE ve MAE değerleri biraz artmış olsa da, iyi bir performans sergilemektedir. ARIMA ve SARIMA modelleri arasındaki fark, XRP için de oldukça belirgindir ve SARIMA modeli yine daha iyi performans göstermektedir. ETH için, İlk Hibrit XGBoost (Set 1) modelinin performansı çok iyi bir şekilde iyileştirilmiştir. Ancak, Optimize Hibrit XGBoost (Set 1) modelinin MSE ve MAE değerleri eşitlenmiş ve modelin ilk set ile devam etmesi gerektiği sonucuna ulaşılmıştır.

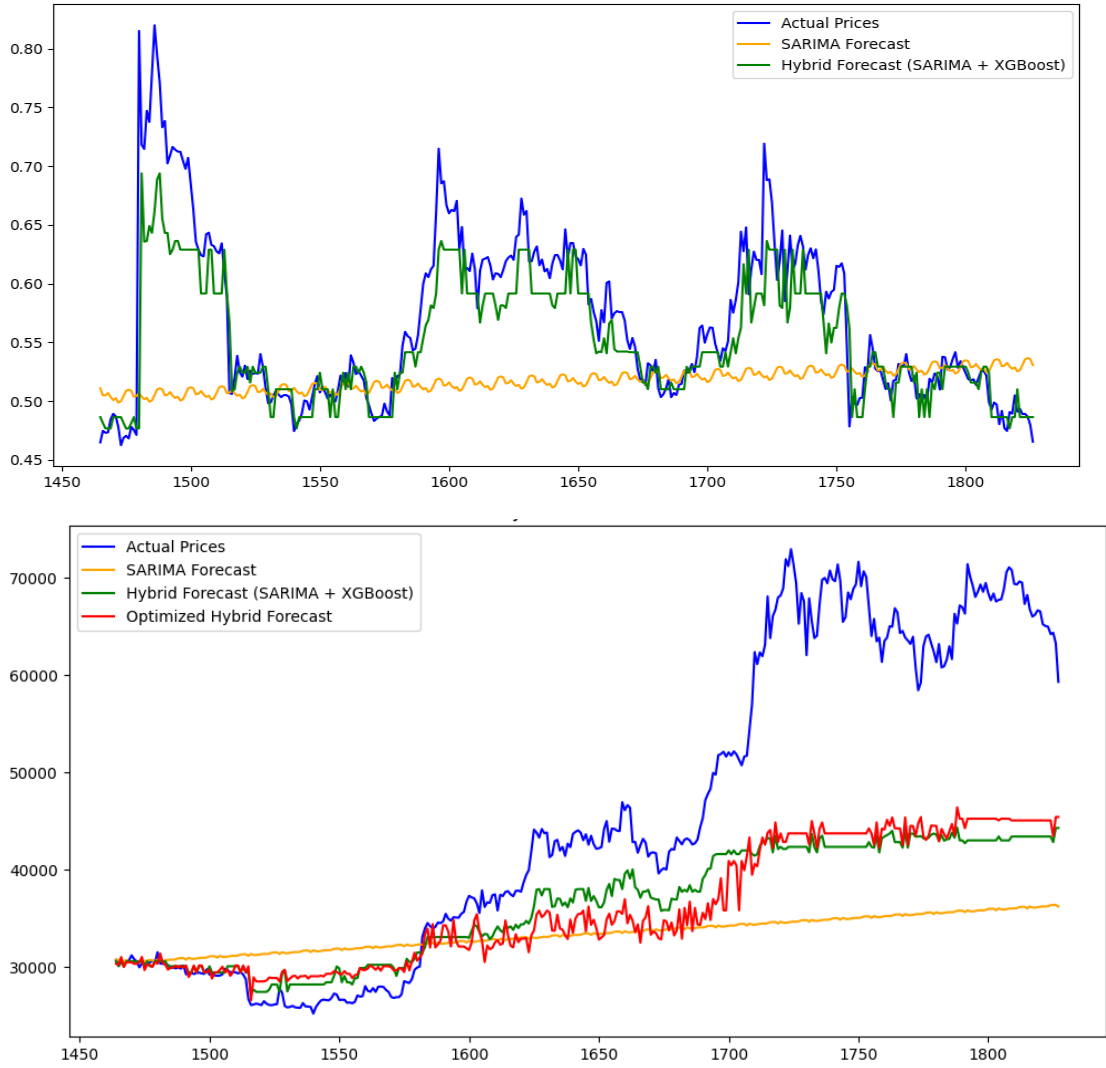
Sonuçlar bir bütün olarak değerlendirildiğinde; Hibrit XGBoost modelleri genel olarak, BTC, XRP ve ETH için daha düşük MSE ve MAE değerleri ile daha iyi performans sergilemektedir. Ancak Optimize Hibrit XGBoost modelleri, bazı durumlarda performansı artırmak yerine düşürmüş olsa da, optimizasyon sürecinin Hibrit modele göre daha iyi sonuç verdiği görülmektedir. Özetle uygulanan modellerin performansını tam olarak görebilmek için grafiklerin karşılaştırmalı olarak değerlendirilmesi gerekmektedir. Aşağıda grafiklerde tüm modellerin sonuçları gösterilmiştir.

Grafik 4. BTC için SARIMA, Hibrit ve Optimize Model Sonuçları

Grafik 4'te İlk grafik, SARIMA modelinin ve hibrit modelin (SARIMA + XGBoost) fiyat tahminlerinin karşılaştırıldığı görselidir. Burada mavi çizgi gerçek fiyatları, yeşil çizgi SARIMA modelinin tahminlerini, turuncu çizgi ise hibrit modelin tahminlerini göstermektedir. Grafikte, SARIMA modelinin fiyat hareketlerini doğru bir şekilde tahmin edemediği, özellikle fiyatlar hızla yükselirken büyük sapmalar yaşadığı görülmektedir. Hibrit model ise fiyatların yükseldiği bölgede daha doğru tahminler yapabilmıştır, ancak yine de bazı durumlarda sapmalar gözlemlenmektedir.

İkinci grafik, aynı zamanda SARIMA modelinin ve hibrit modelin karşılaştırılmasıyla birlikte optimize edilmiş hibrit modelin (kırmızı çizgi) performansını da içermektedir. Burada mavi çizgi yine gerçek fiyatları gösterirken, yeşil çizgi SARIMA tahminlerini, turuncu çizgi hibrit model tahminlerini ve kırmızı çizgi optimize edilmiş hibrit model tahminlerini göstermektedir. Optimize edilmiş hibrit model, SARIMA ve hibrit modele göre daha iyi sonuçlar vermektedir, ancak yine de belirli dönemlerde küçük sapmalar gözlemlenmektedir. Sonuç olarak, optimize edilmiş hibrit model daha başarılı bir tahmin performansı sergilemekle birlikte, hala bazı zorluklarla karşılaşmaktadır.

Grafik 5. XRP için SARIMA, Hibrit ve Optimize Model Sonuçları

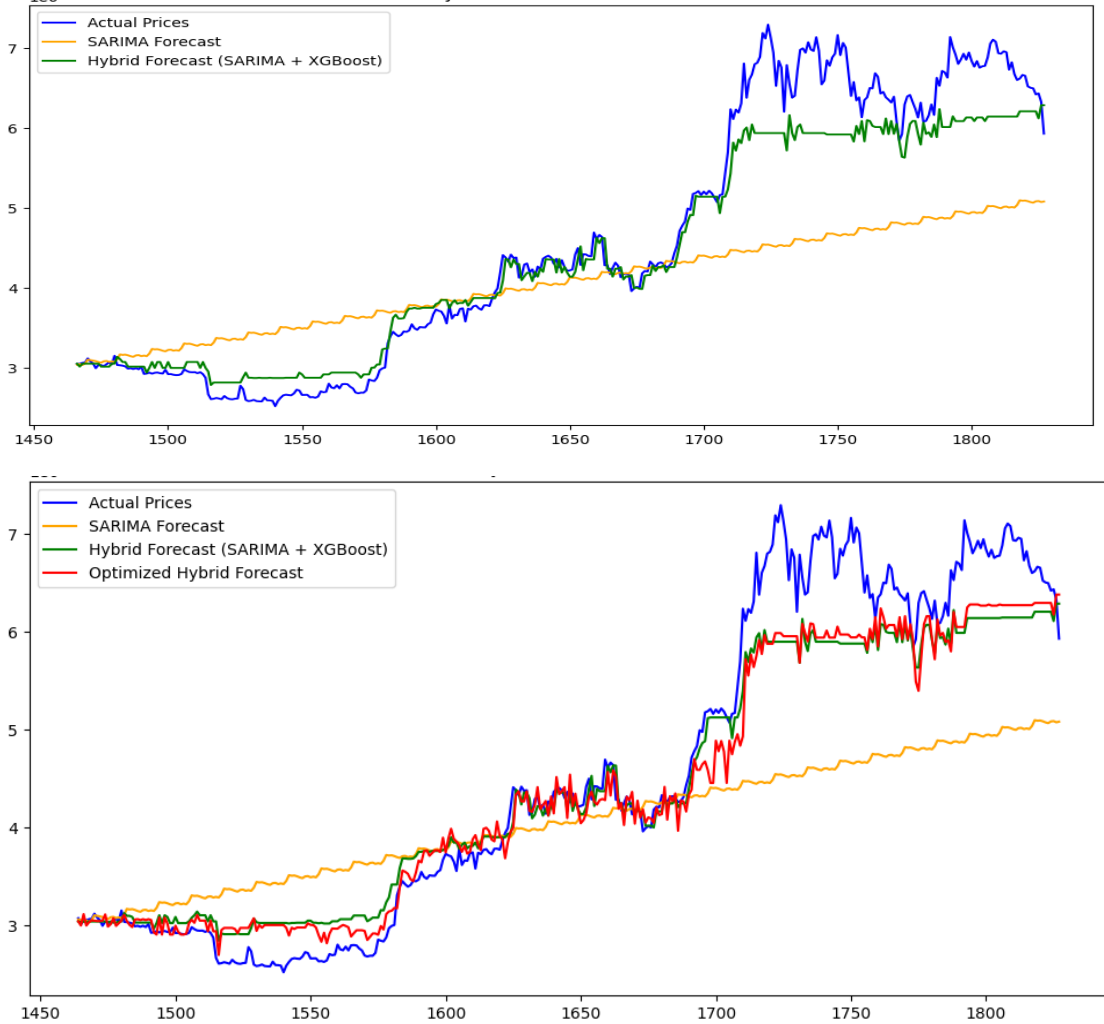


Grafik 5'te ilk grafik, XRP için SARIMA modelinin ve hibrit modelin (SARIMA + XGBoost) fiyat tahminlerinin karşılaştırıldığı görseldir. Burada mavi çizgi gerçek XRP fiyatlarını, turuncu çizgi SARIMA modelinin tahminlerini, yeşil çizgi ise hibrit modelin tahminlerini göstermektedir. Grafikte, SARIMA modelinin XRP fiyat hareketlerini doğru bir şekilde tahmin edemediği, özellikle fiyatların dalgalandığı ve hızla değiştiği bölgelerde büyük sapmalar yaşandığı görülmektedir. Hibrit model ise XRP fiyat hareketlerini daha iyi yakalamış, dalgalı bölgelerde SARIMA'ya kıyasla daha tutarlı tahminler üretmiştir. Ancak, hibrit modelde de bazı dönemlerde sapmalar gözlemlenmiştir.

İkinci grafik, XRP için SARIMA modelinin ve hibrit modelin karşılaştırılmasının yanı sıra optimize edilmiş hibrit modelin (kırmızı çizgi) performansını da içermektedir. Burada mavi çizgi yine gerçek XRP fiyatlarını gösterirken, turuncu çizgi SARIMA tahminlerini, yeşil çizgi hibrit model tahminlerini ve kırmızı çizgi optimize edilmiş hibrit model tahminlerini göstermektedir. Optimize edilmiş hibrit model, SARIMA ve standart hibrit modele göre daha başarılı bir performans sergilemiştir. XRP'nin fiyat trendlerini ve dalgalanmalarını daha doğru bir şekilde yakalamış olsa da, belirli dönemlerde küçük sapmalar yaşandığı görülmektedir. XRP fiyat tahminlerinde optimize edilmiş hibrit model, diğer modellere göre daha üstün bir tahmin performansı sergilemektedir. Ancak XRP'nin yüksek volatilitéye sahip yapısı nedeniyle

bazı zorluklar devam etmektedir. Modelin performansını artırmak için daha fazla optimizasyon yapılması veya modele ek değişkenler dahil edilmesi gerekebilir. XRP gibi değişken piyasa koşullarına sahip varlıklar için bu tür hibrit modeller umut verici bir çözüm sunmaktadır.

Grafik 6. ETH için SARIMA, Hibrit ve Optimize Model Sonuçları



Grafik 6'da ilk grafik, SARIMA modelinin ve hibrit modelin (SARIMA + XGBoost) fiyat tahminlerinin karşılaştırıldığı görseldir. Burada mavi çizgi gerçek fiyatları, turuncu çizgi SARIMA modelinin tahminlerini, yeşil çizgi ise hibrit modelin tahminlerini göstermektedir. Grafikte, SARIMA modelinin fiyat hareketlerini doğru bir şekilde tahmin edemediği, özellikle fiyatların hızla yükseldiği ve dalgalandığı bölgelerde büyük sapmalar yaşadığı görülmektedir. Hibrit model ise fiyatların yükseldiği ve dalgalı olduğu bölgelerde daha doğru tahminler yapabilmiş, ancak yine de bazı dönemlerde sapmalar gözlemlenmektedir.

İkinci grafik, SARIMA modelinin ve hibrit modelin karşılaştırılmasının yanı sıra optimize edilmiş hibrit modelin (kırmızı çizgi) performansını da içermektedir. Burada mavi çizgi yine gerçek fiyatları gösterirken, turuncu çizgi SARIMA tahminlerini, yeşil çizgi hibrit model tahminlerini ve kırmızı çizgi optimize edilmiş hibrit model tahminlerini göstermektedir. Optimize edilmiş hibrit model, SARIMA ve standart hibrit modele göre daha iyi sonuçlar sunmuş ve fiyat hareketlerini daha doğru bir şekilde yakalayabilmiştir. Ancak optimize edilmiş hibrit modelde

de belirli dönemlerde küçük sapmalar gözlemlenmektedir. Optimize edilmiş hibrit model, diğer modellere göre daha başarılı bir tahmin performansı sergilemektedir. Bununla birlikte, ETH'nin yüksek volatilitesi nedeniyle bazı zorluklar hala devam etmektedir.

Sonuç ve Tartışma

Bu çalışma, kripto para birimlerinin fiyat tahmininde SARIMA-XGBoost hibrit modelinin etkinliğini değerlendirmek amacıyla gerçekleştirilmiştir. Bulgular, önerilen modelin hem doğrusal hem de doğrusal olmayan ilişkileri bir araya getirerek geleneksel yöntemlere kıyasla daha düşük hata oranları sunduğunu ortaya koymuştur. SARIMA-XGBoost modeli, özellikle volatilité dönemlerinde üstün bir performans sergilemiş, BTC, XRP ve ETH gibi farklı kripto varlıklar üzerinde tutarlı sonuçlar elde etmiştir.

SARIMA modeli, mevsimsel ve doğrusal desenleri yakalamada başarılı olmasına rağmen, kısa vadeli dalgalanmalar ve yüksek volatilitéyi yeterince temsil edememiştir. Buna karşın, XGBoost algoritması, doğrusal olmayan ilişkileri etkili bir şekilde modelleyerek tahmin performansını artırmıştır. SARIMA modelinden elde edilen artık değerlerin XGBoost'a girdi olarak kullanılması, her iki modelin güçlü yönlerinin birleştirilmesine olanak sağlamıştır. Bu sayede, hem genel trendlerin hem de karmaşık piyasa dinamiklerinin tahmin edilebilmesi mümkün olmuştur.

Araştırma sonuçları, hibrit modelin özellikle BTC ve XRP gibi daha az volatil varlıklarda daha yüksek doğruluk sağladığını, ancak ETH gibi yüksek volatilitéye sahip varlıklarda da belirgin bir iyileşme sunduğunu göstermiştir. Bununla birlikte, ETH gibi varlıkların tahmininde gözlenen daha yüksek hata oranları, modelin optimize edilmesi gerektiğini ortaya koymaktadır. Bu bağlamda, hiperparametre optimizasyonu süreçlerinin model performansını daha da artırabileceği değerlendirilmektedir. Bu bulgular, literatürdeki benzer çalışmalarla da uyumludur. Örneğin, Nichani ve arkadaşları (2023), ARIMA modelini LSTM ve XGBoost ile birleştirerek hisse senedi fiyat tahmininde doğrusal ve doğrusal olmayan desenleri birlikte modellemenin tahmin doğruluğunu artırdığını göstermiştir. Benzer şekilde, Liang (2024), ARIMA, Attention tabanlı CNN-LSTM ve XGBoost modellerini birleştirerek ABD hisse senedi piyasasında yüksek doğrulukta tahminler elde etmiştir. Ayrıca, Bülbül (2024), Genetik Algoritma ve Yapay Sinir Ağı tabanlı hibrit modeller kullanarak Bitcoin fiyat tahmininde yüksek doğruluk oranları elde etmiştir. Ek olarak, Achmadi ve diğerleri (2023), SARIMAX ve LSTM modellerini birleştirerek kripto para piyasasındaki volatilitéyi daha iyi modellemiş ve tahmin doğruluğunu artırmıştır. Elde edilen bulgular literatürdeki çalışmaları destekler niteliktedir.

Ayrıca, modelin performansının yatırım kararları ve risk yönetimi süreçlerinde kullanılabilir olması, çalışmanın uygulama potansiyelini güçlendirmektedir. Bununla birlikte, piyasa dinamiklerinin hızlı değişim göstermesi nedeniyle, modelin gerçek zamanlı verilerle sürekli güncellenmesi gerekliliği öne çıkmaktadır. Ayrıca, yatırımcı duyarlılığı, makroekonomik göstergeler ve sosyal medya etkisi gibi ek faktörlerin modele entegre edilmesi, tahmin doğruluğunu artırabileceği düşünülmektedir. Sonuç olarak, SARIMA-XGBoost hibrit modeli, kripto para piyasalarının karmaşık ve volatil yapısını anlamada ve bu piyasalarda daha iyi tahminler yapmada etkili bir araç olarak öne çıkmaktadır. Ancak, gelecekteki çalışmalar, modelin kısa vadeli tahminlerdeki performansını iyileştirmek ve daha geniş bir veri setiyle test etmek üzere odaklanmalıdır.

Kaynakça

- Achmadi, G. R., Saikhu, A., & Amaliah, B. (2023). Cryptocurrency price movement prediction using the hybrid sarimax-lstm method. 2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA), 711-716. <https://doi.org/10.1109/ICAMIMIA60881.2023.10427973>
- Alessandretti, L., ElBahrawy, A., Aiello, L. M., & Baronchelli, A. (2018). Machine learning the cryptocurrency market. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3183792>
- Alhnaity, B., & Abbod, M. (2020). A new hybrid financial time series prediction model. *Engineering Applications of Artificial Intelligence*, 95, 103873. <https://doi.org/10.1016/j.engappai.2020.103873>
- Bitto, A. K., Mahmud, I., Bijoy, M. H. I., Jannat, F. T., Arman, M. S., Shohug, M. M. H., ... & Jahan, H. (2022). Cryptoar: Scrutinizing the trend and market of cryptocurrency using machine learning approach on time series data. *Indonesian Journal of Electrical Engineering and Computer Science*, 28(3), 1684. <https://doi.org/10.11591/ijeecs.v28.i3.pp1684-1696>
- Bouri, E., Gupta, R., & Roubaud, D. (2019). Herding behaviour in cryptocurrencies. *Finance Research Letters*, 29, 216-221. <https://doi.org/10.1016/j.frl.2018.07.008>
- Bülbül, M. A. (2024). Hybrid optimal time series modeling for cryptocurrency price prediction: feature selection, structure and hyperparameter optimization. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 13(3), 731-743. <https://doi.org/10.17798/bitlisfen.1479725>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: Forecasting and control (5th ed.). Wiley. <https://doi.org/10.1002/9781118619193>
- Caporale, G., Kang, W., Spagnolo, F., & Spagnolo, N. (2019). Non-linearities, cyber attacks and cryptocurrencies. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3409138>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794. <https://doi.org/10.1145/2939672.2939785>
- Derbentsev, V., Kibalnyk, L., & Radzihovska, Y. (2019). Modelling multifractal properties of cryptocurrency market. *Periodicals of Engineering and Natural Sciences (Pen)*, 7(2), 690. <https://doi.org/10.21533/pen.v7i2.559>
- Fleischer, J., Laszewski, G., Theran, C., & Bautista, Y. (2022). Time series analysis of cryptocurrency prices using long short-term memory. *Algorithms*, 15(7), 230. <https://doi.org/10.3390/a15070230>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Gao, M., Yang, H., Xiao, Q., & Goh, M. (2022). Covid-19 lockdowns and air quality: Evidence from grey spatiotemporal forecasts. *Socio-Economic Planning Sciences*, 83, 101228. <https://doi.org/10.1016/j.seps.2022.101228>
- Gupta, H., & Chaudhary, R. (2022). An empirical study of volatility in cryptocurrency market. *Journal of Risk and Financial Management*, 15(11), 513. <https://doi.org/10.3390/jrfm15110513>
- Haykır, Ö., & Yağlı, İ. (2022). Speculative bubbles and herding in cryptocurrencies. *Financial Innovation*, 8(1). <https://doi.org/10.1186/s40854-022-00383-0>
- Hoa, N., Thi, Q., & Ngoan, N. (2023). Time series prediction based on machine learning: A case study, temperature forecasting in Vietnam. *Journal of Military Science and Technology*, 85, 152-162. <https://doi.org/10.54939/1859-1043.j.mst.85.2023.152-162>
- Hossain, M., Ismail, M., & Hossain, M. (2022). Enhancing stock price prediction using empirical mode decomposition, rolling forecast, and combining statistical methods. *International Journal of Computing and Digital Systems*, 12(6), 1343-1356. <https://doi.org/10.12785/ijcds/1201108>

- Huang, H. (2023). Long-term changes in Ethereum prices:a normalized pandemic framework. *BCP Business & Management*, 38, 72-78. <https://doi.org/10.54691/bcpbm.v38i.3672>
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice (2nd ed.). OTexts. <https://doi.org/10.1515/9783110612500>
- Iqbal, M., Iqbal, M. S., Jaskani, F. H., Iqbal, K., & Hassan, A. (2021). Time-series prediction of cryptocurrency market using machine learning techniques. *EAI Endorsed Transactions on Creative Technologies*, 8(28), 170286. <https://doi.org/10.4108/eai.7-7-2021.170286>
- Jabeur, S. B., Khalfaoui, R., & Arfi, W. B. (2021). The effect of green energy, global environmental indexes, and stock markets in predicting oil price crashes: Evidence from explainable machine learning. *Journal of Environmental Management*, 298, 113511. <https://doi.org/10.1016/j.jenvman.2021.113511>
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3146–3154. <https://doi.org/10.5555/3294996.3295074>
- Kim, J., Won, J., Kim, H., & Heo, J. (2021). Machine-learning-based prediction of land prices in Seoul, South Korea. *Sustainability*, 13(23), 13088. <https://doi.org/10.3390/su132313088>
- Li, Z., Han, J., & Song, Y. (2020). On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning. *Journal of Forecasting*, 39(7), 1081-1097. <https://doi.org/10.1002/for.2677>
- Liang, L. (2024). ARIMA with Attention-based CNN-LSTM and XGBoost hybrid model for stock prediction in the US stock market. *SHS Web of Conferences*, 196, 02001. <https://doi.org/10.1051/shsconf/202419602001>
- Liu, Z., Zhu, Z., Gao, J., & Xu, C. (2021). Forecast methods for time series data: A survey. *IEEE Access*, 9, 91896-91912. <https://doi.org/10.1109/access.2021.3091162>
- Nichani, R., Gasmi, L., Laiche, N., & Kabou, S. (2023). Optimizing financial time series predictions with hybrid ARIMA, LSTM, and XGBoost Models. *Studies in Engineering and Exact Sciences*. <https://doi.org/10.54021/seesv5n2-5820js.studiespublicacoes.com.br>
- Papadimitriou, T., Gogas, P., & Athanasiou, A. (2022). Forecasting Bitcoin spikes: A GARCH-SVM approach. *Forecasting*, 4(4), 752-766. <https://doi.org/10.3390/forecast4040041>
- Rampini, L., & Cecconi, F. R. (2021). Artificial intelligence algorithms to predict Italian real estate market prices. *Journal of Property Investment & Finance*, 40(6), 588-611. <https://doi.org/10.1108/jpif-08-2021-0073>
- Song, X. (2023). Predicting cryptocurrency investment suitability using machine learning techniques. *Applied and Computational Engineering*, 29(1), 133-141. <https://doi.org/10.54254/2755-2721/29/20230726>
- Shen, Y., & Wang, H. (2023). Valuation and forecasting of cryptocurrency: analysis of Bitcoin, Ethereum, and Dogecoin. *BCP Business & Management*, 38, 1067-1074. <https://doi.org/10.54691/bcpbm.v38i.3828>
- Takaishi, T. (2018). Statistical properties and multifractality of Bitcoin. *Physica A: Statistical Mechanics and Its Applications*, 506, 507-519. <https://doi.org/10.1016/j.physa.2018.04.046>
- Takaishi, T. (2021). Time-varying properties of asymmetric volatility and multifractality in Bitcoin. *PLoS One*, 16(2), e0246209. <https://doi.org/10.1371/journal.pone.0246209>
- Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price movement prediction of cryptocurrencies using sentiment analysis and machine learning. *Entropy*, 21(6), 589. <https://doi.org/10.3390/e21060589>
- Zheng, M., Guo, F., Zhang, X., & Chang, C. (2023). The transaction behavior of cryptocurrency and electricity consumption. *Financial Innovation*, 9(1). <https://doi.org/10.1186/s40854-023-00449-7>