

Hisse Senedi Fiyatlarının LSTM ve ARIMA Modelleri Kullanılarak Tahmin Edilmesi

Yıldız AYDIN^{1*}, Gizem VAROL², Eyyüb Ensari GÖKDENİZ³, Hakan MANUS⁴

^{1,2,3,4} Yönetim Bilişim Sistemleri Bölümü, İktisadi, İdari ve Sosyal Bilimler Fakültesi, Gelişim Üniversitesi, İstanbul, Türkiye

¹yaydin@gelisim.edu.tr, ²gizemvaroll@gmail.com, ³200326141@ogr.gelisim.edu.tr, ⁴200326059@ogr.gelisim.edu.tr

(Geliş/Received: 11/06/2024;

Kabul/Accepted: 24/09/2024)

Öz: Borsa İstanbul'da işlem gören hisse senedi piyasası, günümüzde en çok rağbet gören yatırım araçlarından biridir. Öyle ki nüfusa göre en çok işlem gerçekleştirilen ilk beş ülkeden biri Türkiye'dir. Teknolojinin gelişimi, yatırımcıların bu gibi yatırım araçlarına yönelmesini kolaylaştırmasının yanı sıra analiz yöntemlerini de kolaylaştırmıştır. Bu çalışma, kullanılan analiz yöntemlerinden olan LSTM ve ARIMA modellerini karşılaştırıp, hangisinin daha iyi performans sağladığını görmek için yapılmıştır. Çalışmada diğer çalışmalardan farklı olarak 2020-2024 yılları arasındaki verilerle gerçekleştirilmiş ve sektörel farklılıklar göz önünde bulundurularak işlem yapılmıştır. Sonuçlar, LSTM'nin karmaşık ve uzun vadeli bağımlılıkları yakalamada daha etkili olduğunu göstermiştir. Bu, yatırımcıların model seçimini yaparken daha bilinçli kararlar almasına yardımcı olacaktır.

Anahtar kelimeler: BIST, Otoresif Entegre Hareketli Ortalama, Uzun-Kısa Süreli Bellek Modeli, LSTM, ARIMA.

Forecasting Stock Prices Using LSTM and ARIMA Models

Abstract: The stock market traded on Borsa İstanbul is one of the most popular investment instruments today. In fact, Turkey is one of the top five countries with the highest number of transactions per capita. The development of technology has not only made it easier for investors to turn to such investment instruments but also facilitated the analysis methods. This study aims to compare the LSTM and ARIMA models, which are among the analysis methods used, to see which one provides better performance. Unlike other studies, this study was conducted using data from the years 2020-2024 and took sectoral differences into consideration. The results showed that LSTM is more effective in capturing complex and long-term dependencies. This will help investors make more informed decisions when choosing a model.

Key words: BIST, Autoregressive Integrated Moving Average, Long Short-Term Memory Model, LSTM, ARIMA.

1. Giriş

Borsa İstanbul (BİST), Türkiye'deki sermaye piyasası araçlarının, kambiyo ve kıymetli madenler ile taşların, SPK'nın (Sermaye Piyasası Kurulu) uygun gördüğü diğer sözleşme, belge ve kıymetlere yönelik alım satım emirlerinin toplanıp eşleştirildiği, işlemler sonucunda oluşan fiyatların tespit ve ilan edildiği pazar yeridir [1]. BİST30, Borsa İstanbul'da işlem gören hem işlem hacmi hem de piyasa değeri en yüksek 30 şirketin hisse senetlerinin performanslarının ölçülmesiyle meydana gelen bir endekstir. Borsa, ülkemizde yaşanan ekonomik sorunların sonucunda pek çok yatırımcının odak noktası haline gelmiş, yatırımcı sayısını artırarak ilerlemiştir. Bu sebeple borsa analizi değer kazanmış ve yatırımcılar için ilgi çekici bir konu olmuştur. Borsa analizi, geçmiş ve güncel piyasa verilerini kullanarak gelecekteki hisse senedi fiyatlarını ve piyasa trendlerini tahmin etme işlemidir. Bu analiz, yatırımcılara bilinçli yatırım kararları vermelerine yardımcı olmak için çeşitli teknikler ve yöntemler kullanır. Borsa analizi için temel ve teknik analiz olmak üzere iki yöntem ele alınır. Temel analiz, bir şirketin içsel değerini belirlemek için kullanılan bir yatırım yöntemidir. Bu analiz, şirketin mali durumunu, sektördeki konumunu, büyüme potansiyelini ve diğer faktörleri inceler. Teknik analiz, fiyat hareketlerini ve grafikleri inceleyerek yatırım kararları vermeyi amaçlayan bir yöntemdir. Bu analiz, trendleri, destek ve direnç seviyelerini ve momentumu belirlemek için kullanılır. Teknik analiz, çoğu zaman kısa vadeli hisse senedi davranışını anlamak için kullanılırken, uzun vadeli tahminler için temel analiz kullanılır [2]. Bu yöntemler yatırımcıların zararını minimum ölçüde tutabilmek ve yatırımlarının tutarlılığını artırabilmek açısından oldukça önemlidir. Bu çalışmanın amacı, BİST30 içerisinde bulunan hisselerin verilerini zaman serisi tahmini için kullanılan ve doğrusal bir model olan ARIMA (AutoRegressive Integrated Moving Average - Otoresif Entegre Hareketli Ortalama) [3] ile yapay sinir ağlarına dayalı bir derin öğrenme modeli olan LSTM (Long Short-Term Memory - Uzun Kısa Dönemli Bellek) [4] modelini kullanarak kısa vadeli fiyat tahminleri elde etmek ve bu iki modeli karşılaştırmaktır. ARIMA

* Sorumlu yazar: gizemvaroll@gmail.com. Yazarların ORCID Numarası: ¹ 0000-0001-9061-0109, ² 0000-0002-2385-2586, ³ 0009-0009-5928-8532, ⁴ 0009-0002-4946-8168

ve LSTM geçmişte farklı çalışmalarda hem karşılaştırma hem de analiz amacıyla kullanılmıştır [5]. Değişken ekonomik gelişmelerin sonucunda artan finansal faaliyetler ve bu faaliyetlerin karmaşık bir hal alması, hisse senedi gibi birden fazla faktöre bağlı yatırım araçlarını da etkileyen bir durum olduğundan, bu çalışma da BİST30 verileri kullanılmıştır.

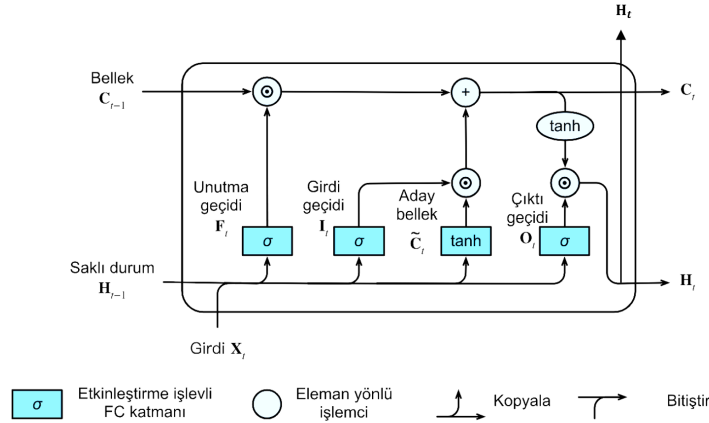
Literatürde Namin ve Namin [6], ekonomik ve finansal zaman serilerinin tahmininde ARIMA ve LSTM modellerini karşılaştırmışlardır. Çalışmada, ARIMA modelinin geleneksel bir zaman serisi tahmin yöntemi olarak doğruluğu ve hassasiyetiyle tanındığı, ancak derin öğrenme tabanlı yeni algoritmaların, özellikle Long Short-Term Memory (LSTM) algoritmasının, bu geleneksel yöntemlere göre üstün performans sergileyip sergilemediği araştırılmıştır. Araştırma bulguları, LSTM modelinin hata oranlarını %84 ile %87 oranında azalttığını ve bu nedenle ARIMA modeline kıyasla daha üstün olduğunu göstermektedir. Ayrıca, çalışmada LSTM modelinin eğitim süresi sayısının (epoch) model performansı üzerinde rastgele bir etki yaptığı ve belirli bir desen izlemediği gözlemlenmiştir. Tanışman, Karcıoğlu, Uğur ve Bulut [7], bitcoin fiyatının tahmininde LSTM sinir ağı ve ARIMA zaman serisi modellerini karşılaştırmıştır. Çalışmada, çok değişkenli LSTM sinir ağı modeli ve klasik ARIMA zaman serisi modeli kullanılarak Bitcoin'in gelecek değerinin tahmini yapılmıştır. Deneysel çalışmalar sonucunda, LSTM modelinin yakın ve uzak gelecekte düşük hata oranı ile tahmin performansı sergilediği, ARIMA modelinin ise sadece yakın gelecek için düşük hata oranı ile tahmin yaptığı tespit edilmiştir. Albayrak ve Saran [8], Borsa İstanbul verilerini kullanarak hisse senedi fiyatlarının tahmininde ARIMA, LSTM, GRU ve dikkat katmanlı LSTM modellerini karşılaştırmıştır. Çalışma, ARIMA modelinin diğer derin öğrenme tabanlı modellerle karşılaştırıldığında daha yüksek hata oranına sahip olduğunu göstermiştir. Deneysel sonuçlar, GRU modelinin en iyi performansı sergilediğini ortaya koymuştur. Çalışma, ARIMA, LSTM ve GRU modellerinin hisse senedi tahminlerindeki performansını değerlendirmiştir. Eyllasov ve Çiçek [9], kripto para getirilerinin tahmininde ARIMA-GARCH ve LSTM yöntemlerini karşılaştırmışlardır. Çalışmada, Bitcoin, Ethereum ve Binance Coin gibi kripto paraların getirileri, geleneksel zaman serisi modellerinden biri olan ARIMA-GARCH modeli ile derin öğrenme tabanlı bir yöntem olan LSTM ile tahmin edilmiştir. Araştırma sonuçları, Bitcoin için ARIMA-GARCH modelinin eğitim verilerinde daha iyi sonuçlar verdiğini, test verilerinde ise LSTM modelinin üstün performans sergilediğini göstermiştir. Ayrıca, Binance Coin'de hem eğitim hem de test verileri için LSTM modelinin daha başarılı olduğu, Ethereum'da ise ARIMA-GARCH modelinin her iki veri setinde de daha iyi sonuçlar elde ettiği tespit edilmiştir. Bu çalışma da, kripto para piyasalarında kullanılan iki farklı yöntem arasındaki performans farklarını ortaya koyarak literatüre katkı sağlamaktadır.

2. Materyal ve Yöntem

Çalışmada kullanılan veriler, BİST30'a dahil olan THYAO (Türk Hava Yolları-Ulaştırma Hizmetleri), AKBNK (Akbank-Bankacılık Hizmetleri), KCHOL (Koç Holding-Yatırım Şirketleri) ve KRDMMD (Kardemir-Metaller ve Madencilik) hisse senetlerinin 24.04.2020 tarihi ile 24.04.2024 tarihi arasında günlük kapanış verilerini içermektedir. Bu hisse senetlerinin de ele alınmasının sebebi her firmanın sektör ve yönetim alanlarında farklılık göstermesidir. Seçili senetler Borsa İstanbul'da yüksek işlem hacmine sahip olup, piyasa likiditesi ve yatırımcı ilgisi açısından önemli göstergelerdir. Bu, modellerin sonuçlarının güvenilirliğini artırmakta ve piyasadaki genel eğilimleri daha iyi temsil etmektedir. Bahsi geçen hisse senetlerine ait veriler Yahoo Finance [10] üzerinden elde edilmiş olup, Jupyter Notebook ortamında Python programlama dili ile işlenmiştir. Bu çalışmada, dört farklı sektörden birer adet hisse senedine ait gözlemlere yer verilmiştir.

2.1. LSTM (Uzun Kısa Süreli Bellek Ağı)

LSTM, 1997 yılında Hochreiter ve Schmidhuber tarafından geliştirilmiştir [11]. Geri yayılım algoritması gibi geleneksel sinir ağları eğitiminde karşılaşılan ve uzun vadeli bağımlılıkları öğrenmeyi zorlaştıran, kaybolan gradyan sorununu çözmeyi amaçlayan çalışmalar sonucu ortaya çıkmıştır. LSTM, bu sorunu girdi kapısı, çıktı kapısı, unutmaya kapısı olmak üzere üç farklı kapı aracılığı ile çözer. Girdi kapısı yeni bilgilere ne kadar izin verileceğini ya da hücrede hangi verilerin depolanacağını, çıktı kapısı hücrede yapılacak güncellemelerin ya da hangi verilerin modelin çıktısında dahil edileceğini belirlemek için kullanılırken unutmaya kapısı ise hücrede hangi verilerin unutulacağını ya da atılacak verileri belirlemek için kullanılır.



Şekil 1. LSTM Mimarisi.

Şekil 1’de verilen LSTM mimarisinde, x_t bilgilerin girişini, σ sigmoid aktivasyonunu, denklem içeriğinde bulunan W parametre matrisi ve b ise vektörü temsil etmektedir. Şekil 1’de gösterilen LSTM yapısındaki kapıların ve katmanların denklemleri sırasıyla aşağıdaki gibidir. Denklem (1)’de belirtildiği gibi, unutma kapısı, önceki gizli durum ve mevcut girdi kullanılarak hesaplanır. Girdi kapısı ise Denklem (2)’de açıklandığı üzere, yine önceki gizli durum ve mevcut girdiyi kullanarak hesaplanır. Denklem (3)’te açıklanan aday hücre durumu, önceki hücre durumu ve mevcut girdinin bir fonksiyonudur. Hafıza hücrelerinin güncellenmesi ise Denklem (4)’te gösterildiği gibi, önceki hücre durumu ile aday hücre durumunun birleşiminden oluşur. Çıkış kapısı, Denklem (5)’te açıklanan bir işlemle hesaplanır ve son olarak, gizli durum Denklem (6)’da belirtildiği gibi çıkış kapısı ve güncellenmiş hücre durumu kullanılarak hesaplanır. Bu denklemler, LSTM ağıının nasıl çalıştığını adım adım açıklayarak hücre durumu ve gizli durumun nasıl güncellendiğini gösterir.

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

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\tilde{C}_t = \tanh(WC \cdot [h_{t-1}, x_t] + bC) \quad (3)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (6)$$

Çalışmada LSTM ağı ile tahminleme gerçekleştirme amacıyla geliştirilen model üç katmana sahiptir ve her katmanda 100 nöron bulunmaktadır. Modelde, dropout oranı 0,2, loss fonksiyonu olarak MSE, optimizör olarak Adam, epoch sayısı 100 ve aktivasyon fonksiyonu olarak ReLU kullanılmıştır. Dropout, overfitting’i önlemek için %20 olarak belirlenmiştir. MSE, regresyon hatalarını minimize etmek için seçilmiştir. Adam, hızlı ve doğru optimizasyon sağlar. ReLU ise pozitif girdilerde etkin çalışarak hesaplama verimliliği sunmaktadır. Modelin tahmin performansı hisse senedi kapanış fiyatları üzerinden değerlendirilmiştir. Modelin istatistiksel performans değerlendirmesi için iki farklı yöntem uygulanmıştır. Bu yöntemler; R-kare veya determinasyon katsayısı ve ortalama mutlak yüzde hata (MAPE) olarak aşağıda denklemleriyle sıralanmıştır. Denklem (7)’de gösterildiği gibi, R-kare, gözlemlenen değerler ile tahmin edilen değerler arasındaki ilişkiyi ölçen bir katsayıdır. Bu katsayı, modelin veri üzerindeki açıklayıcılığını gösterir ve modelin performansını değerlendirmek için kullanılır. R-kare değerinin 1’e yakın olması, modelin tahmin gücünün yüksek olduğunu ifade etmektedir.

Denklem (8)’de verilen ortalama mutlak yüzde hata (MAPE), modelin tahmin edilen değerlerle gerçek değerler arasındaki farkı, yüzdelik hata cinsinden ölçer. MAPE değeri, tahmin edilen değerlerin ne kadar isabetli olduğunu yüzdelik olarak ifade eder. Düşük bir MAPE, modelin tahminlerinin doğru olduğunu gösterir.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2} \quad (7)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} * 100 \quad (8)$$

2.2. ARIMA (Oto regresif Entegre Hareketli Ortalama)

Oto regresif entegre hareketli ortalama (ARIMA), zaman serisi tahmininde kullanılan istatistiksel bir modeldir. Kısa vadeli tahmin yapmak için uygundur ve Box-Jenkins yöntemine dayanır [12]. Bu yöntem, zaman serisinin durağanlığını kontrol etmeyi, modelin AR (Oto regresif) ve MA (Hareketli Ortalama) bileşenlerini belirlemeyi ve modelin parametrelerini tahmin etmeyi içermektedir.

Durağanlık kavramı, zaman serisinin ortalama ve varyansının zamanla sabit kalması anlamına gelir. Durağan olmayan zaman serileri, tahmin yapmayı zorlaştırabilir. Bu nedenle, ARIMA modeli uygulamadan önce zaman serisinin durağan hale getirilmesi gerekir. Durağanlığı kontrol etmek için otokorelasyon fonksiyonu (ACF) ve kısmi otokorelasyon fonksiyonu (PACF) grafikleri kullanılır. AR bileşeni, geçmiş gözlemlerin, güncel gözlem üzerindeki etkisini modeller. MA bileşeni ise, geçmiş hata terimlerinin, güncel gözlem üzerindeki etkisini modeller.

Modelin bileşenleri:

p: AR bileşeninin derecesini göstermektedir. Bu bileşen, geçmiş p sayıda gözlemin, güncel gözlem üzerindeki etkisini belirlemektedir.

d: Zaman serisinin durağan hale getirilmesi gerektiğinden fark alma işleminin derecesini göstermektedir.

q: MA bileşeninin derecesini göstermektedir. Geçmiş hata terimlerinin, güncel gözlem üzerindeki etkisini belirlemektedir.

ARIMA modelinin genel gösterimi Denklem 9'daki gibidir:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (9)$$

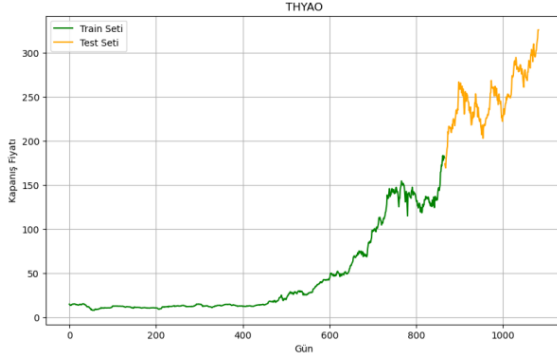
Burada, Z_t , zaman serisinin t. gözlemini, $\phi_1, \phi_2, \dots, \phi_p$, AR parametrelerini, ε_t , hata terimini ve $\theta_1, \theta_2, \dots, \theta_q$ ise MA parametrelerini ifade etmektedir.

3. Bulgular

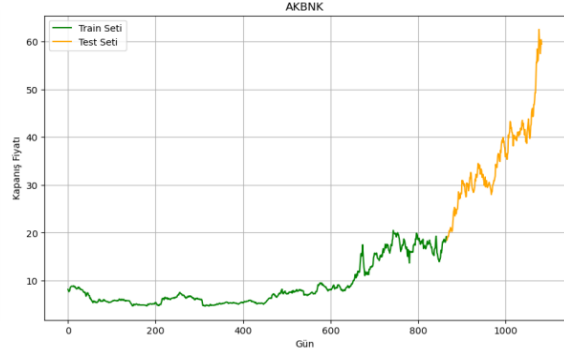
Çalışmada kullanılan LSTM (Uzun Kısa Süreli Bellek) ve ARIMA (Otokorelasyonlu Entegre Hareketli Ortalama) modellerinin sonuçlarına yer verilmiştir. Bu bölümde modellerin performanslarını değerlendirmek ve karşılaştırmak amacıyla, belirli bir veri seti üzerinde yapılan analizler detaylandırılacaktır. Elde edilen sonuçlar, her iki modelin de tahmin yeteneklerini ve doğruluklarını ortaya koymakta olup, hangi modelin belirli senaryolarda daha etkili olduğunu göstermektedir. Aşağıda, LSTM ve ARIMA modellerinin sonuçlarına dair ayrıntılı bilgiler bulunmaktadır.

3.1. LSTM (Uzun Kısa Süreli Bellek Ağı) Bulguları

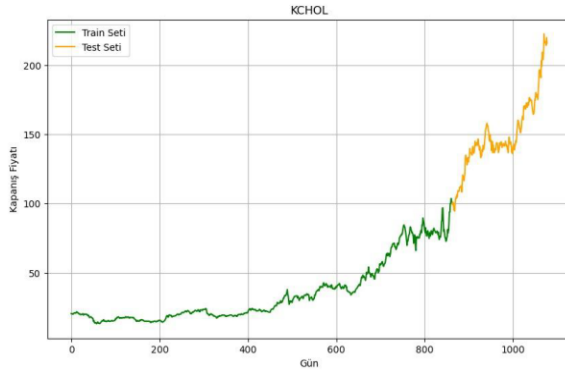
Bu bölümde, çalışmada kullanılan LSTM modelinin sonuçları ele alınmaktadır. Model, BİST30 endeksine dahil olan farklı sektörlerden seçilmiş dört şirketin hisseleri üzerinde test edilmiştir. Test edilen şirketler Türk Hava Yolları (THYAO), Akbank T.A.Ş. (AKBNK), Koç Holding A.Ş. (KCHOL) ve Kardemir Demir Çelik (KRDMD) olarak belirlenmiştir. Her hisse senedine ait veri setinin %80'i modelin eğitimi için ve %20'si test için ayrılmıştır. THYAO hisse senedinin eğitim ile test verilerinin grafiği Şekil 2'de, AKBNK hisse senedinin eğitim ile test verilerinin grafiği Şekil 3'te, KCHOL hisse senedinin eğitim ile test verilerinin grafiği Şekil 4'te, KRDMD hisse senedinin eğitim ile test verilerinin grafiği Şekil 5'te gösterilmiştir.



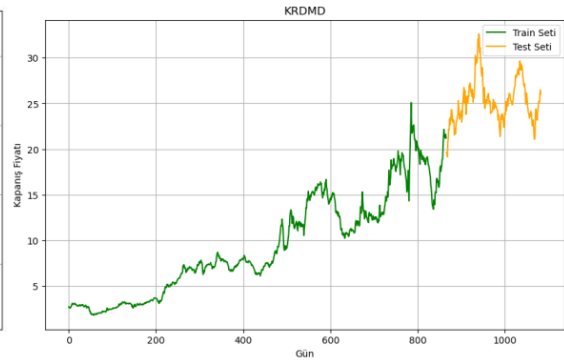
Şekil 2. THYAO Eğitim ve Test Verileri.



Şekil 3. AKBNK Eğitim ve Test Verileri.



Şekil 4. KCHOL Eğitim ve Test Verileri.



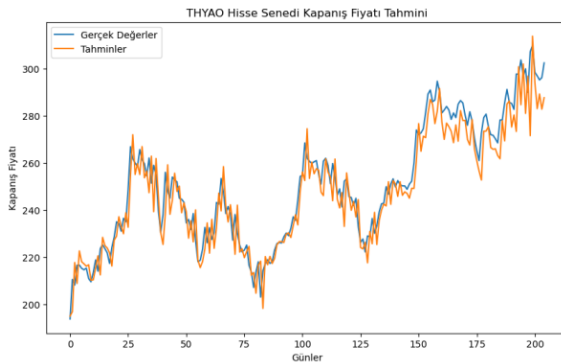
Şekil 5. KRDMMD Eğitim ve Test Verileri.

Oluşturulan LSTM modelinin eğitim ve test aşamasının sonucunda ortaya çıkan performans metrikleri MAPE, R^2 Tablo 1'de gösterilmiştir. Elde edilen R^2 değerlerinin 1'e oldukça yakın olması, MAPE değerlerinin ise 0'a yakın olması tahminin doğruluk oranının yüksek olduğunu fakat minimal sapmalar olabileceğini göstermektedir.

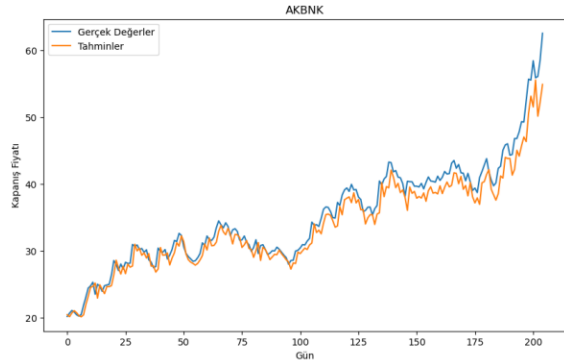
Tablo 1. Performans Metrikleri.

HİSSE SENEDİ	MAPE	R^2
THYAO	1,90	0,9040
AKBNK	1,32	0,9345
KCHOL	1,35	0,9263
KRDMMD	1,09	0,9422

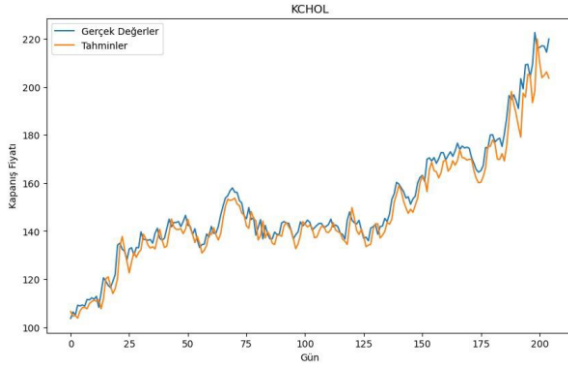
Çalışmada elde edilen, hisse senetlerine ait gerçek ve tahmini değerlerin grafiği Şekil 6-9'da gösterilmiştir. Ortaya çıkan tahmin verilerine ait grafikler, belirtildiği üzere gerçek verilere yakın tahminler üretildiğini göstermektedir.



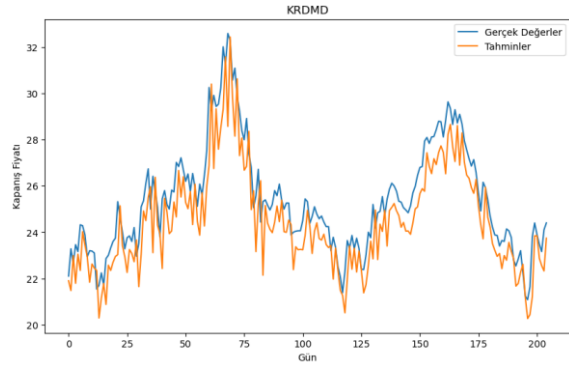
Şekil 6. THYAO Tahmin Verisi.



Şekil 7. AKBNK Tahmin Verisi.



Şekil 8. KCHOL Tahmin Verisi.



Şekil 9. KRDM D Tahmin Verisi.

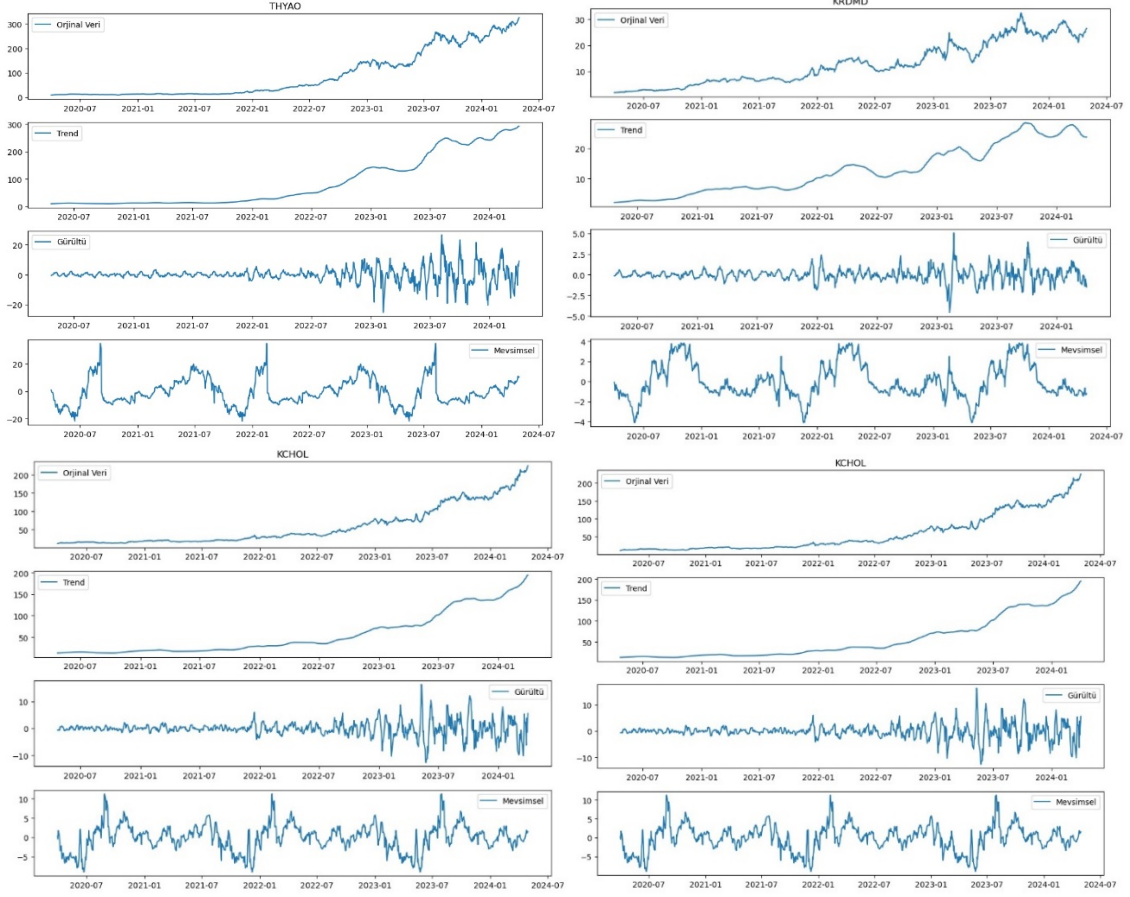
3.2. ARIMA (Oto regresif Entegre Hareketli Ortalama) Bulguları

Çalışmada kullanılan ARIMA modeline ait analizde, zaman serisi verilerinin mevsimsel trend, otokorelasyon ve kısmi otokorelasyon özellikleri modelin yapı taşları olarak ele alınmıştır. Hisse senetlerine yönelik mevsimsel trend grafikleri Şekil 10'da yer almaktadır. Şekil 10, zaman serisinin orijinal durumu ve ilgili bileşenleri (trend, mevsimsel dalgalanmalar, gürültü) ile modelin analiz süreçlerine ışık tutmaktadır. Bu grafikler, zaman serisi verisinin bileşenlerinin ayrıştırılması yoluyla daha derinlemesine analiz edilmesini sağlamaktadır. THYAO grafiğinde, verinin zamanla nasıl arttığını ve özellikle 2022'nin sonlarından itibaren hızlı bir artış gösterdiği görülmektedir. İkinci satır, zaman serisinin trend bileşenini göstermektedir. Trend, veri setinin genel eğilimini ifade etmektedir. Grafiğe göre verinin sürekli bir artış eğiliminde olduğu, genel olarak yükselme eğiliminde olduğu görülmektedir. Üçüncü satır, mevsimsel bileşeni göstermektedir. Bu bileşen, veride belirli periyotlarla tekrar eden değişimleri ifade eder. Grafikte belirgin bir şekilde mevsimsel dalgalanmalar görülmektedir; belirli dönemlerde yüksek ve düşük noktalar arasında düzenli bir geçiş gözlemlenmektedir. Son satırda ise, verinin gürültü bileşeni yer almaktadır. Gürültü, verinin trend ve mevsimsel bileşenlerden arındırıldıktan sonra kalan rastgele dalgalanmaları temsil eder. Bu bileşen, verideki düzensiz ve öngörülemeyen değişiklikleri içerir. Görüldüğü üzere, gürültü bileşeni daha düzensiz ve ani değişiklikler göstermektedir. Bu grafikler, zaman serisi verisinin bileşenlerini ayrı ayrı analiz ederek daha iyi bir anlayış sağlamak için kullanılmaktadır. Orijinal veri ve trend bileşeni genel eğilimleri gösterirken, mevsimsel bileşen düzenli değişimleri, gürültü bileşeni ise öngörülemeyen değişiklikleri göstermektedir. Bu tür bir ayrıştırma, zaman serisi analizinde önemli bir adımdır ve genellikle tahmin modellerinin geliştirilmesinde kullanılmaktadır.

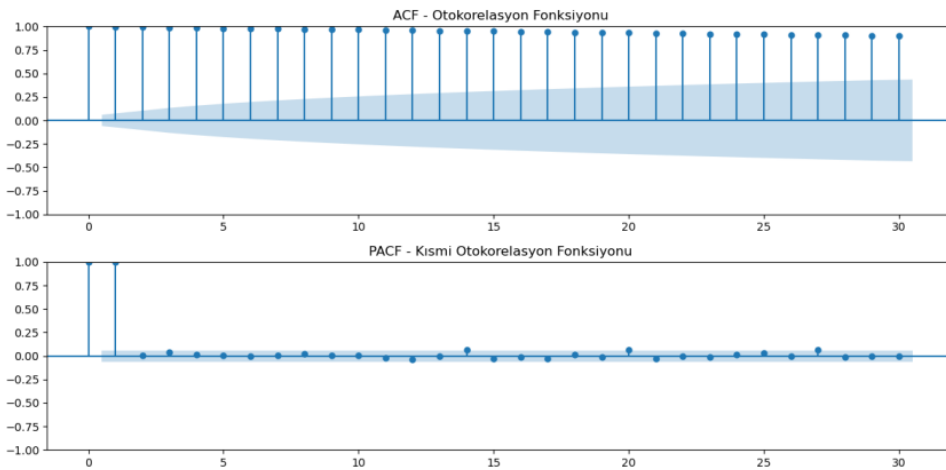
Şekil 11'de bir zaman serisi verisinin otokorelasyon ve kısmi otokorelasyon fonksiyonlarını (ACF ve PACF) göstermektedir. Bu fonksiyonlar, zaman serisinin özelliklerini incelemek ve modelleme için uygun parametreleri belirlemek amacıyla kullanılmaktadır. Üstteki grafik, otokorelasyon fonksiyonunu (ACF) göstermektedir. ACF, zaman serisindeki her bir veri noktasının önceki veri noktalarıyla olan korelasyonunu ölçer. Burada, x eksenini gecikme sayısını (lag) ve y eksenini otokorelasyon katsayısını temsil eder. Grafikte mavi çubuklar otokorelasyon katsayılarını, mavi gölge alan ise güven aralığını göstermektedir. İlk gecikmeler için yüksek pozitif otokorelasyon değerleri görülmektedir, bu da zaman serisinde bir süreklilik olduğunu gösterir. Otokorelasyon katsayıları zamanla azalmaktadır, ancak sifıra yaklaşma eğilimi belirgin bir şekilde gözlenmemektedir. Bu, zaman serisinin durağan olmadığına ve trend içerebileceğine işaret eder. ACF, özellikle 0'dan saparak negatif değerler almaya başladığında, seri içindeki periyodik bileşenler veya dalgalanmalar hakkında bilgi verir. Altındaki grafik, kısmi otokorelasyon fonksiyonunu (PACF) göstermektedir. PACF, bir veri noktasının önceki noktalardan bağımsız olarak sadece bir önceki veri noktasıyla olan korelasyonunu ölçer. Yani, ara etkilerden arındırılmış korelasyon katsayılarını gösterir. İlk gecikmeler için yüksek pozitif kısmi otokorelasyon değerleri vardır. İlk gecikmeden sonra değerler hızla sifıra yaklaşmaktadır. Yalnızca ilk birkaç gecikme için anlamlı kısmi otokorelasyon değerleri gözlenmektedir; bu durum, AR (otomatik regresyon) modelinin düşük bir derece ile (muhtemelen AR(1) veya AR(2)) uygun olabileceğine işaret edebilir. Geri kalan gecikme değerleri için kısmi otokorelasyon sifıra oldukça yakındır, bu da önceki verilerin sonraki verilere olan etkisinin minimal olduğunu gösterir.

Veri setinin durağanlığı Dickey-Fuller [13] birim kök testleri ile analiz edilmiş, elde edilen ADF istatistiği ve p-değeri verilerin durağanlık göstermediğini ortaya koymuştur. Modelin p, d, q değerleri THYAO, AKBNK, KCHOL, KRDM D için sırasıyla (0,2,2), (2,1,2), (3,1,2), (2,1,2) değerlerini almıştır. Performans değerlendirmesi MAPE metriği ile yapılmış olup bahsi geçen bu metriğin çıktıları Tablo 2'de gösterilmiştir. Elde edilen tahmin

verileri Şekil 12-15'te gösterilmiştir. Tablo 2'de, elde edilen MAPE değerlerinin 0'dan oldukça uzakta olması göz ardı edilemeyecek sapmalar olduğunu ve gerçek değerlerden uzak tahminler ortaya çıktığını göstermiştir. Tahmin verilerine ait görseller de bu durumu kanıtlamaktadır.



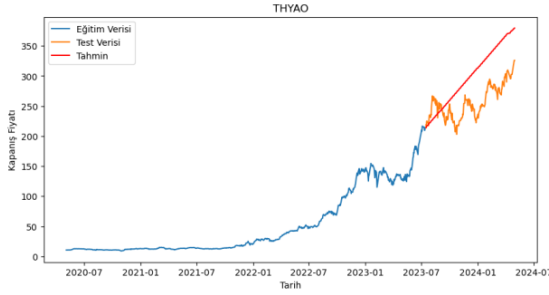
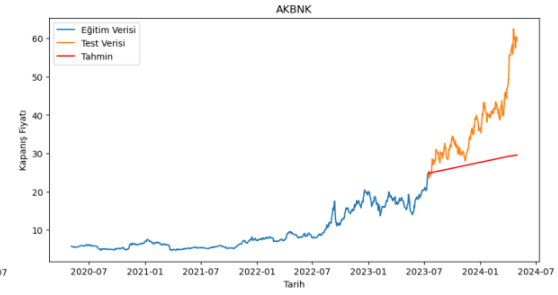
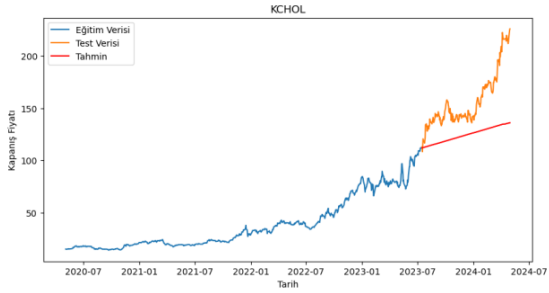
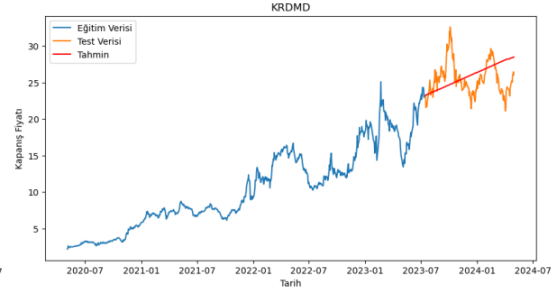
Şekil 10. Hisselere Ait Mevsimsel Trend Grafikleri.



Şekil 11. THYAO ACF-PACF Grafiği.

Tablo 2. ARIMA Performans Göstergesi.

HİSSE SENEDİ	MAPE
THYAO	18,64
AKBNK	23,20
KCHOL	18,89
KRDMD	8,62

**Şekil 12.** THYAO ARIMA Tahmini.**Şekil 13.** AKBNK ARIMA Tahmini.**Şekil 14.** KCHOL ARIMA Tahmini.**Şekil 15.** KRDMMD ARIMA Tahmini.

4. Sonuç ve Tartışma

Hisse senedi tahmininde bulunmak, ülkemizin sürekli değişen diplomatik, siyasi, ekonomik ve daha nice faktörlerinden dolayı oldukça zordur. Zira yalnızca teknik analiz tutarlı bir çıktı almak için yeterli olmadığı gibi yalnızca temel analiz de yeterli olmamaktadır. Bu çalışma, sürekli değişkenliğe sahip, tahmin edilmesi zor olan hisse senedi verilerinin analizinde kullanılan tahmin yöntemlerinin karşılaştırılmasını ve performansı daha yüksek olan yöntemi tespit etmeyi amaçlamıştır. Çalışmada LSTM modeli, ARIMA modelinden daha yüksek performans göstermiştir. ARIMA kısa vadeli, günlük ve saatlik analizler için daha uygundur. Dolayısıyla pek çok dış faktöre bağlı olan ve günlük değişimleri sınırlı olan hisse senedi fiyatlarının tahmininde gösterdiği performansın düşük kaldığı, özellikle uzun vadeli tahminlerde kritik hatalar yaptığı gözlemlenmiştir. LSTM modelinin gerçeğe yakın değerler üretebilmesi ve yüksek performansı, doğru veriler ve parametre girdileri ile hisse senedi analizi için uygun olduğu, olumlu sonuçlar doğurabileceği sonucuna varılmıştır.

Bu sonuçlar, hisse senedi tahmininde kullanılacak modelin seçiminde dikkat edilmesi gereken önemli kriterlerin altını çizmektedir. LSTM modelinin, zaman serilerinin karmaşık ve uzun vadeli bağımlılıklarını yakalama kapasitesi, onu özellikle belirsizliklerin ve ani değişimlerin sıkça görüldüğü finans piyasalarında tercih edilen bir yöntem haline getirmektedir.

Öte yandan, ARIMA modelinin kısa vadeli tahminlerde hala geçerli ve faydalı bir araç olduğu unutulmamalıdır. ARIMA'nın sadeliği ve uygulanabilirliği, kısa vadeli fiyat hareketlerini anlamada ve hızlı tahminlerde bulunmada önemli avantajlar sunar. Ancak, uzun vadeli trendleri ve piyasa dalgalanmalarını

yakalamada sınırlı kalması, bu modelin karmaşık finansal tahminlerde tek başına yeterli olamayacağını göstermektedir.

LSTM modelinin üstün performansının altındaki nedenler arasında, modelin çok katmanlı yapısı ve geri yayılım algoritması sayesinde geçmiş verilerdeki desenleri ve eğilimleri daha etkili bir şekilde öğrenebilmesi bulunmaktadır. Bu nedenle, özellikle uzun vadeli yatırımlar ve stratejik kararlar için LSTM modeli daha sağlam ve güvenilir bir seçenek olarak öne çıkmaktadır.

Bu çalışma, BİST30 endeksindeki hisse senetlerinin fiyat tahmininde LSTM ve ARIMA modellerini karşılaştırmaktadır. Literatürdeki diğer çalışmalardan farklı olarak, bu çalışma farklı sektörlerde ait hisse senetleri üzerinden model performanslarını incelemektedir. Ayrıca, 2020-2024 yılları arasındaki uzun dönem verileri kullanarak modellerin uzun vadeli tahmin performanslarını değerlendirmektedir. Çalışma verilerinin güncel olması çalışmanın literatüre güncel bir katkı sağlamasını ve mevcut piyasa koşullarına uygulanabilirliğini artırmaktadır.

Çalışmada, ulaştırma, bankacılık, yatırım ve metaller-madencilik gibi farklı sektörlerde ait hisse senetleri analiz edilmiştir. Bu, modelin genelleştirilebilirliğini ve güvenilirliğini artırmaktadır. Ayrıca, literatürdeki birçok çalışma yalnızca bir modelin performansını incelerken, bu çalışmada hem ARIMA hem de LSTM modelleri aynı veri seti üzerinde karşılaştırılarak hangi modelin daha üstün performans sergilediği belirlenmiştir. Bu, yatırımcıların hangi modeli tercih etmeleri gerektiği konusunda daha bilinçli kararlar vermelerine yardımcı olmaktadır. Model performansları, ortalama mutlak yüzde hata (MAPE) ve determinasyon katsayısı R^2 gibi metriklerle kapsamlı bir şekilde değerlendirilmiştir. Bu, modelin farklı performans kriterlerine göre incelenmesini sağlamaktadır.

Bir sonraki çalışmalarda hisse senedi tahmininde daha yüksek doğruluk ve güvenilirlik elde etmek için hibrit modellerin geliştirilmesi ve kullanılması da değerlendirilebilir. Örneğin, kısa vadeli tahminler için ARIMA, uzun vadeli tahminler için ise LSTM modelinin kombinasyonu, daha kapsamlı ve dengeli bir analiz sağlayabilir. Bu yaklaşım, farklı modellerin güçlü yanlarını birleştirerek, hisse senedi tahminlerindeki belirsizlikleri ve hataları minimize edebilir. Ayrıca, gelecekteki çalışmaların, piyasa koşullarının değişkenliğini ve çok boyutlu yapısını dikkate alarak, daha sofistike ve uyum sağlayabilen modeller üzerine odaklanması gerektiği sonucuna varılmıştır.

Kaynaklar

- [1] Sermaye Piyasası Kanunu. (2012, 30 Aralık). Resmî Gazete (Sayı: 28513). Erişim Adresi: <https://www.mevzuat.gov.tr/>
- [2] Albayrak E, Saran N. İstatistiksel ve derin öğrenme modellerini kullanarak hisse senedi fiyat tahmini. *Bilgisayar Bilimleri ve Mühendisliği Dergisi* 2023; 16(2): 161-169.
- [3] Hyndman RJ, Athanasopoulos G. *Forecasting: Principles and Practice*. Haziran 2018.
- [4] Gavcar E, Metin H. Hisse senedi değerlerinin makine öğrenimi (derin öğrenme) ile tahmini. *Ekonomi ve Yönetim Araştırmaları Dergisi* 2021; 10(2): 137-154.
- [5] Tanışman S, Karcioğlu AA, Uğur A, Bulut H. Bitcoin fiyatının LSTM ağı ve ARIMA zaman serisi modeli kullanarak tahmini ve karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi* 2021; (32): 514-520.
- [6] Siami Namin S, Siami Namin A. *Forecasting economic and financial time series: ARIMA vs. LSTM*. Texas Tech University, 2018.
- [7] Tanışman S, Karcioğlu AA, Uğur A, Bulut H. Bitcoin fiyatının LSTM ağı ve ARIMA zaman serisi modeli kullanarak tahmini ve karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi* 2021; (32): 514-520.
- [8] Albayrak E, Saran AN. Hisse senedi fiyat tahmini: İstatistiksel ve derin öğrenme modelleri kullanarak hisse senedi fiyat tahmini. *Bilgisayar Bilimleri ve Mühendisliği Dergisi* 2023; 16(2): 161-169.
- [9] Eylasov N, Çiçek M. Kripto para fiyatlarının tahmini: ARIMA-GARCH ve LSTM yöntemlerinin karşılaştırılması. *Finans Ekonomisi Ve Sosyal Araştırmalar Dergisi* 2024; 9(1): 48-62.
- [10] Yahoo Finance. (2024, Nisan 17). Erişim Adresi: <https://finance.yahoo.com/>
- [11] Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Computation* 1997; 9(8): 1735-1780.
- [12] Box GE, Jenkins GM, Reinsel GC, Ljung GM. *Time series analysis: forecasting and control*. 5th ed. New York, NY, USA: John Wiley & Sons, 2015.
- [13] Dickey DA, Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 1979; 74(366a): 427-431.