

GAZİ

JOURNAL OF ENGINEERING SCIENCES

An Empirical Analysis of Stock Price Prediction Using Deep Learning Methods: LSTM, GRU, GAN, and WGAN-GP

Cemal Öztürk^a, Abdulkadir Karacı^b

Submitted: 09.03.2024 Revised: 25.08.2024 Accepted: 19.11.2024 doi:10.30855/gmbd.0705AR03

ABSTRACT

Keywords: Deep learning, LSTM, GRU, GAN, WGAN-GP, stock price prediction

^aPamukkale University,
Faculty of Economics and
Administrative Sciences,
Department of Economics
20150 - Denizli Türkiye
Orcid: 0000-0003-3850-7416
e-mail: cemalo@pau.edu.tr

^bSamsun University,
Faculty of Engineering,
Department of Software Engineering
55420 - Samsun, Türkiye
Orcid: 0000-0002-2430-1372
e-mail:
abdulkadir.karaci@samsun.edu.tr

*Corresponding author:
cemalo@pau.edu.tr

Financial markets are important due to dynamic economic conditions and global interactions. This study compares deep learning methods for stock price prediction, including LSTM, GRU, GAN, and WGAN-GP. Using Google stock data, model performance was assessed via RMSE, MAE, and R^2 metrics. The LSTM model achieved the best performance with the lowest error and highest accuracy, scoring an RMSE of 4.214, MAE of 3.387, and R^2 of 0.9509. The GAN model also delivered strong results, with an RMSE of 4.048, MAE of 3.288, and R^2 of 0.9578 on test data. In contrast, the GRU model performed slightly worse, with an RMSE of 4.303, MAE of 3.42, and R^2 of 0.9490. The WGAN-GP model showed the poorest performance, with an RMSE of 11.523, MAE of 9.015, and R^2 of 0.2452. Unexpected events like the COVID-19 pandemic led to significant deviations in predictions. In conclusion, LSTM and GAN models have proven effective for short- and medium-term forecasting.

Derin Öğrenme Yöntemleri Kullanılarak Hisse Senedi Fiyat Tahmini Üzerine Ampirik Bir Analiz: LSTM, GRU, GAN ve WGAN-GP

ÖZ

Finansal piyasalar, değişen ekonomik koşullar ve küresel etkileşimler nedeniyle büyük bir öneme sahiptir. Bu çalışma, hisse senedi fiyatlarını tahmin etmek için LSTM, GRU, GAN ve WGAN-GP gibi derin öğrenme yöntemlerini karşılaştırmıştır. Google hisse senedi verileri kullanılarak yapılan analizde, modellerin performansı RMSE, MAE ve R^2 metrikleriyle değerlendirilmiştir. LSTM modeli, en düşük hata oranı ve en yüksek doğruluk ile en iyi performansı göstermiştir; RMSE değeri 4.214, MAE değeri 3.387 ve R^2 değeri 0.9509'dur. GAN modeli de güçlü performansı ile dikkat çekmiştir; test verisinde RMSE değeri 4.048, MAE değeri 3.288 ve R^2 değeri 0.9578 olarak bulunmuştur. GRU modeli, LSTM ve GAN'a kıyasla daha düşük doğruluk sergilemiştir; RMSE değeri 4.303, MAE değeri 3.42 ve R^2 değeri 0.9490'dur. WGAN-GP modeli ise diğer modellere göre en yüksek hata oranını ve en düşük doğruluğu göstermiştir; RMSE 11.523, MAE 9.015 ve R^2 0.2452'dir. Ayrıca, COVID-19 gibi beklenmedik olaylar tahminlerde belirgin sapmalara yol açmıştır. Sonuçlar, LSTM ve GAN modellerinin kısa ve orta vadeli tahminlerde başarılı olduğunu göstermektedir.

Anahtar Kelimeler: Derin öğrenme, LSTM, GRU, GAN, WGAN-GP, hisse senedi fiyat tahmini

1. Giriş (Introduction)

Hisse senedi fiyatı tahmini, yatırımcılar için hisse senedi fiyatlarının muhtemel yönü hakkında değerli bilgiler sağlayan önemli bir araçtır ve uzun yıllardır araştırmacıların ilgisini çekmektedir. Hisse senedi fiyatları, iç ve dış ekonomik ortam, uluslararası durum, sektör beklentisi, borsaya kote şirketlerin finansal verileri ve borsa işleyişi gibi çeşitli iç ve dış faktörlerden etkilenmektedir. Hisse senedi fiyatlarının tahmin edilmesi, yatırımcılar, portföy yöneticileri ve politika yapımcılar için büyük önem taşımaktadır. Doğru tahminler, karar verme süreçlerini iyileştirerek ve finansal riskleri azaltarak piyasa katılımcılarının başarısını artırabilir.

Zaman serisi tahmin problemlerini çözmek için kullanılan teknikler ikiye ayrılabilir. İlk kategori, klasik ekonometrik modellerdir. Yaygın yöntemler, otoregresif yöntem (AR), hareketli ortalama modeli (MA), otoregresif hareketli ortalama modeli (ARMA) ve otoregresif bütünleşik hareketli ortalamadır (ARIMA) [1-3]. Bu modeller her yeni sinyali, son birkaç sinyalin ve bağımsız gürültü terimlerinin doğrusal bir kombinasyonu olarak ele almaktadır. Ancak, finansal verilerin karmaşıklığı ve doğrusal olmayan yapısı nedeniyle bu modellerin performansı sınırlı kalmaktadır [4].

İkinci kategori, yapay zeka temelli yöntemlerdir. Bu teknikler arasında yapay sinir ağları (ANN) [5], destek vektör makineleri (SVM) [6], k-en yakın komşu (k-NN) [7], lojistik regresyon (LR) [8] gibi yöntemler yer almaktadır. Derin öğrenme, YSA (Yapay Sinir Ağları)'nın daha gelişmiş bir versiyonu olup, daha fazla katman ve nöron içeren derin sinir ağları (Deep Neural Networks) kullanır. Derin öğrenme, özellikle büyük veri kümeleri ve karmaşık veri yapıları ile çalışırken üstün performans göstermektedirler. Derin sinir ağları, verilerden soyut özellikler çıkarma yeteneği ile zaman serisi tahmin problemlerinde sıklıkla kullanılmaktadır.

Bu çalışmanın amacı, LSTM, GRU, GAN ve WGAN-GP derin öğrenme yöntemlerinin hisse senedi fiyatlarını tahmin etmedeki etkinliğini ve performansını değerlendirmektir. Derin öğrenme modelleri, özellikle uzun vadeli bağımlılıkları öğrenme ve doğrusal olmayan ilişkileri modelleme konularında geleneksel yöntemlere kıyasla daha başarılıdır. Bu nedenle, bu çalışmada derin öğrenme modelleri kullanılmış ve bu modellerin hisse senedi fiyat tahmini üzerindeki performansları karşılaştırılmıştır.

Çalışmanın geri kalan bölümleri şu şekilde düzenlenmiştir: İkinci bölümde ilgili literatür taraması sunulmuştur. Üçüncü bölümde çalışmada kullanılacak veri ve yöntem tanıtılmıştır. Dördüncü bölümde, önerilen model ile deneysel analiz ve elde edilen sonuçların klasik tahmin modellerinin verdiği sonuçlarla karşılaştırması sunulmuştur. Beşinci bölümde sonuçlar ve öneriler tartışılmaktadır.

2. Literatür Araştırması (Literature)

Derin öğrenme yöntemlerinin finansal piyasalarda giderek daha fazla önem kazandığı günümüzde, hisse senedi fiyatlarının doğru tahmin edilmesi yatırımcılar ve finansal kurumlar için büyük öneme sahiptir. Hisse senedi fiyatlarının tahmin edilmesi, geleneksel finans teorilerine ve istatistiksel modellere dayanarak yapılmaktadır; fakat son yıllarda, derin öğrenme teknikleri bu alanda önemli ölçüde ilerlemeler kaydetmiştir. Bu bölümde, hisse senedi fiyat tahmini için kullanılan derin öğrenme yöntemleri üzerine yapılmış literatür çalışmaları incelenmektedir.

Madge ve Bhatt (2015), 34 adet teknoloji hissesi ile SVM'yi kullanarak belirli bir hisse senedinin davranışını tahmin etmeyi amaçlamıştır [9]. Çalışma, kısa vadeli tahminlerin çok düşük doğruluğa sahip olduğu, ancak uzun vadeli tahminlerin %55-60 arasında bir doğruluğa ulaştığı sonucuna varmaktadır. Bu tahmin doğruluğu, EMH'nin uzun vadeli tahmin için mutlaka geçerli olmadığını göstermektedir. Bu sonuç, makine öğrenmesi algoritmalarının hisse senedi fiyatlarını veya hisse senedi fiyat hareketini tahmin etmek için çeşitli borsa teknik göstergelerinin etkisini kullanabileceğini göstermektedir.

Khare ve diğ. (2017), New York Menkul Kıymetler Borsasında kayıtlı 10 hisse senedini dikkate almış ve teknik analizin gücünden yararlanarak bu kısa vadeli fiyatların tahminine odaklanmıştır [10]. Makale, ANN'nin iki farklı türü olan İleri Beslemeli Sinir Ağları (FNN) ve Tekrarlayan Sinir Ağlarını (RNN) dikkate almıştır. FNN'nin bir hisse senedinin kısa vadeli fiyatlarını tahmin etmede LSTM'den daha üstün performans gösterdiğini ortaya koymuşlardır.

Selvin ve diğ. (2017), LSTM, RNN ve Evrişimli Sinir Ağları (CNN) gibi derin öğrenme modelleri kullanılarak hisse senedi fiyatları tahmini incelenmektedir [11]. Makalede, ayrıca kayan pencere yöntemi kullanılarak verilerin hazırlanması da ele alınmaktadır.

Ji ve diğ. (2019), karbon vadeli fiyatını tahmin etmek için bir ARIMA-CNN-LSTM modelini tanıtmıştır [12]. ARIMA-CNN-LSTM modeli, ARIMA modelini ve doğrusal ve doğrusal olmayan veri özelliklerini yakalamak için CNN ve LSTM katmanlarını birleştiren derin sinir ağı yapısını kullanmaktadır. ARIMA-CNN-LSTM model yapısında, doğrusal özellikleri yakalamak için ARIMA kullanılmıştır. CNN, hiyerarşik veri yapısını yakalamak için kullanılırken, LSTM, verilerdeki uzun vadeli bağımlılıkları yakalamak için kullanılmaktadır. Kapsamlı performans değerlendirmesi, haftalık karbon vadeli fiyatları kullanılarak yapılmıştır. Sonuçlar, ARIMA-CNN-LSTM modelinin, Hata Kareler Ortalamasının Karekökü (RMSE) ve Ortalama Mutlak Yüzde Hatası (MAPE) performans ölçümleri açısından kıyaslama modellerinden daha iyi tahmin doğruluğu sağlayabildiğini doğrulamıştır.

Alberto ve Romero (2019), GAN ve LSTM modellerinin etkinliğini inceledikleri çalışmalarında örnekleme döneminden bir gün sonra fiyatın yükselip yükselmeyeceğini belirlemeyi amaçlamışlardır [13]. LSTM üretici, CNN ayırıcı olarak belirlenmiştir. Sonuç olarak GAN ve LSTM modelleri arasında önemli bir fark bulunamamıştır.

Zhang ve diğ. (2019), hisse senedinin kapanış fiyatını tahmin etmek için üretici olarak LSTM ve ayırıcı olarak Çok Katmanlı Perceptron (MLP) ile bir GAN modeli önermiş ve sonucu LSTM modeli ile karşılaştırmıştır [14]. Sonuçlar, GAN'ın geleneksel referans modelinden daha verimli olduğunu göstermiştir. GAN modelinin doğruluğu yaklaşık %75.54 iken LSTM % 68.59'dur.

Raço ve Demirci (2019) tarafından yapılan bir çalışmada, BIST 30 endeks fiyatlarını tahmin etmek için derin öğrenme modeli önerilmiştir. Bu çalışma, teknik göstergeler ve osilatörler kullanarak tahminler yapmış ve diğer benzer önerilerden daha iyi performans göstermiştir. Çalışma, 27 aylık veri seti üzerinde gerçekleştirilmiştir ve modelin ortalama kare hata (MSE) değerleri, BIST 30 endeks fiyatlarının beş işlem günü sonrası tahminlerinde oldukça düşük bulunmuştur [15].

Ta ve diğ. (2020), geçmiş verilere dayalı stok hareketini tahmin etmek için özel bir RNN türü olan uzun bir kısa süreli bellek (LSTM) ağı önermiş, portföy performansını iyileştirmek için eşit ağırlıklı modelleme (EQ), simülasyon modelleme Monte Carlo simülasyonu (MCS) ve optimizasyon modelleme ortalama değişken optimizasyonu (MVO) dahil olmak üzere çoklu portföy optimizasyon teknikleri kullanmıştır [16]. Önerilen LSTM tahmin modelin stok tahmininden yüksek doğruluk elde ederek verimli çalıştığını göstermiştir. LSTM tahmin modeline dayalı olarak oluşturulan portföyler, doğrusal regresyon ve SVM gibi diğer oluşturulmuş portföylere dayalı tahmin modellerinden daha iyi performans göstermiştir.

Lu ve diğ. (2020), uzun kısa süreli hafızaya (LSTM) dayalı olarak, CNN-LSTM'ye dayalı bir hisse senedi fiyatı tahmin yöntemi önermiştir [17]. Ayrıca, hisse senedi fiyatını tahmin etmek için MLP, CNN, RNN, LSTM, CNN-RNN ve diğer tahmin modelleri kullanılmış ve bu modellerin tahmin sonuçları analiz edilmiş ve karşılaştırılmıştır. Bu çalışmada kullanılan veriler 1 Temmuz 1991 - 31 Ağustos 2020 tarihleri arasındaki 7127 işlem günü dahil olmak üzere günlük hisse senedi fiyatlarını içermektedir. Sonuçlara göre, CNN-LSTM en yüksek tahmin doğruluğu ile güvenilir bir hisse senedi fiyatı tahminini vermiştir.

Yang ve diğ. (2020) özellik çıkarımı için bir CNN ile tahmin için bir LSTM ağını birleştirerek, finansal zaman serilerindeki tarihsel bilgilere dayalı olarak fiyat hareketi yönünü tahmin etmek için derin bir öğrenme çerçevesi sunmuştur [18]. Deneysel sonuçlar, çerçevenin hisse senedi fiyatı hareket yönünü tahmin etmede son teknoloji modellerden daha iyi performans gösterdiğini göstermektedir.

Chandra ve diğ. (2021), COVID-19 salgını öncesinde ve sırasında hisse senedi fiyatı tahmini için Bayes sinir ağlarını kullanmıştır [19]. Ayrıca, COVID-19 öncesi veri setlerinin, COVID-19 sırasında hisse senedi fiyatlarının tahmin edilmesine yardımcı olup olmadığını da araştırmışlardır. COVID-19 sırasında hisse senedi fiyatındaki yüksek oynaklık nedeniyle tahmin sağlamanın daha zor olduğunu gösterirken Bayesyen sinir ağlarının, COVID-19 salgınının ilk zirvesi sırasında yüksek piyasa oynaklığına rağmen belirsizlik ölçümü ile makul tahminler sağlayabildiğini bulmuşlardır.

Demirel vd. (2021), çok katmanlı algılayıcılar (MLP), destek vektör makineleri (SVM) ve uzun kısa süreli bellek (LSTM) modelleri kullanarak Borsa İstanbul Ulusal 100 Endeksi'ndeki (BIST100) hisse senedi fiyatlarını tahmin etmeyi amaçlamıştır [20]. Araştırmacılar, 2010'dan 2019'a kadar olan 9 yıllık BIST100 tarihsel verilerini kullanarak tahmin performanslarını hata kareler ortalamasının karekökü (RMSE), ortalama kare hata (MSE) ve R-kare (R²) metrikleri yardımıyla karşılaştırmışlardır. Çalışmanın bulguları, MLP modelinin açılış hisse senedi fiyatlarını tahmin etmede SVM ve LSTM modellerine göre daha iyi olduğunu ve LSTM modelinin kapanış hisse senedi fiyatlarını tahmin etmede SVM ve MLP modellerine göre daha iyi olduğunu göstermiştir.

Jing ve diğ. (2021), hisse senedi fiyat eşzamanlılığı teorisini hisse senedi fiyat trend analizinde uygulayan, stok kümeleri oluşturmak için yakınlık yayılma algoritmasını ve mekansal veri analizi yeteneği sağlamak için CNN modelini kullanan, hisse senedi fiyat eşzamanlılığı ve derin öğrenme yöntemlerine dayalı bir hisse senedi fiyatı tahmin modeli önermiştir [21].

Yan ve diğ. (2021), LSTM derin sinir ağına dayalı ve BP sinir ağı, geleneksel RNN ve geliştirilmiş LSTM derin sinir ağı ile karşılaştırılan yüksek hassasiyetli kısa vadeli bir finansal piyasa zaman serisi tahmin modeli oluşturmuştur [22]. Sonuçlar, LSTM derin sinir ağının yüksek tahmin doğruluğuna sahip olduğunu ve hisse senedi piyasasının zaman serilerini etkili bir şekilde tahmin edebildiğini göstermiştir.

Lin ve diğ. (2021), üretici olarak Geçitli Tekrarlayan Birimler (GRU) ve ayrımcı olarak CNN ile GAN kullanan bir hisse senedi tahmin modeli önermiştir [23]. Hedef fiyat olarak Apple hisse senedi kapanış fiyatını kullanmıştır. Önerilen GAN modeli sonuçlarını temel modellerle karşılaştırmışlardır. Buna göre temel GAN ve WGAN-GP geleneksel modeller olan GRU ve LSTM'den daha iyi performans göstermiştir. Giang ve diğ. (2022), hisse senedi fiyatındaki önemli değişimin analizine dayalı olarak LSTM modelinin tahmin aşamasında dinamik bir kayan pencere uygulamıştır [24]. Amerika Birleşik Devletleri, Almanya ve Vietnam'ın üç hisse senedi veri seti üzerindeki deneysel sonuçlar, modelin hisse senedi fiyat dalgalanmalarının eğilimini tahmin etmede rakiplerinden daha iyi performans gösterdiğini göstermektedir.

Han ve diğ. (2022), finansal varlık fiyatlarının simülasyonunu gerçekleştirmek için Wasserstein Generative Adversarial Network (WGAN) kullanımını incelemektedir [25]. Bu yaklaşım, hisse senetleri, vadeli işlemler ve kripto paralar gibi farklı varlık sınıflarının fiyatlarını modellemeyi amaçlamaktadır. Çalışmanın bulguları, WGAN'nın farklı varlık sınıfları arasındaki bağımlılık yapısını ve fiyat hareketlerini başarıyla yakaladığını göstermektedir. Bu yaklaşım, piyasa riskini ve portföy yönetimi için önemli olan fiyat hareketlerini modellemek ve tahmin etmek amacıyla kullanılabilir.

Liu ve diğ. (2022), CAE ve LSTM sinir ağını birleştiren, CAE'yi hisse senedi fiyat görüntü özellik verilerini çıkarmak için kullanan ve kısa vadeli hisse senedi fiyatlarını tahmin etmek için teknik verileri birleştiren bir model önermiştir [26].

Albayrak ve Saran (2023), ARIMA modelini temel alarak, üç farklı tekrarlayan sinir ağı modeli (LSTM, GRU, dikkat katmanlı LSTM) ile karşılaştırmalar yapmıştır. Bu çalışma, literatürdeki diğer çalışmalardan farklı olarak, Borsa İstanbul verileri üzerinde gün içi tahminler yaparken, 28 farklı finansal indikatör kullanarak, derin öğrenme modellerinin performansını değerlendirmiştir. Sonuçlar, ARIMA modelinin, doğrusal olmayan RNN modellerine kıyasla daha yüksek ortalama hata oranına sahip olduğunu ortaya koymuştur. Özellikle GRU modeli, diğer modellere göre daha iyi performans göstermiştir. Bu çalışmanın bulguları, derin öğrenme modellerinin, geleneksel istatistiksel yöntemlere göre finansal zaman serilerinin tahmininde daha etkili olabileceğini vurgulamaktadır [27].

Çoban ve Hayat (2023), derin öğrenme teknikleri kullanılarak hisse senedi fiyatlarının tahmini üzerine odaklanmış ve özellikle Derin Sinir Ağları (DNN), Evrimsel Sinir Ağları (CNN) ve Tekrarlayan Sinir Ağları (RNN) modelleri karşılaştırmıştır. Çalışma, Borsa İstanbul'da işlem gören Vestel Elektronik Sanayi ve Ticaret A.Ş.'nin 2013-2021 yılları arasındaki günlük kapanış fiyatlarını kullanarak gerçekleştirilmiştir. Elde edilen bulgular, CNN modelinin diğer modellere göre daha düşük hata oranına sahip olduğunu ortaya koymuştur. Bu sonuçlar, literatürde CNN'nin finansal zaman serilerinde güçlü bir tahmin aracı olarak kullanılabilirliğini destekleyen diğer araştırmalarla uyum içerisindedir [28].

Dalkıran ve Ozan (2022), Borsa İstanbul'da işlem gören hisse senetlerinin fiyatlarını tahmin etmek için

derin öğrenme tekniklerini kullanmışlardır. Çalışmada, LSTM ve GRU algoritmaları Adam ve RMSProp optimize edicileri ile kullanılmış ve performansları değerlendirilmiştir. ISCTR hisse senedi verileri üzerine yapılan günlük değer tahminlerinde, VAKBN, GARAN, QNBFB ve AKBNK hisse senetleri ile USD/TRY, BIST30 ve BANKX endekslerinin tarihsel verileri de modelin tahmin performansını artırmak amacıyla dahil edilmiştir. Bu çalışma, özellikle Borsa İstanbul gibi volatilitesi yüksek piyasalarda derin öğrenme algoritmalarının etkinliğini vurgulamakta ve bu alandaki diğer araştırmalar için önemli bir referans teşkil etmektedir [29].

Erdin (2023), Borsa İstanbul endeksinde işlem gören EREGL hisse senedi verilerini kullanarak, ARIMA ve derin öğrenme modellerinin tahmin performanslarını kıyaslamıştır. Çalışmada, ARIMA modeli ile derin öğrenme algoritmalarından LSTM, GRU ve RNN modelleri arasında performans karşılaştırması yapılmıştır. Elde edilen bulgular, RNN algoritmasının diğer yöntemlere kıyasla daha üstün performans sergilediğini ve ortalama %93 doğruluk oranı ile test veri seti üzerinde en başarılı sonuçları verdiğini ortaya koymuştur. Bu çalışma, özellikle finansal zaman serilerinin modellenmesinde derin öğrenme yöntemlerinin etkinliğini vurgulamakta ve geleneksel ARIMA modelleri ile karşılaştırmalı analizler sunmaktadır [30].

Bu çalışmada, literatürdeki bu sınırlılıkları aşmak için çeşitli yenilikçi yaklaşımlar benimsenmiştir. Öncelikle, Google hisse senedi fiyatları üzerine geniş bir veri seti kullanılmıştır. Bu veri seti, modelin performansını daha güvenilir ve genellenebilir kılmaktadır. Çalışmamızda LSTM, GRU, GAN ve WGAN-GP gibi ileri seviye derin öğrenme modelleri kullanılmıştır. Bu modeller, finansal zaman serilerinin karmaşıklığını ve dinamiklerini daha iyi yakalayabilmektedir. Modellerin performansını maksimize etmek için detaylı bir hiper parametre optimizasyonu yapılmıştır. Bu, modellerin doğruluğunu ve güvenilirliğini artırmakta ve gerçek dünya uygulamalarında daha iyi performans göstermelerini sağlamaktadır. Çalışmamız ayrıca, farklı modellerin performanslarını karşılaştırmalı olarak analiz etmiş ve her bir modelin güçlü ve zayıf yönlerini belirlemiştir. Bu, hangi modelin hangi senaryolarda daha etkili olduğunu anlamamıza olanak tanımıştır.

Bununla birlikte, çalışmamızın bazı sınırlılıkları da bulunmaktadır. İlk olarak, sadece Google hisse senedi fiyatları üzerine odaklanılmıştır. Gelecekteki çalışmalarda farklı hisse senetleri ve finansal enstrümanlar üzerinde de testler yapılarak modelin genellenebilirliği artırılabilir. İkinci olarak, modellerin gerçek zamanlı piyasa koşullarında nasıl performans gösterdiği kapsamlı bir şekilde incelenmemiştir. Gerçek zamanlı testler, modelin pratikteki uygulanabilirliğini daha iyi değerlendirmemizi sağlayacaktır. Son olarak, veri ön işleme ve özellik mühendisliği süreçleri daha da geliştirilebilir. Ek özelliklerin eklenmesi ve daha gelişmiş veri ön işleme tekniklerinin kullanılması, modelin performansını artırabilir.

3. Veri ve Metodoloji (Data and Methodology)

Yapay zekâ ve makine öğrenmesi alanında, son yıllarda büyük ilerlemeler kaydedilmiştir. Bu ilerlemelerin temelinde, özellikle yapay sinir ağları ve derin öğrenme gibi güçlü yöntemler ve algoritmalar bulunmaktadır. Bu bölümde, önce yapay sinir ağlarının temel mantığı ve çalışma prensipleri incelenecektir. Ardından, derin öğrenme kavramına ve bu alandaki önemli yöntemlere, yani RNN (Recurrent Neural Networks), LSTM (Long Short-Term Memory), GRU (Gated Recurrent Units) gibi tekrarlayan sinir ağlarına ve GAN (Generative Adversarial Networks) ve WGAN (Wasserstein Generative Adversarial Networks) gibi üretici modellere değineceğiz. Bu yöntemlerin her biri, makine öğrenmesi ve yapay zekâ alanında çeşitli problemlere çözüm sunarak önemli başarılar elde etmeyi mümkün kılmaktadır. Bu bölümde sunulan kavramsal çerçeve, bu yöntemlerin temel prensiplerini ve uygulama alanlarının anlaşılmasına yardımcı olacak ve bu sayede daha derinlemesine çalışmalara zemin hazırlayacaktır.

Yapay sinir ağları (YSA), biyolojik sinir ağlarının birbirine bağlı yapılarını taklit eden hesaplama sistemleridir [31]. Bu ağlar, bilgi işlemek için belirli bir şekilde birbirine bağlanan nöron adı verilen basit işlem birimlerinden oluşur. Bir YSA tipik olarak girdi, gizli ve çıktı katmanlarını içerir ve her katman yapay nöronlar içerir [32]. YSA'ların temel amacı, zaman içinde çıktı doğruluğunu artırmak için öğrenme kurallarına dayalı olarak sürekli olarak çevreden veri toplama ve çıktı doğruluğunu artırmak için ağırlıklar gibi parametreleri optimize etmektir [33].

YSA, temel prensipleriyle birlikte oldukça güçlü ve esnek modeller sunmaktadır. Ancak, karmaşık ve

büyük veri kümeleri ile başa çıkmada ve daha yüksek performans elde etmede bazı sınırlamaları vardır. Derin öğrenme, YSA'nın daha gelişmiş bir versiyonu olup, daha fazla katman ve nöron içeren derin sinir ağları kullanmaktadır. Derin öğrenme, yapay zekâ (AI) alanında önemli gelişmelere yol açan ve özellikle büyük veri kümeleri üzerinde çalışabilen öğrenme algoritmalarına odaklanan bir yöntemdir [34]. Çok katmanlı yapıları sayesinde veriden daha fazla özellik çıkarılabilir ve daha sofistike öğrenme süreçlerini gerçekleştirebilir. Derin öğrenme, özellikle görüntü ve ses işleme, doğal dil işleme ve oyun oynama gibi alanlarda başarılı uygulamalarıyla öne çıkmaktadır [35].

Derin öğrenme, çok katmanlı yapay sinir ağları kullanarak karmaşık veri yapılarını ve özelliklerini öğrenmeye dayanır [36]. Bu sinir ağları, geri yayılım (backpropagation) ve gradyan inişi (gradient descent) gibi öğrenme algoritmaları kullanarak eğitilmektedir [37].

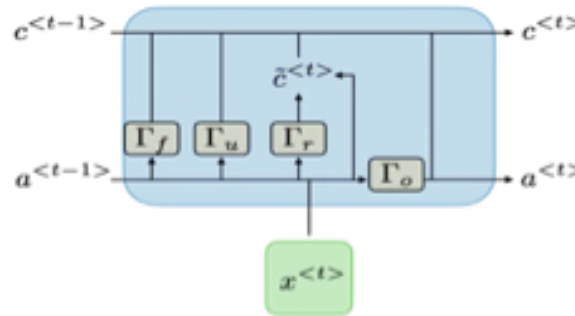
Derin öğrenme kavramının kullanımının son yıllarda yaygınlaşmasının nedeni, geçmiş yıllarda derin mimarileri eğitmek için işlemci gücü olmaması ve yeterli miktarda veri bulunmamasıdır. Günümüzde, işlemci gücünün artması ve birçok alanda sayısallaştırma ile oluşturulan verilerin muazzam boyutlara ulaşmasıyla derin öğrenmenin gelişimi için gerekli altyapı sağlanmıştır. Bu gelişmeler, derin öğrenmenin bilgisayarlı görü, metin işleme, çeviri, zaman serisi tahmini gibi alanlarda yaygın olarak kullanılmasını sağlamıştır.

Derin öğrenme algoritmalarını uygulamak için kullanılan popüler açık kaynaklı kütüphaneler ve araçlar bulunmaktadır. Bunlar arasında TensorFlow [38], Keras [39], PyTorch [40] ve Caffe [41] gibi yaygın olarak kullanılan kütüphaneler yer almaktadır.

Derin öğrenme, yapay zekâ alanında devam eden önemli gelişmelere katkı sağlamaktadır. Gelecekte, derin öğrenmenin daha fazla uygulama alanı keşfedilecektir ve mevcut tekniklerin sınırlarını aşmak için yeni yöntemler ve algoritmalar geliştirilmektedir [34].

3.1. LSTM (Long short-term memory)

LSTM, Schmidhuber ve diğerleri tarafından önerilen bir ağ modelidir [42]. LSTM, RNN'deki gradyan patlaması ve gradyan kaybolması problemlerini çözmek için tasarlanmış bir ağ modelidir [16]. Standart bir RNN'de yalnızca bir tekrar eden modül vardır ve iç yapısı basittir. Bununla birlikte, LSTM modüllerinden dördü standart RNN modüllerine benzer ve etkileşimli bir şekilde çalışmaktadırlar [43,44].



Şekil 1. LSTM mimarisi [45]
(LSTM architecture)

LSTM yapısında, geçmiş ve gelecek bilgileri anlamlandıran öznelikler yinelemeli olarak taşınır. Şekil 1, bir LSTM yapısını göstermektedir. Bu modelde aktivasyon fonksiyonu üç farklı noktada kullanılmaktadır: giriş, geri çağırma/unutma ve çıkış. Giriş ve çıkış katmanlarında, genellikle hiperbolik tanjant fonksiyonu tanh seçilirken hatırla/unutma kapılarında sigmoid fonksiyonu kullanılır.

$\hat{a}^{(t)}$ belleği ifade ederken $\hat{c}^{(t)}$, bellek için aday değeri sembolize etmektedir. Aktivasyon fonksiyonu olarak hiperbolik tanjant kullanılmaktadır. Burada aday değeri şu şekilde hesaplanır:

$$\hat{c}^{(t)} = \tanh(W_c[\Gamma_r * a^{(t-1)}, x^{(t)}] + b_c), \quad (1)$$

Ayrıca, güncelleme ve ilgi düzeyi kapısı değerleri LSTM'de hesaplanmaktadır. Bu değerler için

aktivasyon fonksiyonu olarak sigmoid kullanılır. Güncelleme ağ kapısı Γ_u ve ilgi düzeyi kapısı Γ_r aşağıdaki gibi hesaplanır:

$$\Gamma_u = \sigma(W_u[a^{(t-1)}, x^{(t)}] + b_u), \quad (2)$$

$$\Gamma_r = \sigma(W_r[a^{(t-1)}, x^{(t)}] + b_r). \quad (3)$$

Bu yapının geçitlenmiş tekrarlayan birimlerden (GRU) en önemli farkı, ilgi kapısının unutma (Γ_f) ve çıkış kapısı (Γ_o) olmak üzere iki yeni denklemlerle elde edilen LSTM yapısında uzmanlaşmış olmasıdır. Unutma kapısı sayesinde geçmişten aktarılan ancak zorunlu olmayan bilgilerin ağırlığı azaltılır.

Güncelleme geçişi ve unutma etkisi ile daha etkin bir çıktı üretilir [42].

$$\Gamma_f = \sigma(W_f[a^{(t-1)}, x^{(t)}] + b_f), \quad (4)$$

$$\Gamma_o = \sigma(W_o[a^{(t-1)}, x^{(t)}] + b_o). \quad (5)$$

Eş. (1), (2) ve (4) kullanılarak, yeni $\hat{c}^{(t)}$ değeri şu şekilde elde edilebilir:

$$c^{(t)} = \Gamma_u * \hat{c}^{(t)} + \Gamma_f * \hat{c}^{(t-1)}. \quad (6)$$

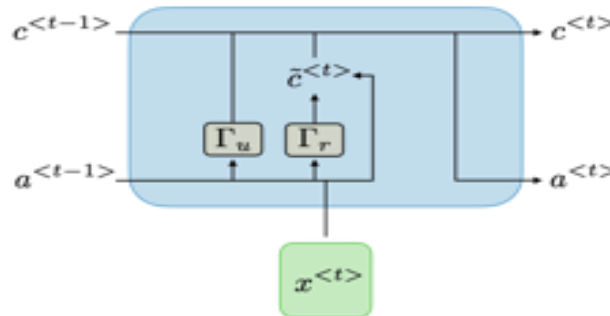
$a^{(t)}$ nin son değeri ise,

$$a^{(t)} = \Gamma_o * c^{(t)}. \quad (7)$$

olarak hesaplanır.

3.2. GRU (Gated recurrent unit)

Geçitlenmiş tekrarlayan birimler (GRU), LSTM'den türetilen ve Cho ve diğ. (2014) tarafından tanımlanan sinir ağındaki hücreler arasındaki bilgi akışını kontrol etmek için geçit mekanizmalarını kullanan bir tür RNN'dir [46]. GRU, güncelleme kapısı ve sıfırlama kapısı olmak üzere iki kapıdan oluşur, bu kapılar hangi bilgilerin kalması ve nelerin atılması gerektiğini filtrelemek için kullanılır. Geleneksel RNN'den farklı olarak GRU'lar, kaybolan ve patlayan gradyan problemlerini çözmektedir. LSTM'den farklı olarak GRU, bir geçidin olmaması nedeniyle LSTM'den daha az parametreye sahiptir. Diğer bir fark, GRU'ların LSTM'den hücre durumu eksikliğine sahip olmalarıdır, böylece GRU yalnızca hem uzun hem de kısa süreli belleği gizli durumda depolayabilmektedir.



Şekil 2. GRU mimarisi [45]
(GRU architecture)

Özyinelemeli sinir ağları a katmanlarındaki aktivasyon fonksiyonları ve zamana bağlı işlemlerin yapılaş şekline göre sınıflandırılabilir. Bu iç birimlerdeki bağlantı ve aktivasyon işlemleri modelin bellek özelliklerini doğrudan etkilemektedir [47]. Hatırlama ve unutma gibi işlemler bu birimlerde gerçekleştirilir [48]. Şekil 2'de bir GRU blok şeması gösterilmektedir.

$\hat{c}^{(t)}$ bellek için aday değeri sembolize eder. Aktivasyon fonksiyonu olarak hiperbolik tanjant kullanılır.

Burada aday değeri şu şekilde hesaplanır:

$$\hat{c}^{(t)} = \tanh(W_c[\Gamma_r * a^{(t-1)}, x^{(t)}] + b_c), \quad (8)$$

Güncelleme ve ilgi düzeyi kapısı değerleri için aktivasyon fonksiyonu olarak sigmoid kullanılır. Güncelleme ağ geçidi Γ_u ve ilgililik geçidi Γ_r aşağıdaki gibi hesaplanır:

$$\Gamma_u = \sigma(W_u[a^{(t-1)}, x^{(t)}] + b_u), \quad (9)$$

$$\Gamma_r = \sigma(W_r[a^{(t-1)}, x^{(t)}] + b_r). \quad (10)$$

Γ_u kullanılarak yeni $\hat{c}^{(t)}$ değeri şu şekilde elde edilebilir:

$$c^{(t)} = \Gamma_u * \hat{c}^{(t)} + (1 - \Gamma_u) * c^{(t-1)}. \quad (11)$$

$a^{(t)}$ nin son değeri ise,

$$a^{(t)} = c^{(t)} \quad (12)$$

olarak hesaplanır.

3.3. GAN (Generative adversarial networks)

Üretici Çekişmeli Ağlar (Generative Adversarial Networks - GAN), görüntü, ses ve video gibi yüksek kaliteli sentetik veriler oluşturmak için popüler hale gelen bir derin öğrenme algoritmaları sınıfıdır. GAN modeli ilk olarak Goodfellow ve diğ. (2014) tarafından tanıtılmış ve bilgisayarlı görü, doğal dil işleme, görüntüden görüntüye çeviri ve tıbbi görüntüleme analizi dahil olmak üzere çok çeşitli uygulamalar için kullanılmıştır [49-51].

GAN modelinin arkasındaki temel fikir, gerçek verilerden ayırt edilemeyen sentetik veriler oluşturmak için bir model eğitmektir. Bunu yapmak için GAN'lar iki sinir ağı kullanır: bir üretici ve bir ayırıcı. Üretici, sentetik veri üretmek için eğitilirken, ayırıcı, belirli bir veri örneğinin gerçek mi yoksa sentetik mi olduğunu belirlemek için eğitilir. İki ağ, üretici ayırıcıyı kandırmaya çalışırken ve ayırıcı da sentetik verileri doğru bir şekilde tanımlamaya çalışırken, çekişmeli bir şekilde eğitilmektedir.

Bir GAN'daki üretici ve ayırıcı, üretilen ve gerçek veriler arasındaki farkı ölçen bir kayıp fonksiyonu kullanılarak eğitilir. Üretici bu kaybı en aza indirmek için eğitilirken, ayırıcı bunu en üst düzeye çıkarmak için eğitilir. Eğitim ilerledikçe, üretici gerçekçi veriler üretme yeteneğini geliştirirken, ayırıcı gerçek verileri sentetik verilerden ayırt etme yeteneğini geliştirir. Eğitim süreci, üretici gerçek verilerden ayırt edilemeyen veriler üretene kadar devam eder.

GAN'ların en güçlü yönlerinden biri, gerçek verilere benzer yüksek kaliteli sentetik veriler üretme yetenekleridir. Bu, onları tıbbi görüntüleme veya uydu görüntüleme gibi büyük miktarda etiketlenmiş verinin bulunmadığı uygulamalar için çok uygun hale getirir. Bu uygulamalarda, diğer makine öğrenimi modellerini eğitmek için kullanılacak sentetik veriler oluşturmak için GAN'lar kullanılabilir.

GAN'ların bir başka avantajı da çok yönlü olmalarıdır. Görüntüler, ses ve video dahil olmak üzere çeşitli veri türleri oluşturabilirler. Bu, gerçekçi yüzler oluşturmak, gerçekçi konuşmaları sentezlemek veya gerçekçi animasyonlar oluşturmak gibi belirli uygulamalar için birçok özel GAN'ın geliştirilmesine yol açmıştır.

GAN'da, kayıp fonksiyonu Jensen-Shannon (JS) Sapmasına dayanır; eğitim sürecinde GAN modeli, iki dağıtım arasındaki farkı en aza indirmek için çapraz entropi kaybını kullanmaktadır; böylece JS sapması en aza indirmeye çalışılır. Ayırıcı, gerçek veri örneklerini doğru sınıflandırmayı ve sahte veri örneklerini sahte olarak sınıflandırmayı amaçlar. Ayırıcı kayıp fonksiyonu şu şekilde tanımlanmaktadır:

$$\mathcal{L}_D = -\mathbb{E}_{x \sim p_{data}(x)} [\log(D(x))] - \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (13)$$

Burada $x \sim p_{data}(x)$, gerçek veri dağılımından örneklenen gerçek veriyi, $z \sim p_z(z)$, gürültü dağılımından (genellikle normal veya uniform dağılım) örneklenen rastgele gürültü vektörünü ifade ederken $G(z)$, üreticinin gürültü vektöründen ürettiği veri örneğidir.

Üretici, ürettiği sahte verilerin ayrıştırıcı tarafından gerçek olarak tanınmasını amaçlar. Üretici kayıp fonksiyonu şu şekilde ifade edilir:

$$\mathcal{L}_G = -\mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))], \quad (14)$$

GAN modelinin, yok olan gradyan, eğitim zorluğu ve zayıf çeşitlilik gibi birçok sorunu vardır [52]. Farklı optimizasyon yöntemleriyle daha iyi GAN'lar elde etmek için birçok çaba sarf edilmiştir. Bu nedenle, GAN ile ilgili üretilen sonuçların kararlılığını ve kalitesini iyileştirmek için birçok yeni GAN tabanlı model önerilmiştir [53].

3.4. WGAN-GP (Wasserstein GAN with gradient penalty)

GAN'daki ayırıcı yeterince güçlü olmadığı ve eğitim sürecinin yavaş ve istikrarsız olduğu bilinmektedir. WGAN, GAN eğitimini stabilize etmeye ve geliştirmeye yardımcı olmak için önerilmiştir [23]. JS sapması, ayrık parçalara sahip dağılımlar arasındaki mesafeyi ölçmek için uygun değildir. Bu nedenle WGAN, Wasserstein mesafesi veya optimal iletim mesafesi olarak da bilinen, olasılık dağılımını q 'dan p 'ye dönüştüren minimum iletim kalitesine atıfta bulunan yeni bir mesafe ölçüm yöntemi olan Earth Moving Distance'ı önermiştir [54]. Wasserstein mesafesinin JS sapmasına göre üstünlüğü, iki dağılım örtüşmese bile Wasserstein mesafesinin mesafelerini yansıtabilmesidir.

Wasserstein mesafesi gerçek veri dağılımı P_r , oluşturulan veri dağılımı P_g olmak üzere matematiksel olarak en büyük alt sınır (infimum) olarak tanımlanır [4].

$$W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} \mathbb{E}_{(x,y) \sim \gamma} [\|x - y\|] \quad (15)$$

$\Pi(P_r, P_g)$, P_r ve P_g arasındaki tüm ortak dağılımların kümesini ifade etmektedir. WGAN'da, bu mesafeyi hesaplamak için Kantorovich-Rubinstein dualitesini kullanarak şu şekilde bir kayıp fonksiyonu elde ederiz:

$$W(P_r, P_g) = \sup_{\|f\|_{L^1} \leq 1} \mathbb{E}_{x \sim P_r} [f(x)] - \mathbb{E}_{x \sim P_g} [f(x)] \quad (16)$$

burada sup en küçük üst sınırdır ve Burada f , 1-Lipschitz sürekli bir fonksiyondur. Bu nedenle, ayırıcı (WGAN terminolojisinde critic) fonksiyonu f olarak adlandırılır.

WGAN-GP, WGAN'ın bir varyasyonudur ve ağırlık kırpması (weight clipping) yerine gradient penalty (gradyan cezası) kullanarak daha stabil bir eğitim süreci sağlar. WGAN-GP, 1-Lipschitz sürekliliğini sağlamak için şu şekilde bir gradyan cezası terimi ekler:

$$\mathcal{L}_{GP} = \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}} [(\|\nabla_{\hat{x}} f(\hat{x})\|_2 - 1)^2], \quad (17)$$

Burada \hat{x} , gerçek veri dağılımı x_r ve üretilmiş veri dağılımı x_g arasındaki doğrusal birleşimdir:

$$\hat{x} = x_r + (1-\epsilon)x_g, \quad 0 \leq \epsilon \leq 1, \quad (18)$$

ve $\epsilon \sim U(0,1)$ bir uniform dağılımdan gelmektedir. Sonuç olarak, WGAN-GP'nin kayıp fonksiyonu şu şekilde yazılır:

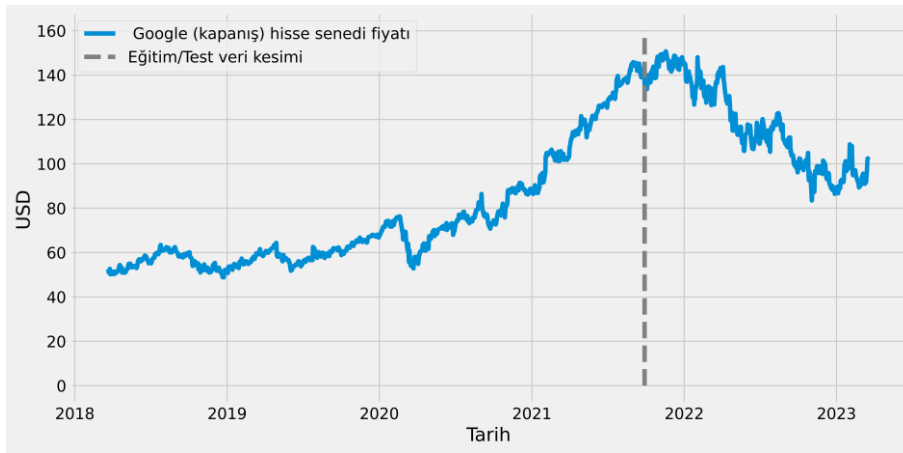
$$\mathcal{L} = \mathbb{E}_{\hat{x} \sim \mathbb{P}_{\hat{x}}} [f(\hat{x})] - \mathbb{E}_{x \sim P_r} [f(x)] + \lambda \mathcal{L}_{GP}. \quad (19)$$

Burada λ gradyan ceza terimidir. Özetle, GAN, WGAN ve WGAN-GP, gerçek veriye benzeyen veri

örnekleri üretmek için farklı teknikler kullanır ve GAN'ların eğitim stabilitesini artırır. GAN'lar Jensen-Shannon sapmasını kullanırken, WGAN'lar Wasserstein mesafesini ve WGAN-GP'ler ise Lipschitz sürekliliğini gradyan cezası ile sağlar.

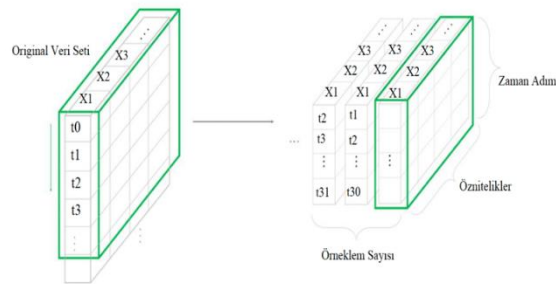
3.5. Veri (Data)

Hisse senedi fiyatı verileri ve hisse senedi endeksi verileri Yahoo Finance [55]'dan alınmıştır. Modeldeki hedef hisse senedi fiyatı Google hisse senedi kapanış fiyatıdır. İstatistiksel veriler hisse senedi kapanış fiyatı kullanılarak hesaplanmıştır. Veri setinde toplam 1258 gözlem bulunmaktadır. Şekil 3'te, Google (GOOG) hisse senedi kapanış fiyatının 22.03.2018 ile 21.03.2023 arasındaki seyri görülmektedir. Eğitim ve test verileri 7:3 oranında bölünmüştür. Grafikte gri kesikli çizgi ile gösterilen Eğitim/Test veri kesimine göre, 2010'dan 2019'a kadar olan günlük veriler eğitim, 2020'den 2024'e kadar olan veriler ise test verisi olarak kullanılacaktır.

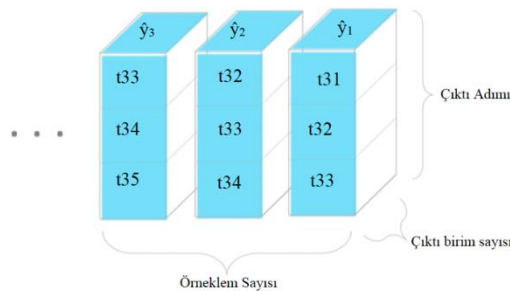


Şekil 3. Google hisse senedi kapanış fiyatları
(Google stock closing prices)

Şekil 4'te görüldüğü gibi orijinal veri seti iki boyutludur ve zaman adımlarına göre 3 boyutlu olarak dönen pencere 1'e eşit olacak şekilde yeniden şekillendirilmiştir. Şekil 5, üreticiden elde edilen veri setinin çıktısını göstermektedir. Burada çıktı birimlerinin sayısı 1'e eşittir. Üç günlük hisse senedi fiyatını tahmin etmek için 30 günlük geçmiş fiyatı kullanılmıştır.



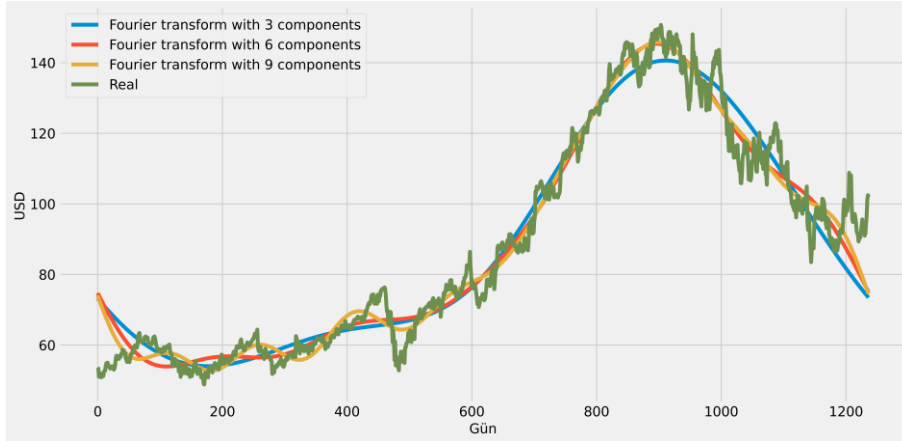
Şekil 4. Giriş verisi [23]
(Input data)



Şekil 5. Çıktı verisi [23]
(Output data)

Google hissesindeki uzun vadeli ve kısa vadeli trendleri çıkarmak için günlük kapanış fiyatıyla birlikte Fourier dönüşümleri oluşturulmuştur. Fourier dönüşümleri sinüs dalgası oluşturarak orijinal fonksiyona yaklaşır ve GRU ve LSTM ağının tahmin trendlerini daha doğru seçmesine yardımcı olmaktadır [56]. Fourier dönüşümü, bir zaman serisinin zaman alanındaki temsilini frekans alanına çevirmeye yarar. Başka bir deyişle, bir işaretin hangi frekans bileşenlerinden oluştuğunu bulmak için kullanılır. Eş. (20) ile zaman serisi verilerine hızlı Fourier dönüşümü uygulanarak veri kümesine yeni özellikler eklenmektedir.

$$G(f) = \int_{-\infty}^{\infty} g(t)e^{-i2\pi ft} dt \quad (20)$$



Şekil 6. Fourier dönüşümü
(Fourier transform)

Hisse senedindeki küresel ve yerel eğilimleri çıkarmak ve aynı zamanda biraz gürültü çıkarmak için Fourier dönüşümler kullanılmıştır. Hisse senedi kapanış fiyatları, Fourier dönüşümü ile analiz edilerek, bu analizin sonuçlarını orijinal veri kümesine yeni özellikler olarak eklemeye yaramaktadır. Bu yeni özellikler, daha sonra derin öğrenme algoritmalarında kullanılmaktadır.

Şekil 6'da görüldüğü gibi, Fourier dönüşümünden ne kadar çok bileşen kullanılırsa, yaklaşıklık fonksiyonu gerçek hisse senedi fiyatına o kadar yakın olmaktadır. Uzun ve kısa vadeli eğilimleri çıkarmak amacıyla Fourier dönüşümleri kullanılmıştır, bu nedenle 3, 6 ve 9 bileşenli dönüşümler kullanılmıştır. 3 bileşenli dönüşüm uzun vadeli trend olarak düşünülebilir.

Finansal piyasaları analiz ederken, birçok yatırımcı çeşitli teknik göstergelere başvurmaktadır. Bu çalışmada da yaygın olarak kullanılan bazı teknik göstergelerden 7 ve 21 günlük hareketli ortalama, üstel hareketli ortalama, momentum, Bollinger bantları ve Hareketli Ortalama Yakınsama İraksama (MACD) gibi göstergeler kullanılmıştır (Şekil 7). Bu göstergeler, yatırımcılara piyasa dinamiklerini daha iyi anlamada ve karar verme süreçlerinde önemli bilgiler sağlamaktadır.

7 ve 21 Günlük Hareketli Ortalama (Simple Moving Average - SMA), fiyat hareketlerini daha pürüzsüz hale getirerek trendleri daha kolay anlaşılır kılar. 7 ve 21 günlük hareketli ortalamalar sıklıkla kullanılır ve belirtilen süreler boyunca ortalama fiyatları gösterir. Örneğin, 7 günlük hareketli ortalama, son 7 günün kapanış fiyatlarının ortalamasını alır. Bu gösterge, kısa vadeli ve orta vadeli trendleri belirlemeye yardımcı olur. Üstel Hareketli Ortalama (Exponential Moving Average - EMA), SMA'ya benzer şekilde, fiyat hareketlerini pürüzsüzleştirir. Ancak, EMA daha yakın zamandaki fiyat hareketlerine daha fazla ağırlık verir. Bu, EMA'nın SMA'dan daha hızlı tepki vermesini sağlar ve böylece trend değişikliklerini daha çabuk yakalar.



Şekil 7. Teknik analiz
(Technical analysis)

Momentum, bir finansal enstrümanın fiyat değişiminin hızını veya gücünü ölçer. Momentum, belirli bir zaman dilimi boyunca fiyat değişimlerinin ortalamasını hesaplayarak elde edilir. Yükselen momentum, artan fiyat hareketleri ve potansiyel yükseliş trendleri için güçlü bir göstergeden, düşen momentum düşüş eğilimlerine işaret edebilir. Bollinger Bantları, bir finansal enstrümanın oynaklığını ve potansiyel aşırı alım veya aşırı satım durumlarını ölçmek için kullanılır. Bollinger Bantları, hareketli bir ortalamanın üzerine ve altına yerleştirilen iki standart sapma bandıyla oluşur. Fiyatlar üst banda yaklaştığında aşırı alım, alt banda yaklaştığında ise aşırı satım olarak kabul edilir. Hareketli Ortalama Yakınsama İraksama (Moving Average Convergence Divergence, MACD), iki hareketli ortalama (genellikle 12 ve 26 günlük) arasındaki farkı ölçer ve bunu bir sinyal çizgisi (genellikle 9 günlük EMA) ile karşılaştırır. MACD çizgisi sinyal çizgisinin üzerine çıktığında, bu alım sinyali olarak yorumlanırken, sinyal çizgisinin altına düştüğünde satış sinyali olarak değerlendirilir. Bu göstergeler, yatırımcıların piyasa eğilimlerini ve potansiyel alım-satım fırsatlarını değerlendirmelerine yardımcı olur. Teknik göstergeler, genellikle yatırımcılar tarafından birbirleriyle birleştirilerek veya temel analizle birlikte kullanılarak daha güçlü ve güvenilir alım-satım sinyalleri üretir. Önemli olan, yatırımcıların bu göstergeleri ve stratejileri kendi risk ve ödül tercihlerine uygun bir şekilde uyarlamalarıdır. Tablo 1'de analizde kullanılan öznitelikler ve açıklamaları görülmektedir.

Tablo 1. Öznitelikler (Features)

Öznitelik adı	Öznitelik açıklaması
Open	İşlem günündeki açılış fiyatı
High	İşlem günündeki en yüksek fiyat
Low	İşlem günündeki en düşük fiyat
Close	İşlem günündeki kapanış fiyatı
Volume	Bir önceki işlem günündeki hacim
MA7	7 günlük basit hareketli ortalama
MA21	21 günlük basit hareketli ortalama
MACD	Hareketli ortalama yakınsama/ıraksama
20SD	Bollinger bantları orta eğri
Upper	Bollinger bandı üst eğri
Lower	Bollinger Bandı alt eğri
EMA	Üstel hareketli ortalama
logmomentum	Logaritmik momentum göstergesi
absolute of 3 comp	3. dereceden yeniden yapılandırma (mutlak)
angle of 3 comp	3. dereceden yeniden yapılandırma (açı)
absolute of 6 comp	6. dereceden yeniden yapılandırma (mutlak)
angle of 6 comp	6. dereceden yeniden yapılandırma (açı)
absolute of 9 comp	9. dereceden yeniden yapılandırma (mutlak)
angle of 9 comp	9. dereceden yeniden yapılandırma (açı)

3.5. Deneysel tasarım (Experimental design)

Bu çalışma, son 30 günün verileriyle önümüzdeki üç gündeki hisse senedi kapanış fiyatını tahmin etmeyi amaçlamaktadır. Veri seti %70 ve %30 olmak üzere eğitim seti ve test seti olarak bölünmüştür. Bu çalışmanın çerçevesi, TensorFlow [38] arka ucuna sahip Keras [39] ile oluşturulmuştur. Aktivasyon

fonksiyonları için üretici ağ lineer (aktivasyonsuz) kullanılırken, çekişmeli ağ 0,01 gibi düşük bir eşik değerine sahip Leaky ReLU deneysel olarak iyi bir performansla sonuçlanmıştır.

Tahmin modellerinin doğruluğunu belirlemek için değerlendirme gereklidir. Modelin tahmin performansını görselleştirmek amacıyla, gerçek fiyatlar ile tahmin edilen fiyatları karşılaştıran bir saçılım grafiği kullanılmıştır. Ayrıca, modelin performansını değerlendirmek için hata metrikleri olarak hata kareler ortalamasının karekökü (RMSE), ortalama mutlak hata (MAE) ve R^2 skorları hesaplanmıştır. RMSE, MAE ve R^2 , literatürde en sık kullanılan performans değerlendirme ölçütlerinden bazılarıdır [57].

Bu çalışmada, her modelin performansı Hata Kareler Ortalamasının Karekökü (RMSE),

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y_i^*)^2}, \quad (21)$$

Ortalama Mutlak hata (MAE),

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - y_i^*|, \quad (22)$$

ve R^2 ,

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - y_i^*)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}, \quad (23)$$

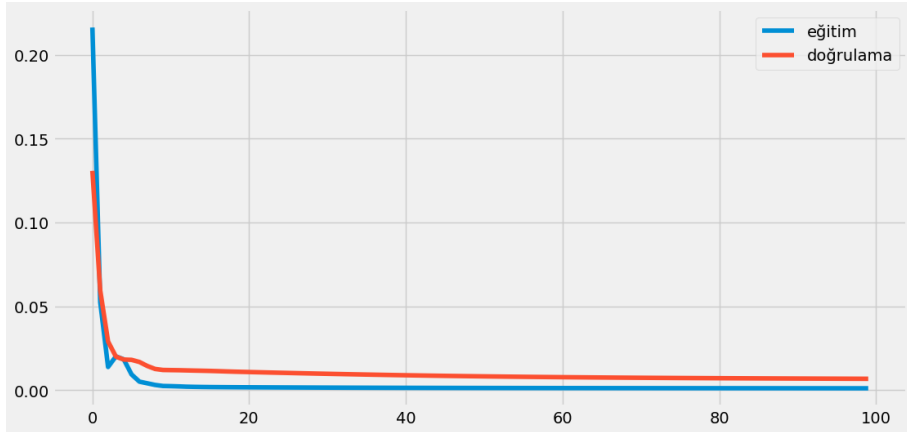
ile değerlendirilmiştir. Burada N veri sayısı, y_i gerçek hisse senedi fiyatı, y_i^* tahmin edilen hisse senedi fiyatı ve \bar{y} ortalama fiyatı belirtmektedir.

4. Bulgular (Results)

Bu bölümde, hisse senedi fiyat tahmini için kullanılan LSTM, GRU, GAN ve WGAN-GP modellerinin detayları ve performanslarının karşılaştırılması sunulmaktadır. Her modelin performansı, eğitim ve test veri kümeleri için Hata Kareler Ortalamasının Karekökü (RMSE), Ortalama Mutlak Hata (MAE) ve R^2 değerleri ile değerlendirilmiştir. Ayrıca, modellerin öğrenme süreci ve tahmin performansı üzerindeki etkilerini gözlemlemek amacıyla çeşitli hiper parametreler ve optimize ediciler kullanılmıştır. Bu bölümde, her modelin uygulanması ve sonuçlarının analizi ayrı ayrı ele alınarak, en iyi performans gösteren modelin belirlenmesine yönelik bir değerlendirme yapılmaktadır.

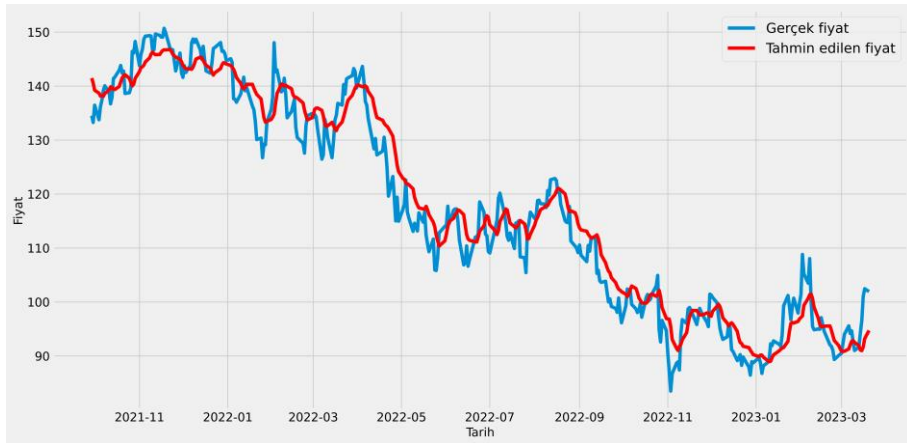
4.1. LSTM tahmin sonucu (LSTM prediction results)

LSTM modelinde ilk katmanda çift yönlü LSTM kullanılmıştır. Çift yönlü LSTM katmanı, zaman serisi verilerinde gelecekteki ve geçmişteki bilgilerin daha etkili bir şekilde öğrenilmesini sağlar. Dense katmanları ise modelin daha derin ilişkileri öğrenmesine olanak tanır. Çalışmamızda, modellerimizi eğitmek için temel olarak öğrenme oranı 0.0001 olan Adam (Adaptive Moment Estimation) optimize edicisi kullanılmıştır. Adam, adaptif öğrenme oranı ve momentum içeren bir optimizasyon algoritmasıdır ve genellikle hızlı ve güvenilir sonuçlar vermektedir. Bu oran, LSTM modelinin hızlı ve etkili bir şekilde öğrenmesini sağlamaktadır. Yığın boyutu 128'dir ve model, veri kümesinde 100 dönem eğitilmiştir. Bu boyut, modelin her iterasyonda yeterli veri görmesini ve daha sık güncellenmesini sağlar. Modelin öğrenme oranı, yığın boyutu ve dönem sayısı, farklı kombinasyonlar deneyerek en iyi sonuçlar veren ayarlar olarak belirlenmiştir. Bu parametrelerin belirlenmesi sırasında modelin eğitim ve doğrulama performansları dikkate alınmıştır. Öğrenme sürecinde kayıp değerlerinin takibi yapılmış ve modelin doğrulama performansı izlenmiştir.

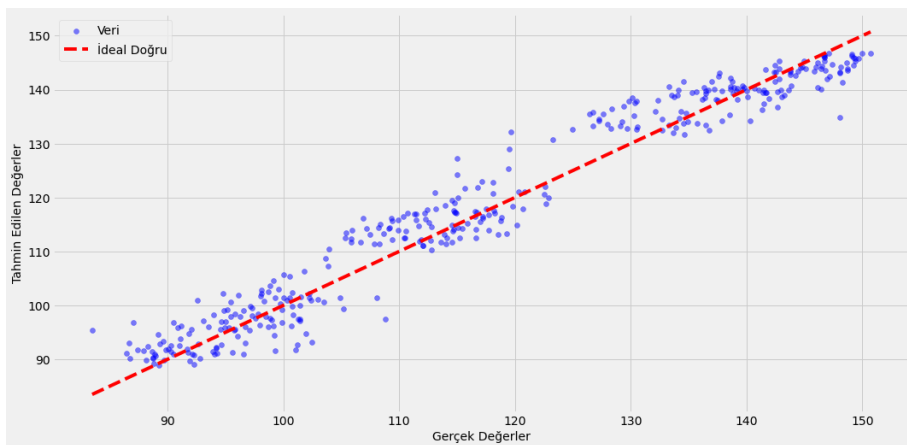


Şekil 8. LSTM modeli için eğitim ve doğrulama kümelerinin kayıp fonksiyonlarının iterasyona göre değişimi
(LSTM model loss for training and test set)

Şekil 8'de LSTM modeli için eğitim ve doğrulama başarımı gösterilmiştir. RMSE, eğitim verisi için 1,74 iken test verisi için 4,21'dir. MAE ise eğitim verisi için 1.29 iken test verisi için 3.39'dur. R2 değeri ise eğitim verisi için %99.53 iken test verisi için %95,09'dur. Şekil 9'da mavi çizgi hisse senedi fiyatını ve kırmızı çizgi tahmin edilen hisse senedi fiyatını göstermektedir. Şekil 10, LSTM modelinin gerçek ve tahmini fiyatların saçılım grafiğini göstermektedir. Buna göre LSTM modeli en iyi performansı göstermektedir.



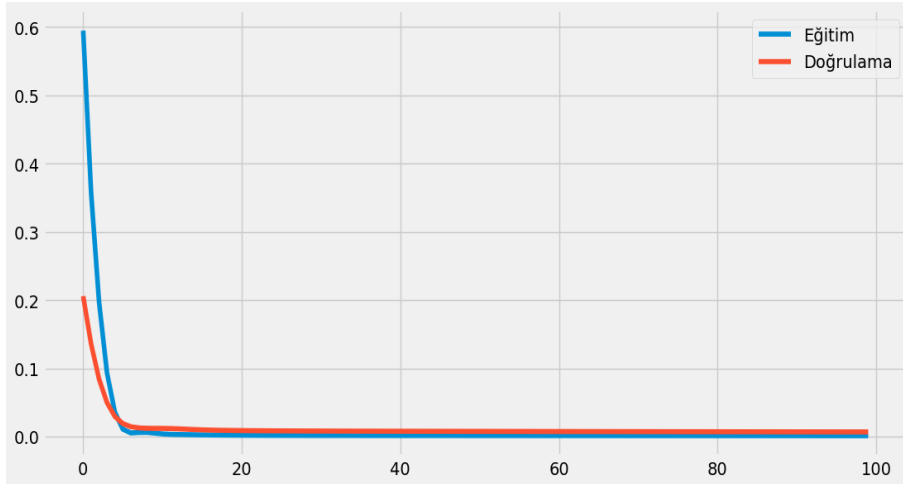
Şekil 9. LSTM tahmin sonucu
(LSTM prediction result)



Şekil 10. LSTM modelinin gerçek ve tahmini fiyatlarının dağılımı
(Scatter of actual and forecast prices of the LSTM model)

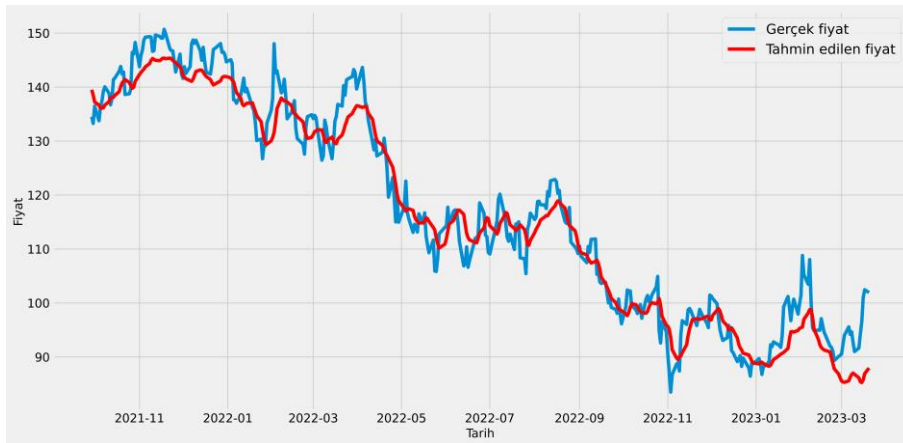
4.2. GRU tahmin sonucu (GRU prediction results)

GRU modelini oluşturmak için 2 GRU katmanı kullanılmıştır ve optimizasyon için öğrenme oranı 0.0001 olan Adam algoritması kullanılmıştır. Yığın boyutu 128'dir ve model 100 dönem eğitilmiştir.

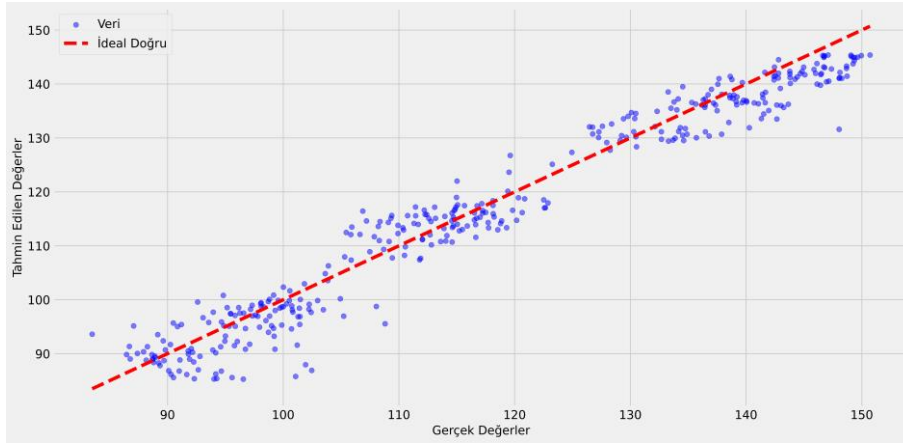


Şekil 11. GRU modeli için eğitim ve doğrulama kümelerinin kayıp fonksiyonlarının iterasyona göre değişimi (GRU model loss for training and test set)

Şekil 11, GRU için eğitim ve doğrulama başarımını, Şekil 12, GRU modelinin tahmin sonucunu, Şekil 13 ise GRU modelinin gerçek ve tahmini fiyatların saçılım grafiğini göstermektedir. RMSE eğitim verisi için 2.02 iken test verisi için 4.30'dur. MAE ise eğitim verisi için 1.51 iken test verisi için 3.42'dir. R2 değeri ise eğitim verisi için %99.36 iken test verisi için %94.40'tır. Performans olarak GAN modelinden sonra gelmektedir. Bu sonuç, tahminin 2020'de gerçek fiyat ile tahmin edilen fiyat arasında büyük bir boşluk oluşmaya başladığını ve gerçek fiyatta beklenmedik bir olay olan COVID-19'dan kaynaklanabilecek ani bir artış olduğunu gösteriyor. Tahmin için GRU'nun beklenmedik olaylar meydana geldiğinde her iki temel modelden daha kötü performans sergilediğini göstermektedir.



Şekil 12. GRU tahmin sonucu (GRU prediction result)

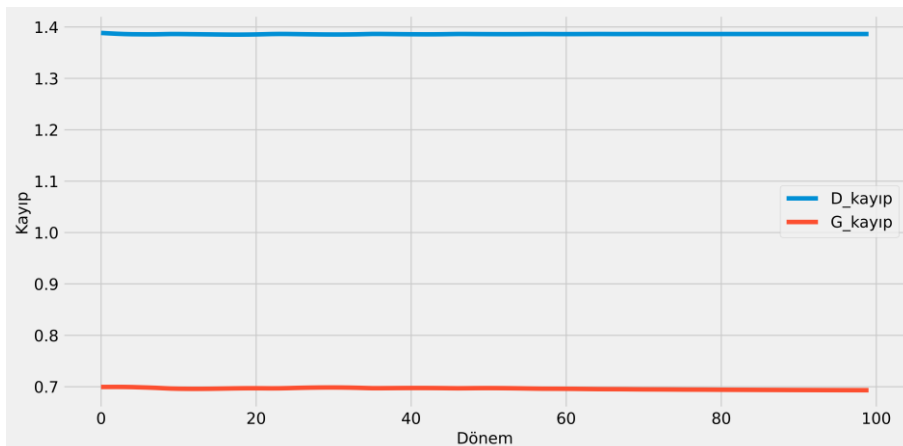


Şekil 13. GRU modelinin gerçek ve tahmini fiyatlarının dağılımı
(Scatter of actual and forecast prices of the GRU model)

4.3. GAN tahmin sonucu (GAN prediction results)

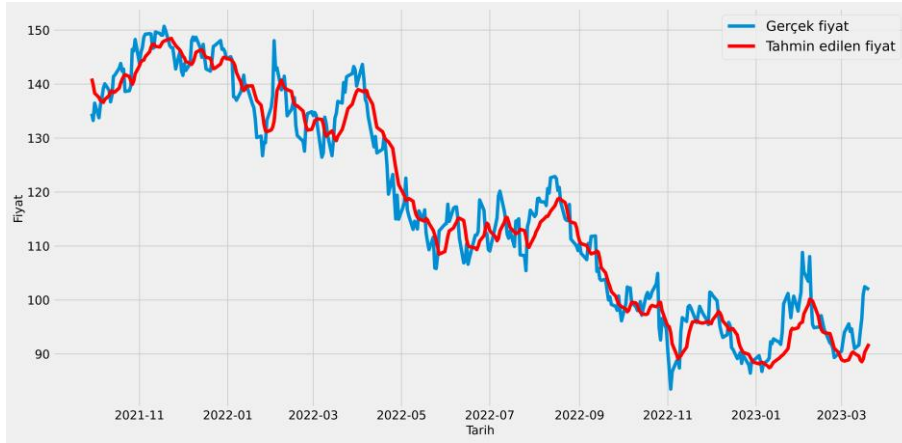
Bu çalışmada kullanılan GAN modeli, üretici ve ayırıcı olmak üzere iki ana bileşenden oluşmaktadır. Üretici modeli, üç katmanlı GRU ve Dense katmanları kullanarak sahte veri üretir. Ayırıcı model ise, üç katmanlı Conv1D ve Dense katmanlarıyla girdi verisinin gerçek mi sahte mi olduğunu belirler. GAN eğitimi sırasında, üretici daha gerçekçi veriler üretmeye çalışırken, ayırıcı bu verileri ayırt etmeye çalışır, böylece her iki model de performansını artırır.

Bu modelde, optimize edici, öğrenme oranı 0.0001 olan Adam algoritmasıdır. Yığın boyutu 128'dir ve bu veri kümesindeki model 100 dönem için eğitilmiştir. Farklı öğrenme oranları ve yığın boyutları, çeşitli denemeler sonucunda belirlenmiştir. Modelin performansını optimize etmek için çeşitli kombinasyonlar denenmiş ve en iyi performansı veren parametreler seçilmiştir. Şekil 14, mavi çizgi ayırıcının kayıp çizgisi ve kırmızı üreticinin kayıp çizgisi olmak üzere temel GAN modelinin kayıp grafiğidir. Başından itibaren, ayırıcının kaybı üreticinin kaybından daha fazladır ve eğitim süreci boyunca her iki kayıp çizgisi de düz hale gelmiştir.

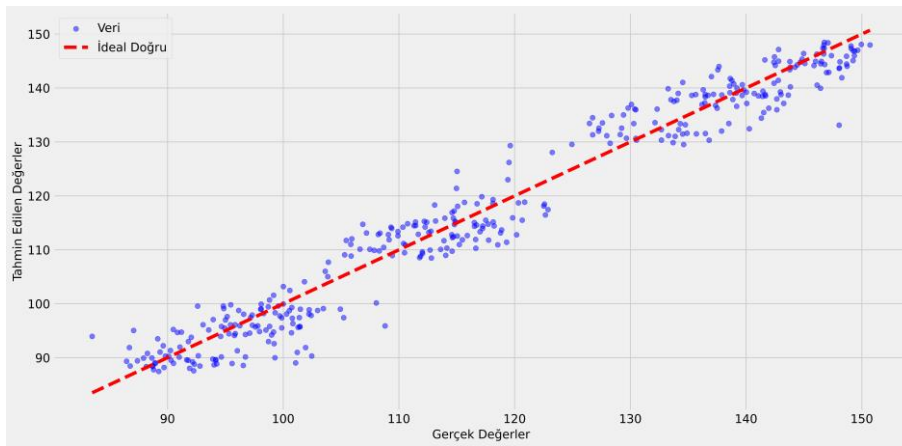


Şekil 14. GAN modeli için eğitim ve doğrulama kümelerinin kayıp fonksiyonlarının iterasyona göre değişimi
(GAN model loss for training and test set)

Şekil 15, GAN modelinin tahmin sonucunu ve Şekil 16 ise GAN modelinin gerçek ve tahmini fiyatların saçılım grafiğini göstermektedir. GAN modelinin RMSE değeri eğitim verisi için 2,30, test verisi için 4,05 iken MAE eğitim verisi için 1,69, test verisi için 3,29'dur. R^2 değeri ise eğitim verisi için %99,09 olurken test verisi için %95,78'dir. Buna göre GAN modeli LSTM modelinden sonra en iyi performansı göstermektedir.



Şekil 15. GAN tahmin sonucu
(GAN prediction result)

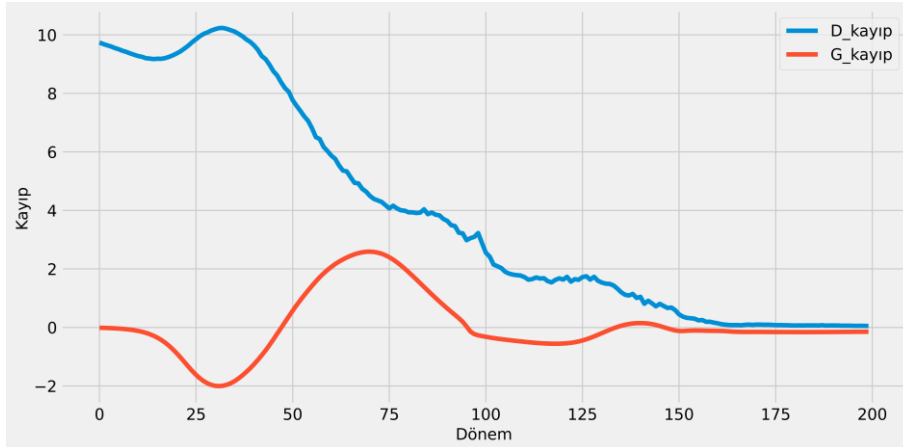


Şekil 16. GAN modelinin gerçek ve tahmini fiyatlarının dağılımı
(Scatter of actual and forecast prices of the GAN model)

4.4. WGAN-GP tahmin sonucu (WGAN-GP prediction results)

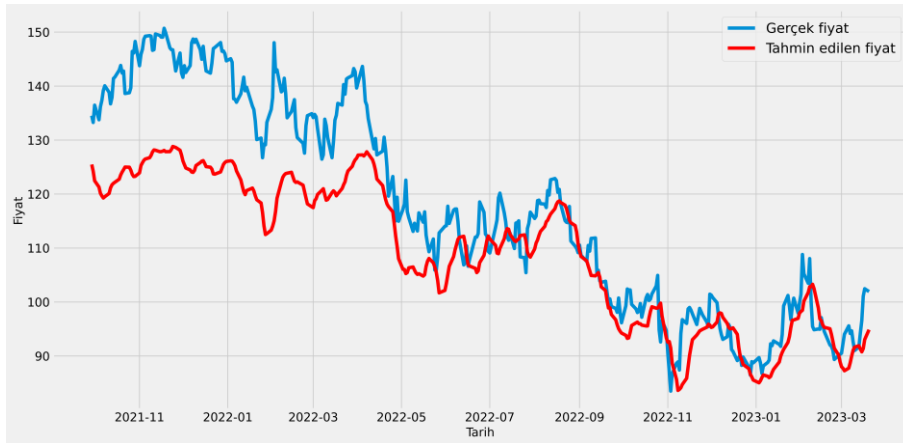
WGAN-GP modeli, üretici ve ayırıcı olmak üzere iki ana bileşenden oluşur. Üretici modeli, GRU ve Dense katmanları kullanarak sahte veri üretirken, ayırıcı model Conv1D ve Dense katmanlarıyla verinin gerçek mi yoksa sahte mi olduğunu belirler. Modelin eğitim sürecinde, gradyan ceza uygulanarak ayırıcının öğrenme süreci dengelenir. Üretici ve ayırıcı modeller, sahte veriler üretip bunları ayırt etmeye çalışarak birbirlerine karşı öğrenir ve böylece her iki modelin de performansı artırılır.

WGAN-GP modelinde optimize edici, öğrenme oranı 0.0001 olan bir Adam algoritmasıdır. Bu, modelin daha yavaş ama daha stabil bir şekilde öğrenmesini sağlamaktadır. WGAN-GP, GAN'den farklı olarak, modelin stabilitesine ve GAN modelinin eğitiminde görülen yaygın problemleri azaltmaya odaklanır. Yığın boyutu 128'dir ve model bu veri seti üzerinde 200 dönem için eğitilmiştir. Bu çalışmada ayırıcıyı bir kez, üretici üç kez eğitilmiştir. WGAN-GP modelinin optimizasyon süreci, GAN'den farklı olarak gradyan cezası kullanılmaktadır. Bu, ayırıcının daha iyi bir şekilde eğitilmesini sağlar ve GAN eğitiminde sıkça karşılaşılan sorunları azaltır. Farklı parametre kombinasyonları denenerek, modelin en iyi performansı göstermesi sağlanmıştır. Şekil 17, WGAN-GP modelinin kayıp grafiğidir, mavi çizgi ayırıcının kayıp yoludur ve turuncu çizgi, üreticinin kayıp yoludur. Ayırıcı kaybı 0'a doğru azalmaktadır. GAN ile karşılaştırıldığında, WGAN-GP'deki ayırıcı daha iyi öğrenmektedir.

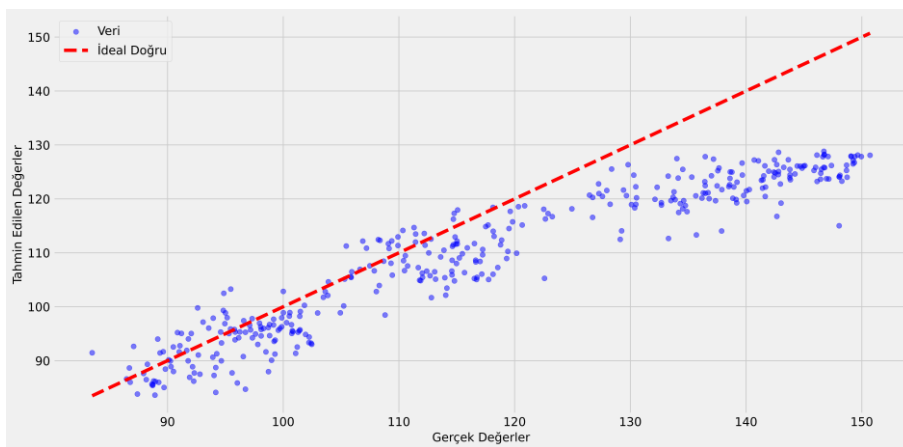


Şekil 17. WGAN-GP modeli için eğitim ve doğrulama kümelerinin kayıp fonksiyonlarının iterasyona göre değişimi (WGAN-GP model loss for training and test set)

Şekil 18, WGAN-GP modelinin tahmin sonucunu, Şekil 19, R^2 grafiğini göstermektedir. RMSE eğitim veri seti için 4,20, test veri seti için 11,52 iken MAE eğitim veri seti için 2,85, test veri seti için 9,02'dir. R^2 değeri ise eğitim verisi için %96,81 olurken test verisi için %24,52'tür. Bu sonuç modellerde en yüksek hata oranıdır. GRU modelinde olduğu gibi, tahminde beklenmedik COVID-19 nedeniyle gerçek fiyat ile tahmin edilen fiyat arasında büyük bir boşluk oluşmaya başlamaktadır.



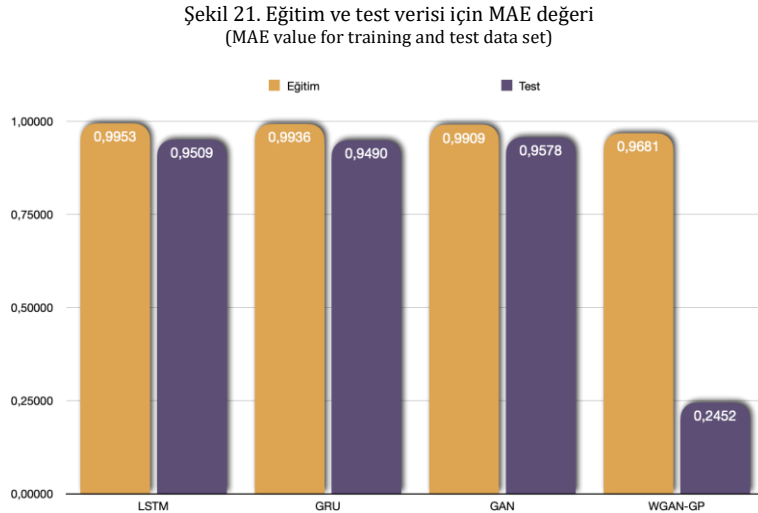
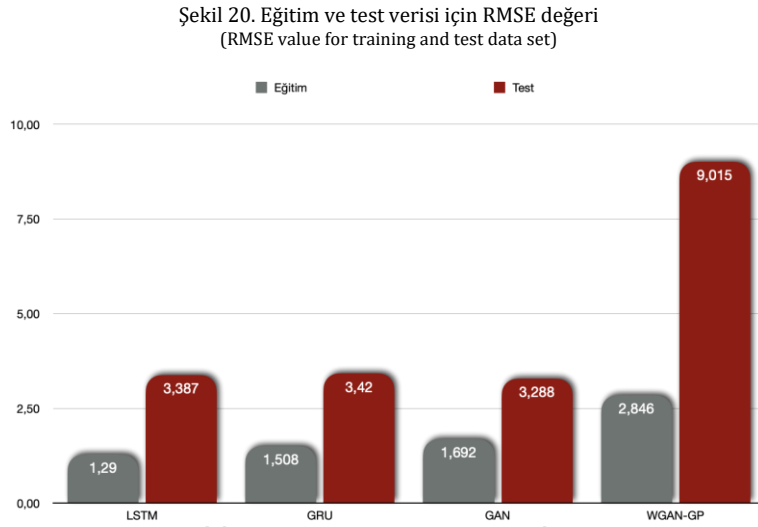
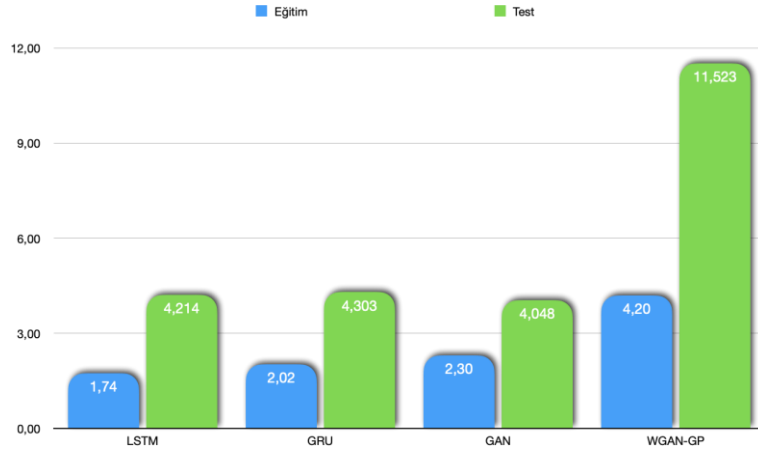
Şekil 18. WGAN-GP tahmin sonucu (WGAN-GP prediction result)



Şekil 19. WGAN-GP modelinin gerçek ve tahmini fiyatlarının dağılımı (Scatter of actual and forecast prices of the WGAN-GP model)

Şekil 20, Şekil 21 ve Şekil 22, deneysel çalışmada kullanılan dört model için sırasıyla RSME, MAE ve R^2 değerlerini karşılaştırmaktadır. Buna göre eğitim ve test veri kümesi için LSTM modeli en iyi performansı sergilerken, GAN modeli ikinci en iyi performansı göstermiştir. GRU modeli ise, LSTM ve GAN'dan daha düşük performans gösterirken, WGAN modeli en düşük performansa sahiptir. Bu

modellerde, gerçek fiyatlar ve tahminler arasında 2020'de COVID-19 nedeniyle beklenmedik bir olay gerçekleştiğinde büyük boşluklar oluşturmaya başlamıştır. Bu da bu tür modellerin beklenmedik olaylarla başa çıkmakta zorlanabileceğini göstermektedir.



Şekil 22. Eğitim ve test verisi için R² değeri (R² value for training and test data set)

4. Tartışma (Discussion)

Bu çalışma kapsamında, hisse senedi fiyatlarının tahmin edilmesi için LSTM, GRU, GAN ve WGAN-GP derin öğrenme yöntemleri kullanılarak ampirik bir analiz gerçekleştirilmiştir. Çalışmanın sonuçları, incelenen modellerin hisse senedi fiyat tahminindeki performanslarını ve zorluklarını ortaya

koymaktadır.

Araştırmamızda kullanılan LSTM ve GAN modelleri, hem eğitim hem de test verisi üzerinde düşük hata ve yüksek doğruluk değerleri ile en iyi performansı göstermiştir. LSTM modelinin eğitim verisindeki RMSE değeri 1.74 ve doğruluk değeri 0.9953, test verisindeki RMSE değeri 4.214 ve doğruluk değeri 0.9509 olup, modelin yüksek doğruluk ve düşük hata ile çalıştığını göstermektedir. Bu sonuç, mevcut literatürde LSTM'nin finansal zaman serilerinin tahmininde etkili bir yöntem olduğunu doğrular niteliktedir [58-60]. Özellikle, Gers ve diğ. (2002), Lu ve diğ. (2020), Lv ve diğ. (2021) gibi çalışmalar, LSTM'nin finansal tahminlerde güçlü bir performans sergilediğini göstermektedir. Bu çalışmada elde edilen RMSE ve MAE değerleri, LSTM modelinin hisse senedi fiyatlarını başarılı bir şekilde tahmin etme yeteneğini göstermektedir [17,61,62].

Benzer şekilde, GAN modelinin eğitim verisindeki RMSE değeri 2.30 ve doğruluk değeri 0.9909, test verisindeki RMSE değeri 4.048 ve doğruluk değeri 0.9578 olarak tespit edilmiştir. Bu sonuçlar, GAN modelinin de yüksek doğruluk ve düşük hata oranlarına sahip olduğunu göstermektedir. [64]. GRU modelinin eğitim verisindeki RMSE değeri 2.02 ve doğruluk değeri 0.9936, test verisindeki RMSE değeri 4.303 ve doğruluk değeri 0.9490 olup, GAN modeline göre biraz daha düşük performans sergilemektedir [44]. GRU'nun beklenmedik olaylar (COVID-19 pandemisi) meydana geldiğinde performansının düştüğü gözlemlenmiştir [63]. Bu durum, GRU modelinin daha dayanıklı hale getirilmesi gerektiğini göstermektedir.

WGAN-GP modelinin eğitim verisindeki RMSE değeri 4.20 ve doğruluk değeri 0.9681, test verisindeki RMSE değeri 11.523 ve doğruluk değeri 0.2452 olup, diğer modellere kıyasla daha yüksek hata oranına ve düşük doğruluğa sahiptir. [65]. Sonuç olarak, LSTM ve GAN modelleri en iyi performansı sergileyen modeller olarak öne çıkmaktadır. GRU modeli de kabul edilebilir performans göstermekte olup, GAN ve LSTM modellerinin ardından gelmektedir. WGAN-GP modeli ise eğitim verisinde makul bir performans gösterse de test verisinde düşük performans sergilemektedir. WGAN-GP'nin başarılı olduğu görüntü sentezi ve diğer alanlardaki performansını finansal tahminlere aktarmak için yapılandırma ve eğitim süreçlerinde önemli iyileştirmeler yapılması gerekmektedir [54,66]. Ayrıca, WGAN-GP gibi modellerin beklenmedik olaylara karşı daha dayanıklı olabilmesi için tahmin süreçlerinin adaptasyon mekanizmalarının geliştirilmesi önemlidir.

Tablo 2. Geçmiş çalışmalar (Past studies)

Çalışma	Yöntem	RMSE
[21]	LSTM	5.41
	GAN	4.10
[23]	CNN	42.97
	RNN	42.96
	LSTM	41.00
	CNN-RNN	40.54
[32]	CNN-LSTM	39.69
	LSTM	6.60
	GRU	-5.33
	GAN	5.36
[67]	WGAN-GP	4.77
	LSTM	1.05
	GRU	0.95
Bu çalışma	WGAN-GP	0.49
	LSTM	4.21
	GRU	4.30
	GAN	4.05
	WGAN-GP	11.52

Tablo 2'de, literatürde yapılmış olan bazı çalışmalara ait sonuçlar özetlenmiştir. Zang ve diğ. (2019), GAN tabanlı yöntemin hisse senedi piyasası tahminlerinde başarılı bir performans sergilediğini göstermektedir. Bu yöntem, diğer geleneksel yöntemlerle karşılaştırıldığında daha iyi sonuçlar elde etmekte ve özellikle kısa vadeli tahminlerde etkilidir [14]. Lu ve diğ. (2020), yöntemler arasında CNN-LSTM'nin performansının en iyi olduğunu göstermektedir [17]. Tahmin doğruluğu açısından RMSE 39.69'dur, bu değerler tahmin modelleri arasında en düşük olanıdır ve yüksek tahmin doğruluğuna sahiptir. Lin ve diğ. (2021), GAN tabanlı modelin hisse senedi fiyat tahminlerinde etkili olduğunu göstermektedir [23]. GAN modeli diğer geleneksel ve makine öğrenimi yöntemlerine göre daha iyi sonuçlar elde etmektedir. Xu ve diğ. (2022), WGAN-GP modelinin RMSE değerinin en düşük olduğunu göstermiş ve sırasıyla LSTM ve GRU modellerinden %61.94 ve %47.42 daha düşüktür. WGAN-GP

modeli, klasik zaman serisi tahmin modellerinden daha yüksek tahmin doğruluğu elde etmiştir [67].

4. Sonuç (Conclusion)

Bu çalışma, incelenen derin öğrenme yöntemlerinin hisse senedi fiyat tahminindeki performanslarını ve zorluklarını gözler önüne sermektedir. Sonuçlar, literatürdeki bulgularla uyumlu olup, LSTM ve GAN modellerinin bu alanda kullanılabilir olduğunu göstermektedir. Bununla birlikte, GRU ve WGAN-GP modellerinin finansal tahminlerde daha güçlü bir performans sergilemesi için yapılandırma, eğitim ve özellikle beklenmedik olaylara adaptasyon süreçlerinde önemli iyileştirmeler yapılması gerekmektedir.

İleride yapılacak çalışmalar, derin öğrenme modellerinin daha geniş finansal veri kümesi üzerinde test edilmesi, farklı parametre ve yapılandırmaların deneyimlerinin incelenmesi ve bu modellerin çeşitli hisse senedi piyasalarında uygulanabilirliğinin değerlendirilmesi üzerine odaklanabilir. Ayrıca, modellerin duyarlılık analizi ve hiper parametre optimizasyonu için daha kapsamlı denemeler de gelecekteki çalışmalarda değerli bilgiler sunabilir.

Ek olarak, daha karmaşık ve gerçek dünya finansal piyasalarına özgü faktörlerin modellere dahil edilmesi, tahmin performansını daha da artırabilir. Örneğin, temel ve teknik analiz verilerinin yanı sıra makroekonomik faktörlerin ve piyasa duyarlılık analizlerinin modelin tahmin yeteneğini nasıl etkilediğini araştırmak ilgi çekici olacaktır. Ayrıca, modellerin portföy yönetimi ve risk yönetimi stratejilerine nasıl entegre edilebileceği de değerlendirilmelidir.

Bu çalışmanın bir başka önemli boyutu, transfer öğrenme ve multitask öğrenme gibi daha gelişmiş öğrenme tekniklerinin finansal tahmin modellerine uygulanmasının araştırılmasıdır. Transfer öğrenme, bir modelin önceden öğrendiği bilgileri farklı ancak ilişkili bir görevde kullanmasını sağlayan bir makine öğrenmesi tekniğidir. Finansal tahmin modellerinde transfer öğrenme, farklı hisse senetleri, piyasa segmentleri veya ekonomik göstergeler arasındaki benzerliklerden faydalanarak tahmin performansını iyileştirebilir. Multitask öğrenme ise aynı anda birden fazla görevi öğrenen ve bu görevler arasında bilgi paylaşımını optimize eden bir makine öğrenmesi yaklaşımıdır. Finansal tahmin modellerinde multitask öğrenme, farklı finansal göstergeleri veya hisse senetlerini aynı anda tahmin ederek model performansını artırabilir.

Sonuç olarak, bu çalışmada, derin öğrenme yöntemlerinin hisse senedi fiyat tahminindeki uygulanabilirliği ve sınırlılıkları araştırılmıştır. Elde edilen sonuçlar, LSTM ve GAN modellerinin finansal tahminlerde etkili olduğunu ve mevcut literatürle uyumlu olduğunu göstermektedir. GRU ve WGAN-GP modelleri, beklenmedik olaylar ve finansal piyasalardaki belirsizlikler nedeniyle daha düşük performans sergilemektedir. Bu modellerin hisse senedi fiyat tahminindeki başarısını artırmak için daha fazla araştırma ve geliştirme yapılması gerekmektedir.

İleride yapılacak çalışmalar, farklı hisse senedi piyasaları ve finansal veriler üzerinde modellerin performansını daha kapsamlı bir şekilde değerlendirebilir. Ayrıca, daha gelişmiş öğrenme tekniklerinin ve veri entegrasyon stratejilerinin modellerin tahmin performansına nasıl katkıda bulunduğu üzerine odaklanabilir. Bu çerçevede, gelecekteki araştırmalar, hisse senedi fiyat tahminindeki derin öğrenme yöntemlerinin potansiyelini daha da geliştirmek için önemli katkılar sağlayacaktır.

Çıkar Çatışması Beyanı (Conflict of Interest Statement)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

Kaynaklar (References)

- [1] P. J. Brockwell, R. A. Davis, *Time Series: Theory and Methods*, Springer Science & Business Media, New York, USA, 2013.
- [2] J. D. Hamilton, *Time Series Analysis*, Vol. 2, Princeton University Press, Princeton, USA, 1994.
- [3] R. H. Shumway, D. S. Stoffer, *Time Series Analysis and Its Applications*, Springer, New York, USA, 2011.
- [4] X. Zhou, Z. Pan, G. Hu, S. Tang, and C. Zhao, "Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets," *Mathematical Problems in Engineering*, vol. 2018, pp. 1–11, 2018. doi:10.1155/2018/4907423

- [5] Y. Kara, M. A. Boyacioglu, and Ö. K. Baykan, "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5311-5319, May 2011. doi:10.1016/j.eswa.2010.10.027
- [6] K. Karagül, "İstanbul Menkul Kıymetler Borsası'nda İşlem Gören Firmaların Destek Vektör Makineleri Kullanılarak Sınıflandırılması," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 20, no. 5, ss. 174-178, 2014. doi:10.5505/pajes.2014.63835
- [7] R. S. Latha, G. R. Sreekanth, R. C. Suganthe, M. Geetha, R. E. Selvaraj, S. Balaji, K. R. Harini, and P. P. Ponnusamy, "Stock Movement Prediction using KNN Machine Learning Algorithm," in *Proceedings of the 2022 Int. Conf. on Computational Intelligence and Communication Networks (CICN)*, Mar. 15-17, 2022, Online [Online]. Available: IEEE Xplore. [Accessed: Mar. 9, 2023]. doi:10.1109/iccci54379.2022.9740781
- [8] G. Kemalbay and B. N. Alkış, "Borsa endeks hareket yönünün çoklu lojistik regresyon ve k-en yakın komşu algoritması ile tahmini," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 27, no. 4, pp. 556-569, 2021.
- [9] S. Madge and S. Bhatt, "Predicting Stock Price Direction using Support Vector Machines," *Independent Work Report*, vol. 45, 2015.
- [10] K. Khare, O. Darekar, P. Gupta, and V. Z. Attar, "Short term stock price prediction using deep learning," in *Proceedings of the 2nd IEEE Int. Conf. on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, May 19-20, 2017, Bangalore, India [Online]. Available: IEEE Xplore. [Accessed: Mar. 9, 2023].
- [11] S. Selvin, R. Vinayakumar, E. A. Gopalakrishnan, V. K. Menon, and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in *Proceedings of the Int. Conf. on Advances in Computing, Communications and Informatics (ICACCI)*, Sept. 13-16, 2017, Udipi, India [Online]. Available: IEEE Xplore. [Accessed: Mar. 9, 2023].
- [12] L. Ji, Y. Zou, K. He, and B. Zhu, "Carbon futures price forecasting based with ARIMA-CNN-LSTM model," *Procedia Computer Science*, vol. 162, pp. 33-38, 2019.
- [13] R. Alberto and C. Romero, "Generative Adversarial Network for Stock Market price Prediction," Unpublished report, Stanford University, San Francisco, USA, 2019.
- [14] K. Zhang, G. Zhong, J. Dong, S. Wang, and Y. Wang, "Stock Market Prediction Based on Generative Adversarial Network," *Procedia Computer Science*, vol. 147, pp. 400-406, 2019.
- [15] H. Raşo and M. Demirci, "Predicting the Turkish Stock Market BIST 30 Index using Deep Learning," *Uluslararası Mühendislik Araştırma ve Gelistirme Dergisi*, pp. 253-265, 2019. doi:10.29137/umagd.425560.
- [16] V. D. Ta, C. M. Liu, and D. A. Tadesse, "Portfolio Optimization-Based Stock Prediction Using Long-Short Term Memory Network in Quantitative Trading," *Applied Sciences*, vol. 10, no. 2, p. 437, 2020.
- [17] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-Based Model to Forecast Stock Prices," *Complexity*, vol. 2020, no. 1, pp. 1-10, 2020.
- [18] C. Yang, J. Zhai, and G. Tao, "Deep Learning for Price Movement Prediction Using Convolutional Neural Network and Long Short-Term Memory," *Mathematical Problems in Engineering*, pp. 1-13, 2020.
- [19] R. Chandra and Y. He, "Bayesian neural networks for stock price forecasting before and during COVID-19 pandemic," *Plos One*, vol. 16, no. 7, pp. 1-25, 2021.
- [20] U. Demirel, H. Çam, and R. Ünlü, "Predicting Stock Prices Using Machine Learning Methods and Deep Learning Algorithms: The Sample of the Istanbul Stock Exchange," *Gazi University Journal of Science*, vol. 34, no. 1, pp. 63-82, 2021. doi:10.35378/gujs.679103
- [21] N. Jing, Q. Liu, and H. Wang, "Stock price prediction based on stock price synchronicity and deep learning," *International Journal of Financial Engineering*, vol. 8, no. 2, 2021.
- [22] X. Yan, W. Weihan, and M. Chang, "Research on financial assets transaction prediction model based on LSTM neural network," *Neural Computing and Applications*, vol. 33, no. 1, pp. 257-270, 2021.
- [23] H. Lin, C. Chen, G. Huang, and A. Jafari, "Stock price prediction using Generative Adversarial Networks," *Journal of Computer Science*, vol. 17, no. 3, pp. 188-196, 2021.
- [24] T. T. H. Giang, T. T. Nguyen, and Q. T. Le, "Dynamic Sliding Window and Neighborhood LSTM-Based Model for Stock Price Prediction," *SN Computer Science*, vol. 3, no. 3, 2022.
- [25] F. Han, X. Ma, and J. Zhang, "Simulating Multi-Asset Classes Prices Using Wasserstein Generative Adversarial Network: A Study of Stocks, Futures and Cryptocurrency," *Journal of Risk and Financial Management*, vol. 15, no. 1, p. 26, 2022.
- [26] H. Liu, L. Qi, and M. Sun, "Short-Term Stock Price Prediction Based on CAE-LSTM Method," *Wireless Communications and Mobile Computing*, vol. 2022, no. 1, pp. 1-7, 2022.

- [27] E. Albayrak ve N. Saran, "İstatistiksel ve Derin Öğrenme Modellerini Kullanarak Hisse Senedi Fiyat Tahmini," *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, no. 16, vol. 2, pp. 161-169, 2023. doi:10.54525/tbbmd.1031017
- [28] Ç. Çoban ve E. Hayat, "Hisse Senedi Piyasası Analizinde Farklı Derin Sinir Ağı Modellerinin Karşılaştırılması," *Adnan Menderes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, no. 10, vol. 2, pp. 120-139, 2023. doi:10.30803/adusobed.1402228
- [29] İ. Dalkıran ve M. Ozan, "Derin Öğrenme Teknikleri Kullanılarak Borsadaki Hisse Değerlerinin Tahmin Edilmesi," *Avrupa Bilim ve Teknoloji Dergisi*, no. 39, oo. 143-148, 2022. doi:10.31590/ejosat.1145949
- [30] C. Erden, "Derin Öğrenme ve ARIMA Yöntemlerinin Tahmin Performanslarının Kıyaslanması: Bir Borsa İstanbul Hissesi Örneği," *Yönetim ve Ekonomi Dergisi*, no. 30, vol. 3, pp. 419-438, 2023. doi:10.18657/yonveek.1208807
- [31] R. Mo, "Review of neural network algorithm and its application in reactive distillation," *Asian Journal of Chemical Sciences*, vol. 9, no. 3, pp. 20-29, 2021. doi:10.9734/ajocs/2021/v9i319073
- [32] K. Saini, N. Kumar, B. Bhushan, and R. Kumar, "Nonlinear complex dynamic system identification based on a novel recurrent neural network," *Springer Nature*, vol. 1, no. 1, 2023. doi:10.21203/rs.3.rs-2894875/v1
- [33] M. Safdar, "Feasibility of wildlife conservation based on artificial neural network," *Nature Environmental Protection*, vol. 1, no. 3, 2020. doi:10.38007/nep.2020.010304
- [34] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.
- [35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [36] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.
- [37] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533-536, 1986.
- [38] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, and M. Wicke, "TensorFlow: A System for Large-Scale Machine Learning," in *Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, Nov. 2-4, 2016, Savannah, GA, USA [Online]. Available: USENIX. [Accessed: Mar. 10, 2023].
- [39] F. Chollet, "Keras," Available: <https://github.com/fchollet/keras> [Accessed: May 10, 2023].
- [40] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: An imperative style, high-performance deep learning library," *Advances in Neural Information Processing Systems*, vol. 8026-8037, 2019.
- [41] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM Int. Conf. on Multimedia*, Nov. 3-7, 2014, Orlando, FL, USA, Eds. ACM, 2014. pp. 675-678.
- [42] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [43] S. Borovkova and I. Tsiamas, "An ensemble of LSTM neural networks for high-frequency stock market classification," *Journal of Forecasting*, vol. 38, no. 6, pp. 600-619, 2019.
- [44] Z. Jin, Y. Yang, and Y. Liu, "Stock closing price prediction based on sentiment analysis and LSTM," *Neural Computing and Applications*, vol. 32, no. 13, pp. 9713-9729, 2020.
- [45] A. Amidi and S. Amidi, "Recurrent Neural Networks cheatsheet," Available: <https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks> [Accessed: March 9, 2023].
- [46] K. Cho, V. Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *ArXiv.org*, 2014.
- [47] M. A. Kızrak and B. Bolat, "Uçak Motoru Sağlığı için Uzun-Kısa Süreli Bellek Yöntemi ile Öngörücü Bakım," *Bilişim Teknolojileri Dergisi*, pp. 103-109, 2019.
- [48] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," *ArXiv.org*, 2014.
- [49] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," *ArXiv.org*, 2014.
- [50] L. Lan, L. You, Z. Zhang, Z. Fan, W. Zhao, N. Zeng, Y. Chen, and X. Zhou, "Generative Adversarial Networks and Its Applications in Biomedical Informatics," *Frontiers in Public Health*, vol. 8, p. 164, 2020.
- [51] A. You, J. K. Kim, I. H. Ryu, and T. K. Yoo, "Application of generative adversarial networks (GAN) for ophthalmology image domains: a survey," *Eye and Vision*, vol. 9, no. 1, 2022.

- [52] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F. Y. Wang, "Generative adversarial networks: introduction and outlook," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 4, pp. 588-598, 2017.
- [53] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved Techniques for Training GANs," *ArXiv.org*, 2016.
- [54] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," *ArXiv.org*, 2017.
- [55] Alphabet Inc. "Alphabet Inc. (GOOG) Stock Price, News, Quote and History - Yahoo Finance," Available: <https://finance.yahoo.com/quote/GOOG/> [Accessed: March 9, 2023].
- [56] B. Banushev, "Using the latest advancements in AI to predict stock market movements," Available: <https://github.com/borisbanushev/stockpredictionai> [Accessed: March 9, 2023].
- [57] R. Dash, P. K. Dash, and R. Bisoi, "A self-adaptive differential harmony search based optimized extreme learning machine for financial time series prediction," *Swarm and Evolutionary Computation*, vol. 19, pp. 25-42, 2014.
- [58] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *European Journal of Operational Research*, vol. 270, no. 2, pp. 654-669, 2018.
- [59] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in Neural Information Processing Systems*, pp. 3104-3112, 2014.
- [60] J. Sirignano and R. Cont, "Universal features of price formation in financial markets: perspectives from Deep Learning," *Quantitative Finance*, vol. 19, no. 8, pp. 1449-1459, 2018.
- [61] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451-2471, 2002.
- [62] J. Lv, C. Wang, W. Gao, and Q. Zhao, "An Economic Forecasting Method Based on the LightGBM-Optimized LSTM and Time-Series Model," *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1-10, 2021.
- [63] F. J. Huang, W. Zhang, and T. Yoshida, "A hybrid model for stock price prediction," *Quantitative Finance*, vol. 18, no. 12, pp. 2001-2016, 2018.
- [64] W. Zhang and T. Yoshida, "Stock price prediction based on generative adversarial network," *Procedia Computer Science*, vol. 169, pp. 101-108, 2020.
- [65] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," in *Advances in Neural Information Processing Systems*, vol. 30, pp. 5767-5777, 2017.
- [66] J. Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE Int. Conf. on Computer Vision*, Oct. 22-29, 2017, Venice, Italy [Online]. Available: IEEE Xplore. [Accessed: Mar. 9, 2023].
- [67] J. Xu, J. He, J. Gu, H. Wu, L. Wang, Y. Zhu, T. Wang, X. He, and Z. Zhou, "Financial Time Series Prediction Based on XGBoost and Generative Adversarial Networks," *International Journal of Circuits, Systems and Signal Processing*, vol. 16, pp. 637-645, 2022.

This is an open access article under the CC-BY license

