

Bellek Tabanlı LSTM ve GRU Makine Öğrenmesi Algoritmaları Kullanarak BIST100 Endeks Tahmini

Yusuf ÇELİK^{1*}

¹ Bilgisayar Mühendisliği, Mühendislik Fakültesi, Munzur Üniversitesi, Tunceli, Türkiye

*¹ celikyusuf@munzur.edu.tr

(Geliş/Received: 19/12/2023;

Kabul/Accepted: 13/05/2024)

Öz: Makine öğrenmesi tabanlı tahmin yaklaşımlarının finansal piyasalarda geliştirilmesi, hızlı ve hassas karar alma, karmaşıklıkla başa çıkma, risk yönetimi, algoritmik ticaret ve duygusal etkilerin azaltılması gibi avantajlar sağlar. Bu yaklaşımlar, sürekli öğrenme ve adaptasyon yetenekleriyle finansal başarı için rekabet avantajı oluşturabilir. Bu makale çalışmada, Borsa İstanbul (BIST) 100 endeks tahmini için bellek tabanlı makine öğrenmesi modellerine dayalı bir yaklaşım sunulmuştur. Bu amaçla, ardışık veri değerlendirmesinde popüler olan uzun kısa-sürelili bellek (LSTM) ve geçitli tekrarlayan birim (GRU) mimarileri kullanılmıştır. Elde edilen model çıktılarına göre bu modellerin, eğitim ve doğrulama aşamalarında düşük kayıplar gösterdiği ve BIST100 endeksinin genel eğilimlerini başarıyla takip ettiği gözlemlenmiştir. Ancak, modeller piyasa dalgalanmaları ve ani değişimlerde gerçek değerlerden sapmalar göstermiş, bu da belirsizlikleri ve genelleme kapasitelerinin sınırlarını ortaya koymuştur. Geleceğe yönelik tahminler, eğitim veri setindeki desenlere dayanarak yapılmış ancak zamanla artan belirsizlik göstermiştir. Çalışma, makine öğrenmesi algoritmalarının finans verileri üzerindeki kullanım alanı konusunda önemli bilgiler sağlayacak potansiyele sahiptir.

Anahtar kelimeler: Makine öğrenmesi, LSTM, GRU, BIST100, finans tahmin.

Predicting the BIST100 Index Using Memory-Based LSTM and GRU Machine Learning Algorithms

Abstract: The development of machine learning-based forecasting approaches in financial markets offers advantages such as rapid and precise decision-making, complexity management, risk mitigation, algorithmic trading, and reduction of emotional biases. These methods can create a competitive edge for financial success through their continuous learning and adaptation capabilities. This article presents a memory-based machine learning approach for predicting the Borsa İstanbul (BIST) 100 index, utilizing popular architectures like Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) for sequential data analysis. The models demonstrated low losses in training and validation phases and successfully followed the general trends of the BIST 100 index. However, they showed deviations from actual values during market volatility, highlighting uncertainties and the limits of their generalization abilities. Predictions were based on patterns in the training dataset but indicated increasing uncertainty over time. This study has the potential to provide significant insights into the application of machine learning algorithms on financial data.

Key words: Machine learning, LSTM, GRU, BIST100, financial forecasting.

1. Giriş

Geleneksel makine öğrenme algoritmaları, spesifik ve sınırlı problemleri çözmekte etkili olsa da, karmaşık ve değişken veri yapılarıyla başa çıkmak konusunda sıklıkla zorlanırlar. Bu durum, özellikle sürekli değişen ve çok boyutlu veri setlerini işlemek gerektiğinde daha belirgin hale gelir. Bu noktada, derin öğrenme devreye girer. Yapay sinir ağları modeli üzerine kurulu olan derin öğrenme, verilerin içerdiği karmaşıklığı ve çeşitliliği anlama ve bu bilgileri öğrenme sürecine dahil etme yeteneği sayesinde çok daha geniş bir uygulama alanına sahiptir [1,2]. Derin öğrenme, günümüzde birçok farklı alanda etkin bir şekilde kullanılmaktadır [3-5]. Derin öğrenme yöntemleri zaman serisi problemlerinin analizinde de kullanılmaktadır. Zaman serisi problemleri, geçmiş verilerin zaman içindeki değişimlerini analiz ederek gelecekteki eğilimleri ve desenleri tahmin etmeyi amaçlar. Bu problemlerde kullanılan yaygın yöntemler arasında LSTM ve GRU bulunur. Bu yöntemler, veri setlerindeki uzun vadeli bağımlılıkları ve karmaşıklıkları modellemekte oldukça etkilidir [6]. LSTM ve GRU doğal dil işleme [7], ses tanıma [8] gibi birçok alanda kullanılmaktadır.

Finans sektöründe, hisse senetleri, döviz kurları ve piyasa trendleri gibi finansal veriler de zaman serisi verileridir. Finans sektöründe, zaman serisi verileri, piyasa trendleri ve ekonomik göstergeler gibi karmaşık veri setlerinin analizi için derin öğrenme yöntemleri sıklıkla kullanılmaktadır [9-11]. Coa ve ark, LSTM modeli,

* Sorumlu yazar: celikyusuf@munzur.edu.tr. Yazarların ORCID Numarası: ¹ 0000-0002-7859-7543

CEEMDAN sinyal ayrıştırma algoritması ile birleştirilerek finansal zaman serileri tahmin modeli (CEEMDAN-LSTM) olarak kullanılmıştır. Çalışmalarında büyük borsa endekslerinin (S&P500, HSI, DAX, SSE) günlük kapanış fiyatlarını tahmin etmek için modellerini kullanmışlardır. Makaledeki modelin performansını doğrulamak için LSTM, SVM, CEEMDAN-SVM, CEEMDAN-MLP ve CEEMDAN-LSTM dahil olmak üzere birkaç modelin tahmin etkisi karşılaştırılmıştır [12]. Siami-Namini ve ark. zaman serisi verilerini tahmin etmek için derin öğrenme tabanlı yeni algoritmaların, özellikle LSTM'nin, geleneksel algoritmalara kıyasla önemli ölçüde daha üstün olduğu vurgulanıyor. LSTM, ARIMA modeline göre hata oranlarını %84 ile %87 oranında azaltarak, zaman serisi tahminindeki doğruluk ve hassasiyet konusunda büyük bir gelişme sağlamıştır [13]. Pirani ve ark. yaptığı çalışmada hisse senedi geçmiş verilerinden oluşturulan bir veri seti kullanarak ARIMA, tek yönlü LSTM, çift yönlü LSTM ve GRU modellerinin performanslarını ve eğitim davranışlarını analiz etmiştir. GRU modeli daha hızlı çalışmıştır. Ayrıca tahmin doğruluk oranında da GRU daha iyi performans göstermiştir [14].

Bu çalışmada, BIST100 endeks tahmini için hem LSTM hem de GRU algoritmalarına dayalı modeller kullanılmıştır. Finans verileri tarihsel olarak geniş kapsamlı bir şekilde çeşitli kaynaklardan elde edilmiş ve model girişlerine uygun bir yapıya dönüştürülmüştür. Tahmin aşamasında günlük kapanış bilgileri üzerinde bellek tabanlı modellerin uygulanması gerçekleştirilmiştir. Böylece sonraki kapanış bilgilerinin modeller tarafından otomatik tahmin edilmesi sağlanmıştır.

2. Bellek Tabanlı Makine Öğrenme Modelleri

Bellek tabanlı makine öğrenme modelleri, özellikle doğal dil işleme (NLP) alanında kullanılan bir tür makine öğrenme modelidir. Bu modeller, büyük miktarda veriyi öğrenmek ve anlamak için geliştirilen derin öğrenme algoritmalarını temel alır. Bellek tabanlı modeller, uzun vadeli bağımlılıkları ele almak için özel olarak tasarlanmış bir dizi mekanizmayı içerir. Bu mekanizma, geçmişteki bilgileri saklama ve hatırlama yeteneğiyle öne çıkar. LSTM [15] ve GRU [16] gibi mekanizmalar, bu tür modellerde sıkça kullanılan yapılarıdır.

2.1. Uzun kısa süreli hafıza (LSTM)

LSTM, Uzun Kısa Süreli Bellek (Long Short-Term Memory) ifadesinin kısaltmasıdır ve derin öğrenme alanında, özellikle zaman serisi verileri ile çalışırken kullanılan bir yapay sinir ağı türüdür. LSTM'ler, tekrarlayan sinir ağları (RNN) sınıfına aittir ve özellikle uzun süreli bağımlılıkları modellemek için tasarlanmıştır. Standart RNN'ler, "vanishing gradient" (kaybolan gradyan) sorunu nedeniyle uzun süreli bağlantıları öğrenmekte zorlanırken, LSTM'ler bu sorunu, hücre durumu ve kapılar adı verilen özel yapılar aracılığıyla aşar. Bu kapılar, ağın bilgiyi ne kadar süre boyunca saklayacağını veya unutacağını kontrol eder, böylece önemli bilgilerin korunmasını ve önemsiz bilgilerin atılmasını sağlar. Bu özellikleri sayesinde, LSTM'ler, dil modelleme, konuşma tanıma, el yazısı tanıma ve karmaşık zaman serisi tahminleri gibi pek çok alanda başarıyla kullanılmaktadır [7,8,17-19]. LSTM'nin ana prensibi, hücre durumunun korunması ve kontrol edilmesidir. Bu durum, yalnızca özel kapılar aracılığıyla bilgi eklenebilir veya çıkarılabilir şekilde tasarlanmıştır. Bu kapılar, dış etkilerden bağımsız olarak hücre durumunun sabit kalmasını sağlar. Bir LSTM ünitesi genellikle üç ana kapı içerir: unut kapısı, giriş kapısı ve çıkış kapısı. Bu kapılar, hücre içindeki bilgi akışını düzenleyerek LSTM'nin zamanla ilgili veri öğrenme yeteneğini artırır [20].

Giriş kapısı, mevcut girdiden hangi bilgilerin hücre durumuna eklenmesi gerektiğini belirler. Denklem 1 ve 2'de giriş kapısına ait matematiksel gösterimler sunulmuştur.

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

$$\tilde{C}_t = \tanh(x_t U^g + h_{t-1} W^g) \quad (2)$$

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

Unut kapısı, geçmiş hücre durumu C_{t-1} 'den hangi bilgilerin unutulacağını belirler. σ sigmoid fonksiyonudur ve 0 ile 1 arasında değerler alır. x_t mevcut girdidir. h_{t-1} ise önceki hücre çıktısıdır. U ve W ağırlık matrisleridir. Unut kapısına ait hesaplamalar Denklem 3'de verilmiştir.

$$f_t = \sigma(x_t U^f + h_{t-1} W^f) \quad (3)$$

Çıkış kapısı, sonraki katmana ve hücrenin kendisine gönderilecek çıktıyı belirleyen kapıdır. Denklem 4 ve 5'de çıkış kapısına ait matematiksel gösterimler sunulmuştur.

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \quad (4)$$

$$h_t = \tanh(C_t) * o_t \quad (5)$$

Bu denklemlerde, o_t çıkış kapısının aktivasyonu, sigmoid fonksiyonu σ ile hesaplanır. Bu, kapının ne kadar "açık" olduğunu belirler ve böylece çıktının ne kadarının geçeceğini düzenler. Denklemde yer alan h_t elde edilen hücre çıktısıdır ve bu değer sonraki katmana iletilir. Hücre durumu, uzun süreli hafızayı temsil eder ve geçmiş durum ile mevcut girdinin etkisine göre güncellenir. Denklem 6'da hücre durumunun nasıl güncellendiğini gösteren hesaplamalar verilmiştir.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

Bu denklemde yer alan f_t unutmaya kapısının çıktısıdır ve geçmiş hücre durumunun hangi kısmının korunacağını belirler. i_t giriş kapısının çıktısıdır ve mevcut girdinin hücre durumuna hangi kısmının eklenmesi gerektiğini belirler. \tilde{C}_t mevcut girdinin hücre durumuna potansiyel katkısını temsil eden ve tanh fonksiyonu ile normalleştirilen değerdir [21].

2.2. Geçitli tekrarlayan birim (GRU)

Geçitli Tekrarlayan Birim (GRU), Chung ve arkadaşları tarafından 2014 yılında önerilmiştir [22]. GRU, LSTM'nin kapı yapısını basitleştirerek iki kapılı bir model sunar. Bu kapılar güncelleme ve sıfırlama kapılarıdır [23]. GRU, her bir tekrarlayan birimin farklı zaman ölçeklerine bağlı olarak adaptif bir şekilde bağımlılıkları yakalamasını sağlamak amacıyla tasarlanmıştır. LSTM birimine benzer şekilde, GRU'nun da birim içindeki bilgi akışını modüle eden kapı birimleri bulunmaktadır, ancak ayrı hafıza hücrelerine sahip değildir [24]. GRU da LSTM gibi zaman serisi problemlerinde sıklıkla kullanılmaktadır.

$$z_t = \sigma(x_t U^z + h_{t-1} W^z) \quad (7)$$

Denklem 7, güncelleme kapısının fonksiyonu göstermektedir. z_t değeri ne kadar yüksekse, o kadar fazla önceki bilgi korunur ve yeni bilgiyle o kadar az güncelleme yapılır. Tersine de geçerlidir: z_t değeri ne kadar düşükse, o kadar fazla yeni bilgi eklenir ve önceki bilgi o kadar az korunur. Bu işlem, GRU'nun her adımda neyi saklayıp neyi unutacağına karar vermesine olanak tanır ve bu sayede karmaşık veri dizilerini etkili bir şekilde işleyebilir.

Denklem 8, sıfırlama kapısının aktivasyonunu ifade eder. Eğer sıfırlama kapısının aktivasyonu düşükse, geçmiş bilgi büyük ölçüde yok sayılır ve model mevcut girdiye daha fazla odaklanır. Eğer aktivasyon yüksekse, önceki gizli durum yeni girdiyle birleştirilir ve bu, modelin önceki bilgiyi daha fazla korumasını sağlar.

$$r_t = \sigma(x_t U^r + h_{t-1} W^r) \quad (8)$$

r_t sıfırlama kapısının aktivasyonunu ifade eder ve 0 ile 1 arasında bir değer alır. σ , sigmoid fonksiyonudur ve her türlü sayıyı 0 ile 1 arasında bir değere normalleştirir. x_t , mevcut zaman adımındaki girdidir. h_{t-1} önceki zaman adımındaki gizli durumdur. U ile W mevcut girdi ve önceki gizli duruma uygulanan ağırlık matrisleridir ve model eğitimi sırasında öğrenilir.

$$\tilde{h}_t = \tanh(x_t U^h + (r_t * h_{t-1}) W^h) \quad (9)$$

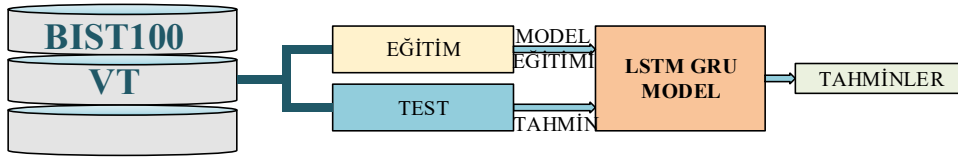
Denklem 9, yeni gizli durumun adayını hesaplama işlemini gösterir. GRU'da yeni gizli durumun adayını, mevcut girdi ve önceki gizli durumun bir kombinasyonu ile hesaplanır. \tilde{h}_t yeni gizli durumun adayını ifade eder. Bu işlem, sıfırlama kapısının etkisi altında gerçekleşir. Yeni gizli durum adayını, girdi ve önceki gizli durum arasındaki ilişkinin ne kadarının korunacağını belirleyen bir bileşendir.

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (10)$$

Denklem 10, yeni gizli durumun nasıl hesaplandığını açıklar. h_t mevcut zaman adımıdaki yeni gizli durumu ifade eder. z_t güncelleme kapısının aktivasyonunu ifade eder ve önceki gizli durumun yeni gizli durumla ne kadar güncelleneceğini belirler. $(1 - z_t)$ önceki gizli durumun korunma miktarını ifade eder. \tilde{h}_t yeni gizli durum adayını ifade eder. Yeni gizli durum, güncelleme kapısının aktivasyonuna göre önceki gizli durum ile yeni gizli durum adayının birleşiminden oluşur. Bu, GRU hücrenin zaman serisi veri akışında bilgiyi nasıl güncelleyeceğini ve saklayacağını belirler [21]. GRU, mevcut durum ile yeni hesaplanan durum arasında lineer bir toplam alınması süreci açısından LSTM birimine benzerdir. Ancak, GRU'nun durumunun ne dereceye kadar açığa çıkarılacağını kontrol eden bir mekanizması yoktur ve her seferinde tüm durumu açığa çıkarır [24].

3. Materyal ve Metot

BIST100 endeksinin geçmiş günlük kapanış değerleri bir veri tabanında toplanmıştır. Bu veri tabanı, farklı tahmin modellerinin uygulanması için temel oluşturmuştur. Söz konusu modeller, piyasanın geçmişteki hareketlerini analiz ederek öğrenme modeli oluşturulmuş daha sonra test verileri üzerinde tahminler elde edilmiştir. Şekil 1'de, BIST100 endeks değerleri tahmininin blok gösterimi sunulmaktadır. Bu gösterimde, veri tabanında (VT) toplanan BIST100 verilerinin eğitim ve test setlerine ayrılması ve bu verilerin LSTM ve GRU modellerine uygulanarak tahminlerin elde edilmesi süreci yer almaktadır.



Şekil 1. BIST100 endeks değerleri tahmini blok gösterimi.

3.1. BIST100 veri seti

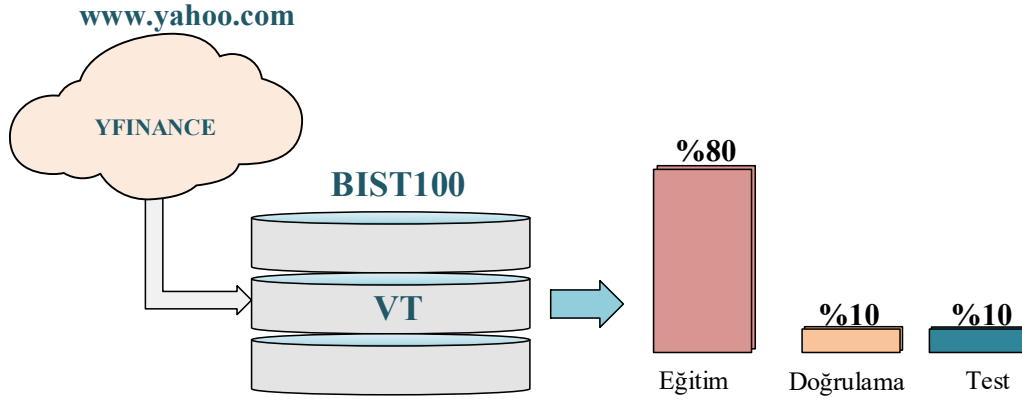
Bu çalışmada, analiz için Borsa İstanbul verileri temel alınmıştır. BIST100 endeksine ait geçmiş veriler, veri setinin oluşturulmasında kullanılmıştır. Veri seti, Yahoo Finance web servisi tarafından finansal verilerin elde edilmesi için geliştirilen yfinance kütüphanesi aracılığıyla derlenmiştir. BIST100'e ait tüm geçmiş veriler, 1 Temmuz 1997'den 1 Aralık 2023'e kadar olan süreyi kapsayacak şekilde toplanmıştır. yfinance kütüphanesi bu tarihten itibaren BIST100 için veri sağlamaktadır.

Tablo 1. yfinance kütüphanesi aracılığıyla çekilen BIST100 geçmiş verilerine ait örnekler.

Tarih	Açılış	Yüksek	Düşük	Kapanış	Düzeltilmiş K.	Hacim
1997-07-01	1857.0	1864.0	1826.0	1851.0	1851.0	2342600
1997-07-02	1851.0	1877.0	1849.0	1873.0	1873.0	2946200
...
2023-11-29	8115.39	8125.10	7993.10	8009.29	8009.29	2351023100
2023-11-30	8007.79	8049.29	7912.0	7948.60	7948.60	2475424100

Tablo 1, Borsa İstanbul BIST 100 endeksinin yfinance kütüphanesi kullanılarak elde edilen günlük verilerini içermektedir. Bu veriler 1 Temmuz 1997'den itibaren başlamaktadır. Çalışmada BIST 100'ün günlük kapanış fiyatları (close) kullanılmıştır.

13 Haziran 2017'de BIST 100 endeksi 100.000 puan seviyesine ulaşmıştır, 27 Haziran 2020'de endeksin değerlerini uluslararası borsalarla uyumlu hale getirmek amacıyla iki sıfır atılmıştır [25]. Bu değişiklikten dolayı, veri setindeki uyumluluğu sağlamak için 27 Haziran 2020'ye kadar olan veriler 100'e bölünmüştür. Bu işlem, tüm veri setinin birbiriyle uyumlu olmasını sağlamaktadır.



Şekil 2. Veri setinin dağılımı: eğitim için %80 doğrulama için %10 ve test için %10 kullanıldı.

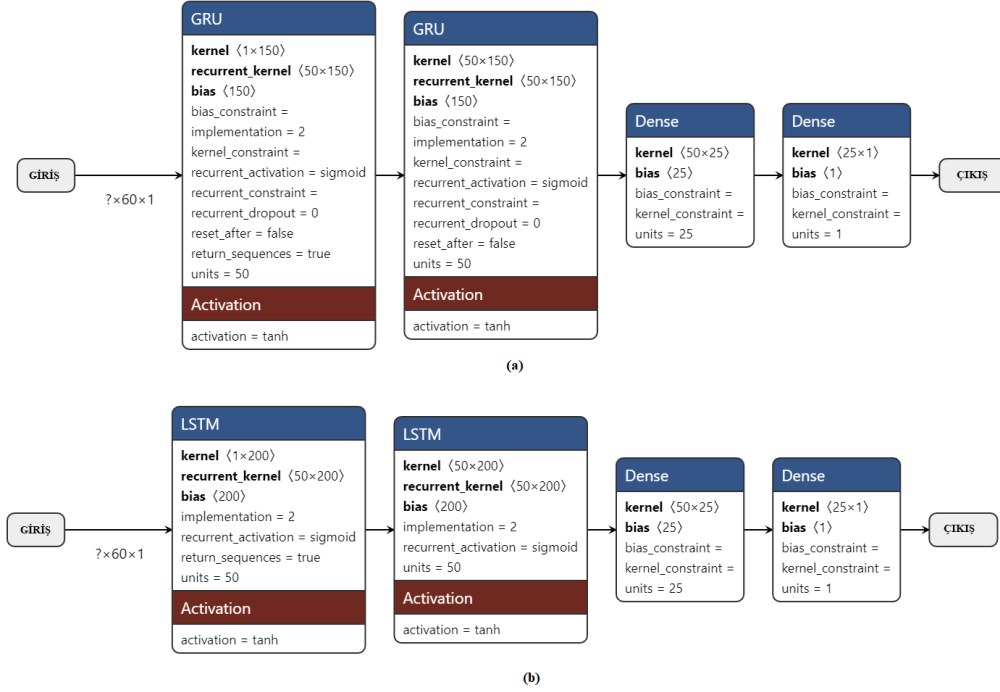
Çalışmadaki toplam veri seti 6600 kayıttan oluşmaktadır. Şekil 2'de gösterildiği üzere, bu verilerin %80'i (5280 kayıt) eğitim, %10'u (660 kayıt) doğrulama ve %10'u (660 kayıt) test için ayrılmıştır. Test seti için ayrıca, tahminlerin daha doğru yapılabilmesi için önceki 60 günlük veriler de dahil edilmiştir. Bu eklemelerle birlikte, test için kullanılan toplam veri sayısı 720'ye ulaşmıştır. Bu dağılım, veri setinin etkin bir şekilde eğitim, doğrulama ve test amaçları için kullanılmasını sağlamaktadır.

3.2. LSTM ve GRU modelleri

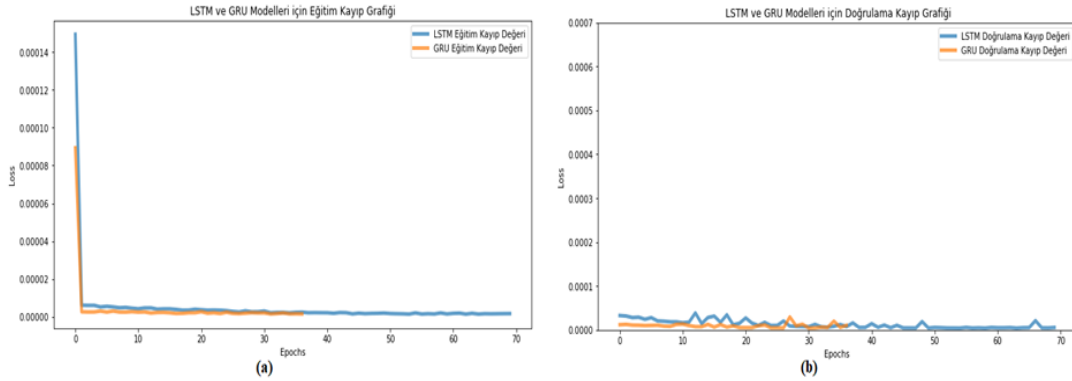
Bu çalışmada, BIST 100 endeksi değerlerinin tahmini için derin öğrenme tabanlı iki farklı tekrarlayan sinir ağı modeli incelenmiştir: LSTM ve GRU. Her iki model de zaman serisi verilerinin özelliklerini yakalayabilme ve uzun süreli bağımlılıkları öğrenebilme kapasiteleri ile seçilmiştir. LSTM modeli, ardışık veri noktaları arasındaki uzun vadeli ilişkileri tanımlamak için tasarlanmış iki katmanlı bir yapı kullanır. Her bir katman, 60 günlük geçmiş verileri temel alarak gelecekteki değerleri tahmin etme üzere eğitilmiştir. GRU modeli ise, benzer şekilde iki katmanlı bir yapıya sahiptir. Her iki modelin ağ mimarisi, BIST 100 endeksi verilerinin zaman içindeki karmaşık desenlerini modelleyebilmek ve tahmin performansını maksimize etme üzere optimize edilmiştir.

Şekil 3'te BIST 100 endeksi değer tahminleri için kullanılan derin öğrenme modellerinin ağ mimarileri. Şekilde (a) iki GRU katmanı barındıran GRU modeli; (b) iki LSTM katmanı içeren LSTM modeli gösterilmektedir. Her iki model de 60 günlük zaman dilimindeki verileri işleyerek gelecekteki endeks değerlerini tahmin etme üzere eğitilmiştir. Bu süreçte, 60 günlük ardışık veri dizilerini alır ve bu verileri temel alarak 61. gün için BIST 100 endeksinin kapanış değerini tahmin eder. Model eğitimleri sırasında, hem LSTM hem de GRU mimarileri için maksimum 200 epoch belirlenmiştir. Erken durdurma mekanizması (early stopping) devreye alınmış ve doğrulama kaybında iyileşme olmaması durumunda eğitimin durdurulması için sabır (patience) parametresi 6 epoch olarak ayarlanmıştır. Sonuç olarak, GRU modeli 37 epoch ve LSTM modeli ise 70 epoch sonrasında daha fazla iyileşme göstermeyince erken durdurma kriteri tarafından eğitim sonlandırılmıştır.

Şekil 4, LSTM ve GRU modellerinin eğitim sürecindeki performansını kayıp değerleri üzerinden karşılaştırmalı olarak sunmaktadır. LSTM ve GRU modellerinin eğitim ve doğrulama kayıp değerlerinin epoch boyunca nasıl değiştiğini gösteren iki grafik bulunmaktadır. Şekil 4 (a)'da, her iki modelin eğitim kaybının epoch ilerledikçe benzer bir düşüş trendi gösterdiği ve son derece düşük kayıp değerlerine ulaştığı görülmektedir. Bu durum, modellerin eğitim veri setinden öğrenme sürecinin başarılı bir şekilde gerçekleştiğini ve kayıp değerlerinin minimum seviyelere indirildiğini gösterir. Şekil 4 (b)'de, her iki modelin doğrulama seti üzerinde benzer performans sergilediğini ve eğitim süresince istikrarlı bir doğrulama kaybı sergilediğini vurgulamaktadır, bu da öğrenmenin gerçekleştiğini göstermektedir.

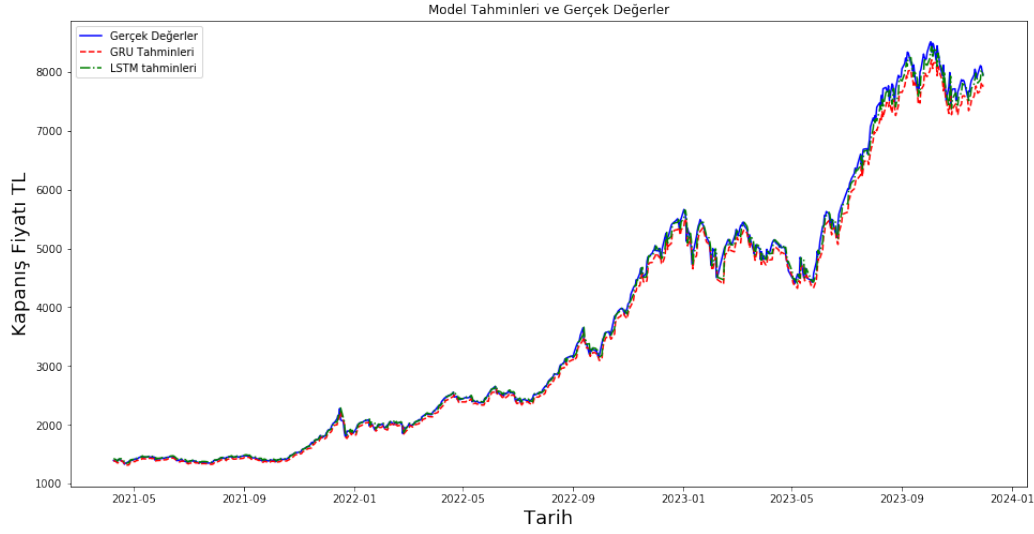


Şekil 3. Tahmin modeli olarak kullanılan ağların mimarisi (a) GRU modeli (b) LSTM modeli.



Şekil 4. LSTM ve GRU modelleri için (a) Eğitim kayıp grafiği (b) Doğrulama kayıp grafiği.

Şekil 5 BIST100 için test veri seti ile LSTM ve GRU modellerin tahminlerini göstermektedir. Mavi renk gerçek değerleri GRU model tahminleri kırmızı renkte kesik çizgili ve LSTM model tahminleri mavi renkte ve noktalı çizgili olarak gösterilmiştir. Grafik, gerçek verilerle model tahminleri arasında yakın uyum olduğunu gösteriyor; her iki modelin tahmin çizgisi de gerçek değerlerin çizgisini genel olarak iyi takip ediyor. Her iki model de veri setinin genel eğilimini ve dalgalanmalarını başarıyla yakalamış. Ancak, özellikle ani değişimlerde, model tahminlerinin gerçek değerlerden biraz sapma gösterdiği noktalar var gibi görünüyor. Bu, modellerin bazı özel durumlarda gerçek değerlerin hassasiyetini tam olarak tahmin edemediğini gösterir.



Şekil 5. Gerçek BIST100 değerleri ile LSTM, GRU tahminleri.

Tablo 2. 30 günlük BIST100 günlük değerleri ile LSTM/GRU Tahminleri Karşılaştırması.

Tarih	Gerçek Değer	LSTM Tahmin	GRU Tahmin
2023-10-20	7510.700195	7669.972168	7565.794922
2023-10-23	7750.399902	7419.475098	7357.185059
2023-10-24	7990.399902	7643.312988	7474.110840
2023-10-25	7425.000000	7930.264160	7698.537109
2023-10-26	7662.100098	7393.363770	7354.715820
2023-10-27	7706.399902	7539.312012	7402.899902
2023-10-30	7712.899902	7630.809570	7470.582520
2023-10-31	7514.100098	7647.412598	7495.452148
2023-11-01	7528.600098	7445.541016	7355.254395
2023-11-02	7663.600098	7433.399902	7319.752930
2023-11-03	7706.000000	7580.813965	7415.254395
2023-11-06	7868.000000	7646.384277	7479.120605
2023-11-07	7844.799805	7804.164062	7609.477539
2023-11-08	7860.899902	7788.689453	7627.725098
2023-11-09	7841.500000	7785.067871	7634.780273
2023-11-10	7771.299805	7759.842773	7620.055664
2023-11-13	7568.299805	7686.708984	7563.953613
2023-11-14	7671.100098	7481.522461	7401.311523
2023-11-15	7665.899902	7569.920410	7429.474121
2023-11-16	7780.700195	7596.593262	7446.645508
2023-11-17	7853.399902	7714.648438	7534.709961
2023-11-20	7946.399902	7797.120117	7613.264648
2023-11-21	8046.600098	7885.582520	7697.936523
2023-11-22	7987.000000	7980.756348	7789.391113
2023-11-23	7898.600098	7918.560059	7767.724121
2023-11-24	7960.000000	7810.809570	7689.947754
2023-11-27	8107.200195	7860.246094	7710.458984
2023-11-28	8093.500000	8021.414551	7828.037109
2023-11-29	8009.299805	8026.615723	7854.217773
2023-11-30	7948.600098	7931.341309	7793.922363

Tablo 2, test sürecinde ele alınan son 30 günlük veri setine ait gerçek değerleri ve bu veriler üzerinde yapılan LSTM ve GRU modellerinin tahmin sonuçlarını içermektedir.

4. SONUÇ

Bu çalışmada, BIST 100 endeksi üzerinde LSTM ve GRU modelleri kullanılarak yapılan tahminlerin performansı değerlendirilmiştir. Analiz sonuçları, her iki modelin de eğitim ve doğrulama aşamalarında düşük kayıp değerlerine ulaştığını ve gerçek zaman serisi verilerinin genel eğilimlerini yakalayabildiğini göstermektedir. Özellikle, modellerin endeks değerlerinin yükseliş ve düşüşlerini büyük ölçüde takip edebildiği ve uzun vadeli eğilimleri başarıyla modelleyebildiği gözlemlenmiştir. Ancak, tahminlerin gerçek değerlerle tam bir uyum içinde olmadığı ve özellikle piyasa dalgalanmalarının yoğun olduğu dönemlerde gerçek verilerden sapmalar olduğu belirlenmiştir. Bu sapmalar, modellerin ani piyasa hareketlerine karşı duyarlılığının sınırlı olabileceğine işaret etmektedir. Geleceğe yönelik yapılan tahminler, modelin eğitim veri setinde öğrendiği desenlerden yola çıkarak yapılmasına rağmen, tahminlerin zamanla artan bir belirsizlik gösterdiği tespit edilmiştir. Bu durum, modelin gerçek dünya koşullarında karşılaşılabileceği yeni veri noktalarına adaptasyonunun ve genelleme kapasitesinin sınırlarını ortaya koymaktadır.

Kaynaklar

- [1] Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press, 2016.
- [2] Sze V, Chen YH, Yang TJ, Emer JS. Efficient processing of deep neural networks: A tutorial and survey. Proc IEEE, 2017; 105(12): 2295-2329.
- [3] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med, 2020; 121: 103792.
- [4] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proc IEEE Conf Comput Vis Pattern Recognit, 2016; pp. 770-778.
- [5] Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst, 2015; 28.
- [6] Huang J, Chai J, Cho S. Deep learning in finance and banking: A literature review and classification. Front Bus Res China, 2020; 14(1): 1-24.
- [7] Singh V, Chen SS, Singhania M, Nanavati B, Gupta A. How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries—A review and research agenda. Int J Inf Manag Data Insights, 2022; 2(2): 100094.
- [8] Ahmed S, Alshater MM, El Ammari A, Hammami H. Artificial intelligence and machine learning in finance: A bibliometric review. Res Int Bus Finance, 2022; 61: 101646.
- [9] Gamboa JCB. Deep learning for time-series analysis, 2017; arXiv preprint arXiv:1701.01887.
- [10] Sutskever I, Vinyals O, Le QV. Sequence to sequence learning with neural networks. Adv Neural Inf Process Syst, 2014; 27.
- [11] Amodei D, Ananthanarayanan S, Anubhai R, Bai J, Battenberg E, Case C, Zhu Z. Deep speech 2: End-to-end speech recognition in English and Mandarin. In: Int Conf Mach Learn, 2016; pp. 173-182. PMLR.
- [12] Cao J, Li Z, Li J. Financial time series forecasting model based on CEEMDAN and LSTM. Physica A, 2019; 519: 127-139.
- [13] Siami-Namini S, Namin AS. Forecasting economics and financial time series: ARIMA vs. LSTM, 2018; arXiv preprint arXiv:1803.06386.
- [14] Pirani M, Thakkar P, Jivrani P, Bohara MH, Garg D. A comparative analysis of ARIMA, GRU, LSTM and BiLSTM on financial time series forecasting. In: 2022 IEEE Int Conf Distributed Comput Electr Circuits Electron (ICDCECE), 2022; pp. 1-6. IEEE.
- [15] Lindemann B, Maschler B, Sahlab N, Weyrich M. A survey on anomaly detection for technical systems using LSTM networks. Comput Ind, 2021; 131: 103498.
- [16] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014; arXiv preprint arXiv:1412.3555.
- [17] Dey R, Salem FM. Gate-variants of gated recurrent unit (GRU) neural networks. In: 2017 IEEE 60th Int Midwest Symp Circuits Syst (MWSCAS), 2017; pp. 1597-1600. IEEE.
- [18] Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014; arXiv preprint arXiv:1412.3555.
- [19] Yu Y, Si X, Hu C, Zhang J. A review of recurrent neural networks: LSTM cells and network architectures. Neural Comput, 2019; 31(7): 1235-1270.
- [20] Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput, 1997; 9(8): 1735-1780.
- [21] Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation, 2014; arXiv preprint arXiv:1406.1078.

- [22] Merity S, Keskar NS, Socher R. Regularizing and optimizing LSTM language models, 2017; arXiv preprint arXiv:1708.02182.
- [23] Graves A, Jaitly N, Mohamed AR. Hybrid speech recognition with deep bidirectional LSTM. In: 2013 IEEE Workshop Autom Speech Recognit Understanding, 2013; pp. 273-278. IEEE.
- [24] Carbune V, Gonnet P, Deselaers T, Rowley HA, Daryin A, Calvo M, Gervais P. Fast multi-language LSTM-based online handwriting recognition. Int J Document Anal Recognit (IJ DAR), 2020; 23(2): 89-102.
- [25] Wikipedia. (2023, Mayıs 5). Borsa İstanbul. https://tr.wikipedia.org/wiki/Borsa_İstanbul