

İstatistiksel ve Derin Öğrenme Modellerini Kullanarak Hisse Senedi Fiyat Tahmini

Stock Price Prediction Using Statistical and Deep Learning Models

Emre ALBAYRAK
Çankaya Üniversitesi,
Bilgisayar Mühendisliği Bölümü,
Ankara, Türkiye
c1771010@student.cankaya.edu.tr
ORCID: 0000-0003-1352-6776

A. Nurdan SARAN
Çankaya Üniversitesi,
Bilgisayar Mühendisliği Bölümü,
Ankara, Türkiye
buz@cankaya.edu.tr
ORCID: 0000-0002-7148-0457

Öz

Borsa analizi, geleceğe yönelik tahminler yapmak için finansal, politik ve sosyal göstergeleri göz önünde bulundurarak borsayı inceler ve değerlendirir. Büyük veri ve derin öğrenme teknolojilerindeki gelişmelerin çığır açan sonuçları, araştırmacıların ve endüstrinin dikkatini bilgisayar destekli borsa analizine çekmektedir. Geleneksel makine öğrenimi ve derin öğrenme modellerini kullanarak borsa analizi konusunda çeşitli çalışmalar bulunmaktadır. Bu çalışmada, temel model olarak Otoregresif Entegre Hareketli Ortalama (ARIMA) yöntemini tekrarlayan sinir ağlarının üç farklı modeliyle karşılaştırılmıştır; Uzun Kısa Süreli Bellek (Long Short Term Memory- LSTM) ağları, Geçitli Tekrarlayan Birim (Gated Recurrent Unit- GRU), dikkat katmanlı LSTM modeli. Bu çalışmada literatürdeki diğer çalışmalardan farklı olarak 28 tane finansal indikatör kullanılarak Borsa İstanbul verileri üzerinde gün içi tahminler yaparken dört farklı modelin sonuçları karşılaştırılmıştır. İstatistiksel ve doğrusal bir model olan ARIMA, zaman serileri tahmini için doğrusal olmayan RNN modelleri ile karşılaştırılmıştır ancak 3 sinir ağı modelinden de yüksek ortalama hata oranına sahip olduğu görülmüştür. LSTM sonuçları GRU modeline çok yakın olsa da GRU diğerlerinden biraz daha iyi performans göstermektedir. Dikkat mekanizmalı sinir ağı diğer temel sinir ağlarından daha iyi sonuç vermemektedir.

Anahtar Sözcükler: Otoregresif Entegre Hareketli Ortalama, BIST, Geçitli Tekrarlayan Birim Modeli, Uzun Kısa Süreli Bellek Modeli

Abstract

The stock market analysis examines and evaluates the

Gönderme, düzeltme ve kabul tarihi: 02.11.2022 - 10.01.2023 – 13.06.2023

Makale türü: Derleme

stock market by considering the financial, political, and social indicators to make future predictions. Breakthrough results of advancements in big data and deep learning technologies attract the attention of researchers and

traders to computer-assisted stock market analysis. There are several studies on stock market analysis using conventional machine learning and deep learning models. In this paper, we used Autoregressive Integrated Moving Average (ARIMA) as a base model and compared it with three different models of recurrent neural networks: Long Short-Term Memory (LSTM) networks, Gated Recurrent Unit (GRU), LSTM with an attention layer model. While making intraday forecasts on Borsa İstanbul data, the results of four different models have been compared. The statistical model, ARIMA, is used as a baseline model for comparison with neural networks, but it has higher mean absolute error than other neural network models. Even though the LSTM results are very close to the GRU model, GRU slightly outperforms the others. The attention neural network model does not give better results than other basic neural networks.

Keywords: ARIMA, BIST, GRU, LSTM

1. Giriş

Borsa analizi, kripto para birimleri, yüksek performanslı uygulamalar vb. ile günümüzün en ilgi çeken konularından biridir. İki çeşit analiz yönteminden bahsetmek mümkündür: temel ve teknik analiz. Temel analiz, borsa analiz sürecindeki verilerden yararlanmak için çok çeşitli veri kaynaklarını birleştirir. Veri kaynakları, siyasi haberler, hükümet politikaları gibi birçok farklı kaynak olabilir. Teknik analiz, hisse senedi çizelgeleri, mevcut piyasa eğilimleri, ticaret hacimleri vb. gibi finansal göstergelere dayanır. Teknik analiz, kısa vadeli hisse senedi davranışını anlamak için kullanılırken, uzun vadeli tahminler için temel analiz kullanılır [1]. Hisse

senedi tahmini için öncelikle istatistiksel yöntemler önerilmiştir; otoregresif bütünlümlü hareketli ortalama (ARIMA) [22] ve genelleştirilmiş otoregresif koşullu değişen varyans (GARCH) [23] gibi. ARIMA modeli, son teknoloji etkinliği ve sağlamlığı nedeniyle finansal tahmin uygulamalarında yaygın olarak kullanılan istatistiksel modellerden biridir [9]. ARIMA'nın tahmin ve karşılaştırma aşamalarında ayırt edici kullanımının yanı sıra veri işleme aşamalarında melez öznelilikler olarak derin öğrenme ile birleştirildiği çalışmalar vardır [12] [13] [14].

Ancak doğrusal olmayan, yüksek gürültülü finansal zaman serilerinde geleneksel istatistik modelleri bunları yüksek hassasiyetle tahmin etmekte güçlük çekmektedir [24]. Makine öğrenmesi modellerinin geliştirilmesiyle birlikte, regresyon ve destek vektör makinesi gibi yöntemler finansal zaman serileri tahmini için de uygulanmaktadır [25] [5] [6]. Destek vektör makineleri her ne kadar sınıflandırma yöntemleri için tasarlanmış olsa da zaman serileri tahmini içinde yayınlarda kullanılmıştır [26].

Günümüzde sinir ağları olarak da bilinen derin öğrenme algoritmaları, doğrusal olmayan yaklaşım yeteneği ve adaptif kendi kendine öğrenmesi sayesinde güncel bir tahmin aracı haline gelmiştir. Bu algoritmalar, bilgisayarla görme, doğal dil işleme ve konuşma tanıma alanlarında çığır açan sonuçların arkasında olmuştur. Hisse senedi fiyatlarını doğru bir şekilde tahmin etmenin karmaşıklığı ve ortaya çıkan büyük veri ve derin öğrenme teknolojileri, borsa analizlerinde bilgisayarların kullanılmasına yol açmıştır [2] [3]. Uygun fiyatlı emtia donanımı, hızla gelişen bilgi erişim teknolojileri ve artan veri hacimleri, anlamlı örüntüler ve ilişkiler çıkarmak için büyük hacimli verilerin işlenmesini mümkün kılar [4]. Derin öğrenmede en çok çalışılan yöntemlerden Evrişimli Sinir Ağları (CNN) zaman serisi verilerine az da olsa uygulanarak hisse senedi tahminleri için kullanılmıştır [1]. Ancak tekrarlayan sinir ağları (LSTM [7] [2], GRU gibi özellikle zaman serileri için tasarlandığı için hisse senedi tahminleri için literatürde sıklıkla kullanılmıştır. RNN ağları, zaman serisi gibi sıralı verileri işlemeye özellikle uygundur. Derin öğrenme yöntemleri, hisse senedi fiyatları hakkında doğru tahminler elde etmek için kullanılan hisse senedi verilerinin içindeki içsel doğrusal olmayan ilişkileri ve gizli faktörleri anlamamıza yardımcı olur [8].

2014'te Bahdanau ve ark. tarafından dikkat mekanizması [33] makine çevirisi için önerilmiş ancak sonra zaman serisi tahmininde ve hisse senedi tahmini için de kullanılmıştır [3]. Hisse senedi fiyatları gibi veri dizilerinin durağan olmama özelliği, zaman serisinin istatistiksel özelliklerinin (örneğin, ortalama ve varyans gibi) zaman içinde sabit olmaması olarak açıklanabilir [28]. Bu, zaman serisinin sürekli değiştiği anlamına gelir ve geçmiş verilere dayanarak uzun vadeli tahminler yapmak zorlaştırmaktadır. Dikkat mekanizması, modele tahmin yaparken girdi dizisinin belirli kısımlarına odaklanmasına izin verdiğinden zaman serilerinin durağan olmayan özelliğine rağmen modelin zaman içinde verilerdeki değişen desenlere uyum sağlamasına izin vermektedir [31]-[32].

Bu çalışmada temel amacımız, BIST verilerini kullanarak gün içi tahminler yapmak için istatistiksel modelleri ve tekrarlayan sinir ağları ile öğrenme yöntemlerini karşılaştırmaktır. Karşılaştırmalı bir çalışma ortaya çıkarmak için istatistiksel bir model olan ARIMA modeliyle, dikkat mekanizmalı LSTM gibi en popüler ve üst düzey derin öğrenme yöntemleri kullanılmaktadır. ARIMA hem istatistiksel hem de derin öğrenme modelleriyle karşılaştırma amacıyla yaygın olarak bir temel model olarak birçok çalışmada kullanılmıştır [10] [11]. Hisse senedi tahmini için ham fiyat ve hacim kullanımının yanı sıra, fiyat hareketlerinin tahminine temel ve teknik göstergeler kullanılarak yaklaşılabılır [26]-[27]-[29]. Dikkat mekanizmalı LSTM yönteminin finansal piyasa tahminine uygulanıp uygulanamayacağına ilişkin yayınlarda az çalışma gösterilmiştir. Bu nedenle, bu çalışma, Borsa İstanbul verileri üzerinde gün içi hisse senedi piyasasını tahmin etmek için 28 tane finansal göstergeler kullanılarak dikkat mekanizmalı LSTM yöntemi, GRU gibi derin öğrenme yaklaşımlarını istatistiksel bir yöntem olan ARIMA ile karşılaştırmaktadır.

Bu makale 4 bölümden oluşmaktadır. Bölüm 1, stok analizi, derin öğrenme konularına kısa bir giriş sağlar. Ayrıca ilgili çalışmaları özetler. Bölüm 2, modellerin ayrıntılı açıklamasını verir. Veri alma, ön işleme operasyonları ve özellikle mühendisliği süreçlerini açıklar. Bölüm 3'te model oluşturma aşamaları, sonuçlar ve ayrıntılı grafik ve sonuçları içeren tartışma bölümleri yer almaktadır. 4. bölüm çalışmamızın sonuç bölümüdür.

2. Yöntem

2.1. Veri

Veriler, Borsa İstanbul'daki 100 farklı hisse senedinin saatlik verileri olup, veriler Matrics Bilgi Dağıtım Hizmetleri A.Ş. aracılığıyla elde edilmiştir. Veriler 2001-2020 tarih aralığında olup, toplam 772533 satırdır.

Her dosyanın beş veri sütunu vardır: tarih saat, açılış fiyatı, yüksek fiyat, düşük fiyat, kapanış fiyatı ve hacim. Ön işleme sırasında 28 teknik gösterge ile hesaplama yapılmış ve ayrıca USD/TRY para birimi verilerini de eklenmiştir. Bu çalışmada, yer kaplamaması açısından sonuçları gösterirken alfabetik sıraya göre Borsa İstanbul'dan üç farklı hisse senedi seçilmiştir; ASELS, AKBNK ve AKENR. Ön işleme safhası sonrasında elde edilen teknik göstergeler ve formülleri Çizelge-1'de yer almaktadır.

Çizelge-1 Teknik Göstergeler ve Açıklamaları

Gösterge	Açıklama
SMA	Basit Hareketli Ortalama
RSI	Göreceli Güç Endeksi
CCI	Emtia Kanal Endeksi
DMI	Yönel Hareket Endeksi
ADX	Ortalama Yönel Endeksi
VPT	Hacim Fiyat Trendi
EFI	Elder Güç Endeksi
WOBV	Ağırlıklı Denge İşlem Hacmi
VZO	Hacim Bölgesi Osilatörü
PZO	Fiyat Bölgesi Osilatörü
TP	Tipik Fiyat

ADL	Birikim Dağıtım Göstergesi
SMMA	Yuvarlatılmış Hareketli Ortalama
TR	Gerçek Aralık
SAR	Bitiş ve Dönüş Göstergesi
VWAP	Hacim Ağırlıklı Ortalama Fiyat
SSMA	Yuvarlatılmış Basit Hareketli Ortalama
DEMA	Çift Üstel Hareketli Ortalama
TEMA	Üçlü Üstel Hareketli Ortalama
TRIX	Üçlü Üstel Hareketli Ortalama Osilatörü
CURRENCY	Dolar/TL Paritesi

Çizelge-1'deki teknik göstergelerin açıklamaları ve hesaplama yöntemleri aşağıda verilmiştir:

Göreceli Güç Endeksi (RSI): Hisse senedinin değişen fiyatının gücünü değerlendiren bir göstergedir. Hisse senedinin şimdiki ve zaman aralığındaki güç ve zayıflığını ölçer:

$$RSI = 100 - \frac{100}{H_t - L_t}$$

Burada, H: t zaman aralığında en yüksek ortalama, L: t zaman aralığında en düşük ortalama

Basit Hareketli Ortalama (SMA): En çok kullanılan momentum göstergelerindendir, belirli bir zaman aralığında endeksin kapanış fiyatları toplanarak zaman sayısına bölünmesiyle hesaplanır:

$$SMA = \frac{\sum_{i=0}^n C_i}{n}$$

Burada n zaman aralığı, C_i kapanış değerleri

Yuvarlatılmış Basit Hareketli Ortalama (SMMA): Basit hareketli ortalamanın benzeri olan, ancak daha fazla ağırlığı son veri noktalarına veren bir teknik göstergedir. Önceki n veri noktalarının toplamı alınarak, sonra n+1 ile bölünerek hesaplanır.

$$SSMA_i = \frac{SSMA_{i-1} * (n - 1) + C_i}{n + 1}$$

Üstel Hareketli Ortalama (EMA): Basit hareketli ortalamanın üstel değişimidir, son fiyat değişikliklerine basit hareketli ortalamadan daha hızlı tepki verir [30]:

$$EMA_t = C_t * \frac{2}{n + 1} + EMA_y * \left(1 - \frac{2}{n + 1}\right)$$

Burada C_t şimdiki kapanış, EMA_t şimdiki değer, EMA_y önceki değer (burada EMA₀ olarak SMA kullanılır). DEMA (çift üstel hareketli hesaplama), TEMA (üçlü üstel hesaplama) ve TRIX (üçlü üstel hareketli ortalama osilatörü) benzer şekilde hesaplanır.

Emtia Kanal Endeksi (CCI): Fiyatın mevcut ve geçmiş arasındaki ortalama farkıdır. [30].

$$CCI = \frac{(C - SMA)}{0.015 * Standard_deviation}$$

Denge İşlem Hacmi (OBV): Endeks hacim akışını dikkate alan bir momentum göstergesidir.

$$OBV_t = \begin{cases} OBV_y + V_t & , if C_t > C_{t-1} \\ OBV_y & if C_t = C_{t-1} \\ OBV_y + V_t & if C_t < C_{t-1} \end{cases}$$

Burada C_t, V_t şimdiki kapanış ve hacim, C_{t-1} önceki kapanış

Tipik Fiyat (TP): en yüksek, en düşük ve kapanış fiyatlarının ortalaması hesaplanır.

$$TP = \frac{H_t + L_t + C_t}{3}$$

Direksiyonel Hareket İndeksi (DMI) bir trendin yönünü belirlemek için kullanılan bir teknik göstergedir: pozitif yönel hareket indeksi (DI+) yükseliş eğilimini gösterirken, negatif yönel hareket indeksi (DI-) düşüş eğilimini gösterir.

$$DI+ = 100 * (H_t - H_{t-1}) / (H_t - L_t)$$

$$DI- = 100 * (L_t - L_{t-1}) / (H_t - L_t)$$

Ortalama Yönel Endeksi (ADX): Bir finansal enstrümanın eğiliminin güçlülüğünü ölçmek için kullanılan bir teknik analiz aracıdır.

$$ADX = (DI+ - DI-) / (DI+ + DI-)$$

Hacim Fiyat Trendi (VPT): Bir hisse senedinin fiyatı ve hacminin yüzdesel değişimlere dayalı olarak trend gücünü gösterir. Örneğin, bir enstrümanın fiyatı yükselirken hacim de artarsa, VPT değeri yükselecektir.

$$VPT_t = VPT_{t-1} + (C_t - C_{t-1}) * V_t$$

Hareket Kolaylığı (EFI): VPT gibi bir enstrümanın fiyatının hareket ettiği yönde hacim değişimlerini gösterir.

$$EFI = \frac{(C_t - C_{t-1}) * V_t}{(H_t - L_t) * \sum_t V_t}$$

Birikim Dağılım Çizgisi (ADL): Hisse senedi fiyatının birikimli ya da dağılımlı olup olmadığının bir göstergesidir.

$$ADL_t = ADL_{t-1} + ((C_t - L_t) - (H_t - C_t)) * V_t / (H_t - L_t)$$

Gerçek Aralık (TR): Bir varlığın volatilitasını ölçmek için kullanılan bir teknik göstergedir.

$$TR = \max(H_t - L_t, \text{abs}(H_t - C_{t-1}), \text{abs}(L_t - C_{t-1}))$$

Bitiş ve Dönüş Göstergesi (SAR): Potansiyel eğilim dönüşlerini belirlemek için kullanılan teknik bir göstergedir.

$$SAR_t = SAR_{t-1} + AF * (SAR_{t-1} - SAR_{t-2})$$

SAR(t) enstrümanın t. zaman için tahmini fiyatı gösterir ve AF ise hızlandırma faktörüdür ve genellikle 0,02 ile 0,20 arasında değişir.

Hacim Ağırlıklı Ortalama Fiyat (VWAP): Hacmi dikkate alarak, belirli bir zaman diliminde ortalama fiyatını belirlemek için kullanılan teknik bir göstergedir. Menkul kıymetteki tüm işlemlerin toplam değerinin menkul kıymetin toplam hacmine bölünmesiyle hesaplanır.

Osilatör: Finansal gösterge olarak kullanımları çok fazladır; henüz başlamamış olası bir eğilim değişimini işaret edebilecek öncü ve gecikmeli göstergeler olarak kabul edilirler.

Hacim Bölgesi Osilatörü (VZO): Şimdiki kapanış fiyatı önceki kapanış fiyatından yüksekse, hacim pozitif bir değere sahip olacaktır; aksi takdirde negatif bir değere sahip olacaktır.

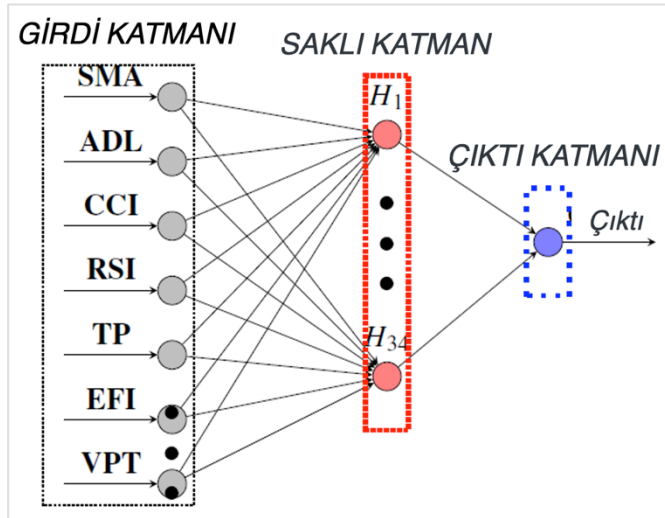
Fiyat Bölgesi Osilatörü (PZO): Mevcut fiyatı ortalama geçmiş fiyatlara karşı ölçen teknik bir göstergedir. Gösterge, iki üstel hareketli ortalamayı hesaplar ve aralarındaki orandır.

Tüm göstergelerde de varsayılan periyot 14 alınmıştır [34].

2.2. Modeller

Bu çalışmada BİST verileri üzerinde teknik analiz yapmak ve gün içi frekansta başarılı tahminler yapmak hedefi ile hem istatistiksel hem de derin öğrenme modelleri uygulanmıştır. İstatistiksel modeller açısından temel model olarak ARIMA, yapay sinir ağları açısından ise LSTM, GRU, dikkat mekanizmalı LSTM modelleri kullanılmıştır. ARIMA modeli, tüm verilerdeki alt kümelerin ortalamalarının hesaplanmasına ve gelecekteki eğilimler hakkında tahminlerde bulunulmasına dayanmaktadır. Sinir ağı modelleri, zaman serisi veri ile ilgili uygulamalarda yaygın olarak kullanılan tekrarlayan sinir ağları modelleridir. LSTM, Vanilla RNN'nin bir gelişimi olarak kabul edilebilir ve GRU, LSTM 'in bir çeşididir. Dikkat mekanizmasına sahip LSTM, dikkat mekanizmasını kullanan bir kodlayıcı-kod çözücü modelidir.

Yapay sinir ağları, insan beynini taklit eden ve bilgisayarlarla görme sorunları, konuşma tanıma, görüntü işleme vb. alanlarda yaygın olarak kullanılan özel makine öğrenme algoritmaları sınıfıdır. Yapay sinir ağlarının bir çeşidi olan tekrarlayan sinir ağları zaman içinde meydana gelen bilgileri anlamlandırabilmektedir. Sinir ağları, derin ağları eğitmeyi çok daha kolay hale getiren bir yeniliktir. Şekil 1 Yapay Sinir Ağlarının Genel Yapısı, bir YSA'nın giriş ve gizli katmanlarını göstermektedir.



Şekil-1: Yapay Sinir Ağlarının Genel Yapısı

Ölçünlü bir ileri beslemeli sinir ağında, bir nöron aldığı girdiyi toplar ve ardından bir çıktı üretmek için girdiyi bir etkilenim fonksiyonundan (ReLU, Sigmoid, Tanh, vb.) geçirir. Tekrarlayan bir sinir ağında, bir nöron çıktığı kendisine geri gönderir. Tekrarlayan bir nöron açılırsa, tekrarlayan nöronun bir önceki zamandaki (t-1) girişi t-1'de çıktı verir ve daha sonra t zamanındaki durumunda nörona geçirilir ve ardından t

zamanında bir çıktısı olur ve bu böyle devam eder. Tekrarlayan nöron, hem önceki bir zaman adımından girdileri hem de mevcut zaman adımından girdileri alır. Tekrarlayan bir sinir ağında, önceki zaman adımlarından girdi alan bu nöronlar "bellek hücreleri" olarak da bilinir. LSTM ve GRU, tekrarlayan bir sinir ağlarının alt sınıflarıdır ve aşağıdaki bölümlerde bu iki türü ayrıntılı olarak açıklanacaktır.

Otoregresif Entegre Hareketli Ortalama (ARIMA)

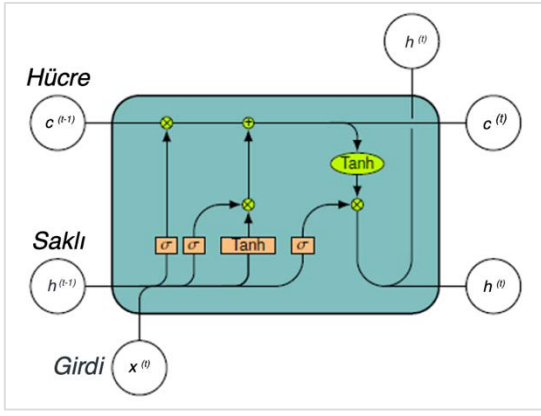
Bu model, gelecekteki eğilimleri tahmin etmek için geçmiş verileri kullanan zaman serisi veri analizinde kullanılan istatistiksel bir modeldir. ARIMA modeli, ARMA (Otomatik Regresif Hareketli Ortalama) modeline dayanmaktadır. ARMA modeli, AR (Otomatik Regresif) ve MA (Hareketli Ortalama) bölümlerinden oluşur. Oto regresif, mevcut değer geçmiş değerlere dayandığını düşünen bir modeli ifade eder. Hareketli ortalama, tüm veri setinin farklı alt kümelerinin ortalamalarından oluşan bir hesaplamadır. ARIMA ve ARMA arasındaki temel fark, ARIMA modelinin durağan olmayan verileri üzerinde çalışmaya başlamadan önce durağan hale getirmesidir. Akaike Bilgi Kriterleri (AIC), model performanslarını karşılaştırmak için istatistiksel bir modelde kullanılan bir ölçüm metriğidir [15].

Uzun Kısa Süreli Bellek Ağları (LSTM)

Derin ağlarda, bazen kaybolan veya patlayan türevler (gradient) olarak adlandırılan geri yayılımdan kaynaklanan sorunlar ortaya çıkabilir. Alt seviye katmanlara veya giriş katmanlarına daha yakın olan ön katmanlara geri döndüğümüzde, türevler genellikle küçülür ve sonuçta nöronların ağırlıklarının bu alt seviyelerde asla değişmemesine neden olur. Farklı bir etkinleştirme işlevi veya toplu normalleştirme teknikleri, bir ölçekte kaybolan türev sorunlarına çözüm olarak önerilebilir. Ancak, uzun zaman serisi verileri gibi dizi girişinin uzunluğu nedeniyle, önerilen bu çözümler eğitimi yavaşlatabilir.

Tekrarlayan sinir ağlarının karşılaştığı bir başka sorun da, bu ağlarda geçen her adımda bilgi kaybolduğu için bir süre sonra ağın ilk girdileri "unutmaya" başlamasıdır. Bu noktada, tekrarlayan sinir ağları için uzun süreli bir bellek çözümüne ihtiyaç duyulmaktadır. Uzun Kısa Süreli Bellek hücresi, 1997'de kaybolan türev sorununun ele alınmasına yardımcı olmak için oluşturulmuştur [16].

LSTM hücresi, giriş kapısı, unutma kapısı ve çıkış kapısı olmak üzere üç farklı kapı içermektedir. Hücre durumunda neyin depolanacağına karar vermek için giriş kapısı, hücre durumundan hangi bilgilerin unutulacağını veya atılacağını belirlemek için unutma kapısı, hücre durumunda neyin güncelleneceğine karar vermek için çıkış kapıları kullanılmaktadır [16] [17].



Şekil-2 LSTM in Genel Yapısı

Şekil 2'de x_t geçerli giriş, c_{t-1} son LSTM biriminin belleği olan hücre durumu, h_{t-1} son LSTM biriminin çıkışı olan gizli durumdur. h_t , mevcut çıktıdır. c_t yeni güncellenen bellektir. Bir sonraki hücre durumu olarak kabul edilebilir. σ ve \tanh , sigmoid ve hiperbolik tanjant fonksiyonlarını, x , ölçekleme bilgisi işlemini ve $+$ bilgi ekleme işlemini ifade eder.

Denklemlerde i_t giriş kapısı, f_t unut kapısı ve o_t çıkış kapısıdır. W , önceki gizli durum ile mevcut gizli durum arasındaki bağlantıdır. U , girdiler ve gizli katman arasındaki ağırlık matrisidir. \tilde{C}_t , mevcut girdinin ve önceki gizli durumun hesaplanması olan aday gizli durumdur. C_t , hücrenin dahili hafızasıdır. C_t , önceki belleğe dayalı olarak hesaplanır, unut kapısı, yeni gizli durum ve giriş kapısı ile çarpılır.

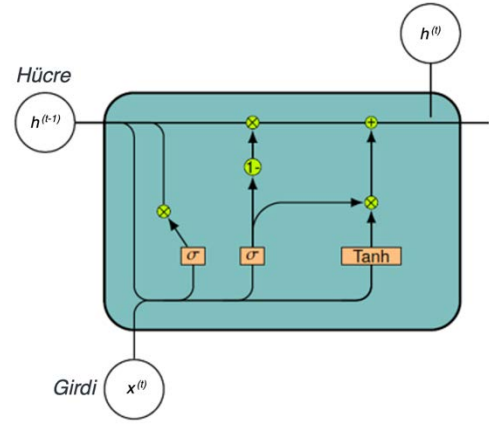
Ayrıntılı LSTM denklemleri aşağıdaki gibidir:

$$\begin{aligned} i_t &= \sigma(x_t U^i + h_{t-1} W^i) \\ f_t &= \sigma(x_t U^f + h_{t-1} W^f) \\ o_t &= \sigma(x_t U^o + h_{t-1} W^o) \\ \tilde{C}_t &= \tanh(x_t U^g + h_{t-1} W^g) \\ C_t &= \sigma(f_t \times C_{t-1} + i_t \times \tilde{C}_t) \\ h_t &= \tanh(C_t) \times o_t \end{aligned}$$

Kullandığımız LSTM modeli 3 katman LSTM ve dropout(0,2) fonksiyonları ve en on katmanda dense katmanından oluşturulmuştur.

Geçitli Tekrarlayan Birimler (GRU)

LSTM hücresinin bir varyantı, 2014'te tanıtılan GRU olarak adlandırılır [18]. GRU, unut ve giriş kapılarını "güncelleme kapısı" adı verilen tek bir geçitte birleştirerek LSTM' in iç yapısını basitleştirir. Ayrıca hücre durumunu ve gizli durumu birleştirir. Güncelleme kapısı, hangi bilgilerin iletileceğine karar vermek için kullanılır, sıfırlama kapısı, hangi bilgilerin unutulacağını belirlemek için kullanılır. Bu kapılar sırasıyla uzun süreli ve kısa süreli belleği temsil eder.



Şekil 3 GRU nun Genel Yapısı

GRU, mevcut hücre durumunu ve LSTM' in gizli durumunu önceki aşamanın çıktısı olarak tek bir durumda birleştirir. Şekil 3'te girişler, mevcut giriş ve hücre durumudur. x^t mevcut girdi, h^{t-1} önceki aşamanın çıktısı, h^t mevcut çıktıdır. GRU, LSTM' in bir varyantı olarak LSTM' in işlevlerini ve işlemlerinin çoğunu korur.

Ayrıntılı denklemlerde z_t , güncelleme kapısını ve r_t , sıfırlama kapısı, \tilde{h}_t aday gizli durumu ve h_t gizli durumu, W ağırlık matrisidir. Ayrıntılı GRU denklemleri aşağıdaki gibidir:

$$\begin{aligned} z_t &= \sigma(W_z \cdot [h_{t-1}, x_t]) \\ r_t &= \sigma(W_r \cdot [h_{t-1}, x_t]) \\ \tilde{h}_t &= \tanh(W \cdot [r_t \times h_{t-1}, x_t]) \\ h_t &= (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \end{aligned}$$

Kullandığımız GRU modeli 3 katman GRU ve dropout(0.2) fonksiyonları ve en on katmanda dense katmanından oluşturulmuştur.

Dikkat mekanizmalı LSTM

Derin öğrenmede dikkat mekanizması, eğitim sürecinde verilerin belirli özelliklerine odaklanır. Dikkat mekanizmasının en sık kullanılan durumlarından biri, kodlayıcı-kod çözücü modeli olarak da bilinen seq2seq modelleridir [18]. Sıralı derin öğrenme modelleri, doğal dil işleme, tanıma ve zaman serisi verilerini kullanan uygulamalarda yaygın olarak kullanılan tekrarlayan sinir ağları olarak kabul edilir. Kodlayıcılar ve Kod Çözücüler, LSTM veya GRU modelleri olabilir. Bu mekanizmada, kodlayıcı, dahili durumların girdi verilerini kod çözücüyeye iletir ve kod çözücü, kodlayıcıdan alınan girdi verilerinin dahili durumlarına dayalı olarak çıktı üretir. Dikkat mekanizmalı seq2seq modelinde, kod çözücü çıktı üretirken standart seq2seq model davranışı olan kodlayıcının son durumu yerine tüm dizinin iç durumlarını dikkate alır. Dikkat mekanizması, çıktıyı üretmek için belirli özelliklere odaklanırken kod çözücünün tüm verilere erişmesine izin verir.

[19]'da yazarlar, dikkat temelli sinir ağlarını global ve yerel dikkat modelleri olmak üzere iki geniş kategoriye ayırmaktadır. Bu iki model arasındaki fark, kodlayıcının çıktısını işleme biçimine bağlıdır. Global dikkatte, kodlayıcının tüm gizli durumları dikkate alınır, bu da tüm özelliklere önem verildiği anlamına gelir. Ancak, yerel dikkatte, çıktının üretilmesinde gizli durumların seçilmiş alt kümeleri kullanılır.

Bahdanau ve Luong, zaman serisi veri uygulama problemlerini ele almak için yaygın olarak kullanılan iki yaygın dikkat modelidir [3] [20] [18]. Bu iki mekanizma arasındaki temel fark, puan benzerliklerinin hesaplanma şeklidir. Luong'un dikkati, basit matris çarpımlarını kullanır ve böylece daha hızlı ve alan açısından daha verimli hale getirir. Her iki model de dikkat mekanizmasında "bağlam vektörü" olarak adlandırılan çıktıyı üretirken tüm gizli durum bilgilerini kullanan global dikkat modelleri olarak kabul edilebilir [19]. Buna ek olarak, Luong dikkati yerel dikkat mekanizması olarak da uygulanabilir.

Aşağıdaki denklemler, Bahdanau ve Luong'un puan benzerliklerini hesaplama yöntemi arasındaki farkı göstermektedir. h_s , kodlayıcıdan kod çözücüye iletilen giriş verilerinin gizli durumlarını belirtir. h_t , kod çözücünün önceki çıktısı, c_t bağlam vektörü ve W_t ağırlık matrisidir. Bu puanları, kodlayıcının tüm gizli durumları ile önceki kod çözücünün çıktısı arasındaki ilişkinin düzeyi olarak çıkarabiliriz. Bu puanlar, dikkat ağırlıklarının aşağıdaki denklemleri hesaplamak için kullanılır:

$$score(h_t, \tilde{h}_s) = \begin{cases} h_t^T W \tilde{h}_s & \text{Luong} \\ v_a^T \tanh(W_1 h_t + W_2 \tilde{h}_s) & \text{Bahdanau} \end{cases}$$

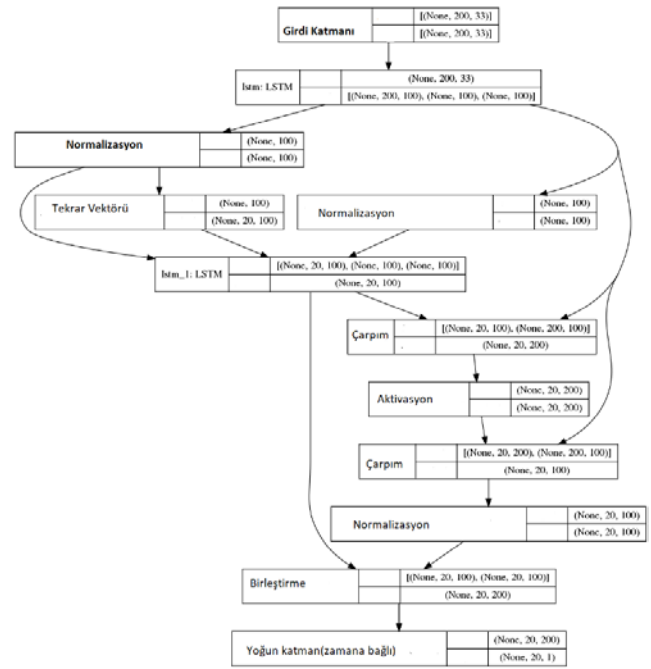
Bağlam vektörü, aşağıdaki gibi dikkat ağırlıkları kullanılarak hesaplanır:

$$a_{ts} = \frac{\exp(score(h_t, \tilde{h}_s))}{\sum_{s'=1}^S \exp(score(h_t, \tilde{h}_{s'}))}$$

Puan benzerliklerinin hesaplanması dışındaki aşamalar hem Bahdanau hem de Luong dikkat mekanizmaları için aynıdır.

$$c_t = \sum_s a_{ts} \tilde{h}_s$$

Bu çalışmada, 2015 yılında tanıtılan Luong küresel dikkat mekanizmasını kullanmıştır [19]. Dikkat mekanizması, ağdaki nöronlara ağırlık atarken diğer modellerden temel olarak farklıdır. Ağırlıkları ayarlayarak ağır verilerin özneliklerinin dikkate almasını sağlar [19]. 33 öznelik (adım sayısı:20) ile kullanılan dikkat mekanizmalı LSTM modeli Şekil 4 'te verilmiştir.



Şekil 4 Dikkat Mekanizmalı LSTM Modeli

Burada her katmanın ilk satırındaki 3-lü girdi, ikinci satırındaki 3-lü çıktı büyüklüğünü vermektedir. Normalizasyon için toplu normalizasyon, çarpım için nokta çarpım, birleştirme uç uca ekleme, yoğun katman çıktıya göre şekillendirme (dense) katmanı ifade etmektedir.

3. Sonuçlar ve Tartışmalar

Hiper Parametre Seçimi:

Arima için parametre seçimi: Çizelge 2 de verildiği üzere grid arama sonuçlarına göre görece uzun sürmesine rağmen en iyi performans veren ARIMA modeli ARIMA (2,1,0) olmuştur.

Burada ARIMA parametreleri (p, d, q) sırasıyla; p, verilerin geçmişteki değerleriyle ilişkisini belirtir, d (değişim miktarı), verilerdeki trendleri düzeltmek için kullanılır (seriyi durağan hale getirmek için gereken fark sayısıdır), q (Hareketli Ortalama) verilerdeki dalgalanmaları (hata terimleri) ölçen bir değişkendir.

Çizelge-2: Arima için Parametre Seçimi

Model Parametresi	AIC	Zaman
(0,1,0)	21592.2	0.20
(2,1,0)	21638.4	0.84
(1,1,0)	21590.2	0.25
(3,1,0)	21637.2	0.48
(2,1,1)	21637.0	0.45
(1,1,1)	21588.7	0.32
(3,1,1)	21635.2	0.55

Yinelenen Sinir Ağları için Parametre Seçimi: Yapay sinir ağlarında hiperparametrelerin farklı kombinasyonları eğitimin hızı ile verilerin geliştirilmesi arasında bir ödünleşime neden olur, bir başka deyişle, ağır çalışma hızını arttıran parametreler sistemin geliştirme yetkinliklerini azaltabilir. Bu nedenle hiperparametreler, test süreçlerimizde farklı

hiperparametre kombinasyonları dikkate alınarak ve en iyi sonuç veren parametreler elde edilmiştir. Bu çalışmada, Adam/Rmsprop optimizasyonu kullanılmıştır ve öğrenme katsayısı 0,001 seçilmiştir.

Çizelge- 3: ASELS Hisse Senedi için Farklı Dönem Sonuçları

Model	Dönem	Toptan	Adım Sayısı	MSE
LSTM	50	20	30	0,0014
GRU	50	20	30	0,0013
Dikkat Mekanizmalı	50	20	30	0,0471
LSTM	100	20	30	0,0011
GRU	100	20	30	0,0010
Dikkat Mekanizmalı	100	20	30	0,0460
LSTM	150	20	30	0,0009
GRU	150	20	30	0,0008
Dikkat Mekanizmalı	150	20	30	0,0456

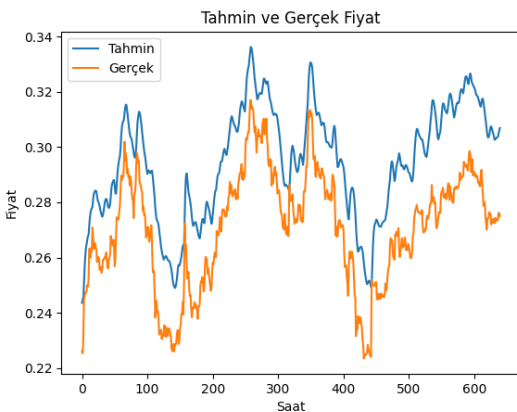
Çizelge-3'te farklı dönem (epoch) 50, 100, 150 değerleri için test sonuçları verilmiştir [21].

Çizelge-4'de görüldüğü gibi LSTM modelinde RELU, sigmoid fonksiyona göre daha iyi sonuçlar verdiği gözlemlenmiştir ve hiperbolik tanjant fonksiyonu en iyi sonuçları vermiştir.

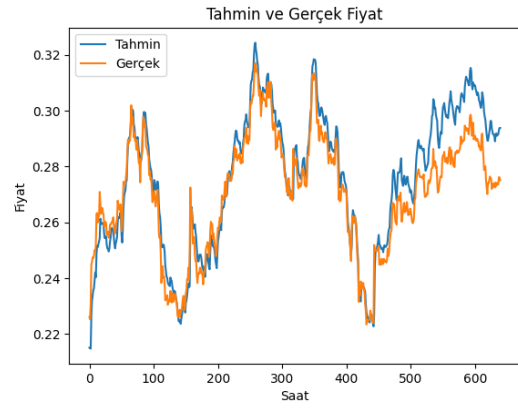
Çizelge- 4 Tüm Hisse Senetleri için Aktivasyon Fonksiyonu Seçimi

Aktivasyon	Hisse Senedi	MSE
Sigmoid	AKBNK	0,0097
RELU	AKBNK	0,0049
Tanh	AKBNK	0,0042
Sigmoid	AKENR	0,0080
RELU	AKENR	0,0042
Tanh	AKENR	0,0036
Sigmoid	ASELS	0,0444
RELU	ASELS	0,0274
Tanh	ASELS	0,0266

Modellerimizi BİST verilerinden üç farklı hisse senedine uygulanmıştır. ASELS hisse verileri için LSTM, GRU sonuçları (150 dönem-epoch), ARIMA (2,1,0) sırasıyla Şekil 5, Şekil 6 ve Şekil 7'da verilmektedir.



Şekil 5 ASELS hisse senedi için LSTM sonuçları



Şekil 6 ASELS hisse senedi için GRU sonuçları



Şekil 7 ASELS hisse senedi için ARIMA (2,1,0) sonuçları

Çizelge-5'te 3 hisse senedi için ARIMA sonuçlarını göstermektedir.

Çizelge-5 ARIMA Karşılaştırma Çizelgesi

Stock	MSE	MAE
ASELS	0,0105	0,0836
AKBNK	0,0194	0,1066
AKENR	1,4056	0,9692

Modellerin ortalama kare hata (MSE) ve ortalama mutlak hata (MAE) karşılaştırması Çizelge-6'da ASELS hisse senedi için, Çizelge-7 AKENR hisse senedi için,

Çizelge-8 AKBNK hisse senedi için verilmiştir. Burada ARIMA modeli için en iyi performans veren ARIMA (2,1,0) kullanılmış, derin öğrenme modelleri içinde 150 dönem, 20 toptan, 30 adım sayısı, Adam optimizasyonu 0.0001 öğrenme katsayısı kullanılmıştır.

Çizelge-6 ASELS verileri için 4 modelin karşılaştırılması

Model	MSE	MAE
ARIMA	0,0105	0,0836
LSTM	0,0009	0,0213
GRU	0,0008	0,0205
ATTENTION	0,0405	0,1670

Çizelge-7 AKENR verileri için 4 modelin karşılaştırılması

Model	MSE	MAE
ARIMA	1,4056	0,9692
LSTM	0,0019	0,0290
GRU	0,0017	0,0284
ATTENTION	0,0530	0,1959

Çizelge- 8 AKBNK verileri için 4 modelin karşılaştırılması

Model	MSE	MAE
ARIMA	0,0194	0,1066
LSTM	0,0021	0,0317
GRU	0,0014	0,0257
ATTENTION	0,0484	0,1817

Çizelgede görüldüğü üzere üç hisse senedi içinde derin öğrenme modeli olan LSTM ve GRU modelleri, istatistiksel model olan ARIMA modeline göre daha iyi sonuç vermektedir ancak GRU her üç hisse senedinde de en iyi sonuçlar vermektedir. Dikkat katmanına sahip LSTM, veri dizisinin uzun olmasına rağmen diğer modellerden daha iyi performans göstermediği görülmektedir.

Bu makalede, hata değerlerini göz önünde bulundurarak üç model karşılaştırılmıştır ve BIST100 deki tüm hisse senetleri içinde genel olarak en iyi sonuçları GRU modelinin verdiği gözlemlenmiştir.

4. Sonuç

Deneylerimizi BİST' ten AKBNK, AKENR, ASELS olmak üzere üç farklı hisse senedi üzerinde gösterilmiştir. Verilerimizi dört modele beslenmektedir: LSTM, GRU, LSTM ile dikkat katmanı ve ARIMA. Model oluşturmada, dönem sayısı, parti boyutu, öğrenme oranı, etkilenim fonksiyonu, kayıp fonksiyonu, zaman adımı vb. seçimimize karar vermek için birçok farklı parametre kombinasyonu ile deneysel çalışmalar yapılmıştır. Etkilenim fonksiyonu olarak hiperbolik tanjant fonksiyonunu ve kayıp fonksiyonu olarak MAE ve MSE fonksiyonlarını kullanılmıştır. BIST verilerinin teknik analizini yaparak, seçilen hisse senetleri için bir sonraki saatin açılış fiyatını tahmin etmeyi amaçlanmıştır.

Çalışmamızda, üç farklı hisse senedi için dört model karşılaştırdık ve deneysel sonuçlar, GRU modeli BİST hisse senedi fiyatlarını tahmin etmede diğer modellerden daha iyi performans göstermektedir. ARIMA, istatistiksel modellerin etkilerini sinir ağları ile karşılaştırmak için bir temel model olarak kullanılmıştır. Hem LSTM hem de GRU, ARIMA modelinden daha iyi performans göstermiştir. Dikkat katmanı performansına sahip LSTM, diğer modellerin gerisinde kalmıştır. Bunun nedeni, uyguladığımız küresel dikkat modeli

olabilir. Çıktıyı üretecek iç yapısı nedeniyle uzun dizileri işlemede zorluklar yaşayabildiği gözlemlenmiştir. Genel olarak, GRU' nun performansı hem istatistiksel hem de diğer sinir ağı modellerinde en iyisidir.

Hisse senedi tahminlerinde verinin durağan olmamasından dolayı diğer makine öğrenmesi veri setlerinde olduğu gibi normalizasyon uygulanmamaktadır. Bu çalışmada çalışma da birçok finansal indikatör hesaplanmış olsa da özellik ayırıştırma optimizasyonu yapılması mümkündür. Çalışmanın ilerleyen versiyonlarında global dikkat mekanizması yerine yerel dikkat mekanizması kullanılarak dikkat modelinin performansı değerlendirilebilir.

Kaynakça

- [1] Gunduz, H., Yaslan, Y., And Cataltepe, Z., *Intraday prediction of borsa istanbul using convolutional neural networks and feature correlations*. Knowledge- Based Systems, 137:138–148, 2017.
- [2] Boronovkova, S. And Tsiamas, I., *An ensemble of lstm neural networks for high-frequency stock market classification.*, Journal of Forecasting, 38(6):600–619, 2019.
- [3] Qiu, J., Wang, B., And Zhou, C. , *Forecasting stock prices with long-short term memory neural network based on attention mechanism*. PLOS ONE, 15(1):1–15, 2020.
- [4] Hasan, A., Kalipsiz, O., And Akyoku, S., *Predicting financial market in big data: Deep learning*, International Conference on Computer Science and Engineering (UBMK), pages 510–515, 2017.
- [5] Kara, Y., Acar Boyacioglu, M., And Ömer Kaan 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, 38(5):5311–5319, 2011
- [6] Rezaei, H., Faaljou, H., And Mansourfar, G., *Stock price prediction using deep learning and frequency decomposition*, Expert Systems with Applications, 169:114332, 2021.
- [7] Fischer, T. And Krauss, C., *Deep learning with long short-term memory networks for financial market predictions*, European Journal of Operational Research, 270(2):654–669, 2018.
- [8] Nguyen, T.-T. And Yoon, S., *A novel approach to short-term stock price movement prediction using transfer learning*, Applied Sciences, 9(22):4745, 2019.
- [9] Ariyo, A. A., Adewumi, A. O., And Ayo, C. K., *Stock price prediction using the ARIMA model*, 16th International Conference on Computer Modelling and Simulation. IEEE, 2014.
- [10] Ma, Q., *Comparison of ARIMA, ANN and LSTM for stock price prediction*, E3S Web of Conferences, 218:01026, 2020.
- [11] Adebisi, A. A., Adewumi, A. O., And Ayo, C. K., *Comparison of ARIMA and artificial neural networks models for stock price prediction*, Journal of Applied Mathematics, 2014:1–7, 2014.
- [12] Wang, J.-H. And Leu, J.-Y., *Stock Market Trend Prediction Using Arima-Based neural networks*, In Proceedings of International Conference on Neural Networks (ICNN'96), volume 4, pages 2160–2165 vol.4,1996.
- [13] Kumar, M. And Thenmozhi, M., *Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models*. 5(3):284,2014.
- [14] Rathnayaka, R. M. K. T., Seneviratna, D., Jianguo, W., And Arumawadu, H. I., *A hybrid statistical approach for stock market forecasting based on artificial neural network and ARIMA time series models*. IEEE, 2015.

- [15] Yang, Y., *Can the strengths of aic and bic be shared? a conflict between model identification and regression estimation.* Biometrika, 92(4):937–950,2005.
- [16] Banerjee, D., *Forecasting of indian stock market using time-series ARIMA model.* IEEE, 2014.
- [17] Almasarweh, M. And Wadi, S. A., *ARIMA model in predicting banking stock market data*,12(11):309.2018.
- [18] Cho, K., Merrienboer, B. V., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., And Bengio, Y., *Learning phrase representations using rnn encoder–decoder for statistical machine translation.* Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.
- [19] Luong, T., Pham, H., And Manning, C. D., *Effective approaches to attention- based neural machine translation,* Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, 2015.
- [20] Chen, S. And Ge, L., *Exploring the attention mechanism in lstm-based Hong Kong stock price movement prediction,* Quantitative Finance, 19(9):1507–1515, 2019.
- [21] Albayrak E. , *Stock Price Prediction Using Deep Learning Methods In High-Frequency Trading* ,MSc. Thesis, Çankaya University, 2021.
- [22] Hyndman, R. J., And Athanasopoulos G., *Forecasting: principles and practice.* Otexts, 2018.
- [23] Bollerslev, T., *Generalized autoregressive conditional heteroskedasticity.,* Journal of econometrics ,31.3, 307-327, 1986.
- [24] Bontempi, G., Souhaib B. T., And Yann-Aël L. B., *Machine learning strategies for time series forecasting.,* European business intelligence summer school, 62-77, Springer, Berlin, Heidelberg, 2012.
- [25] Alpaydin, E., *Introduction to machine learning,* MIT press, 2020.
- [26] Tay, F. E., and Lijuan C., *Application of support vector machines in financial time series forecasting.,*Omega 29.4, 309-317,2001.
- [27] Lo, A. W., Harry M., And Jiang W., *Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation,* The journal of finance 55.4: 1705-1765,2000.
- [28] Shintate, T., And Lukáš Pi., *Trend prediction classification for high frequency bitcoin time series with deep learning,* Journal of Risk and Financial Management 12.1: 17,2019.
- [29] Bao, W., Jun Y., And Yulei R., *A deep learning framework for financial time series using stacked autoencoders and long-short term memory.,* PloS one 12.7: e0180944,2017.
- [30] Raşo H., And Demirci M. *Predicting the turkish stock market bist 30 index using deep learning.,* International Journal of Engineering Research and Development 11.1: 253-265,2019.
- [31] Li Y., Zhu Z., Kong D., And Han H., & Zhao, Y., *EA-LSTM: Evolutionary attention-based LSTM for time series prediction,* Knowledge-Based Systems, 181, 104785, 2019.
- [32] Zhang, X., Liang X., Zhiyuli A., Zhang S., Xu R., And Wu, B., *AT-LSTM: An attention-based LSTM model for financial time series prediction,* IOP Conference Series: Materials Science and Engineering (Vol. 569, No. 5, p. 052037). IOP Publishing, 2019.
- [33] Bahdanau, D., Cho, K., And Bengio, Y., *Neural machine translation by jointly learning to align and translate,* arXiv preprint arXiv:1409.0473, 2014.
- [34] Wilder, J.W., *New concepts in technical trading systems,* Trend Research, 1978.