

Orijinal Makale/Research Article

## Petrol fiyatlarının zaman serileri ve LSTM modeli kullanılarak tahminlenmesi

Tuncay Yiğit<sup>1</sup>, Bekir Aksoy<sup>2</sup>, Mevlüt Ersoy<sup>1\*</sup>, Ramazan Şenol<sup>3</sup>, Osamah Khaled Musleh Salman<sup>2</sup>

<sup>1</sup> Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, 32200, Isparta, Türkiye

<sup>2</sup> Isparta Uygulamalı Bilimler Üniversitesi, Teknoloji Fakültesi, Mekatronik Mühendisliği Bölümü, 32200, Isparta, Türkiye

<sup>3</sup> Isparta Uygulamalı Bilimler Üniversitesi, Teknoloji Fakültesi, Elektrik – Elektronik Mühendisliği Bölümü, 32200, Isparta, Türkiye

### Anahtar Kelimeler

Petrol Fiyatları  
Zaman Serisi Analizi  
LSTM Modeli

**Öz:** Veri depolama altyapılarının kapasitelerinin artması ile uzun yıllar boyunca elde edilen verilerin çeşitliliği de artış göstermiştir. Bu verilerin zaman verisi ile birleştirilmesi sayesinde büyük verileri üzerinde zaman serisi analizi yapılması sağlanmaktadır. Dünya ekonomisinde çok önemli yeri olan ve “siyah altın” olarak kabul edilen ham petrol; sanayi, ulaşım, otomobil, kozmetik, enerji, kimya, ilaç sektörleri gibi birçok alanda kullanılmaktadır. Çalışmada açık erişimli internet sitesinden alınan Brent Petrol fiyatlarına ait 1987 ile 2020 yıllarına ait toplam 8267 veri alınmıştır. Kullanılan veri setindeki veri sayısının yüksek ve zamana bağlı olmasından dolayı istatistiksel yöntem olan ARIMA ile derin öğrenme yöntemi olan LSTM modelleri çalışmada kullanılmıştır. Mevcut veri seti kullanılarak ARIMA ve LSTM modelleri ile eğitilerek ileriye dönük olarak petrol fiyatlarının 180 günlük olası fiyatları tahmin edilmiştir. Elde edilen sonuçların zaman serisi analizleri ve LSTM modelinin Brent Petrolün ileriye dönük fiyat tahminlemede önemli bir başarı sağladığı görülmüştür.

### Makale geçmişi:

Geliş Tarihi: 25.11.2020  
Kabul Tarihi: 08.04.2021

### Atıf için/To Cite:

Yiğit T. Aksoy B. Ersoy M. Şenol R. Salman O.K.M. Petrol Fiyatlarının Zaman Serileri ve LSTM Modeli Kullanılarak Tahminlenmesi. Uluslararası Teknolojik Bilimler Dergisi, 13(1), 34-38, 2021.

## Forecasting oil prices using time series and LSTM model

### Keywords

Oil Prices  
Time Series Analysis  
LSTM Model

**Abstract:** In parallel with the increase in the capacity of data storage infrastructures, the variety of data obtained has also increased. By combining these data with time data, it is possible to perform time series analysis on big data. By combining these data with time data, it is possible to perform time-series analysis on big data. Crude oil, which has a very important place in the world economy and is considered "black gold"; is used in many fields such as industry, transportation, automobile, cosmetics, energy, chemistry and pharmaceutical industries. In the study, a total of 8267 data belonging to the years 1987 and 2020 of Brent Oil prices obtained from the open access website. Since the number of data in the data set used is high and depends on time, the statistical method ARIMA and the deep learning method LSTM models were used in the study. Using the existing data set, the ARIMA and LSTM models have been trained to estimate the 180-day possible prices of oil prices prospectively. It has been observed that the time series analysis of the results and the LSTM model have achieved a significant success in the forward price estimation of Brent Oil.

### Article history:

Received: 25.11.2020  
Accepted: 08.04.2021

## 1. Giriş

Veri depolama kapasitelerinin giderek artış göstermesi elektronik ortamda tutulabilen verilerin çeşitliliğini artırmaktadır. Depolama birimlerinde zaman bilgisi ile tutulan verilerin değerlendirme açısından daha anlamlı olduğu görülmektedir. Bu nedenle zamana bağlı

verilerin değerlendirmesinde büyük veriler üzerinde zaman serisi analizi yapılması popülerlik kazanmıştır. Mühendislik, istatistik, jeofizik, sismoloji, ekonometri gibi uygulama alanlarında zaman serisi verilerinin incelenmesi ve analizi yapılmaktadır [1],[2]. Zaman serisi analizinde gelecek verilerin tahmin edilmesi

\* İlgili yazar/Corresponding author: mevlutersoy@sdu.edu.tr

gerçekleştirilmektedir. Bu tahminleri yapabilmek için geçmiş verilerden analiz edilmesi gerekmektedir.

Dünya ekonomisinde çok önemli yeri olan ve “siyah altın” olarak kabul edilen ham petrol; sanayi, ulaşım, otomobil, kozmetik, enerji, kimya, ilaç sektörleri gibi birçok alanda kullanılmaktadır. Petrolün ekonomik değerinin birçok sektör tarafından değerli olması nedeniyle maliyetlerin ortaya çıkarılmasında önemli rol oynamaktadır. Petrol, pek çok ülkede ekonomik değerinden dolayı birçok ülkeden ekonomik krizlerin temel sebebi olmuştur. Bu nedenle ülkeler ekonomik krize maruz kalmamak için petrol fiyatlarının geçmiş verilerini analiz edilmesi oldukça önemlidir. Bu analizleri gerçekleştirebilmek için birçok matematiksel, istatistiksel ve son dönemlerde popüler olan yapay zekâ teknikleri kullanılmaktadır.

LSTM, ARIMA gibi modeller zaman serisi analizlerinde gelecek verilerin tahmin edilmesinde kullanılabilen istatistiksel ve yapay zekâ yöntemleri olarak görülmektedir. Bu yöntemler matematiksel ve istatistiksel yöntemlerde genellikle bir önceki zaman verilerinin bir sonraki zaman verileri ile karşılaştırılması sonucu oluşan parametrelerle analizi gerçekleştirilmektedir. Yapay zeka tekniklerinde ise veriler, eğitim ve test verisi olarak ayrılarak zaman serisi verileri arasındaki ağırlıklar en optimum seviyede belirlenir. Gelecek verilerin tahmin edilmesi elde edilen ağırlık değerleri ile hesaplanmasına dayanmaktadır [3].

Zaman serisi analizi ile yapılan çalışmalar incelendiğinde farklı veriler üzerinde kullanılan istatistiksel ve yapay zeka yöntemlerinin kullanıldığı bir çok çalışma yer almaktadır [4]-[8]. Nie vd. yapmış oldukları çalışmada, ARIMA ve SVM modellerini kullanarak kısa dönemli zaman serisi analizi yapmışlardır [9]. Zhi ve Li çalışmalarında, hisse senedi piyasası fiyat tahminlerinin doğruluğunu artırmak için, ampirik mod ayrışımını, uzun kısa vadeli bellekle birleştiren hibrit model önermişlerdir [10]. Diğer bir çalışmada, gürültülü finansal veriler üzerinde eğilim tahminlerini derin öğrenme teknikleri ile gerçekleştirmişlerdir [11]. Siami-Namini vd., Uzun Kısa Süreli Bellek (Long Short-Term Memory LSTM) gibi zaman serisi verilerini tahmin etmek için yeni geliştirilen derin öğrenme tabanlı algoritmaların geleneksel algoritmalarından üstün olup olmadığı araştırmışlardır. LSTM gibi derin öğrenmeye dayalı algoritmaların ARIMA modeli gibi geleneksel tabanlı algoritmalarından daha iyi performans gösterdiğini bildirmişlerdir [12].

Bu çalışmada, Brent Petrol fiyatları kullanılarak istatistiksel yöntem olan ARIMA ve derin öğrenme tekniği olan LSTM teknikleri kullanılarak zaman serisi

analizi yapılmıştır. Zaman serisi analizinde 6 aylık verinin tahmin edilmesi sağlanmıştır. Elde edilen sonuçlara göre uzun yıllara ait veri kullanımından dolayı LSTM modeli diğer modellere göre daha iyi tahminlerde bulunmuştur.

## 2. Zaman Serisi Analiz Metotları

Zaman serisi analizinde verilerin zamana bağlı değişimleri incelenmektedir. Bu verilerdeki değişimlerin analiz edilmesinde aykırı verilerin ortaya çıkarılması, ileriye yönelik olası olayların tahmin edilmesi, eksik verilerin tamamlanması ve ortaya çıkarılan aykırı verilerin düzeltilmesi gibi analizlerde kullanılmaktadır [13]. Bu çalışma kapsamında ileriye yönelik olarak Brent petrol fiyatlarının tahmini yapılmıştır. Bu tahminleri gerçekleştirebilmek için 2 farklı yöntem önerilmiştir. Bu makalede kullanılan ve çalışılan karşılaştırmalı zaman serisi tekniklerinin performans analizleri karşılaştırılmıştır. Bu kapsamda Otoregresif Entegre Hareketli Ortalama (ARIMA-Autoregressive Integrated Moving Average) ve derin öğrenmeye dayalı teknik, Uzun Kısa Süreli Bellek (LSTM-Long Short-Term Memory) hakkında bilgiler sunulmuştur. Bu yöntemler; istatistiksel ve derin öğrenme yöntemi olarak bilinen ARIMA ve LSTM' dir. Bu yöntemler alt başlıklar şeklinde verilmiştir.

### 2.1. Otoregresif Entegre Hareketli Ortalama (ARIMA) Yöntemi

Otoregresif Entegre Hareketli Ortalama Modeli (ARIMA), Otoregresif (AR) ve Hareketli Ortalama (MA) süreçlerini birleştiren ve zaman serilerinin birleşik bir modeli ile elde edilmiştir. ARIMA modelin 3 ana bileşeni (p, d, q) bulunmaktadır. Burada p; bir gözlem ile birkaç gecikmeli gözlem arasındaki bağımlılıkları elde ederken, d; farklı zamanlarda gözlemlerin farklılıklarını ölçerek zaman serilerini durağan hale getirmekte ve q ise; gecikmiş gözlemler için hareketli bir ortalama model kullanarak, gözlemler ve kalan hata terimleri arasındaki bağımlılıkları hesaplamaktadır [14].

$$x_t = c + \sum_{i=1}^p \phi_i x_{t-i} + \epsilon_t \quad (1)$$

Denklem 1' e göre,  $x_t$  durağan değişkendir,  $c$  sabittir, içindeki terimler  $\phi_i$  gecikmelerdeki otokorelasyon katsayılarıdır.  $\epsilon_t$  ise Gauss beyaz gürültü serisidir [15].

$$x_t = \mu + \sum_{i=0}^q \theta_i \epsilon_{t-i} \quad (2)$$

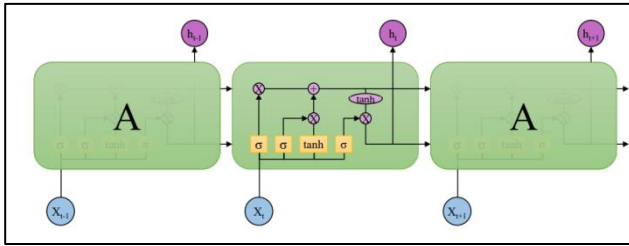
Denklem 2' ye göre  $x_t$ ' ye ait  $\mu$  genellikle sıfıra eşit olduğu varsayılır.  $\theta$  ise zaman serisindeki bir verinin mevcut ve önceki değerlerine uygulanan ağırlık

değeridir.  $\epsilon_{t-1}$  ise ortalama sıfır varyanslı bir Gauss beyaz gürültü serisidir [16].

Mevsimsel zaman serisi verileriyle, kısa vadeli mevsimsel olmayan bileşenlerin modele katkıda bulunması muhtemeldir. Bu nedenle, çarpımsal bir modelde hem mevsimsel olmayan hem de mevsimsel faktörleri içeren mevsimsel ARIMA modelini tahmin edilmesi uygun olacaktır. Mevsimsel ARIMA modelini tahmin etmenin en önemli adımı, (p, d, q) değerlerini belirlemektir. Verinin zaman grafiğine bağlı olarak, örneğin varyans zamanla büyürse, varyans dengeleyici dönüşümler ve farklılaşma kullanılmalıdır. Ardından, bir gecikme ile ayrılmış bir zaman serisindeki gözlemler arasındaki doğrusal bağımlılık miktarını ölçmek için genellikle otokorelasyon fonksiyonunu (ACF) kullanılmaktadır [16]. Çalışmada (p, d, q) değerleri verilerin durağanlığı açısından analiz edilmiş ve en uygun değer olarak her üç parametre için 1 değeri belirlenmiştir.

## 2.2. LSTM Yöntemi

LSTM yöntemi uzun kısa süreli bellek ağlarında kullanılan Tekrarlayan Sinir Ağlarının (Recurrent Neural Networks (RNN)) bir çeşididir. İlk olarak, 1997'de Hochreiter ve Schmidhuber tarafından tanıtılmıştır [17]. LSTM'ler, uzun vadeli bağımlılık problemlerinden kaçınmak için tasarlanmıştır. LSTM sıklıkla sıralı zaman serisi verilerini işlemeden kullanılmaktadır. Şekil 1'de LSTM modeli verilmiştir. [18-19]



Şekil 1. LSTM'deki Etkileşimli Dört Katman

Şekil 1'de görüldüğü gibi sarı kutular katmanları, pembe daireler ise vektör eklenmesi gibi noktasal işlemleri temsil etmektedir. LSTM şekildedeki gibi giriş, unutmaya, çıkış olmak üzere üç kapıya ve çıkış aktivasyon fonksiyonlarına sahiptir (blok giriş fonksiyonu). Aktivasyon fonksiyonları sayesinde LSTM 0 veya 1 değerini alarak bellekteki bilgi ile ne yapılacağına karar vermektedir. Çıkış değeri 1 ise, "bilgiyi tut", 0 ise "bilgiyi sil" anlamına gelmektedir. LSTM'nin ilk adımında sigmoid( $\sigma$ ) tabaka tarafından hangi bilginin bloktan çıkarılacağına karar veren unutmaya katmanı görev yapar. İkinci adımda ise bellekte hangi yeni bilginin depolanacağına karar verilmektedir. Kararı vermek için önce, "giriş katmanı" olarak adlandırılan bir sigmoid

katmanı hangi değerlerin güncelleneceğine karar verir. Sonra *tanh* katmanı yeni aday değer vektörlerini oluşturur. Sonra durumu güncellemek için bu katmanlardaki bilgiler birleştirilir ve yeni bilgi hafızaya eklenmek için bekletilir. Son olarak da çıkışa hangi bilgilerin gönderileceğine karar verilmektedir [17]. Bu çalışmada kullanılan LSTM modeli ile ilgili özellikleri Tablo 1' de verilmiştir.

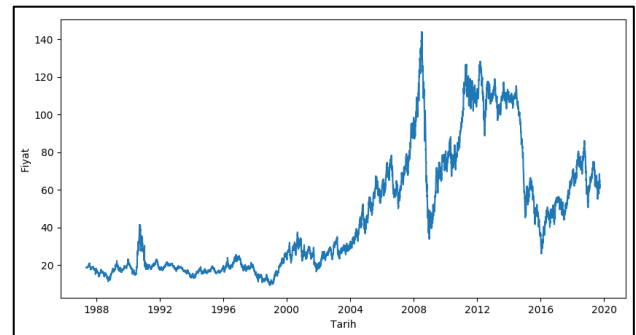
Tablo 1. Çalışmada kullanılan LSTM modelinin Özellikleri

Katman Türü	Değer
Giriş Katmanı	60 Nöron
Ara Katman Sayısı	16 Nöron
Çıkış Katmanı	1 Nöron
Aktivasyon fonksiyonu	"relu"
İterasyon Sayısı	10
Batch Boyutu	15
Optimizasyon	"Adam"

Tablo 1 incelendiğinde çalışmada kullanılan LSTM modelinin giriş, ara ve çıkış katmanı olmak üzere üç katmanda olduğu görülmektedir. Giriş katmanında 60 nöron, ara katmandan 16 nöron ve çıkış katmanında da 1 nöron olmak üzere toplam 77 nöron kullanılmıştır. Çalışmada aktivasyon fonksiyonu olarak "relu" kullanılmasının sebebi verimli bir hesaplama yükü sağlamak için kullanılmıştır. İterasyon sayısının 10 ile sınırlı olmasını sebebi 10. iterasyondan sonra hata değerinin çok fazla değişim göstermemesidir. Çalışmada Adam optimizasyon yönteminin tercih edilmesinin sebebi RMSProp ve Momentum'un bir kombinasyonu olmasıdır. Adam optimizasyon yöntemi ile her parametre için uyarlanabilir öğrenme oranını hesaplanabilmektedir.

## 3. Sonuçlar

Gerçekleştirilen zaman serisi analizinde ARIMA modeli ile LSTM modeli karşılaştırılmıştır. Modellerde kullanılan veriler 21-05-1987 ile 09-09-2019 tarihleri arasındaki günlük Brent Petrol fiyatları kullanılmıştır. Eğitimde kullanılan Brent Petrol fiyatları veri setinin zamana bağlı değişimi Şekil 2'de verilmiştir.



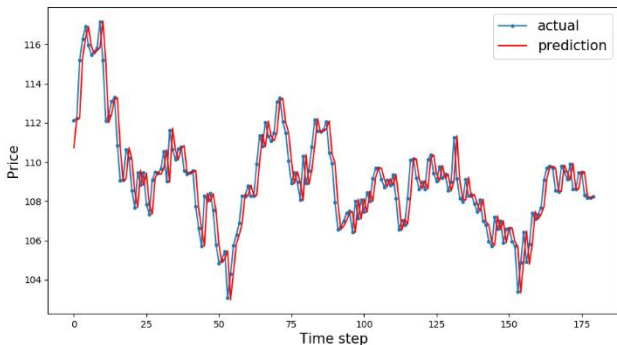
Şekil 2. Eğitimde kullanılan veri setinin zaman bağlı olarak değişimi

Çalışmada, Brent Petrol' ün 1987 ile 2020 yılları arasındaki işlem gören günlere ait toplam 8216 adet veri kullanılmıştır. Bu verilerin %80' i (6573) eğitim için kullanılırken %20' si (1643) test verisi olarak kullanılmıştır. Veri Seti için kullanılan LSTM modelinde giriş, ara ve çıkış olmak üzere 3 katmanda Python programlama dilinde hazırlanan bir ara yazılım ile 15873 parametre eğitilmiştir. Eğitim sırasında kayıp herhangi bir parametre tespit edilmemiştir. Eğitim esnasında her bir iterasyon 20 sn altında gerçekleşmiştir. LSTM modelinden korelasyon katsayısı performans değerlendirme ölçütüne göre %99 doğruluk oranı elde edilmiştir. ARIMA modelinde verilerde herhangi bir ön işleme süreci yapılmamıştır. LSTM modelinde ise veriler (-1,1) arasında minimum maksimum skaler fonksiyonu ile ön işlemden geçirilmiştir. Sonrasında Adam optimizer fonksiyonu ile veriler optimize edilmiştir. LSTM modelinde relu aktivasyon fonksiyonu kullanılmıştır. Her iki model ile 180 günlük ileriye dönük tahminler yapılarak sonuçlar elde edilmiştir. Modellerin verimliliği Ortalama Hata Karelerinin Karekökü (RMSE), Mutlak Hataların Ortalaması (MAE) ve Korelasyon katsayısı (R) performans değerlendirme ölçütlerine göre değerlendirilerek sonuçlar Tablo 2' de verilmiştir.

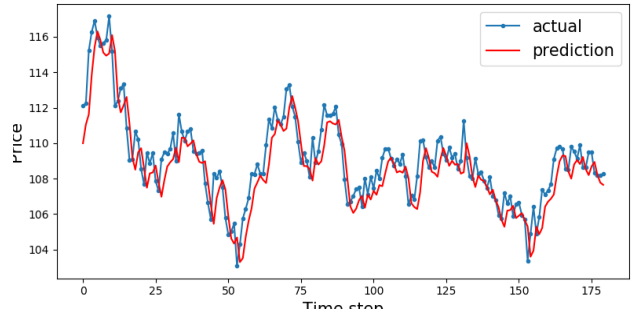
Tablo 2. Önerilen modellerin performansları

Modeller	RMSE	MAE	R
LSTM	1.189	0.886	0.99
ARIMA	1.512	1.164	0.87

Tablo 2' de görülmekte olan ARIMA ile elde edilen sonuçlara göre, RMSE ve MAE değeri sırasıyla 1.189 ve 0.886 olarak elde edilmiştir. LSTM ile elde edilen sonuçlara göre RMSE ve MAE değerleri sırasıyla 1.512 ve 1.164 olarak elde edilmiştir. Elde edilen RMSE ve MAE değerleri çalışmada kullanılan veri setinin LSTM modeli ile uyumlu olduğunu göstermektedir. Uygulanan modellere göre hesaplanan ve tahmin edilen değerleri gösteren grafik Şekil 3 ve Şekil 4' de gösterilmiştir.



Şekil 3. LSTM Modeli 180 günlük zaman serisi analizi



Şekil 4. ARIMA Modeli 180 günlük zaman serisi analizi

Şekil 4' ye göre ARIMA modelinde mevcut değerler ile tahmin edilen değerler LSTM modeline göre aralıklı olarak elde edilmiştir. Her iki modelde de Brent Petrol fiyatlarının trendi yakalandığı görülmektedir. Petrol fiyatlarının yükseldiği dönemlerde her iki modelinde fiyat artışı ve düşüşlerini tahmin ettiği şekillere göre söylenebilir.

Gelecek çalışmalarda kullanılan ARIMA ve LSTM modelinden farklı yapay zekâ modelleri kullanılarak sonuçların tartışılacağı çalışmaların yapılması düşünülmektedir.

## Teşekkür

Çalışmada kullanılan veri seti açık erişimli İnternet sitesi (Kaggle) ekleyen Mustafa Abusalah' a teşekkür ederiz.

## Kaynaklar

- [1] Sezer ÖB. *Zaman serisi verilerinin derin yapay sinir ağları ile analizi ve eniyilemesi: finansal tahmin algoritmaları*. PhD Thesis, TOBB Ekonomi ve Teknoloji Üniversitesi, Ankara, Turkey, 2018.
- [2] Tsay RS. *Analysis of Financial Time Series*. Third Edition, A John Wiley & Sons, Inc., Publication, Chicago, 2010.
- [3] Sulistiyo, MD., Dayawati, RN. Evolution strategies for weight optimization of Artificial Neural Network in time series prediction. *International Conference on Robotics, Biomimetics, Intelligent Computational Systems*, 143-147. 2013.
- [4] Saab S, Badr E, Nasr G. Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon. *Energy*, 26(1), 1-14, 2001.
- [5] Kaytez F, Taplamacioglu MC, Cam E, Hardalac F. Forecasting electricity consumption: A comparison of regression analysis, neural networks and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, 67, 431-438, 2015.
- [6] Kankal M, Akpınar A, Kömürcü Mİ, Özşahin TŞ. Modeling and forecasting of Turkey's energy consumption using socio-economic and

- demographic variables. *Applied Energy*, 88(5), 1927-1939, 2011.
- [7] Şimşek O, Gümüş V, Soydan NG, Yenigün K, Kavşut ME, Topçu E. Hatay İlinde Bazı Meteorolojik Verilerin Gidiş Analizi. *Uluslararası Teknolojik Bilimler Dergisi*, 5(2), 132-144, 2013.
- [8] Sen PK. Estimates of the Regression Coefficient Based on Kendall's Tau. *Journal of the American Statistical Association*, 63(324), 1379 – 1389, 1968.
- [9] Nie H, Liu G, Liu G, Wang. Hybrid of ARIMA and SVMs for Short-Term Load Forecasting. *Energy Procedia*, 16(C), 1455-1460, 2012.
- [10] Cao J, Li Z, Li J. Financial time series forecasting model based on CEEMDAN and LSTM. *Physica A: Statistical Mechanics and its Applications*, 519 (2019), 127-139, 2019.
- [11] Moews B, Herrmann JM, Ibikunle G. Lagged correlation-based deep learning for directional trend change prediction in financial time series. *Expert Systems with Applications*, 120, 197-206, 2019.
- [12] Siami-Namini S, Tavakoli N, Siami Namin A. A Comparison of ARIMA and LSTM in Forecasting Time Series. 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, 2018.
- [13] Seker SE. Zaman Serisi Analizi (Time Series Analysis). *YBS Ansiklopedi*, 2(4), 23-31, 2015.
- [14] Kaynar O, Taştan S. Zaman Serisi analizinde Mlp Yapay Sinir Ağları ve Arıma Modelinin Karşılaştırılması. *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, (33), 161-172, 2009.
- [15] Uyanık, Ü. (2014). Zaman serilerinde yapısal kırılma ve vergi affı üzerine bir uygulama (Doctoral dissertation, DEÜ Sosyal Bilimleri Enstitüsü).
- [16] Tortum, A., Gözcü, O., & Çodur, M. Y. (2014). Türkiye'de Hava Ulaşım Talebinin Arıma Modelleri ile Tahmin Edilmesi. *İğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 4(2), 39-54.
- [17] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [18] Olah, C. (2015, August 27). Understanding LSTM Networks. [colah.github.io: colah.github.io/posts/2015-08-UnderstandingLSTMs/](https://colah.github.io/posts/2015-08-UnderstandingLSTMs/) Erişim tarihi 18.11.2020.
- [19] Temür, A. S. (2019). İşletmelerin satış bütçelerinin oluşturulmasında arıma, lstm ve hibrit modellerin karşılaştırılması: üretim işletmesi örneği.