

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

COVID-19 PANDEMİSİ ETKİSİNDEKİ HAVAYOLU YOLCU
TALEBİNİN YAPAY ZEKA KULLANILARAK TAHMİN
EDİLMESİ

Nalan YÜCE
YÜKSEK LİSANS TEZİ
İŞLETME ANABİLİM DALI

GEBZE
2022

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ

COVID-19 PANDEMİSİ ETKİSİNDEKİ
HAVAYOLU YOLCU TALEBİNİN YAPAY
ZEKA KULLANILARAK TAHMİN
EDİLMESİ

Nalan YÜCE
YÜKSEK LİSANS TEZİ
İŞLETME ANABİLİM DALI

Tez Danışmanı
Prof. Dr. Bülent SEZEN

GEBZE

2022

	YÜKSEK LİSANS JÜRİ ONAY FORMU
---	--------------------------------------

GTÜ Sosyal Bilimler Enstitüsü Yönetim Kurulu'nun 10/02/2022 tarih ve 2022/06 sayılı kararıyla oluşturulan jüri tarafından 25/02/2022 tarihinde tez savunma sınavı yapılan Nalan Yüce'nin tez çalışması İşletme Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Bülent Sezen

ÜYE

: Mehmet Şahin Gök

ÜYE

: Özge Nalan Bilişik

ONAY

Gebze Teknik Üniversitesi Sosyal Bilimler Enstitüsü Yönetim Kurulu'nun

...../...../..... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

ÖZET

2019 yılında Çin’de ortaya çıkan Covid-19 salgını kısa sürede tüm Dünya’da yayılmıştır. Salgın sebebiyle küresel çapta birçok sektör olumsuz etkilenmiştir. Havayolu yolcu taşımacılığı da Covid-19 salgınından en yoğun etkilenen sektörlerden biridir. Havayolu yolcu mili havacılık sektöründe sıkça kullanılan bir metrik olup toplam uçulan mil ile toplam yolcu sayısının çarpımı ile elde edilir. Havayolu yolcu mili metriği ile sektördeki hareketlilik ölçülebilmektedir. Bu çalışmada Amerika Birleşik Rezerv Ekonomik Veri (FRED) sisteminden alınan Amerika havayolu yolcu mili metriğine ait 2000 ile 2021 yılları arasında toplam 259 veri kullanılmıştır. Kullanılan veri seti yukarı doğru artış eğilimi barındırdığı için durağan özellik göstermemektedir. Bunun yanında yılın bazı mevsimlerinde artan, bazı mevsimlerinde de tam tersine azalan bir yapıya yani, mevsimselliğe sahip olduğu gözlemlenmiştir. Derin öğrenme metodlarından Tekrarlayan Sinir Ağlarının (RNN) LSTM (Uzun Kısa Dönem Hafıza) mimarisinin durağan olmayan veri seti ile çalışabilmesi ve mevsimsellik etkilerini hesaba katabiliyor olması sebebiyle bu çalışmada tercih edilmiştir. Covid-19 döneminde veri setinin eğiliminde meydana gelen ani değişimin LSTM mimarisinin performansına etkisini gözlemleyebilmek amacıyla hem Covid-19 dönemi verilerini içeren veri seti hem de Covid-19 dönemini içermeyen veri seti ile iki ayrı tahmin yapılmış ve sonuçlar kıyaslanmıştır. Bulgulara göre, Covid-19 dönemini içermeyen veri seti ile yapılan tahminlerde LSTM mimarisinin performansının çok daha yüksek olduğu görülmüştür. Aynı veri setinin AR (Oto regresif) ile de tahmini yapılmış ve LSTM mimarisinin performansı ile kıyaslanmıştır. Son olarak LSTM mimarisi ile 1960-2020 yılları arasında Türkiye’ye ait yolcu sayısı verileri ile tahmin yapılmıştır.

Anahtar Kelimeler: Havayolu Yolcu Mili, LSTM Modeli, Zaman Serileri.

SUMMARY

The Covid-19 epidemic, which emerged in China in 2019, spread all over the world in a short time. Due to the epidemic, many sectors around the world have been adversely affected. Airline passenger transport is also one of the sectors most heavily affected by the Covid-19 outbreak. Airline passenger mile is a frequently used metric in the aviation industry and is obtained by multiplying the number of miles flown by the number of passengers. With the airline passenger mile metric, the mobility in the sector can be measured. In this study, a total of 259 data for the American airline passenger mile metric taken from the United Reserve Economic Data (FRED) system between 2000 and 2021 were used. Since the data set used has an upward trend, it does not show a static feature. In addition, it has been observed that it has a seasonality that increases in some seasons of the year and decreases in some seasons, on the contrary. LSTM (Long Short-Term Memory) architecture of RNN (Recurrent Neural Networks), which is one of the deep learning methods, was preferred in this study because it can work with non-stationary data set and can take seasonal effects into account. In order to observe the effect of the sudden change in the trend of the data set during the Covid-19 period on the performance of the LSTM architecture, two separate estimations were made with both the data set containing the Covid-19 period data and the data set not containing the Covid-19 period, and the results were compared. According to the findings, it was seen that the performance of the LSTM architecture was much higher in the predictions made with the data set that did not include the Covid-19 period. The same data set was also estimated with the AR (Autoregressive) and compared with the performance of the LSTM architecture. Finally, with the LSTM architecture, estimations were made with the number of passengers in Turkey between the years 1960-2020.

Key Words: Airline Passenger Mile, LSTM Model, Time Series.

TEŞEKKÜR

Yüksek lisans eğitimimde desteğini ve yardımlarını hiçbir zaman esirgemeyip bilgisi ile bu çalışmanın oluşmasının yolunu açan danışmanım Prof. Dr. Bülent Sezen'e teşekkürlerimi sunarım.



İÇİNDEKİLER

ÖZET	V
SUMMARY	VI
TEŞEKKÜR	VII
İÇİNDEKİLER	VIII
SİMGELER VE KISALTMALAR DİZİNİ	IX
ŞEKİLLER DİZİNİ	X
TABLolar DİZİNİ	XI
1. GİRİŞ	1
2. LİTERATÜR TARAMASI	4
3. ZAMAN SERİSİ TAHMİNİ	6
4. YAPAY SİNİR AĞLARI	8
4.1 Yapay Sinir Ağır Mimarileri	9
4.1.1 İleri Beslemeli Yapay Sinir Ağları	9
4.1.2 Geri Beslemeli Yapay Sinir Ağları	9
4.2 Yapay Sinir Ağı Öğrenme Kuralları	10
4.2.1 Danışmansız Öğrenme	10
4.2.2 Danışmanlı Öğrenme	11
4.3 Aktivasyon Fonksiyonları	11
5. RNN MODELLERİ VE LSTM	13
6. AR SÜRECİ	17
7. UYGULAMA	18
8. BULGULAR VE TARTIŞMA	22
9. SONUÇLAR VE ÖNERİLER	29
KAYNAKLAR	31
ÖZGEÇMİŞ	34
EKLER	35

SİMGELER VE KISALTMALAR DİZİNİ

Simgeler ve Açıklamalar

Kisaltmalar

LSTM	: Uzun Kısa Dönemli Hafıza (Long-Short Term Memory)
RNN	: Tekrarlayan Sinir Ağları (Recurrent Neural Network)
MLP	: Çok Katmanlı Algılayıcı (Multilayer Perceptron)
FRED	: Amerika Birleşik Rezerv Ekonomik Veri
AR	: Otoregresif (Autoregressive)
p	: İstatiksel p değeri
TÜİK	: Türkiye İstatistik Kurumu
CNN	: Evrişimli Sinir Ağları
ARDL	: Otoregresif dağıtılmış gecikmeler
AB	: Avrupa Birliği
OECD	: İktisadi İşbirliği ve Kalkınma Teşkilatı (Organisation for Economic Co-operation and Development)
TOPSIS	: İdeal Çözüme Benzerliğe Göre Sıra Tercihi Tekniği (Technique For Order Preference by Similarity to An Ideal Solution)
MA	: Hareketli Ortalama
FFNN	: İleri Beslemeli Sinir Ağları (Feed Forward Neural Network)
RMSE	: Kök Ortalama Kare Hata (Root Mean Squared Error)
YSA	: Yapay Sinir Ağları
ARIMA	: Hareketli Ortalama ile Birleştirilmiş Otoregresif
MSE	: Ortalama Kare Hata

ŞEKİLLER DİZİNİ

<u>Sekil No:</u>	<u>Sayfa</u>
4.1: Biyolojik nöron yapısı	8
4.2: İleri beslemeli ağ yapısı	9
4.3: Geri beslemeli ağ yapısı	10
4.4: Danışmansız öğrenme yapısı	11
4.5: Danışmanlı öğrenme yapısı	11
4.6: Lineer aktivasyon fonksiyonu	12
4.7: Sigmoid aktivasyon fonksiyonu	12
4.8: Tan-sigmoid aktivasyon fonksiyonu	12
5.1: RNN yapısı	14
5.2: LSTM yapısı	15
7.1: Covid-19 dönemini içeren veri seti-Amerika	19
7.2: Verinin trend grafiği	19
7.3: Verinin mevsimsellik grafiği	20
7.4: Türkiye havayolu yolcu sayısı grafiği	21
8.1: Covid dönemi hariç modelin kayıp değerleri	22
8.2: Covid dönemi hariç modelin tahmini-Amerika	23
8.3: Covid dönemi dahil modelin kayıp değerleri-Amerika	24
8.4: Covid dönemi dahil modelin tahmini-Amerika	24
8.5: Covid dönemi hariç-12 aylık hareketli ortalama	25
8.6: Covid dönemi hariç modellerin tahmin değerleri	25
8.7: Covid dönemi dahil-12 aylık hareketli ortalama	26
8.8: Covid dönemi dahil modellerin tahmin değerleri	26
8.9: Covid dönemi hariç modelin tahmini-Türkiye	27
8.10: Covid dönemi dahil modelin tahmini-Türkiye	28

TABLÖLAR DİZİNİ

<u>Tablo No:</u>	<u>Sayfa</u>
4.1: Terminoloji kıyaslaması	8
7.1: Genel veri seti-Amerika	18
7.2: Genel veri seti-Türkiye	20
8.1: Oluşturulan LSTM mimarisi-Amerika	22
8.2: Covid dönemi hariç modelin tahmin değerleri-Amerika	23
8.3: Covid dönemi dahil modelin tahmin değerleri-Amerika	24
8.4: Oluşturulan LSTM mimarisi-Türkiye	27
8.5: Covid dönemi hariç tahmin değerleri-Türkiye	27
8.6: Covid dönemi dahil tahmin değerleri-Türkiye	28
9.1: Yapılan tahminlerin Kök Ortalama Kare Hata değerleri	30

1. GİRİŞ

Covid-19 salgını 2019 Aralık ayından itibaren kısa sürede yayılarak, tüm Dünyada etkisini göstermiştir. Covid-19 ile en çok etkilenen sektörlerden birisi olan havayolu yolcu taşımacılığı bu çalışmada incelenmiştir. Havayolu yolcu mili verileri zamana bağlı değişen verilerdir. Zamana bağlı değişen veriler literatürde zaman serileri verileri olarak adlandırılmaktadır. Zaman serisi tahmini, aynı değişkenin geçmiş gözlemlerinin toplandığı ve altta yatan ilişkiyi tanımlayan bir model geliştirmek için analiz edildiği önemli bir tahmin alanıdır (Zhang, 2003). Son zamanlardaki gelişmeler ile zaman serileri tahmininde sık sık YSA (Yapay Sinir Ağları) kullanılmaktadır (Abbasimehr et al., 2020).

LSTM, derin öğrenme alanında kullanılan yapay bir RNN mimarisidir. Sıralı bir zaman serisi modeli oluşturmak için verimli bir algoritmadır (Luo et al., 2021). LSTM daha önceki verileri hafızaya alarak gelecek tahmini yapabilen bir mimaridir (Chimmula and Zhang, 2020). LSTM mimarileri, verilerin zamansal yönlerini modellemedeki doğal yeteneği nedeniyle makine çevirisi, dil modelleme, el yazısı tanıma, resim yazısı, konuşma tanıma, süreç tahmini ve tıbbi teşhiste yaygın olarak kullanılmaktadır (Somu et al., 2020). LSTM mimarisi literatürde farklı alanlarda kullanılarak zaman serisi tahminlerinde iyi bir yöntem olduğunu göstermiştir.

Tüm tahmin modelleri parametrik, parametrik olmayan ve hibrit modeller olarak sınıflandırılabilir. LSTM mimarisi parametrik olmayan modeller sınıfına girmektedir. Bu mimarinin diğer yöntemlere göre avantajları arasında arka plan için ön koşul gerektirmemesi, veriye dayalı otomatik prosedürler, durağan olmayan verileri iyi analiz etme, güçlü genelleme kapasitesi, karmaşık problemlerin çözümüne yakınsamadaki başarısı sayılabilir. Dezavantajları arasında da aşırı uydurma, yerel optimumlara kolayca düşme, kapsamlı veri gerektirmesi sayılabilir (Cao et al., 2021).

Daha önce Choi ve arkadaşı en yoğun havalimanlarından biri olan Hartsfield–Jackson Atlanta Uluslararası Havalimanı'nın saatlik kalkış ve varış kapasitesini tahmin etmek için YSA algoritmaları (MLP–Çok Katmanlı Algılayıcı, RNN ve LSTM) kullanmıştır. Tahmin için kurulan modeller, havalimanının 1 Ocak 2013 ile 31 Aralık 2017 arasında saatlik kapasite ve hava durumu verileri kullanılarak eğitilmiştir (Choi and Kim, 2021). Brown ve arkadaşı yaptıkları çalışmada Amerikan havayolu sektörünün gelecek tahmininde dış faktörler karşısında yetersiz kaldıklarını ve Covid-19 salgınında da başarısız süreç yönetimi yaptıklarını göstermiştir (Brown and Kline,

2020). Hotle ve arkadaşı Covid-19 döneminde Amerika havalimanlarındaki operasyonlarının nasıl etkilendiğini incelemiş, Mayıs 2020'deki uçuşların 2019 uçuşlarına göre %70 azaldığını göstermiştir (Hotle and Mumbower, 2021).

Literatür incelediğinde ani değişimlerde havayolu sektörünün hazırlıksız yakalandığı ve büyük kayıplara uğradığı görülmektedir. Bu doğrultuda geleceğe yönelik tahmin metotları ile veriler incelenerek havayolu sektöründe gelecek projeksiyonunun daha iyi yapılması mümkün kılınabilir ve sektörün uğradığı hasar minimuma indirilebilir.

Bu çalışmada kapsamında LSTM mimarisinin Amerika havayolu yolcu milinde tahmin performansı incelenerek havayolu sektöründe gelecek projeksiyonu yapılması için doğru yöntemlerin bulunmasına destek sağlanmıştır. Veri seti FRED sisteminden alınmıştır. FRED 96 kaynak ve 765.000'den fazla ekonomik zaman serisine sahip St. Louis Federal Rezerv Bankası'nın Araştırma bölümü tarafından sağlanan bir veri tabanıdır. 2000 ve 2021 yılları arasındaki Amerika havayolu yolcu mili sayısına ait 259 veri, bu çalışma kapsamında incelenmiştir. Havayolu yolcu mili, toplam uçulan mil ile toplam yolcu sayısının çarpımı ile elde edilir. Havayolu yolcu taşımacılığı için önemli bir metriktir. Bu çalışmada Covid-19 etkisindeki havayolu sektörü yapay sinir ağlarından LSTM mimarisi aracılığıyla incelenmiştir. Havayolu yolcu mili verisi yukarı yönde artış trendi göstermektedir. Yılın belli dönemlerinde artış belli dönemlerinde azalış göstermesiyle de mevsimsel etkiler barındırmaktadır. Veri seti Augmented Dickey-Fuller test ile analiz edilmiş, p değeri 0,97 olarak hesaplanmış ve durağan olmadığı görülmüştür. Mevsimsel ayrıştırma (decompose) ile veri ayrıştırılıp trend, mevsimsellik ve artık değerlerin grafikleri çizdirilmiştir. Bu grafikler sayesinde veri doğru yorumlanarak kurulacak modele temel oluşturulmuştur.

Covid-19 dönemini içeren ve Covid-19 dönemini içermeyen veri seti incelenerek bir yıla ait tahmin yapılmış ve sonuçlar kıyaslanmıştır. Çalışmanın amacı zaman serilerinde ani değişimlerde LSTM mimarisinin performansını incelemektir. RMSE (Kök Ortalama Kare Hata) değerleri Amerika Covid-19 dönemini içeren veri setinde $3,4E+07$ olarak, Covid-19 dönemini içermeyen veri setinde $2,3E+6$ olarak hesaplanmıştır. Buradan açıkça görülmektedir ki veri setinde ani değişim olduğunda LSTM mimarisinin başarı performansı büyük oranda düşmektedir. AR modeli ile de veri setinde tahmin yapılmış ve sonuçlar LSTM mimarisi ile kıyaslanmıştır. Sonrasında LSTM mimarisi ile Türkiye için de yolcu sayısı tahmini yapılmıştır.

Çalışmanın devamında havayolu sektörüne ve zaman serileri tahmin yöntemlerine ilişkin literatüre, üçüncü bölümde zaman serisi tahminine, dördüncü, beşinci ve altıncı bölümde yönteme, yedinci bölümde problemin uygulamasına, sekizinci bölümde bulgular ve tartışmaya, dokuzuncu bölümde ise sonuç ve öneriler kısmına yer verilmiştir.



2. LİTERATÜR TARAMASI

Yapılan çalışma Amerika havayolu yolcu mili sayısının gelecek dönem tahmini üzerine yoğunlaşmıştır. Covid-19 döneminde ani değişim gösteren havayolu sektörü için LSTM yöntemi kullanılmış olup ek olarak AR modeli ile sonuçlar kıyaslanmıştır. Sonrasında LSTM mimarisi ile Türkiye için de tahmin çalışmaları yürütülmüştür. Literatürde birçok farklı yöntem ile zaman serileri analizi yapılmış ve geleceğe yönelik tahminlerde başarı gösterilmiştir.

Yang ve arkadaşlarının yaptıkları çalışmada hava sahası yoğunluk indeksi tahmini için yeni bir model geliştirmiştir. Regresyon modelleri için, grafik evrişim ve LSTM mimarisi, uzay-zamansal özellik çıkarma ve hava sahası yoğunluk indeksi tahmini için entegre edilmiştir. Sınıflandırma modelleri için, XGBoost yöntemi ayrıca hava sahası yoğunluğu tanımlaması için dahil edilmiştir (Yang et al., 2021). Karami ve arkadaşı akıllı ulaşım planlaması için zaman serileri tahmininde derin öğrenme, ARIMA (Hareketli Ortalama ile Birleştirilmiş Otoregresif), Holt Winters' Exponential Smoothing gibi yöntemleri kıyaslamıştır. Gelecek yıllarda akıllı ulaşım planlamasında kullanılabilecek çeşitli veri kaynakları, kurulabilecek modeller ve yeni iş fırsatları çalışma kapsamında belirtilmiştir (Karami and Kashef, 2020). Bi ve arkadaşları yaptıkları çalışmada görüntü işlemede derin öğrenmenin faydalarından yararlanarak turizmde talep tahmini için farklı modeller kurarak kıyaslama yapmıştır. LSTM ve CNN (Evrişimli Sinir Ağı) yöntemleri birlikte kullanılarak, görüntü işleme de talep tahmini konusuna dahil edilmiştir (Bi et al., 2021). Cao ve arkadaşları yüksek hızlı tren taşımacılığı için talep tahmini yaparken artık bileşen imhasına dayalı olarak deterministik ve olasılıksal tahmin kapasitelerine sahip yeni bir sistem sunmaktadır (Cao et al., 2021). Abdar ve arkadaşları yaptıkları çalışmada derin öğrenmede belirsizliklerin nicelleştirilmesi için literatür taraması yapmış, teknikleri, uygulamaları ve zorlukları değerlendirmiştir (Abdar et al., 2021). Alassafi ve arkadaşları yaptıkları çalışmada Malezya, Morokko ve Suudi Arabistan'daki Covid-19 vakalarının tahminini LSTM ve RNN yöntemleri ile yapmış ve sonuçları kıyaslamıştır. LSTM mimarisi ile çok daha doğru sonuçlar ortaya çıkartılmıştır (Alassafi et al., 2021). Hanson ve arkadaşları ekonomik aktivitelerdeki değişikliklerin havayolu yolcu trafiğini nasıl etkilediğini incelemiştir. Kazanç ve havayolu taşımacılığı talebinin arasındaki nedensellik araştırılmıştır. ARDL (Otoregresif dağıtılmış gecikmeler) sınır testi ile kısa-uzun vadeli gelir esnekliklerinin talebe etkisi üzerine çalışılmıştır (Hanson et al.,

2021). Singh yaptığı çalışmada havayolu taşımacılığına duyulan bağlılığı literatürde bulunan etmenler aracılığıyla tahmin etmek için model geliştirmiştir. Havayolu yolcularından anket çalışmasıyla veri toplanmış ve yapay sinir ağları teorisi ile kıyaslamasını yapmıştır (Singh, 2021). Truong sinir ağları ve Monte Carlo simülasyonunu kullanarak Covid-19'un havayolu taşımacılığının orta ve uzun vadeli tahmini üzerine çalışmıştır (Truong, 2021). Grosche ve arkadaşları havayolu yolcu talebinin tahmini için ağırlık (gravity) modelini kullanmıştır. Modelde havayolu hizmet değişkenleri yerine uçulan şehirlerin genel ekonomik faaliyetlerini ve coğrafi özelliklerini değerlendirilmiştir (Grosche et al., 2007). Solvoll ve arkadaşları Norveç havalimanlarındaki hava trafiğini tahmin etmek üzere farklı modeller denemiş ve kıyaslamasını yapmıştır. Değişkenlik, analoji ve taşıma yöntemleri ile tahmin yapılarak ekonomik yararları kıyasla bölgesel etkilerin daha önemli olduğu vurgulanmıştır. Analoji yöntemlerinin esneklik yaklaşımına dayalı daha geleneksel yöntemlerden daha uygun olduğu görülmüştür (Solvoll et al., 2020). Iacus ve arkadaşları Covid-19 dönemindeki havayolu yolcu trafiğini incelemiş ve gelecek dönem tahmini yapmıştır. 2010 yılından 2019 yılına kadar olan havayolu verisi homojen olmayan Poisson süreci ile işlenmiş ve 2020 yılı için tahmin yapılmıştır (Iacus vd., 2020). İç ve arkadaşı yaptıkları zaman serileri analizinde etmensele deney tasarım yöntemi ile gayrisafi yurtiçi hasılasının tahmini üzerinde çalışmıştır (İç ve Civelek, 2021). Ünsal ve arkadaşı zaman serisi ve küme analizi kullanarak G20, AB (Avrupa Birliği) ve OECD (İktisadi İşbirliği ve Kalkınma Teşkilatı) ülkelerindeki Covid-19 vakalarını incelemiştir (Ünsal ve Kasap, 2020).

Tüm bu çalışmalar incelendiğinde havayolu taşımacılığının ileriye dönük projeksiyonunun tahminin kritik öneme sahip olduğu görülmüştür. Bu noktada da LSTM yöntemi ile başarılı tahmin sonuçları elde edilmeye çalışılmış ve havayolu sektörünün büyük zararlara uğramadan krizleri atlmasına destek olacak modeller geliştirilmiştir.

3. ZAMAN SERİSİ TAHMİNİ

Zaman serisi analizi, bir zaman aralığında toplanan bir dizi veri setinin analizinin özel bir yoludur. Zaman serisi analizinde, analistler, veri noktalarını yalnızca aralıklı veya rastgele kaydetmek yerine, belirli bir süre boyunca tutarlı aralıklarla kaydeder. Ancak bu tür bir analiz, yalnızca zaman içinde veri toplama eylemi değildir. Zaman serisi verilerini diğer verilerden ayıran şey, analizin değişkenlerin zaman içinde nasıl değiştiğini gösterebilmesidir. Başka bir deyişle, zaman çok önemli bir değişkendir, çünkü verilerin nihai sonuçların yanı sıra veri noktaları boyunca nasıl ayarlandığını gösterir. Ek bir bilgi kaynağı ve veriler arasında belirli bir bağımlılık sırası sağlar. Zaman serisi analizi, tutarlılık ve güvenilirlik sağlamak için tipik olarak çok sayıda veri noktası gerektirir. Kapsamlı bir veri seti, temsili bir örneklem büyüklüğüne sahip olmayı ve bu analizin gürültülü verilerini kesebilmesini sağlar. Ayrıca, keşfedilen trendlerin veya modellerin aykırı değerler olmamasını ve mevsimsel değişiklikleri hesaba katabilmesini sağlar. Ek olarak, zaman serisi verileri, geçmiş verilere dayalı olarak gelecekteki verileri tahmin etmek için kullanılabilir.

Bir zaman serisinin gelecekteki değerlerinin evrimini bilmek, birçok uygulamada önemli avantajlar sağlayabilir. Çünkü bu tahmin, kararların önceden alınmasını ve daha iyi sonuçlar elde etmek için uygun şekilde uyarlanmasını sağlar. Tahminler yeterince kesin ve güvenilir ise, maliyet ve tüketimde tasarruf, emisyonların azaltılması veya lojistik optimizasyon gibi birçok faktörde önemli iyileştirmeler elde etmek mümkündür (Chatigny et al., 2021). Halihazırda, istatistiksel modeller, en yakın komşular veya sinir ağları gibi farklı yaklaşımlar kullanarak zaman serilerini analiz ederek verilerin gelecekteki tahmini hakkında kesin sonuçlar elde etmek için çok çeşitli tahmin yöntemleri bulunmaktadır (Talavera-Llames et al., 2016; Torres et al., 2019; Galicia et al., 2018). Bu nedenle, analiz edilecek zaman serisinin özelliklerine ve kullanılan modelin eğitim veya tahmin süresi gibi gereksinimlerine bağlı olarak her uygulama için en uygun olanı seçmek gerekir.

Son yıllarda, hava kirliliği veya elektrik tüketimi gibi farklı alanlarda zaman serilerinde yapılan tahminlerin davranışını ve kalitesini analiz eden çok sayıda çalışma yayınlanmıştır (Aznarte-Mellado et al., 2007; Torres et al., 2021).

70'lerin sonlarında Box ve Jenkins, matematiksel doğrusal modellerden oluşan uygulamaları incelemek için önemli bir çalışma yaptılar (Chen and Chang, 2009). Bu modeller, AR ve MA (Hareketli Ortalamalar) süreçleri ile temsil edilir. AR

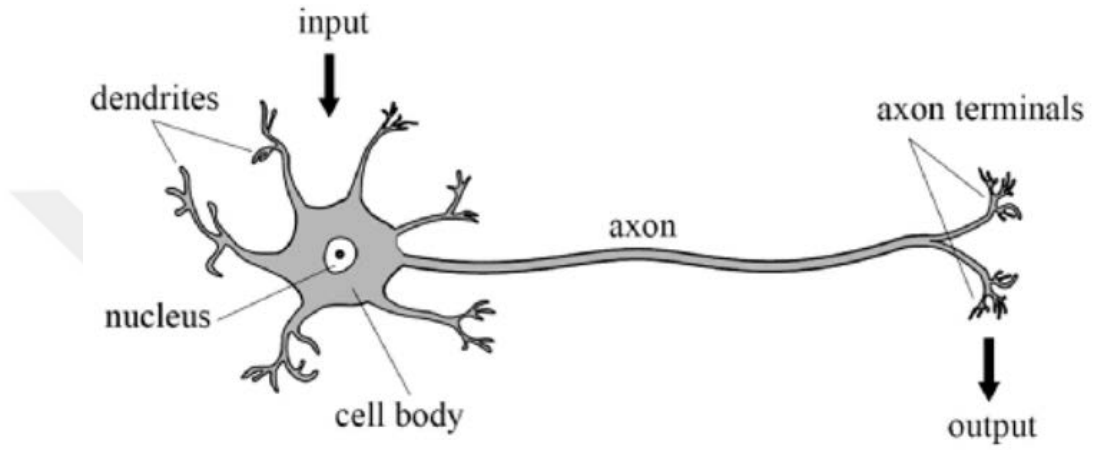
süreçlerinde, zaman serisinin mevcut deęerinin, gemiř deęerlerinin doęrusal bir birleřimi olduęu varsayılır. MA'daki süreçler, mevcut deęerin, seriyi etkileyen rastgele giriřim geiřlerinin veya bozulmaların bir fonksiyonu olduęu varsayılır. Birok pratik deney, bu yaklařımın birok gerek zaman serisinin dinamiklerini temsil etmeyi mümkün kıldığını gstermiřtir. Bu modeller hem akademik hem de profesyonel alanlarda olduka poplerdir (Tealab, 2018).

Bununla birlikte, birok gerek zamanlı serinin doęrusal olmayan davranıřı takip ettięi ve Box-Jenkins'in yaklařımının dinamiklerini temsil etmede yetersiz olduęu da bulunmuřtur (Clements et al., 2004; Granger and Tersvirta, 1993). Doęrusal olmayan modellerde birok model denenmiř ve Yapay Sinir Aęları n plana ıkmıřtır.



4. YAPAY SİNİR AĞLARI

Derin öğrenme teknikleri makine öğrenimini, Şekil 4.1’te de görülen insan beyninin sinir ağlarından esinlenerek, makinelerin görevleri ayırt etmeyi öğrenebileceği yepyeni bir düzeye taşımıştır. YSA insan beyninin çalışma prensibini taklit eder. Bu noktada yapılan hesaplamalar derin öğrenme kavramını ortaya çıkarmıştır. Derin öğrenmenin çıkış noktası algoritmalara dayalı sinir ağlarına bir matematiksel model oluşturulmasıdır.



Şekil 4.1: Biyolojik nöron yapısı.

Bir YSA birbirine ağırlıklı bağlantılar ile bağlı çok sayıda işlemciden oluşur. İşlem düğümleri nöronlar olarak adlandırılabilir. Her düğüm çıkışı, yalnızca düğümde yerel olarak mevcut olan ya dahili olarak depolanan ya da ağırlıklı bağlantılar yoluyla gelen bilgilere bağlıdır. Her birim diğer birçok düğümden girdi alır, çıktısını başka bir düğüme iletir. Tek başına, tek bir işleme ögesi çok güçlü değildir; girdilerinin doğrusal olmayan basit bir fonksiyonu olan tek sayısal değere sahip skaler bir çıktı üretir (Dongare vd., 2012).

Biyolojik ve YSA terminolojisi incelendiğinde birbirine karşılık gelen terimler aşağıda Tablo 4.1’te gösterilmektedir.

Tablo 4.1: Terminoloji kıyaslaması.

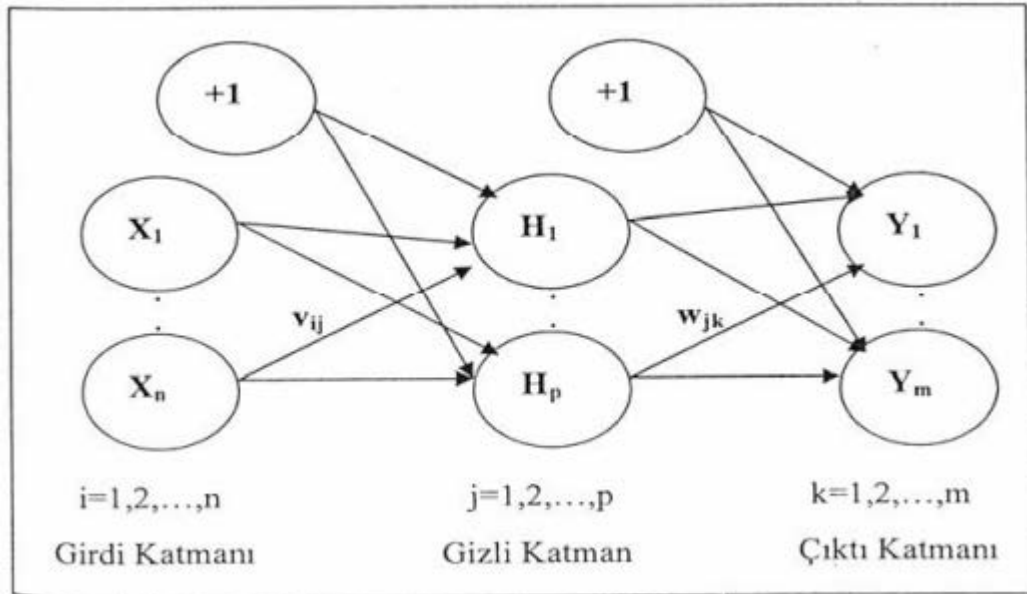
Biyolojik Terminoloji	YSA Terminolojisi
Nöron	Düğüm/birim/hücre
Sinaps	Bağlantı
Sinaptik yeterliliği	Bağlantı ağırlığı
Atış sıklığı	Çıkış düğümü

YSA yöntemi ile tahmin yapılırken modelin yapısı problemi en iyi şekilde çözebilecek performansa sahip olmalıdır. Gizli katman sayısı, katmanlardaki hücre sayısı, öğrenme algoritması, aktivasyon fonksiyonları vb. faktörler değerlendirilerek başarılı bir model kurulur. Gizli katman ve bu katmandaki hücre sayısının belirlenmesi için deneme yanılma ile sonuçlar bulunur, performanslar kıyaslanır.

4.1 Yapay Sinir Ağır Mimarileri

4.1.1 İleri Beslemeli Yapay Sinir Ağları

İleri beslemeli sinir ağları, giriş ve çıkış katmanlarına ek olarak bir veya daha fazla gizli katmana sahiptir. Her katmanın nöron adı verilen hesaplama birimleri vardır. Sinir ağlarının hesaplama yönü girdiden çıktıya doğrudur (Şahin vd., 2020). Bu ağ yapısında geri besleme yoktur. Şekil 4.2’de X değerleri girdi nöronlarını, H değerleri gizli nöronları, Y değerleri çıktı nöronlarını, v_{ij} değeri; i girdi nöronundan j gizli nöronuna olan bağlantının ağırlığını ve w_{jk} değeri; j gizli nöronundan k çıktı nöronuna olan bağlantının ağırlığını göstermektedir. +1 olarak gösterilen birimler bias değerleridir.

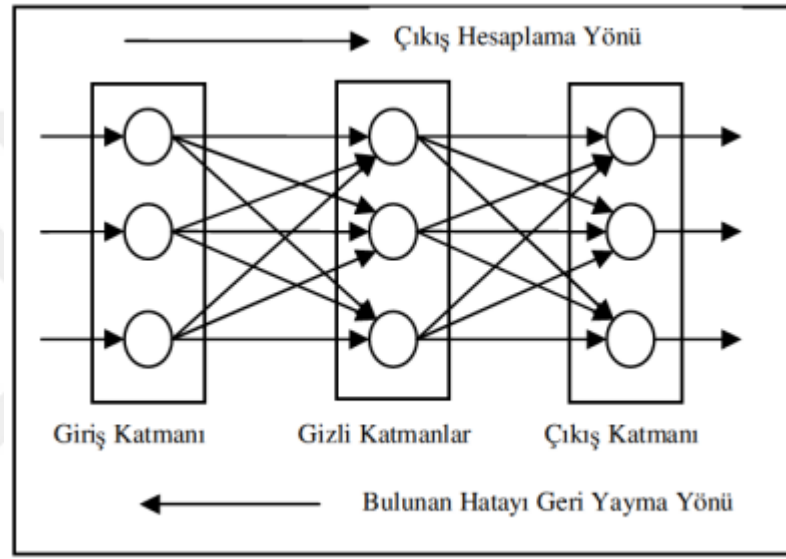


Şekil 4.2: İleri beslemeli ağ yapısı.

4.1.2 Geri Beslemeli Yapay Sinir Ağları

Geri Yayımlı YSA’larda girdi değerlerinin ağa tanıtılması aracılığıyla çıktı değerleri üretilir. Üretilen çıktı değerleri ve gerçek verilerin kıyaslanması sonucunda

ortaya çıkan fark ağda bulunan ağırlıklara geriye doğru iletilir. Bu sayede hata ağa dağıtılmış olur. Hata istenilen değere ulaşınca kadar bu işlem tekrarlanır. Bu tekrarlama sayesinde gerçek değerlere yakın sonuç elde edilir. İleri beslemeli yapay sinir ağlarından farklı bağlam katmanı bulunur ve çift yönlüdür. Bağlam katmanı, gizli katman çıktılarını ağırlıklandırarak tekrar aynı katmana giriş yapmasını sağlar. Bu mimaride bir yapay sinir hücresinin çıkışı tekrar kendisine, başka bir hücreye veya giriş katmanına girdi olabilmektedir. Şekil 4.3'te geri beslemeli ağ yapısına ait giriş katmanı, gizli katmanlar ve çıkış katmanı görülmektedir.

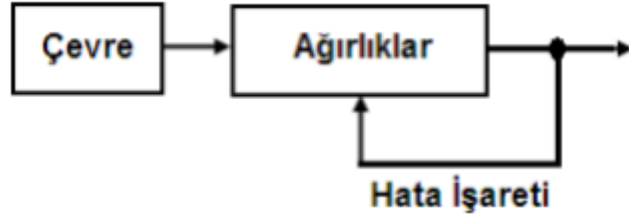


Şekil 4.3: Geri beslemeli ağ yapısı.

4.2 Yapay Sinir Ağı Öğrenme Kuralları

4.2.1 Danışmansız Öğrenme

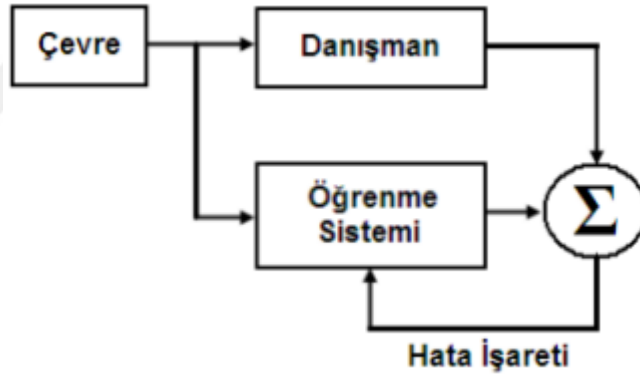
Çıktıların istenen değerleri ağa tanıtılamıyorsa, bu tip öğrenme şekli danışmansız öğrenme olarak adlandırılır. Danışmansız öğrenme algoritmaları sistemin geçmişte karşılaştığı veri setinin içerdiği istatistiksel bilgilerin elde edilmesini amaçlar. Çok elemanlı veri setleri içerisinde deneme-yanılma yoluyla bilgi genelleştirilmesi yapılabilir. Danışmansız öğrenmede girdiler aynı zamanda çıktı görevi görmektedir. Hem çıktı hem de girdi olan veriler arasındaki kural ve ilişkilerin araştırılması ve en uygununun bulunması ağın eğitilmesi anlamına gelmektedir. (Efe, 2000). Şekil 4.4'te danışmansız öğrenme yapısı görsel olarak ifade edilmiştir.



Şekil 4.4: Danışmansız öğrenme yapısı.

4.2.2 Danışmanlı Öğrenme

Bu öğrenme türünde eğitim süreci bulunmaktadır. Ağ eğitilirken veri setindeki gerçek veriler kullanılır, giriş verilerinden çıkış değerleri üretilir. Üretilen çıkış değerleri ve gerçek değerler arasındaki fark, öncesinde belirlenen hata değerinden küçük oluncaya kadar ağırlıklar değiştirilir. Hata değeri istenilen değerin altına düştüğünde eğitime son verilir. Bu aşamada ağıdaki ağırlıklar sabit değerlerine ulaşır. Eğer istenilen hata değerine ulaşılamazsa, eğitim gerçek değerler ile çıktı değerleri arasındaki hatayı minimuma indiren bağlantı ağırlıklarının bulunması ile sonlandırılır (Çelik, 2011). Şekil 4.5'te danışmanlı öğrenme yapısı görsel olarak ifade edilmiştir.

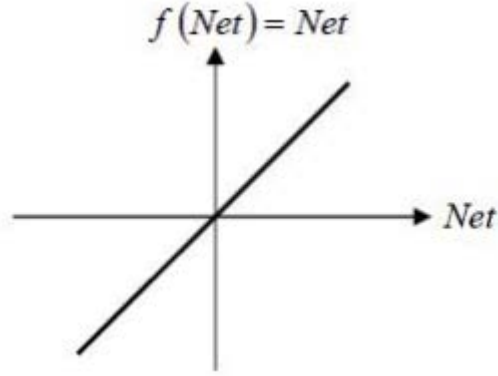


Şekil 4.5: Danışmanlı öğrenme yapısı.

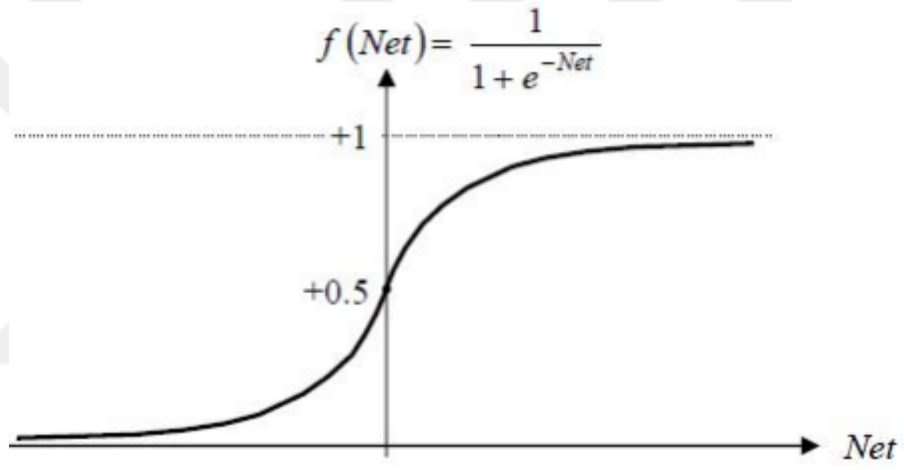
4.3 Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları, öğrenme eğrisi olarak da adlandırılmaktadır. Aktivasyon fonksiyonları bir yapay sinir ağında nöronun çıkış değerini, istenilen değerler arasında tutar. Bu değerler çoğunlukla $[0,1]$ veya $[-1,1]$ arasındadır. Ek olarak, yapay sinir ağına bir bias (eşik) değeri uygulanarak da aktivasyon fonksiyonu artırılabilir. Yapay sinir ağlarında kullanılacak olan fonksiyonların türevi alınabilir ve süreklilik göstermesi gerekmektedir. YSA'nın kullanım amacına uygun olarak tek ya da çift kutuplu aktivasyon fonksiyonları kullanılabilir. Lineer veya doğrusal olmayan aktivasyon fonksiyonlarının yapay sinir ağlarında kullanılması karmaşık problemlerde

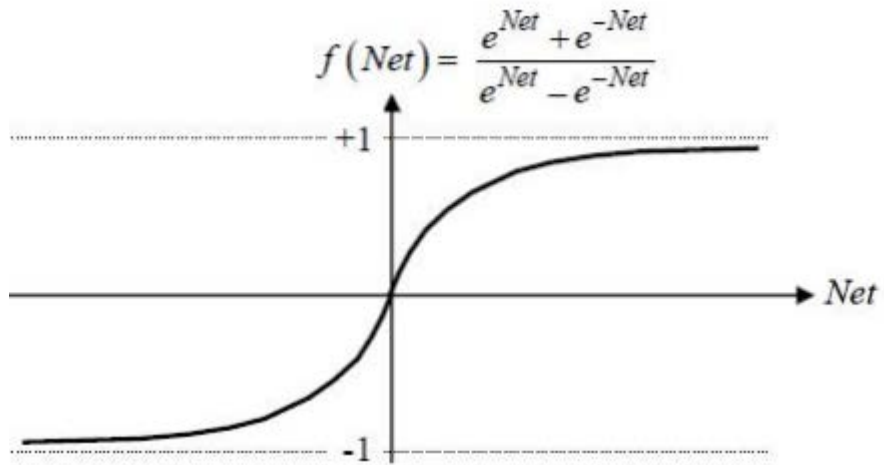
kullanılmasına olanak sağlamıştır. Şekil 4.6, Şekil 4.7 ve Şekil 4.8’de en çok kullanılan aktivasyon fonksiyonları gösterilmiştir.



Şekil 4.6: Lineer aktivasyon fonksiyonu.



Şekil 4.7: Sigmoid aktivasyon fonksiyonu.



Şekil 4.8: Tan-Sigmoid aktivasyon fonksiyonu.

5. RNN MODELLERİ ve LSTM

Derin öğrenme yöntemleri, hava durumu verileri, hisse senedi fiyatları, elektrokardiyogram (EKG) kayıtları ve ham petrol fiyatları vb. gibi doğrusal olmayan veri kümelerini tahmin etmek için yaygın olarak kullanılmaktadır. Yaygın olarak kullanılan iki tür derin öğrenme tekniği vardır: FFNN (İleri Beslemeli Sinir Ağları) ve RNN. Ancak FFNN metodu, zaman serisi verilerindeki eğilimleri yakalayamadığı için zaman serisi verilerini tahmin etmek için uygun değildir. Öte yandan, RNN metodu, belirli bir süre boyunca gelecekteki verileri tahmin etmek için mevcut zaman serisi verilerini kullanan sağlam yapay sinir ağlarıdır. RNN'nin dahili belleği, onu çok başarılı bir teknik haline getirir, çünkü girdi sıralı verilerinin geleceği doğru bir şekilde tahmin etmelerini sağlayan önemli özelliklerini hatırlamayı destekler. FFNN ve RNN arasındaki en büyük fark, FFNN'lerde bilginin kesinlikle katmandan katmana tek yönde akması, RNN'lerde ise mevcut zaman adımından gelen girdinin ve önceki zaman damgasından gelen çıktının RNN hücrelerine beslemesidir. Böylece modelin mevcut durumu önceki durumlarından etkilenir.

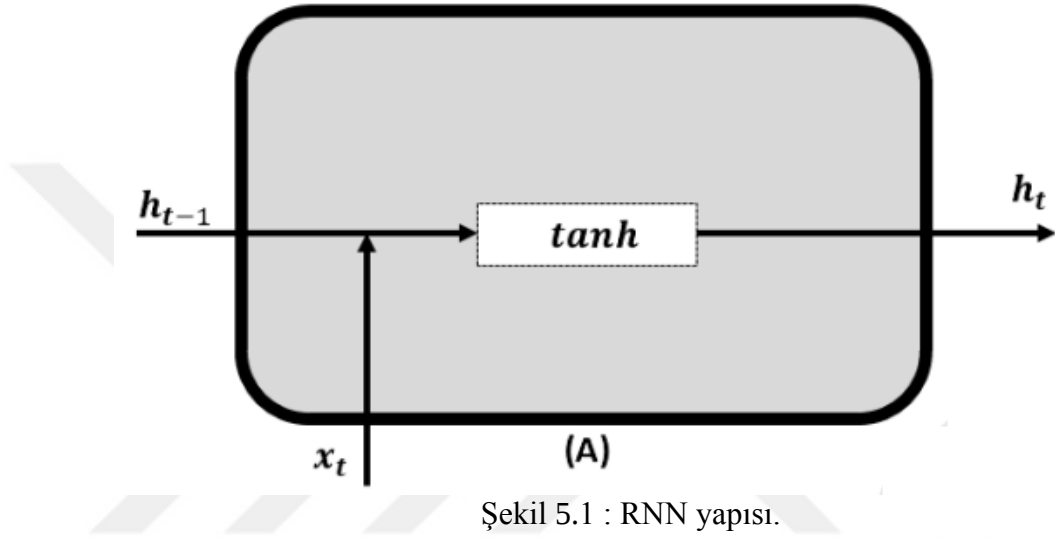
Tek RNN hücresi matematiksel olarak aşağıdaki 5.1 numaralı denklemde gösterilmektedir:

$$h_t = \tan_h(W[h_{t-1}, x_t] + b) \quad (5.1)$$

Bu denklemde b bias matrisini, W ağırlık matrisini, h_t ve h_{t-1} şu andaki ve bir önceki zamandaki gizli katmanı göstermektedir. Şekil 5.1'te görüldüğü üzere, giriş dizisinin her bir ögesi için RNN hücrelerindeki hesaplama, ağırlıklar, bias değerleri ve aktivasyon fonksiyonları kullanıldığında aynı sonucu verecektir. Temel olarak RNN modelinin bir nöronu, h_{t-1} ve x_t 'nin birleştirildiği, ardından bir ağırlık matrisi ile çarpıldığı ve ardından bir biasın eklendiği tek bir hiperbolik tanjant işlevi kullanır. Daha sonra bu değer, h_t 'yi geri veren hiperbolik tanjant fonksiyonundan geçirilir. Hiperbolik tanjant fonksiyonu ($\tanh$), giriş verilerini, değerleri -1 ile +1 aralığında olacak şekilde dönüştürür. Aşağıdaki 5.2 numaralı denklemde yapay sinir ağlarına doğrusal olmayanlığı (non-linearity) tanıtmak için kullanılan sigmoid fonksiyonunun matematiksel olarak gösterilmiştir. Bu fonksiyon RNN hücresinin çıktısını güncellemek için her zaman adımında kullanılır.

$$f(x) = \frac{1}{1+e^{-x}} \quad (5.2)$$

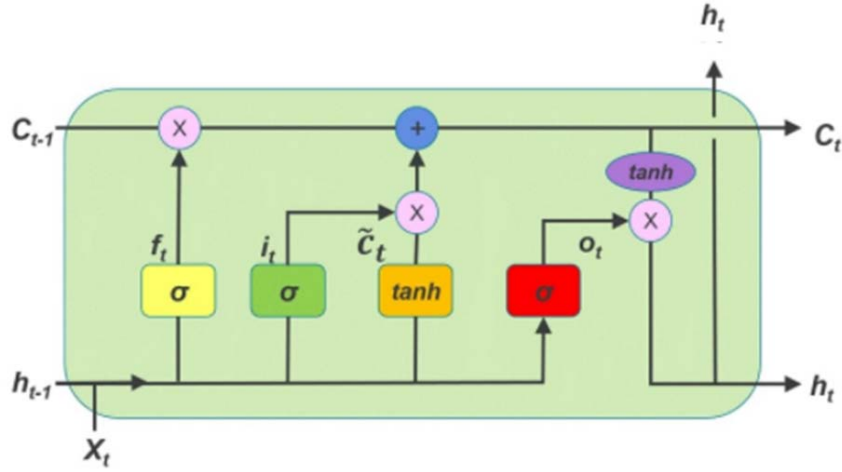
RNN modelleri en son bilgileri hatırlayabilir fakat mevcut zaman adımından çok uzakta olan daha önceki bilgileri hatırlayamaz. RNN, kaybolan gradyan sorunu nedeniyle uzun giriş dizileri için eğitilemez. Ayrıca, RNN mimarisi özellikleri hatırlamak için daha kısa belleğe sahiptir ve kaybolan-patlayan gradyanlar sorunuyla karşı karşıyadır. Bu nedenle, RNN'ler, bu dezavantajların üstesinden gelmek için genellikle LSTM hücreleri ve GRU'lar ile birleştirilir (ArunKumar vd., 2022).



Şekil 5.1 : RNN yapısı.

Kaybolan ve patlayan gradyanlar sorununun üstesinden gelmek için Hochreiter ve Schmidhuber LSTM'yi önerdi (Hochreiter ve Schmidhuber, 1997). LSTM'nin avantajı, giriş hücresi belleğini depolayan ve çıkış hücresi durumuna dönüştüren bir hücre durumuna sahip olmasıdır. Şekil 5.2'de temel olarak giriş kapısı, çıkış kapısı, unutma kapısı ve güncelleme kapısından oluşan LSTM hücresinin genel mimarisi gösterilmektedir. Unutma kapısı, önceki bellek birimlerinden alınan bilgilerden nelerin unutulacağına karar verir, giriş kapısı nörona hangi bilgilerin kabul edileceğine karar verir, çıkış kapısı yeni uzun süreli belleği oluşturur ve güncelleme kapısı hücreyi günceller. Aşağıda 5.3 numaralı denklemde, hücreye hangi bilgilerin aktarılması gerektiğini belirleyen giriş kapısının matematiksel ifadesi gösterilmektedir:

$$i_t = \sigma(W_i * [h_{t-1}, x_t] + b_i) \quad (5.3)$$



Şekil 5.2: LSTM yapısı.

Aşağıdaki 5.4 numaralı denklemde, hangi bilgilerin ihmal edileceğini belirleyen unutma kapısının matematiksel ifadesi gösterilmektedir:

$$f_t = \sigma(W_f * [h_{t-1}, x_t] + b_f) \quad (5.4)$$

Güncelleme kapısı, aşağıdaki 5.5 ve 5.6 numaralı denklemlerle matematiksel olarak ifade edilen hücre durumunu günceller:

$$\tilde{c}_t = \tanh(W_c * [h_{t-1}, x_t] + b_c) \quad (5.5)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (5.6)$$

Çıkış kapısı, çıkışın güncellenmesinden ve önceki zaman adımının gizli katmanını güncellemekten sorumludur. Matematiksel ifade aşağıdaki 5.7 ve 5.8 numaralı denklemlerde gösterilmektedir:

$$o_t = \sigma(W_o * [h_{t-1}, x_t] + b_o) \quad (5.7)$$

$$Loss_1 = \frac{1}{n} \sum_{i=1}^n (y_i' - y_i)^2 + \lambda \sum_k w_k^2 \quad (5.8)$$

Regresyon modeli için ortalama kare hatası üzerinde kayıp fonksiyonu olarak L2 düzenleme yöntemi uygulanmaktadır. Kayıp fonksiyonun matematiksel gösterimi 5.9 numaralı denklemdir (Yang vd., 2021).

$$Loss_1 = \frac{1}{n} \sum_{i=1}^n (y_i' - y_i)^2 + \lambda \sum_k w_k^2 \quad (5.9)$$

Bu 5.9 numaralı denklemde y_i ve y_i' değerleri tahmin edilen ve gerçek değerleri, λ toplam kayıptaki düzenleme oranını, w ise regresyon modelinin eğitilebilir ağırlıklarıdır.

Kurulan modelin tahmin performansı RMSE değeri hesaplanarak incelenebilmektedir. 5.10 numaralı denklemde RMSE değerinin matematiksel ifadesi görülmektedir. y_i ve y'_i değerleri tahmin edilen ve gerçek değerleri, n değeri örneklem büyüklüğünü göstermektedir.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y'_i - y_i)^2} \quad (5.10)$$



6. AR SÜRECİ

Hareketli bir ortalama modelinde, doğrusal bir tahmin edici kombinasyonu kullanarak ilgilenilen değişkeni tahmin edilmektedir. Bir otoregresyon modelinde, değişkenin geçmiş değerlerinin doğrusal bir kombinasyonu kullanılarak tahmin yapılmaktadır. Otoregresyon terimi, değişkenin kendisine karşı bir regresyonunu tanımlar. p düzeyindeki bir dizi gecikmeli değere karşı bir otoregresyon çalıştırılır. Aşağıdaki 6.1 nolu denklemde c değeri sabiti, ϕ_1 ve ϕ_2 değerleri p mertebesine kadar olan gecikme katsayılarını, ε_t değeri ise beyaz gürültüyü göstermektedir.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (6.1)$$

Örneğin bir AR(1) modeli aşağıdaki 6.2 numaralı denklemdeki gibi ifade edilebilmektedir.

$$y_t = c + \phi_1 y_{t-1} + \varepsilon_t \quad (6.2)$$

AR(2) modeli ise aşağıdaki 6.3 numaralı denklemdeki gibi ifade edilebilmektedir. Denklemler bu şekilde devam etmektedir.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t \quad (6.3)$$

Genellikle otoregresif modelleri durağan verilerle sınırladığımız için gecikme katsayılarının genellikle birden küçük olduğuna dikkat edilmesi gereklidir. Örneğin AR(1) modelinde 6.4 numaralı denklem sağlanmalıdır.

$$-1 < \phi_1 < 1 \quad (6.4)$$

AR(2) modelinde ise 6.5 numaralı denklem gerekli şartları belirtmektedir.

$$-1 < \phi_2 < 1, \phi_1 + \phi_2 < 1, \phi_2 - \phi_1 < 1 \quad (6.5)$$

AR(3) modeli matematiksel olarak çok daha karışık ifade edilmektedir fakat mevcut istatistik modelleri tüm bu matematiksel hesaplamaları yapabilmektedir.

7. UYGULAMA

Bu çalışmada Amerika havayolu yolcu mili tahmininde derin öğrenme modellerinin Covid-19 döneminde nasıl performans sergilediği araştırılmıştır. Genel tahmin metotlarından AR modeli ile de tahmin yapılarak sonuçlar kıyaslanmıştır.

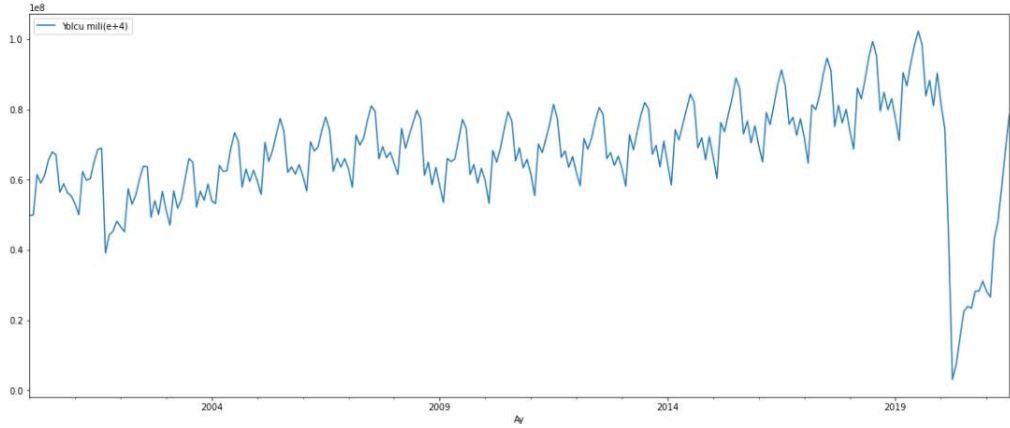
Derin öğrenme yöntemlerinde tahmin yapmak için çeşitli mimarilerden faydalanılmaktadır. Bunlardan biri de LSTM mimarisidir. LSTM hafızasını kullanarak zaman serisi verilerinin tahmininde kullanışlıdır. LSTM mimarisi RNN modelinden türetilmiştir. LSTM mimarisinde, RNN'deki standart ileri beslemeli bağlantılara ek olarak, ünitelerde geri besleme bağlantıları bulunur. Bir LSTM birimi, bir hücre, bir giriş kapısı, bir çıkış kapısı ve bir unutma kapısından oluşur. Hücre keyfi zaman aralıklarında değerleri hatırlar ve bu üç kapı hücreye giren ve çıkan bilgi akışını düzenler (Karami ve Kashef, 2020).

Bu çalışma kapsamında derin öğrenme mimarisi LSTM, zaman serilerinin incelenmesinde kullanılmıştır. Veri seti FRED sisteminden alınmış olup Amerika Havayolu Yolcu Mili'nin 2000 ve 2021 yılları arasını kapsamaktadır. Öncelikli olarak veri setinin trend ve mevsimsellik durumları incelenmiştir. Covid-19 salgını sebebiyle yolcu milindeki ani değişimin LSTM mimarisi üzerindeki değişimini incelemek için veri seti ikiye ayrılmıştır. Birinci uygulamada veri seti Covid-19 dönemini içeren şekilde, ikinci uygulamada Covid-19 dönemini içermeyen şekilde incelenmiştir. Her bir uygulamada son 12 ay test verisi olarak kullanılmıştır. Aylık bazlı değişen yolcu mili sayısı bu sayede iki farklı uygulama ile incelenmiştir. Veri setinin ilk 5 ayı Tablo 7.1'de gösterilmektedir.

Tablo 7.1: Genel veri seti-Amerika.

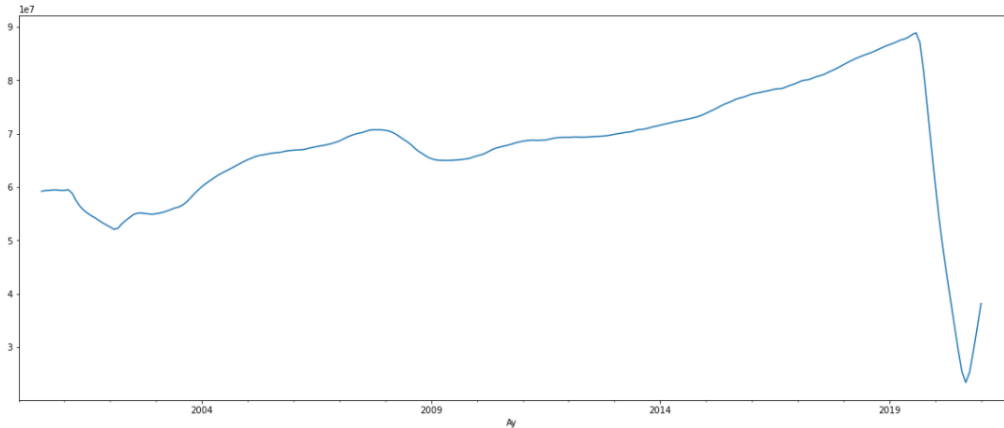
Ay	Yolcu mili (E+4)
2000-01-01	49843099
2000-02-01	49931931
2000-03-01	61478163
2000-04-01	58981617
2000-05-01	61223861

Veri setindeki özellikler ve değerleri Şekil 7.1, Şekil 7.2 ve Şekil 7.3'te görülmektedir. Şekil 7.1'de Covid-19 dönemini içeren yolcu mili veri seti aylık periyotlarla gösterilmiştir.



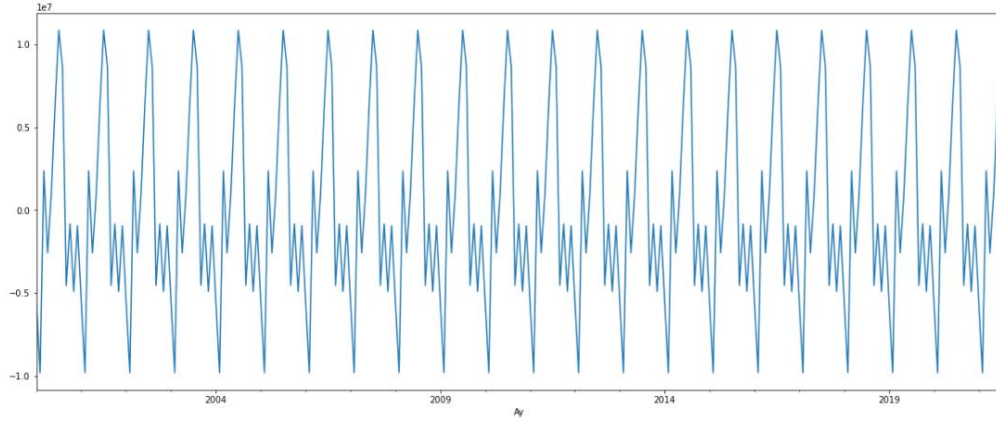
Şekil 7.1: Covid-19 dönemini içeren veri seti-Amerika.

Şekil 7.2’de bu veri setine ait trend grafiği sunulmuştur. Trend grafiği incelendiğinde 2001-2002 yıllarında ve 2008-2009 yıllarında artan trendin yavaşladığını hatta azaldığını görüyoruz. Bu durumun o yıllarda yaşanan global ekonomik krizlerle ilgili olduğu düşünülmektedir. 2020 yılında görülen sert düşüş ise bu çalışmanın temelini oluşturmaktadır.



Şekil 7.2: Verinin trend grafiği.

Şekil 7.3’te de veri setinin mevsimsellik grafiği görülmektedir. Mevsimsellik grafiğinde görüldüğü üzere veride bir döngüsellik söz konusudur. Veri seti incelendiğinde yaz aylarında yolcu milinde artış gözlemlenmiştir.



Şekil 7.3: Verinin mevsimsellik grafiği.

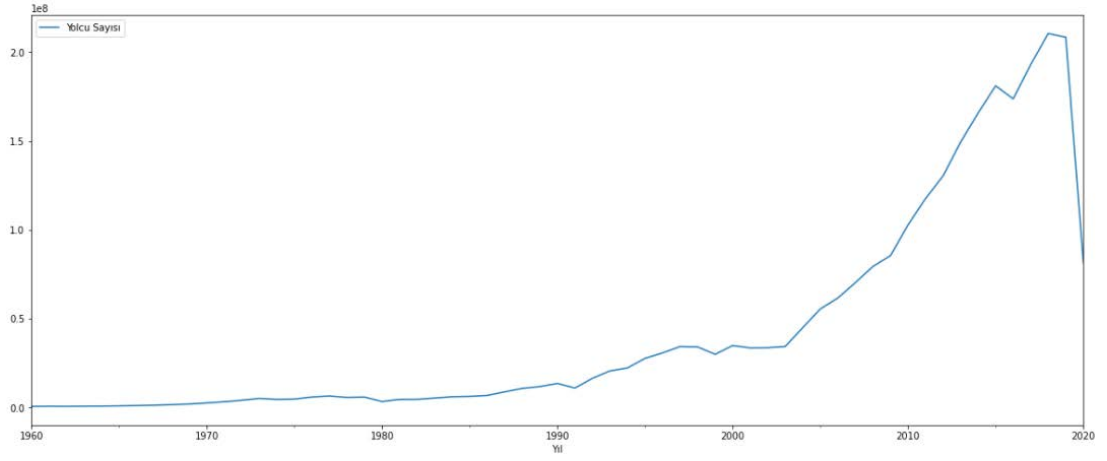
AR modeli genel tahmin modelleri kapsamında kullanılan bir yöntemdir. Bir otoregresyon modelinde, değişkenin geçmiş değerlerinin doğrusal bir kombinasyonu kullanılarak tahmin yapılır. Oto regresyon terimi, değişkenin kendisine karşı bir regresyonunu tanımlar. p düzeyindeki bir dizi gecikmeli değere karşı bir otoregresyon çalıştırılır.

LSTM ve AR modelinin tahmin performansları kıyaslandıktan sonra Türkiye için de Devlet Hava Meydanları İşletmesi Genel Müdürlüğü, Sivil Havacılık Genel Müdürlüğü'nden alınan 1960-2020 yılları arasındaki yıllık havayolu yolcu sayısı verilerine göre tahmin yapılmıştır. Birinci uygulamada veri seti Covid-19 dönemini içeren şekilde, ikinci uygulamada Covid-19 dönemini içermeyen şekilde incelenmiştir. Her bir uygulamada son 3 yıl test verisi olarak kullanılmıştır. Veri setinin ilk 5 yılı Tablo 7.2'de görülmektedir.

Tablo 7.2: Genel veri seti-Türkiye.

Yıl	Yolcu Sayısı
1960-12-31	713217
1961-12-31	788636
1962-12-31	729996
1963-12-31	821489
1964-12-31	842841

Şekil 7.4'te de 60 yıla ait Türkiye havayolu yolcu sayısı grafiği görülmektedir. Yolcu sayısındaki artış trendi grafikte açıkça görülmektedir.



Şekil 7.4: Türkiye havayolu yolcu sayısı grafiği.

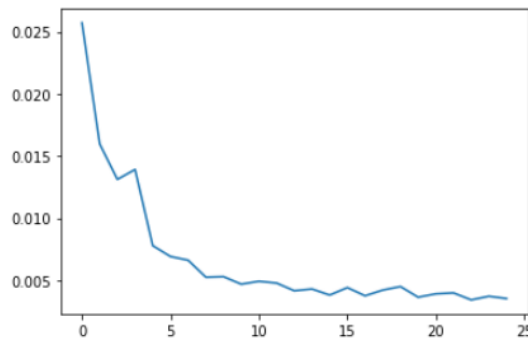
8. BULGULAR ve TARTIŞMA

Bu çalışmada, LSTM mimarisi ile 2000-2021 yılları aralığından oluşan Amerika havayolu yolcu mili zaman serisi verilerinin tahmini yapılmış ve AR modeli ile sonuçlar kıyaslanmıştır. Türkiye havayolu yolcu sayısı için de LSTM mimarisi ile tahmin çalışması yürütülmüştür. LSTM mimarisi ile zaman serisi verilerinin analizine yönelik literatürde çalışmalar mevcuttur. Bu çalışmada Python programlama diliyle Pandas, Numpy, Matplotlib, Stasmodels, Sklearn, Keras ve TensorFlow kütüphaneleri kullanılmıştır. Amerika için kurulan LSTM modeli tek katmanlı olup 150 nörondan oluşmakta ve çıkış katmanında tek bir çıkışı bulunmaktadır. Oluşturulan LSTM mimarisi aşağıdaki Tablo 8.1’de gösterilmiştir.

Tablo 8.1: Oluşturulan LSTM mimarisi-Amerika.

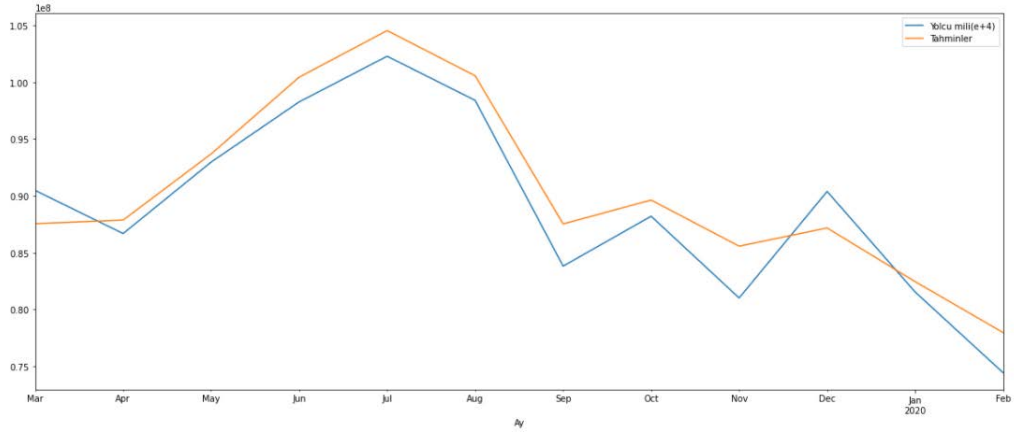
Model: "sequential"		
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 150)	91200
dense (Dense)	(None, 1)	151
Total params: 91,351		
Trainable params: 91,351		
Non-trainable params: 0		

Amerika havayolu yolcu mili değerleri Covid-19 dönemini içeren ve Covid-19 dönemini içermeyen veri seti için 25 epoch ve 1 batch size değerleri kullanılarak model eğitilmiştir. Bu parametre değerlerine göre sonuçlar aşağıdaki grafiklerde gösterilmiştir. Şekil 8.1’de görüldüğü üzere ilk 10 epoch değerine kadar hata dramatik olarak azalmakta fakat sonrasında büyük bir değişiklik göstermemektedir. Yani epoch sayısının daha fazla arttırılmasının modelin doğruluğu üzerinde fark edilir olumlu bir etkisi olmayacaktır.



Şekil 8.1: Covid dönemi hariç modelin kayıp değerleri.

Şekil 8.2’de Covid dönemini kapsamayan veri seti ile yapılan tahmin sonuçları gösterilmiştir. Grafikte mavi çizgi orijinal veriyi gösterirken, turuncu çizgi ise tahmin sonucunu göstermektedir. Yapılan tahminin gerçek veriye paralel ve yakın sonuçlar verdiği anlaşılmaktadır.



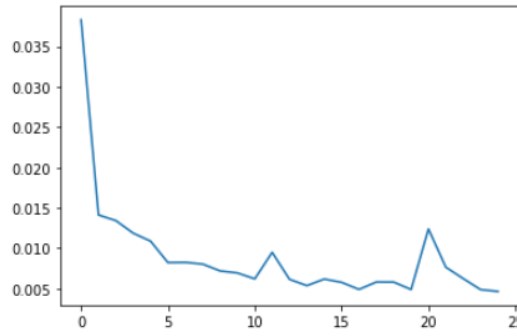
Şekil 8.2: Covid dönemi hariç modelin tahmini-Amerika.

Şekil 8.2’yi oluştururken kullanılan tahmin ve test verileri Tablo 8.2’de aşağıda paylaşılmıştır.

Tablo 8.2: Covid dönemi hariç modelin tahmin değerleri-Amerika.

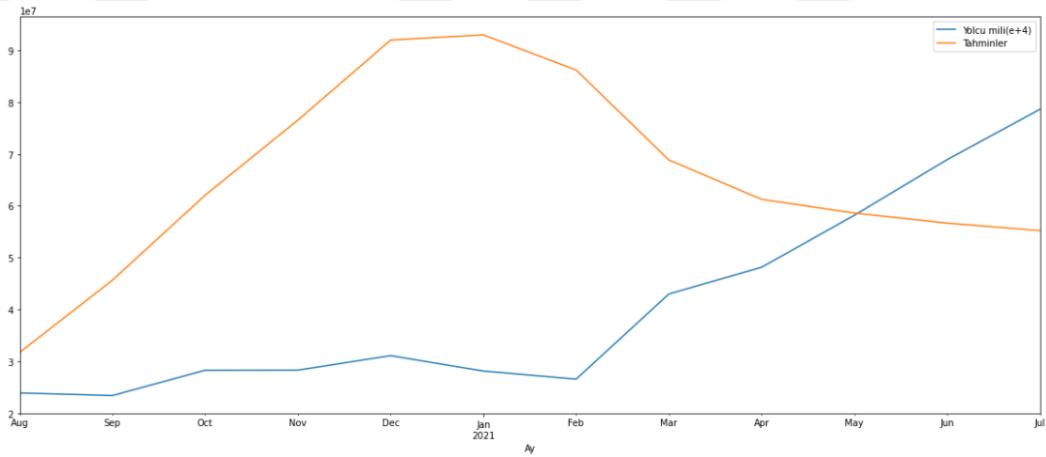
Ay	Yolcu mili(E+4)	Tahminler
2019-03-01	90478107	8.791079E+07
2019-04-01	86684142	8.680529E+07
2019-05-01	92977220	9.273014E+07
2019-06-01	98273616	9.974016E+07
2019-07-01	102295052	1.046480E+08
2019-08-01	98411628	1.009051E+08
2019-09-01	83814263	8.635440E+07
2019-10-01	88208329	8.969266E+07
2019-11-01	81010197	8.505415E+07
2019-12-01	90389595	8.695119E+07
2020-01-01	81537477	8.147397E+07
2020-02-01	74428909	7.633474E+07

Tablo 8.1’de gösterilen LSTM mimarisi Covid-19 dönemini de içeren veri seti için tekrar kullanılmıştır. Bu çalışma da 25 epoch değeri kullanılarak gerçekleştirilmiştir. Şekil 8.3’te de görüldüğü gibi ilk 10 epoch değerine kadar hata değeri hızlı bir düşüş göstermekte sonrasında bu düşüş azalmaktadır. Fakat Covid-19 dönemini içermeyen veri setinde olduğu gibi düzgün bir azalış eğilimi göstermediği saptanmıştır.



Şekil 8.3: Covid dönemi dahil modelin kayıp değerleri-Amerika.

Şekil 8.4'te Covid-19 veri setini içeren modelin yaptığı tahminlerin gerçek veri ile karşılaştırmalı grafiği sunulmuştur. Mavi çizgi orijinal veriyi göstermekte iken, turuncu çizgi yapılan tahmini göstermektedir. Grafikte orijinal veri ile tahmin edilen veri arasında ciddi farklar olduğu görülmektedir.



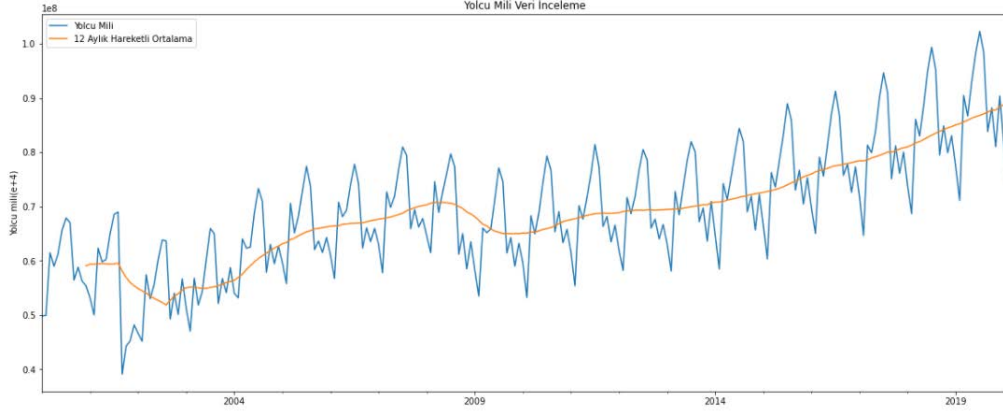
Şekil 8.4: Covid dönemi dahil modelin tahmini-Amerika.

Şekil 8.4'ü oluştururken kullanılan tahmin ve test verileri Tablo 8.3'te aşağıda paylaşılmıştır.

Tablo 8.3: Covid dönemi dahil modelin tahmin değerleri-Amerika.

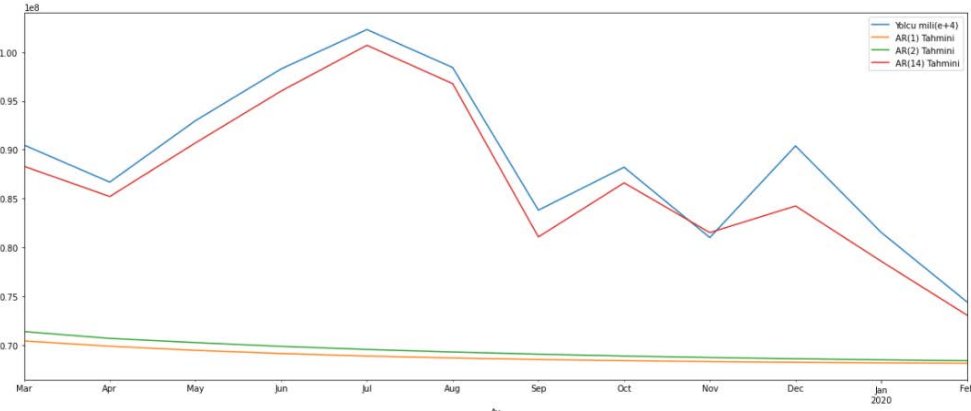
Ay	Yolcu mili(E+4)	Tahminler
2020-08-01	23886437	4.901438E+07
2020-09-01	23384462	6.787354E+07
2020-10-01	28273454	9.003389E+07
2020-11-01	28292601	9.921048E+07
2020-12-01	31081758	1.095932E+08
2021-01-01	28114490	1.027146E+08
2021-02-01	26550396	9.321919E+07
2021-03-01	42982241	7.896106E+07
2021-04-01	48137353	6.956822E+07
2021-05-01	58172779	6.367094E+07
2021-06-01	68903179	5.953058E+07
2021-07-01	78610304	5.676194E+07

AR modeli, ilk olarak Covid-19 dönemini içermeyen veri setinde maksimum olabilirlik (mle) yöntemi aracılığıyla çalıştırılmış ve farklı gecikme (lag) değerleri ile tahmin yapılmıştır. Test değerleri için son 1 yıllık veri kullanılmıştır. Bu tahminden önce de Şekil 8.5'te görüleceği üzere 12 aylık hareketli ortalama ile veri incelenmiştir.



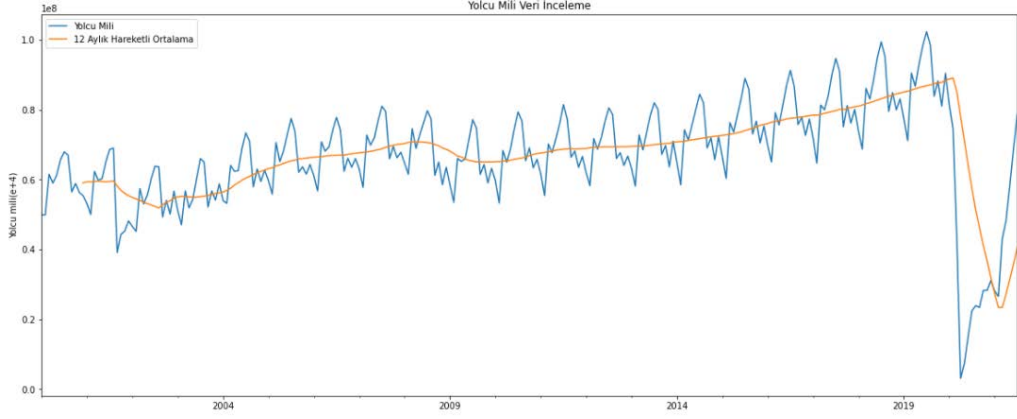
Şekil 8.5: Covid dönemi hariç-12 aylık hareketli ortalama.

Maksimum gecikme değeri 1 ve 2 olarak model çalıştırılmış ardından da T-istatistiği aracılığıyla model çalıştırıldığında maksimum gecikme değeri 14 olarak görülmüştür. Şekil 8.6'da da görüleceği üzere mavi renkli çizgi ana veri setini, kırmızı renkli çizgi maksimum gecikme değeri 14 olan modelin tahminini, yeşil renkli çizgi maksimum gecikme değeri 2 olan modelin tahminini, turuncu renkli çizgi de maksimum gecikme değeri 1 olan modelin tahminini göstermektedir. Test verisine en yakın sonucun maksimum gecikme değeri 14'te olduğu görülmektedir.



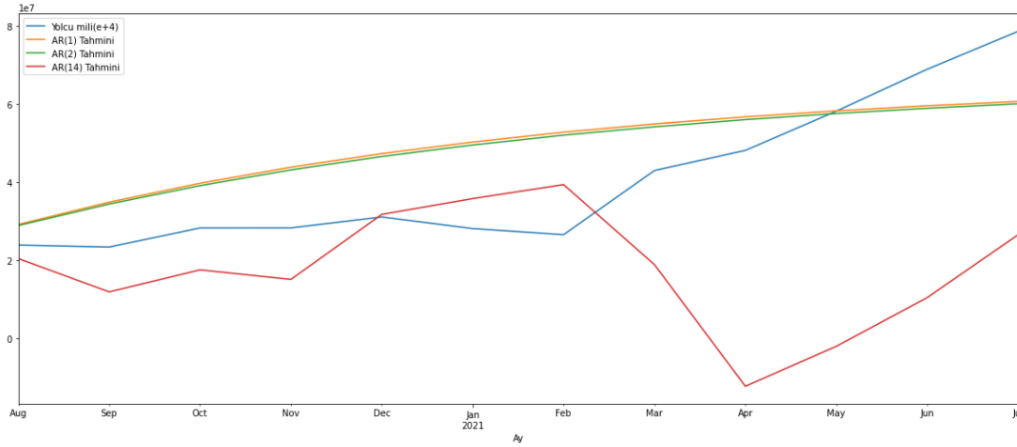
Şekil 8.6: Covid dönemi hariç modellerin tahmin değerleri.

AR modeli, ikinci olarak Covid-19 dönemini içeren veri setinde maksimum olabilirlik (mle) yöntemi aracılığıyla çalıştırılmış ve farklı gecikme (lag) değerleri ile tahmin yapılmıştır. Test değerleri için son 1 yıllık veri kullanılmıştır. Şekil 8.7'de Covid döneminin dahil edildiği veri setinin 12 aylık hareketli ortalaması incelenmiştir.



Şekil 8.7: Covid dönemi dahil-12 aylık hareketli ortalama.

Maksimum gecikme değeri 1 ve 2 olarak model çalıştırılmış ardından da T-istatistiği aracılığıyla model çalıştırıldığında maksimum gecikme değeri 14 olarak görülmüştür. Şekil 8.8’de görüleceği üzere görüleceği üzere mavi renkli çizgi ana veri setini, kırmızı renkli çizgi maksimum gecikme değeri 14 olan modelin tahminini, yeşil renkli çizgi maksimum gecikme değeri 2 olan modelin tahminini, turuncu renkli çizgi de maksimum gecikme değeri 1 olan modelin tahminini göstermektedir. Maksimum gecikme değeri 14’te test verisi ile büyük farklar mevcuttur.

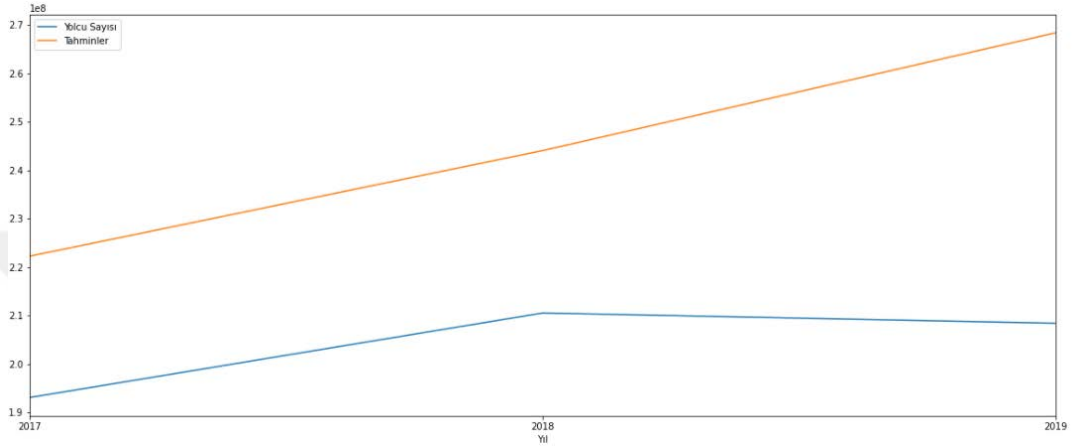


Şekil 8.8: Covid dönemi dahil modellerin tahmin değerleri.

Türkiye’ye ait havayolu yolcu sayısı Covid-19 dönemini içeren ve Covid-19 dönemini içermeyen 2 ayrı veri seti olarak, 25 epoch ve 1 batch size değerleri kullanılarak eğitilmiştir. Tablo 8.4’te görüleceği üzere kurulan LSTM modeli tek katmanlı olup 150 nörondan oluşmakta ve çıkış katmanında tek bir çıkışı bulunmaktadır. Şekil 8.9’da Covid-19 dönemini içermeyen veri seti kullanılarak son 3 yıla ait tahmin değerleri ve test verileri görülmektedir. Mavi renkli çizgi orijinal veriyi, turuncu renkli çizgi tahmin değerlerini göstermektedir. Modelin tahmin performansı incelenmiş ve hata değerinin verinin standart sapmasından ($5,98E+7$) düşük olduğu tespit edilmiştir. Modelin RMSE değeri $4,3E+7$ olarak hesaplanmıştır.

Tablo 8.4: Oluşturulan LSTM mimarisi-Türkiye.

Model: "sequential"		
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 150)	91200
dense (Dense)	(None, 1)	151
Total params: 91,351		
Trainable params: 91,351		
Non-trainable params: 0		



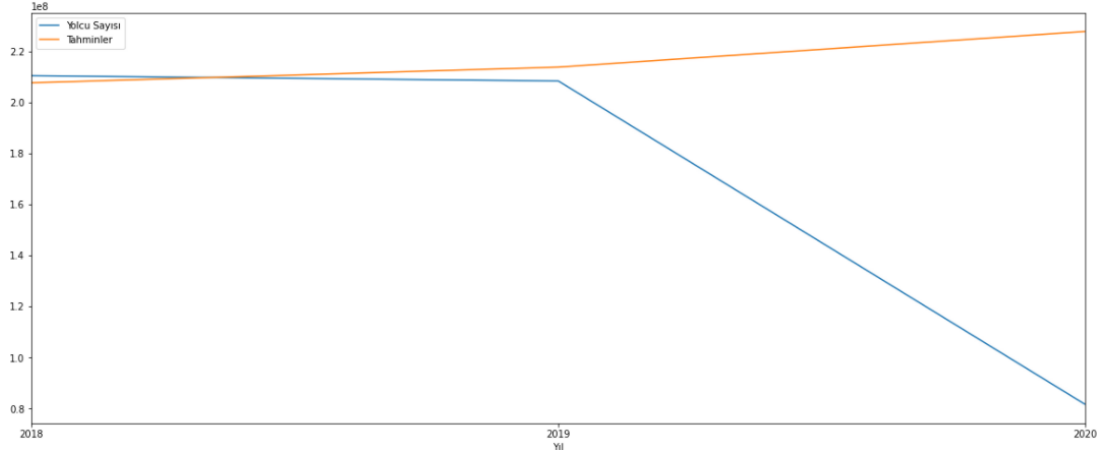
Şekil 8.9: Covid dönemi hariç modelin tahmini-Türkiye.

Şekil 8.9'daki grafiği oluşturan tahmin ve test verileri Tablo 8.5'te görülmektedir.

Tablo 8.5: Covid dönemi hariç tahmin değerleri-Türkiye.

Yıl	Yolcu Sayısı	Tahminler
2017-12-31	193045343	2.222415E+08
2018-12-31	210498164	2.440476E+08
2019-12-31	208373696	2.683765E+08

Şekil 8.10'da Covid-19 dönemini içeren veri seti kullanılarak son 3 yıla ait tahmin değerleri ve test verileri görülmektedir. Mavi renkli çizgi orijinal veriyi, turuncu renkli çizgi tahmin değerlerini göstermektedir. 2019 yılına kadar modelin tahmin performansı iyi görünmektedir fakat sonrasında hata değeri artmıştır. Hata değerinin verinin standart sapmasından ($5,95E+7$) yüksek olduğu tespit edilmiştir. Modelin RMSE değeri $8,4E+7$ olarak hesaplanmıştır.



Şekil 8.10: Covid dönemi dahil modelin tahmini-Türkiye.

Şekil 8.10'daki grafiği oluşturulan tahmin ve test verileri Tablo 8.6'da görülmektedir.

Tablo 8.6: Covid dönemi dahil tahmin değerleri-Türkiye.

Yıl	Yolcu Sayısı	Tahminler
2018-12-31	210498164	2.076688E+08
2019-12-31	208373696	2.138289E+08
2020-12-31	81616140	2.277837E+08

9. SONUÇLAR ve ÖNERİLER

Bu çalışma ile derin öğrenme yaklaşımlarından LSTM mimarisi kullanılarak Covid-19 etkisindeki Amerika havayolu yolcu mili değerinin tahmini için bir model geliştirilmiştir. Geliştirilen modelde 2000 yılından 2021 yılına kadar olan veriler kullanılmış olup Covid-19 etkisinde LSTM mimarisinin çalışma performansı incelenmiştir. Kullanılan Amerika havayolu yolcu mili Covid dönemi dahil edilen veri 259 satırdan oluşmakta, $6,68E+7$ ortalama değerine ve $1,5E+7$ standart sapma değerine sahiptir. Diğer yandan Amerika havayolu yolcu mili Covid dönemi hariç tutulan veri 242 satırdan oluşmakta, $6,91E+7$ ortalama değerine ve $1,14E+7$ standart sapma değerine sahiptir. Türkiye için yolcu mili verilerine ulaşılamamış olup, bunun yerine benzer nitelikli 1960 yılından 2020 yılına kadar olan yolcu sayısı verisi kullanılmıştır. Türkiye için Covid dönemi dahil edilen veri 61 satırdan oluşmakta, $4,29E+7$ ortalama değerine ve $5,95E+7$ standart sapma değerine sahiptir. Bunun yanında Covid dönemi hariç tutulan veri 58 satırdan oluşmakta, $4,22E+7$ ortalama değerine ve $5,98E+7$ standart sapma değerine sahiptir. Amerika'ya ait veriler aylık bazda olup, Türkiye'ye ait veriler ise yıllık bazdadır. Aylık çözünürlüğe sahip veri sayesinde Amerika için mevsimsel etkilerin de hesaba katılma fırsatı bulunmuş, fakat yıllık çözünürlükte olan Türkiye verisi için bu mümkün olmamıştır. Türkiye verisinin yıllık bazda olması daha az gürültüye sahip bir veri olmasını da sağlamıştır. Şekil 7.1 ve Şekil 7.4 kıyaslandığında bu durum açıkça görülmektedir. Amerika'ya ait veriler 2000 yılından itibaren başlamakta olup, Türkiye'ye ait veriler ise 1960 yılından başlamaktadır.

LSTM haricinde farklı bir model ile karşılaştırma amacıyla AR modeli ile de tahmin yapılmış ve bulgular LSTM ile elde edilen sonuçlar ile kıyaslanmıştır. Sonrasında LSTM mimarisi ile Türkiye havayolu yolcu sayısı için de tahmin yapılmıştır. LSTM ile yapılan çalışmada modelin performans metriği olarak RMSE (Kök Ortalama Kare Hata) seçilmiş olup tek katmanlı 150 nöronlu oluşan bir mimari kullanılmıştır. AR modeli ise maksimum olabilirlik yöntemi ile kurulmuş olup 14 maksimum gecikme sayısına sahiptir.

Amerika havayolu yolcu mili, LSTM mimarisi ile yapılan çalışma sonuçlarına göre Covid-19 dönemini içeren veri setinde RMSE değeri $3,49E+7$ olarak hesaplanmıştır. Covid-19 dönemini içermeyen veri setinde ise, RMSE değeri $2,3E+6$

olarak hesaplanmıştır. Veriler açıkça göstermektedir ki LSTM mimarisi ani değişimlerde düşük performans sergilemektedir.

Geçmiş verilerle eğitilen LSTM mimarisi trend ve mevsimsellik etkisi gösteren veri setlerinde yüksek performans sergileyerek, Covid-19 dönemini içermeyen Amerika Havayolu Yolcu Mili örneğinde olduğu gibi başarılı tahminler yapabileceğini kanıtlamıştır.

AR modeli ile yapılan çalışma sonuçlarına göre Covid-19 dönemini içeren veri setinde 2 maksimum gecikme sayısında RMSE değeri $1,4E+7$, Covid-19 dönemini içermeyen veri setinde 14 maksimum gecikme sayısında ise RMSE değeri $2,6E+6$ olarak hesaplanmıştır. Tablo 9.1’de görüleceği üzere, Covid-19 dönemini içeren veri setinde AR modeli ve içermeyen veri setinde de LSTM modeli çok daha iyi sonuç vermiştir. Türkiye havayolu yolcu sayısı verileri için çalıştırılan LSTM mimarisi RMSE değerleri çok daha yüksek çıkmıştır. Veri sayısının az olması burada etken olmuştur. Literatürde doğrusal olmayan veri setlerinde LSTM başarılı olarak düşünülürken AR modeli ani veri değişiminde daha başarılı sonuçlar göstermiştir.

Tablo 9.1: Yapılan tahminlerin RMSE değerleri.

RMSE Değerleri	LSTM mimarisi	AR modeli
Covid Dönemi Hariç Veri Seti	$2,3E+6$	$2,6E+6$
Covid Dönemi Dahil Veri Seti	$3,49E+7$	$1,4E+7$

Gelecek çalışmalarda veri setlerinde ani değişimlerde hangi derin öğrenme yaklaşımlarının başarılı olacağının araştırılması önerilmektedir. Farklı mimariler denenerek en yüksek performansı sergileyen yöntem ortaya çıkartılabilir. Hiper parametre optimizasyonu yapılarak mevcut veri seti için en uygun parametreler belirlenebilir. Buna ek olarak mevcut veri setini etkileyebilecek dış faktörler ile karşılaştırmalı, aynı anda birden fazla girdi kullanılarak da analizler yapılarak modeller zenginleştirilebilir. İnsanların havayolu yolcu taşımacılığına talebi ve ülkelerin ekonomik büyümesi arasındaki nedenselliği test etmek için Granger nedenselliği kullanılabilir.

KAYNAKLAR

- Abbasimehr, H., Shabani, M., Yousefi, M. (2020), "An optimized model using LSTM network for demand forecasting", *Computers & Industrial Engineering*, 143, 106435.
- Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., Fieguth, P., Cao, X., Khosravi, A., Acharya, U. R., Makarencov, V., Nahavandi, S. (2021), "A review of uncertainty quantification in deep learning: Techniques, applications and challenges", *Information Fusion*, 76, 243-297.
- Alassafi, M. O., Jarrah, M., Alotaibi, R. (2021), "Time series predicting of COVID-19 based on deep learning", *Neurocomputing*, 468, 335-344.
- ArunKumar, K.E., Kalaga, D.V., Sai Kumar, M., Kawaji, M., Brenza, T.M. (2022), "Comparative analysis of Gated Recurrent Units (GRU), Long Short-TermMemory (LSTM) cells, Autoregressive Integrated Moving Average (ARI-MA), Seasonal Autoregressive Integrated Moving Average (SARIMA) for forecasting COVID-19 trends", *Alexandria Engineering Journal*, 168.
- Aznarte-Mellado, J., Benitez-Sanchez, J., Nieto-Lugilde, D. (2007), "Forecasting airborne pollen concentration time series with neural and neuro-fuzzy models", *Expert Systems with Applications*, 32 (4), 1218-1225.
- Bi, J-W., Li, H., Fan, Z-P. (2021), "Tourism demand forecasting with time series imaging: A deep learning model", *Annals of Tourism Research*, 90, 103255.
- Brown, R. S., Kline, W. A. (2020), "Exogenous shocks and managerial preparedness: A study of U.S. airlines' environmental scanning before the onset of the COVID-19 pandemic", *Journal of Air Transport Management*, 89, 101899.
- Cao, W., Sun, S., Li, H. (2021), "A new forecasting system for high-speed railway passenger demand based on residual component disposing", *Measurement*, 183, 109762.
- Chatigny, P., Patenaude, J. M., Wang, S. (2021), "Spatiotemporal adaptive neural network for long-term forecasting of financial time series", *International Journal of Approximate Reasoning*, 132, 70-85.
- Chen, Y., Chang, F-J. (2009), "Evolutionary artificial neural networks for hydrological systems forecasting", 367(12), 125137.
- Chimmula, V. K. R., Zhang, L. (2020), "Time series forecasting of COVID-19 transmission in Canada using LSTM networks", *Chaos, Solitons and Fractal*, 153, 109864.
- Choi, S., Kim, Y. J. (2021), "Artificial neural network models for airport capacity prediction", *Journal of Air Transport Management*, 97, 102146.

Clements, M.P., Franses, P.H., Swanson, N.R. (2004), "Forecasting economic and financial time-series with non-linear models", *International Journal of Forecasting*, 20(2), 169-183.

Çelik, E. (2011), "Görüntü İşlemeye Dayalı Avuç İçi İzinin Yapay Sinir Ağı ile Tanınması", Yüksek Lisans Tezi, Marmara Üniversitesi.

Dongare, A.D., Kharde, R.R., Kachare, A.D. (2012), "Introduction to Artificial Neural Network", *International Journal of Engineering and Innovative Technology (IJEIT)*, 2(1).

Efe, M.Ö. (2000), "Yapay sinir ağları ve uygulamaları", Boğaziçi Üniversitesi Yayinevi, (İstanbul), 1-25.

Galicía, A., Torres, J.F., Martínez-Álvarez, F., Troncoso, A. (2018), "A novel spark-based multi-step forecasting algorithm for big data time series", *Information Sciences*, 467, 800-818.

Granger, C.W.J., Terasvirta, T. (1993), "Modeling nonlinear economic relationships", Oxford University Press.

Grosche, T., Rothlauf, F., Heinzl, A. (2007), "Gravity models for airline passenger volume estimation", *Journal of Air Transport Management*, 13, 175-183.

Hanson, D., Delibasi, T. T., Gatti, M., Cohen, S. (2021), "How do changes in economic activity affect air passenger traffic? The use of state-dependent income elasticities to improve aviation forecast", *Journal of Air Transport Management*, 98, 102147.

Hochreiter, S. and J. Schmidhuber (1997), "Long short-term memory", *Neural computation*, 9(8), 1735-1780.

Hotle, S., Mumbower, S. (2021), "The impact of COVID-19 on domestic U.S. air travel operations and commercial airport service", *Transportation Research Interdisciplinary Perspectives*, 9, 100277.

Iacus, S. M., Natale, F., Santamaria, C., Spyrtos, S., Vespe, M. (2020), "Estimating and projecting air passenger traffic during the COVID-19 coronavirus outbreak and its socio-economic impact", *Safety Science*, 129, 104791.

Karami, Z., Kashef, R. (2020), "Smart transportation planning: Data, models, and algorithms", *Transportation Engineering*, 2, 100013.

Luo, J., Zhang, Z., Fu, Y., Rao, F. (2021), "Time series prediction of COVID-19 transmission in America using LSTM and XGBoost algorithms", *Results in Physics*, 27, 104462.

Singh, B. (2021), "Predicting airline passengers' loyalty using artificial neural network theory", *Journal of Air Transport Management*, 94, 102080.

Solvoll, G., Mathisen, T. A., Welde, M. (2020), "Forecasting air traffic demand for major infrastructure changes", *Research in Transportation Economics*, 82, 100873.

Somu, N., Raman M R, G., Ramamritham, K. (2020), "A hybrid model for building energy consumption forecasting using long short term memory networks", *Applied Energy*, 261, 114131.

Şahin, F., Işık, G., Şahin, G., Kara, M. K. (2020), "Estimation of PM10 levels using feed forward neural networks in Igdir, Turkey", 34, 100721.

Talavera-Llames, R.L., Pérez-Chacón, R., Martínez-Ballesteros, M., Troncoso, A., and Martínez-Álvarez, F. (2016), "A nearest neighbours-based algorithm for big time series data forecasting", In *Hybrid Artificial Intelligent Systems*, 174–185.

Tealab, A. (2018), "Time series forecasting using artificial neural networks methodologies:A systematic review", *Future Computing and Informatics Journal*, 3, 334-340.

Torres, J.F., Troncoso, A., Koprinska, I., Wang, Z., Martínez-Álvarez, F. (2019), "Deep learning for big data time series forecasting applied to solar power", *International Conference on Soft Computing Models in Industrial and Environmental Applications*, 123-133.

Truong, D. (2021), "Estimating the impact of COVID-19 on air travel in the medium and long-term using neural network and Monte Carlo simulation", *Journal of Air Transport Management*, 96, 102126.

Ünsal, M. G., Kasap, R. (2020), "Investigating Covid 19 data for G20, EU and OECD countries via using time series analysis & cluster analysis", *Journal of Turkish Operations Management*, 4 (2), 424-432.

Yang, Z., Tang, R., Zeng, W., Lu, J., Zhang, Z. (2021), "Short-term prediction of airway congestion index using machine learning methods", *Transportation Research Part C*, 125, 103040.

Zhang, G. P. (2003), "Time series forecasting using a hybrid ARIMA and neural network model", *Neurocomputing*, 50, 159–175.

ÖZGEÇMİŞ

2012 yılında Metin Nuran Çakallıklı Anadolu Lisesinden mezun olduktan sonra, 2016 yılında Yıldız Teknik Üniversitesi Makine Fakültesi Endüstri Mühendisliği bölümünden mezun olmuştur. Lisans eğitiminden sonra Gebze Teknik Üniversitesi Sosyal Bilimler Enstitüsü İşletme Anabilim dalında yüksek lisans eğitime başlamıştır. Halen Getir Lojistik firmasında Lojistik Uzmanı unvanı ile görev yapmaktadır.



EKLER

Ek A: Tez Çalışması Kapsamında Yapılan Yayınlar

Yüce, N., Sezen, B., (2022), “Amerika Demiryolu Yolcu Milinin Lstm Modeli Kullanılarak Tahmini”, 2. ISPEC Ulusal Bilimsel Araştırmalar Kongresi, Adana, Türkiye, 17-18 Ocak.

Ek B: Açıklamalar

Amerika’ya ait yolcu mili veri seti aşağıda görülmektedir.

Tarih	Yolcu mili(E+4)
2000-01-01	49843099
2000-02-01	49931931
2000-03-01	61478163
2000-04-01	58981617
2000-05-01	61223861
2000-06-01	65601574
2000-07-01	67898320
2000-08-01	67028338
2000-09-01	56441629
2000-10-01	58834210
2000-11-01	56283261
2000-12-01	55380280
2001-01-01	53129922
2001-02-01	49992995
2001-03-01	62323049
2001-04-01	59801567
2001-05-01	60246478
2001-06-01	64987598
2001-07-01	68573410
2001-08-01	69003617
2001-09-01	39106905
2001-10-01	44271037

2001-11-01	45245063
2001-12-01	48167518
2002-01-01	46587854
2002-02-01	45157584
2002-03-01	57423396
2002-04-01	53013179
2002-05-01	55663570
2002-06-01	60224150
2002-07-01	63828500
2002-08-01	63657944
2002-09-01	49270915
2002-10-01	54019748
2002-11-01	50106815
2002-12-01	56656595
2003-01-01	51197174
2003-02-01	47040806
2003-03-01	56766580
2003-04-01	51857453
2003-05-01	54335598
2003-06-01	60272900
2003-07-01	65962215
2003-08-01	64989766
2003-09-01	52121480
2003-10-01	56724551
2003-11-01	54128776
2003-12-01	58739845
2004-01-01	53979787
2004-02-01	53179693
2004-03-01	64035864
2004-04-01	62340117
2004-05-01	62530704
2004-06-01	68866397
2004-07-01	73335889

2004-08-01	70961522
2004-09-01	57881041
2004-10-01	63021143
2004-11-01	59453942
2004-12-01	62680310
2005-01-01	59629608
2005-02-01	55795165
2005-03-01	70595861
2005-04-01	65145551
2005-05-01	68268899
2005-06-01	72952959
2005-07-01	77432998
2005-08-01	73778918
2005-09-01	62046798
2005-10-01	63640895
2005-11-01	61518984
2005-12-01	64289340
2006-01-01	61035027
2006-02-01	56729212
2006-03-01	70799794
2006-04-01	68120559
2006-05-01	69352606
2006-06-01	74099238
2006-07-01	77798621
2006-08-01	74227051
2006-09-01	62345744
2006-10-01	66082164
2006-11-01	63548732
2006-12-01	65980782
2007-01-01	63016014
2007-02-01	57793832
2007-03-01	72700241
2007-04-01	69836157

2007-05-01	71933108
2007-06-01	76926452
2007-07-01	80988340
2007-08-01	79439827
2007-09-01	65913314
2007-10-01	69445349
2007-11-01	66250428
2007-12-01	67782533
2008-01-01	64667106
2008-02-01	61504425
2008-03-01	74575531
2008-04-01	68906882
2008-05-01	72725750
2008-06-01	76162104
2008-07-01	79707545
2008-08-01	77300568
2008-09-01	61198749
2008-10-01	65018204
2008-11-01	58512623
2008-12-01	63503416
2009-01-01	58373786
2009-02-01	53506580
2009-03-01	66027342
2009-04-01	65166301
2009-05-01	65868255
2009-06-01	71350228
2009-07-01	77136799
2009-08-01	74614772
2009-09-01	61435354
2009-10-01	64274740
2009-11-01	59011015
2009-12-01	63231672
2010-01-01	59651062

2010-02-01	53240066
2010-03-01	68307090
2010-04-01	64953251
2010-05-01	68850905
2010-06-01	74474551
2010-07-01	79304441
2010-08-01	76741308
2010-09-01	65341293
2010-10-01	69071626
2010-11-01	63334293
2010-12-01	65797943
2011-01-01	61630362
2011-02-01	55391206
2011-03-01	70158268
2011-04-01	67683559
2011-05-01	71711448
2011-06-01	76057910
2011-07-01	81423231
2011-08-01	77247896
2011-09-01	66345318
2011-10-01	68177919
2011-11-01	63492941
2011-12-01	66595644
2012-01-01	61940192
2012-02-01	58243799
2012-03-01	71696205
2012-04-01	68669276
2012-05-01	71887545
2012-06-01	76760765
2012-07-01	80499331
2012-08-01	78609005
2012-09-01	66008348
2012-10-01	67677764

2012-11-01	64039210
2012-12-01	66702987
2013-01-01	63139217
2013-02-01	58109870
2013-03-01	72764478
2013-04-01	68453537
2013-05-01	73574534
2013-06-01	78467134
2013-07-01	81974582
2013-08-01	80007788
2013-09-01	67203059
2013-10-01	69727370
2013-11-01	63611195
2013-12-01	70967040
2014-01-01	64743622
2014-02-01	58474896
2014-03-01	74237229
2014-04-01	71278694
2014-05-01	75712694
2014-06-01	80063355
2014-07-01	84404540
2014-08-01	81957560
2014-09-01	69028870
2014-10-01	71887841
2014-11-01	65654285
2014-12-01	72233639
2015-01-01	66538437
2015-02-01	60331344
2015-03-01	76271051
2015-04-01	73607258
2015-05-01	78502373
2015-06-01	83253406
2015-07-01	88951664

2015-08-01	85855967
2015-09-01	73017733
2015-10-01	76678109
2015-11-01	70460064
2015-12-01	75327877
2016-01-01	69778658
2016-02-01	65026219
2016-03-01	79121758
2016-04-01	75617435
2016-05-01	81226986
2016-06-01	87042101
2016-07-01	91256833
2016-08-01	86708521
2016-09-01	75735172
2016-10-01	77790770
2016-11-01	72625059
2016-12-01	77310983
2017-01-01	71876269
2017-02-01	64666236
2017-03-01	81311996
2017-04-01	79935852
2017-05-01	83939006
2017-06-01	90130486
2017-07-01	94641128
2017-08-01	90980861
2017-09-01	75092290
2017-10-01	81168039
2017-11-01	76124199
2017-12-01	80037671
2018-01-01	73795634
2018-02-01	68674043
2018-03-01	86103037
2018-04-01	83009406

2018-05-01	88525173
2018-06-01	94917275
2018-07-01	99373376
2018-08-01	95268669
2018-09-01	79470511
2018-10-01	84871382
2018-11-01	79928484
2018-12-01	83062296
2019-01-01	77340229
2019-02-01	71123685
2019-03-01	90478107
2019-04-01	86684142
2019-05-01	92977220
2019-06-01	98273616
2019-07-01	102295052
2019-08-01	98411628
2019-09-01	83814263
2019-10-01	88208329
2019-11-01	81010197
2019-12-01	90389595
2020-01-01	81537477
2020-02-01	74428909
2020-03-01	43175580
2020-04-01	3086778
2020-05-01	7373339
2020-06-01	15138037
2020-07-01	22438942
2020-08-01	23886437
2020-09-01	23384462
2020-10-01	28273454
2020-11-01	28292601
2020-12-01	31081758
2021-01-01	28114490

2021-02-01	26550396
2021-03-01	42982241
2021-04-01	48137353
2021-05-01	58172779
2021-06-01	68903179
2021-07-01	78610304

Türkiye'ye ait yolcu sayısı veri seti aşağıda görülmektedir.

Yıl	Toplam	İç hat	Dış hat
1960	713 217	528 846	184 371
1961	788 636	573 877	214 759
1962	729 996	484 061	245 935
1963	821 489	560 679	260 810
1964	842 841	587 814	255 027
1965	977 913	681 623	296 290
1966	1 161 329	804 773	356 556
1967	1 377 609	951 919	425 690
1968	1 696 923	1 173 139	523 784
1969	2 037 467	1 343 984	693 483
1970	2 679 139	1 661 890	1 017 249
1971	3 301 143	2 166 447	1 134 696
1972	4 179 900	2 583 010	1 596 890
1973	5 191 093	3 236 997	1 954 096
1974	4 651 344	2 615 368	2 035 976
1975	4 800 902	2 599 373	2 201 529
1976	5 979 087	3 670 020	2 309 067
1977	6 545 487	4 177 259	2 368 228
1978	5 752 304	3 338 722	2 413 582
1979	5 935 276	3 485 847	2 449 429
1980	3 458 165	1 621 998	1 836 167
1981	4 641 772	2 638 294	2 003 478
1982	4 669 648	2 631 692	2 037 956
1983	5 387 791	2 954 497	2 433 294

1984	6 132 361	3 277 079	2 855 282
1985	6 323 448	3 061 822	3 261 626
1986	6 869 986	3 263 967	3 606 019
1987	8 903 699	4 049 182	4 854 517
1988	10 840 179	4 329 890	6 510 289
1989	11 843 563	4 696 520	7 147 043
1990	13 629 965	5 347 723	8 282 242
1991	11 019 464	4 009 724	7 009 740
1992	16 495 118	5 445 081	11 050 037
1993	20 674 531	7 403 941	13 270 590
1994	22 334 286	8 784 310	13 549 976
1995	27 767 379	10 347 528	17 419 851
1996	30 780 662	10 862 539	19 918 123
1997	34 396 334	12 413 720	21 982 614
1998	34 199 679	13 238 832	20 960 847
1999	30 011 658	12 931 771	17 079 887
2000	34 972 534	13 339 039	21 633 495
2001	33 620 448	10 057 808	23 562 640
2002	33 755 452	8 700 839	25 054 613
2003	34 424 340	9 128 124	25 296 216
2004	45 034 589	14 438 292	30 596 297
2005	55 545 473	20 502 516	35 042 957
2006	61 684 203	28 799 878	32 884 325
2007	70 352 867	31 970 874	38 381 993
2008	79 438 289	35 832 776	43 605 513
2009	85 508 508	41 226 959	44 281 549
2010	102 800 392	50 575 426	52 224 966
2011	117 620 469	58 258 324	59 362 145
2012	130 351 620	64 721 316	65 630 304
2013	149 430 421	76 148 526	73 281 895
2014	165 720 234	85 416 166	80 304 068
2015	181 074 531	97 041 210	84 033 321
2016	173 743 537	102 499 358	71 244 179

2017	193 045 343	109 511 390	83 533 953
2018	210 498 164	112 911 108	97 587 056
2019	208 373 696	99 946 572	108 427 124
2020	81 616 140	49 740 303	31 875 837

