



**TÜRKİYE İÇİN YAPAY SİNİR AĞLARI VE
ZAMAN SERİSİ YÖNTEMLERİYLE
YENİLENEBİLİR ENERJİ ÜRETİMİ
TAHMİNLERİ**

Mehmet Berke ÇOLAK

Yüksek Lisans Tezi

**Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Dr. Öğr. Üyesi ERKAN ÖZHAN**

2021

T.C.
TEKİRDAĞ NAMIK KEMAL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

TÜRKİYE İÇİN YAPAY SİNİR AĞLARI VE ZAMAN SERİSİ YÖNTEMLERİYLE
YENİLENEBİLİR ENERJİ ÜRETİMİ TAHMİNLERİ

Mehmet Berke ÇOLAK

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DANIŞMAN: Dr. Öğretim Üyesi ERKAN ÖZHAN

TEKİRDAĞ-2021

Her hakkı saklıdır

Bu tez (TÜBİTAK / SANTEZ / NKÜBAP vb.) tarafından
..... numaralı proje ile desteklenmiştir.



ÖZET

Yüksek Lisans Tezi

TÜRKİYE İÇİN YAPAY SİNİR AĞLARI VE ZAMAN SERİSİ YÖNTEMLERİYLE YENİLENEBİLİR ENERJİ ÜRETİMİ TAHMİNLERİ

Mehmet Berke ÇOLAK

Tekirdağ Namık Kemal Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Dr. Öğr. Üyesi Erkan ÖZHAN

Zaman içerisinde artmakta olan nüfus ve gelişmekte olan sanayi ve yaşam şartlarındaki değişimler sebebiyle sürekli veya çevre kirliliğini en aza indirdiği için yenilenebilir enerji kaynakları tercih edilir. Yapılan çalışmada Türkiye'deki yenilenebilir enerji kaynaklarından elektrik üretim oranı, gelecek yıllar içerisinde gereksiz enerji kaybını önlemek için uygun bir şekilde belirlenmelidir. Bu kaybı önlemek için 1960-2019 yılları arasında Türkiye'deki yenilenebilir enerji kaynaklarından elektrik üretim oranları veri kümesi zaman serisi analizi ve yapay sinir ağları tahminleme teknikleri kullanılarak işlenmiştir. Ayrıca 2020'den 2025 yılına dahil olmak üzere tahminleme işlemi de yapılmıştır. Kullanılan tahminleme tekniklerin performans değerlendirmesi için veri kümesinin içerisinde 1960-2014 arasındaki veriler eğitim verisi, 2015-2019 arasındaki veriler ise test verisi olarak alınarak ortalama mutlak hata (MAE) ve kök ortalama kare hatası (RMSE) değerleri üzerinde incelenmiştir. En iyi sonucu veren tekniğin sinir ağı otoregresyon (NNAR) olduğu ortaya çıkmasına rağmen yapılan tahminleme işleminde elde edilen değerlerin sabit kalmasından dolayı tercih edilmemiştir. Onun yerine, ARIMA (2,2,1) modeli ve kübik düzeltme spline teknikleri tercih edilebileceği ortaya çıkmıştır. Bu çalışmada yenilenebilir enerji kaynaklarını önceden öngörerek bunun için, daha verimli planlamalar ve organizasyon olması gerektiği belirlenmiştir.

Anahtar kelimeler: Yenilenebilir enerji, Yapay Sinir Ağları, Zaman Serisi Analizi, ARIMA, Kübik Düzeltme Spline, Sinir Ağı Otoregresyon

2021, 77 sayfa

ABSTRACT

MSc. Thesis

FORECASTS OF RENEWABLE ENERGY PRODUCTION FOR TURKEY BY ARTIFICIAL NEURAL NETWORKS AND TIME SERIES METHODS

Mehmet Berke ÇOLAK

Tekirdağ Namık Kemal University
Graduate School of Natural and Applied Sciences
Department of Computer Engineering

Supervisor: Assist. Prof. Dr. Erkan ÖZHAN

Renewable energy sources are preferred because they minimize continuous or environmental pollution due to the increasing population and developing industry and changes in living conditions over time. In the study, the rate of electricity generation from renewable energy sources in Turkey should be determined appropriately in order to prevent unnecessary energy loss in the coming years. In order to prevent this loss, the dataset of electricity generation rates from renewable energy sources in Turkey between 1960-2019 was processed using time series analysis and artificial neural network estimation techniques. In addition, estimation was made from 2020 to 2025. For the performance evaluation of the estimation techniques used, the data between 1960-2014 was taken as training data and the data between 2015-2019 was taken as test data and examined on mean absolute error (MAE) and root mean square error (RMSE) values. Although it turned out that the technique that gave the best results was neural network autoregression (NNAR), it was not preferred because the values obtained in the estimation process remained constant. Instead, it turned out that ARIMA (2,2,1) model and cubic smoothing spline techniques can be preferred. In this study, it has been determined that renewable energy sources should be foreseen and more efficient planning and organization is required for this.

Key words: Renewable Energy, Artificial Neural Networks, Time Series Analysis, ARIMA, Cubic Smoothing Spline, Neural Network Autoregulation Method

2021, 77 pages

İÇİNDEKİLER

ÖZET.....	ii
ABSTRACT	iii
İÇİNDEKİLER.....	iii
ÇİZELGE DİZİNİ.....	iv
ŞEKİL DİZİNİ.....	v
SİMGELER ve KISALTMALAR.....	vi
TEŞEKKÜR.....	viix
1. GİRİŞ.....	10
2. KAYNAK ÖZETLERİ.....	12
3. MATERYAL VE YÖNTEM.....	19
3.1. MATERYAL.....	19
3.2. YÖNTEM	24
3.2.1. YAPAY SİNİR AĞLARI	24
3.2.2. ZAMAN SERİSİ ANALİZİ	29
3.2.2.1. ARIMA MODELLEME YÖNTEMİ	31
3.2.2.2. ÜSTEL DÜZELTME YÖNTEMİ	34
3.2.2.2.1. BASİT ÜSTEL DÜZELTME YÖNTEMİ.	35
3.2.2.2.2. HOLT WINTERS YÖNTEMİ	35
3.2.2.3. DOĞRUSAL REGRESYON ANALİZİ	37
3.2.2.3.1. BASİT DOĞRUSAL REGRESYON ANALİZİ	38
3.2.2.3.2. ÇOKLU DOĞRUSAL REGRESYON ANALİZİ	39
3.2.2.4. TORBALAMA YÖNTEMİ	40
3.2.2.5. SIRALI MİNİMUM OPTİMİZASYON REGRESYON ALGORİTMASI	40
3.2.2.6. YEREL AĞIRLIKLİ ÖĞRENME YÖNTEMİ.....	41
4. ARAŞTIRMA BULGULARI.....	42
4.1. TEST VE EĞİTİM VERİSİ AYRIMI İLE YAPILAN DENEYLER	54
5. TARTIŞMA VE SONUÇ	60
KAYNAKLAR.....	61
EKLER.....	77

ÇİZELGE DİZİNİ

Çizelge 2.1. 2010-2019 Yılları Arasında Yapılan Önceki Çalışmalar	18
Çizelge 2.2. 2020-2021 Yılları Arasında Yapılan Önceki Çalışmalar	19
Çizelge 4.1. İkinci Aşamada Kullanılan Yöntemlerin Tahminleme Öncesi Performans Ölçütleri Karşılaştırılması.....	46
Çizelge 4.2. Kullanılan Yöntemlerin Tahminleme Sonrası Performans Karşılaştırılması.....	52
Çizelge 4.3. Kullanılan Tahminleme Yöntemlerin Tahmin Değerleri.	53
Çizelge 4.4. En İyi Sonucu Veren Üç Tane Yöntemin Eğitim ve Test Verilerinin Performans Karşılaştırılması	54
Çizelge 4.5. En İyi Sonucu Veren Üç Tane Yöntemin Tahmin Değerleri ve Gerçek Değerleri.....	55
Çizelge 4.6. ARIMA (2,2,1) Modelleme Yönteminin Tahmin Değerleri.....	57
Çizelge 4.7. Kübik Düzeltme Spline Yöntemi Tahmin Değerleri.....	58

ŞEKİL DİZİNİ

Şekil 3.1. Weka GUI Seçme Ekranı	20
Şekil 3.2. Weka Uygulamasında Filtre Algoritmaları	21
Şekil 3.3. Weka Uygulamasında Kullanılan Modelin Performans Ölçütleri	22
Şekil 3.4. R Programlama Dili Veri Çerçevesi Komutları	22
Şekil 3.5.R-Studio Uygulamasının Grafikselsel Kullanıcı Ara Yüzü.....	23
Şekil 3.6. YSA Modelinin Şematik Gösterimi	25
Şekil 3.7.Yapay Bir Nöron Modeli.....	25
Şekil 3.8. Çok Katmanlı Algılayıcı (MLP) Yönteminin Şematik Gösterimi	27
Şekil 3.9. Aşırı Öğrenme Makinesi (ELM) Yönteminin Şematik Gösterimi.....	28
Şekil 3.10. NNAR Modelinin Şematik Gösterimi	29
Şekil 3.11. Torbalama Yönteminin Çalışma Şeması	40
Şekil 4.1.Ana Materyaldeki Veri Setinin Gösterimi	42
Şekil 4.2.Yapay Sinir Ağları Yöntemlerinin Tahminleri	44
Şekil 4.3. Birinci Aşamada Kullanılan Yöntemlerin Tahminleri	45
Şekil 4.4. İkinci Aşamada Kullanılan Yöntemlerin Tahminleri.....	46
Şekil 4.5. Doğrusal Regresyon Yöntemindeki Verilerin Gerçek ve Tahmin Değerleri.....	47
Şekil 4.6. Kübik Düzeltme Spline Yöntemindeki Verilerin Gerçek ve Tahmin Değerleri	48
Şekil 4.7. Üçüncü Aşamada Kullanılan Yöntemlerin Tahminleri.....	49
Şekil 4.8. Durağan Hale Getirilen Verilerin ACF Grafiği.....	50
Şekil 4.9. Durağan Hale Getirilen Verilerin PACF Grafiği	50
Şekil 4.10. Dördüncü Aşamada Kullanılan Yöntemlerin Grafik Gösterimi	51
Şekil 4.11. En İyi Sonucu Veren Üç Tane Yöntemin Tahmin Değerleri ve Gerçek Değerleri.....	56
Şekil 4.12. ARIMA (2,2,1) Modelleme Yöntemin Tahminleri.....	57
Şekil 4.13. Kübik Düzeltme Spline Yönteminin Tahminleri.....	58

SİMGELER VE KISALTMALAR

IEA	: Uluslararası Enerji Ajansı
MW	: Megawatt
RE	: Yenilenebilir Enerji
GW	: Gigawatt
TWh	: Terawatt Saat
ktoe	: Ton Petrol Eşdeğeri
Mtoe	: Milyonlarca Ton Petrol Eşdeğeri
MTA	: Maden Tetkik ve Arama Enstitüsü
EİE	: Elektrik İşleri Etüt İdaresi Genel Müdürlüğü
REPA	: Rüzgâr Enerjisi Potansiyel Atlası
DSİ	: Devlet Su İşleri Genel Müdürlüğü
YSA	: Yapay Sinir Ağları
MLP	: Çok Katmanlı Algılayıcı
NRMSE	: Normalleştirilmiş Kök Ortalama Kare Hatası
RMSE	: Kök Ortalama Kare Hatası
ARIMA	: Otoregresif Entegre Hareketli Ortalama
ME	: Ortalama Hata
MSE	: Ortalama Kare Hatası
MAE	: Ortalama Mutlak Hata
ELM	: Aşırı Öğrenme Makinesi
GSYİH	: Gayri Safı Yurt İçi Hâsıla
EMD	: Ampirik Mod Ayrışımı
ISRLU	: Ters Karekök Doğrusal Birimi
BFOA	: Bakteriyel Toplayıcılık Optimizasyon Algoritması
TEDA	: Tipiklik ve Eksantriklik Veri Analitiği
GPL	: Genel Kamu Lisansı

Weka	: Bilgi analizi İçin Waikato Ortamı
GUI	: Grafiksel Kullanıcı Ara Yüzü
ARFF	: Nitelik İlişki Dosya Formatı
CLI	: Komut Satırı Ara Yüzü
SMOreg	: Sıralı Minimum Optimizasyon Regresyonu
LWL	: Yerel Ağırlık Öğrenmesi
NNAR	: Sinir Ağı Otoregresyonu
SLFN	: Tek Gizli Katman İleri Beslemeli Sinir Ağı
ARMA	: Otoregresif Hareketli Ortalama
MA	: Hareketli Ortalama
AR	: Otoregresif
I	: Entegre
ACF	: Otokorelasyon Fonksiyonu
PACF	: Kısmı Otokorelasyon Fonksiyonu
SES	: Basit Üstel Düzeltme
HW	: Holt Winters
MLR	: Çoklu Doğrusal Regresyon
QP	: İkinci Dereceden Programlama
α	: Veriler İçin Üstel Düzeltme Parametresi
β	: Trend İçin Üstel Düzeltme Parametresi
γ	: Mevsimsellik İçin Üstel Yumuşatma Parametresi
λ	: Lambda
w	: Ağırlık vektörü
x	: Girdi vektörü
b	: Ön Yargı değeri
R	: Hatanın Hesapladığı Toplam Gözlem Sayısı
$\bar{X}$	: Değişkenlerin veya Zaman Serisi Verilerinin Ortalama Değeri

ε	: Hata Deęeri
M	: Toplam Deęişken Sayısı
R^2	: Belirleme Sayısı
F	: LWL Fonksiyonu
y_i	: Baęımlı Deęişken
x_i	: Baęımsız Deęişken
y_k	: Rastgele Deęişkenin Deęeri
X_t	: Güvenilir Deęişken
ϕ_p	: Otoregresif Parametresi
θ_q	: Hareketli Ortalama Parametresi
β_i	: Regresyon Katsayısı
γ	: Mevsimsellik Üstel Yumuşatma Parametresi
L	: Mevsimsel uzunluk
I_t	: Mevsimsel Faktör Düzeltme
m	: Tahmin Edilecek Dönem Sayısı
ε_k	: Rastgele Hata Deęeri
y_k	: Rastgele Deęişken Deęeri
ξ_t	: Hata Deęeri
P_k	: Otokorelasyon Fonksiyon Deęeri (k gecikme)
$F(x)$	: Sigmoid Fonksiyonu
sum	: Aęırlık Deęeri ve Giriş Deęerinin Çarpılarak Elde Edilen Toplam Deęer
r_n	: Test Verisindeki Gerçek Deęer
Q	: Asimptotik Daęılım
α_{K0}	: Regresyon Katsayısı
e_i	: İstatistiksel Hata
pyFITS	: Python Esnek Görüntü Taşıma Sistemi

TEŞEKKÜR

Yüksek lisans tez çalışmamın; planlamasında, araştırılmasında, yürütülmesinde ve oluşumunda her türlü desteği veren, engin bilgi ve tecrübeleriyle bana yol gösteren, tezimi bilimsel temeller ışığında şekillendiren sayın hocam Dr. Öğretim Üyesi Erkan ÖZHAN 'a (Tekirdağ Namık Kemal Üniversitesi) sonsuz teşekkürlerimi sunarım.

Aralık, 2021

Mehmet Berke ÇOLAK
Yüksek Lisans Öğrencisi



1. GİRİŞ

Giriş bölümünde enerji kavramı, yenilenebilir enerji kavramı ile ilgili bilgiler, Türkiye ve dünya için yenilenebilir enerjinin öneminden ve potansiyel gücünden bahsedilmiştir. Ek olarak yaptığımız tez çalışmasının amacı giriş bölümünde bahsedilmiştir.

İnsan yaşamı boyunca, enerji, insanların vücutlarında kullandığı ve depoladığı enerjiden, yaşamı daha kolay ve daha iyi hale getirmek için ürettikleri enerjiye kadar insanın varoluşunun hayati bir yönü olmuştur ve her zaman olacaktır [1]. Birçok kurum ve kuruluş, enerji taleplerini tahmin etmek için çalışmaktadır [2]. Küresel enerji talebi, artan dünya nüfusu ve teknolojiye gelişmeler sayesinde yaşam standartlarının iyileştirilmesine paralel olarak artmıştır ve dünyanın enerji talebi öncelikle petrol ve doğal gaz gibi fosil yakıtlarla karşılanmaktadır [3]. Enerji kullanımının tarihini değerlendiren ülkeler, nükleer ve fosil yakıtlardan yenilenebilir enerji kaynaklarına genel enerji geçişi olmadan dünyada sürdürülebilir bir gelecek olmadığını fark etmişlerdir [4]. Yenilenebilir enerji, doğal kaynaklardan gelen bir enerji olup bu kaynaklar okyanus, rüzgar, güneş, hidroelektrik ve dalgalar olarak adlandırılmış olup yenilenebilir kaynakların bazı dezavantajları olsa da çevreye daha az zararlıdır ve çeşitli yenilenebilir enerji kaynaklarından gelen toplam enerji çok büyüktür [5].

Yenilenebilir enerjinin (yani jeotermal, rüzgar, güneş ve biyokütle) toplam enerji tüketimindeki payını artırmaya yönelik enerji politikası, sürdürülebilir kalkınmanın temel gerekliliği olarak kabul edilmektedir çünkü yenilenemeyen enerji (yani kömür, petrol ve doğalgaz) küresel ısınma ve iklim değişikliğinin tetikleyicisi olarak kabullenilmiştir [6]. Yenilenebilir enerjinin doğa döngüsü içerisinde bir sonraki gün aynen mevcut kalabilmektedir. Dahası karbon emisyonunu azaltmasıyla çevre dostudur. Son birkaç on yıl, yenilenebilir enerji kaynaklarının ve enerji olarak hizmet veren teknolojilerin geliştirilmesinde küresel ısınmanın hızlandırılmış olduğunu göstermiştir [7]. Yenilenebilir enerji (RE) teknolojileri, daha fazla sürdürülebilir ve kendi kendine yeterli enerji arzı için çözümler sağlar [8]. 2012 yılında Birleşmiş Milletler tüm dünya için Uluslararası Sürdürülebilir Enerji Yılı'nı ilan etmiştir ve şu andan itibaren 2030'a kadar, bu duyuruda gerçekleştirilecek üç hedef belirlenmiştir bunlar çeşitli kaynaklarda enerji verimliliğinin teşvik edilmesi, yenilenebilir ve yeni enerji kaynaklarının kullanımının artırılması modern enerji özelliklerinin dünya çapında kabul edilmesini desteklemek hedefleri belirlenmiştir[9]. Yenilenebilir enerji kaynakları arasında güneş enerjisi, rüzgar enerjisi, jeotermal enerji, okyanus enerjisi, geri dönüştürülebilir özellikli ve düşük düzeyde çevre kirliliği bulunur [10].

Türkiye zengin bir çeşitlilik ve yenilenebilir enerji kaynakları potansiyeline sahip olmasına rağmen yenilenebilir enerji potansiyelinin %1'inden daha azını kullanmaktadır ve jeotermal enerji potansiyelinin %8'ine ve yüksek miktarda hidro enerji ve rüzgar enerjisi potansiyeline sahiptir ayrıca coğrafi konumundan dolayı büyük miktarda güneş enerjisi almaktadır [11]. Türkiye yükselen bir ekonomi ve Türkiye'nin enerji tüketimi özellikle son yıllarda hızla artmaktadır. Örneğin Uluslararası Enerji Ajansı (2018) (bundan böyle IEA) verilerini baz olarak aldığımızda Türkiye'nin enerji tüketimi 1990'da 40.169 ktoe iken, 2000'de 57.908 ktoe ve 2014'te 85.545 ktoe' ye ulaşmıştır. Demografik eğilimler, ekonomik büyüme ve artan kişi başı gelir. Türkiye'nin enerji tüketiminin önümüzdeki on yıla kadar yılda %4–6 artacağı beklenmektedir. Öte yandan, ülkenin enerji üretimi, enerji tüketimine göre zayıf görünmektedir [12]. Türkiye'nin hidroelektrik potansiyeli 2020 yılına kadar elektrik enerjisi talebinin %46'sını karşılayabilmekte olup hidroelektrik potansiyel basit ve ekonomik olarak geliştirilebilmektedir [13]. Türkiye'nin hidroelektrik potansiyeli 26 ana nehir bölgesine dağılmıştır ve 135'i geliştirilmekte olan hidroelektrik santral inşası için uygun 678 saha bulunmaktadır [14]. Bu sahaların toplam brüt potansiyeli yaklaşık 37 GW ve toplam enerji üretim kapasitesi 127 TWh/yıl 'dır. Şu anda toplam hidroelektrik güç potansiyelinin (yaklaşık 13.000 MW) %35'i faaliyette olup ulusal kalkınma planı, bunu 2020 yılına kadar %100'e ulaştırmayı hedeflemektedir [15].

Türkiye, genç nüfusu ve kişi başına artan enerji talebi, hızla büyüyen kentleşmesi ve ekonomik gelişimiyle son yirmi yıldır dünyanın hızla büyüyen güç pazarlarından biridir [16]. Biyoenerji, Türkiye'deki yenilenebilir enerji üretiminin yaklaşık üçte ikisini temsil etmektedir [17]. Toplam geri kazanılabilir biyoenerji potansiyeli 2000' de 16,8 Mtoe ve 2008'de 14,2 Mtoe olarak tahmin edilmiştir [18]. Yapılan tahminler; tarım atıklarından, ormancılıktan ve ahşap işleme artıklarından geri kazanılabilir enerji potansiyeline dayanmaktadır [19]. Türkiye jeotermal enerji potansiyeli ise dünyada ilk yedi ülke arasında yer almasına rağmen 28,3 MW' e kurulu kapasitesiyle (20,4 MW Denizli-Kızıldere ve 7,9 MW Aydın-Salavatlı) son sırada yer almaktadır [20]. MTA tarafından 1962 yılından bu yana yapılan çalışmalarda ise Ege Bölgesi, Kuzeybatı Anadolu ve Orta Anadolu'da 170 jeotermal saha tespit edilmiştir [21]. Türkiye'de rankine ve flash çevrimleri ile elektrik üretimi için yeterli olan bilinen 11 jeotermal saha bulunmaktadır [22]. Bu alanların sadece %7'si yüksek enerjiye sahiptir ve elektrik üretimi için yeterlidir [23]. Türkiye'nin rüzgar enerji potansiyeli ile ilgili Türkiye'deki teknik rüzgar enerjisi potansiyelinin 88.000 MW, ekonomik potansiyelin teknik duruma bağlı olarak yaklaşık 10.000 MW olduğu belirlenmiştir [24]. Türkiye'deki kara ve deniz alanlarının rüzgar gücü

potansiyelini, rüzgar özelliklerini sergilemek, enerji planlamasına katkıda bulunmak ve rüzgar enerjisinin gelecekteki gelişimi için olanakları değerlendirmek doğrultusunda; Elektrik İşleri Etüt İdaresi Genel Müdürlüğü (EİE) tarafından Türkiye Rüzgar Enerjisi Potansiyeli Atlası (REPA) geliştirilmiştir [25]. Türkiye'nin güneş enerji potansiyeli durumu incelendiğinde Türkiye'nin jeopolitik konumu sayesinde birçok ülkeye göre güneş enerjisi konusunda daha avantajlı olduğu için Türkiye'nin yıllık güneş enerjisi miktarı 1015 kWh ve buda enerji üretimimizin 10.000 eklenmesiyle yıllık radyasyon miktarı 1400-2000kWh/m² değerindedir [26]. Ortalama direk güneş ışığı süresi 7,5 saat olduğu için özellikle fotovoltaik sistemle lisanssız olarak elektrik enerjisi üretimi popüler hale gelmektedir. Çünkü bu sistem istikrarlı bir şekilde enerji sağlayabilmektedir ve Türkiye'nin toplam güneş enerjisi potansiyeli ise 10.000 MW olup önceki güneş enerjisi teknolojileri ise aktif ısıtma sistemi, termal güç ve fotovoltaik sistemdir [27]. Türkiye'nin hidroelektrik enerji potansiyeli ile ilgili Türkiye'nin yıllık ortalama yağış 643 mm olması Türkiye'nin yıllık ortalama yağış 500 km³ debi olup coğrafi ve jeolojik kısıtlamalar dikkate alındığı zaman toplam tüketilebilir su hacmi 107 km³ değerine düşmesi sonucuyla DSİ'ye göre mevcut potansiyelin %35'i şu anda kullanımdadır[28].Yapılan çalışmada Türkiye'nin yenilenebilir enerji üretim oranının yıllar içinde tahmin edilerek yenilenebilir enerjinin üretiminde yaşanan olası durumların eksiklerini görerek buna göre mevcut durumun daha iyi nasıl olacağıyla ilgili yapay sinir ağları ve zaman serisi analizi tahminleme yöntemleri kullanılmıştır. Türkiye'nin bu alan hakkında gelecek yıllarda izleyeceği yol hakkında karar verme sürecinde yardımcı olabilmesi ve hangi yöntemin daha iyi olduğunun tespit edilmesi amaçlanmıştır.

2. KAYNAK ÖZETLERİ

Önceki çalışmalar bölümünde, kronolojik olarak Türkiye'de ve dünyada yenilenebilir enerji kaynakları ile ilgili tahminleme yöntemleri kullanılarak yapılan çalışmaların literatür taramaları kısa özetler halinde tanıtılmıştır. Ayrıca yapılan çalışmaların yılı, adı, yapanlar ve kullandıkları tahminleme yöntemleri tablo halinde gösterilmiştir.

Paoli ve ark. [29] yaptıkları çalışmada yenilenebilir enerji alanında Yapay Sinir Ağlarının (YSA) bir uygulamasını sunmuştur. Hem yenilenebilir enerji alanında hem de zaman serisi tahminlerinde YSA mimarilerinin en çok kullanıldığı Çok Katmanlı Algılayıcı (MLP) ağına özellikle bakmıştır. Yatay bir yüzeyde küresel güneş radyasyonunun günlük tahmini için bir metodoloji geliştirmek için bir MLP ve bir geçici zaman serisi ön işleme kullanmıştır. NRMSE yaklaşık olarak %21 ve RMSE yaklaşık olarak 3,59 MJ/m² ile elde edilen değerler

doğrultusunda ilk sonuçlar umut vermiştir. Optimize edilmiş MLP; ARIMA teknikleri, bayesci çıkarım, markov zincirleri ve k-en yakın komşular gibi geleneksel ve referans yöntemlere benzer veya hatta onlardan daha iyi tahminler sunmuştur. Ayrıca, önerilen veri ön işleme yaklaşımının, markov zincirleri veya bayes çıkarım gibi geleneksel tahmin yöntemlerine kıyasla yaklaşık %6'lık tahmin hatalarını önemli ölçüde azaltabileceğini ortaya çıkarmıştır.

Hocaoğlu ve Karanfil [30] yaptıkları çalışmada yenilenebilir enerji kaynaklarına ilişkin tahmin çalışmalarında önerilen modellerin girdileri olarak seçilen değişkenler arasındaki nedensellik ilişkisinin araştırılmadığından bahsedilmiştir ve bu çalışmanın konusu ise rüzgar veya sıcaklık tahmin modellerinin tahmini girdi değişkenlerine karar vermek için yeni bir yaklaşım önerilmiştir. Bu yaklaşım daha spesifik olarak bazı meteorolojik değişkenler arasındaki Granger nedensellik ve dürtüsel yanıt analizleri zaman serisi teknikleri kullanılmıştır. Çalışmamızı yürütmek için Türkiye'nin farklı bölgelerinden elde edilen rüzgâr hızı, sıcaklık ve basınç verileri kullanılmıştır. Yaptıkları çalışmada elde edilen analiz sonucunda bu değişkenler arasında çift yönlü nedensel ilişkilerin var olduğunu ve kısa rundinamiğin lokasyona göre farklılık gösterdiği (iç bölgelere karşı kıyı alanı) ortaya çıkmıştır. Tahmin doğruluğunu iyileştirmek için yenilenebilir enerji modellerinin buna göre inşa edilmesi gerektiği sonucuna varıldığı ortaya çıkmıştır.

Cadenas ve Rivera [31] yaptıkları çalışmada Baja California'daki Isla de Cedros, Zacatecas'taki Cerro de la Virgen ve QuintanaRoo'daki Holbox'taki rüzgar hızı tahmini sunulmuştur. Zaman serileri için yaklaşık bir ay boyunca farklı sahalarda gerçekleştirilen ölçümlerden doğrudan elde edilen ortalama saatlik rüzgâr hızı verileri kullanılmıştır. Rüzgâr hızı tahmini yapabilmek için otoregresif bütünleşmiş hareketli ortalama (ARIMA) modelleri ve Yapay Sinir Ağı (YSA) modellerinden oluşan hibrit modeller geliştirilmiştir. ARIMA modelleri ilk olarak zaman serilerinin rüzgâr hızı tahminlerini yapmak için kullanılmıştır, daha sonra elde edilen hatalar ile ARIMA tekniğinin tanımlayamadığı doğrusal olmayan eğilimler dikkate alınarak YSA oluşturulmuştur ve bu sayede nihai hatalar azaltılmıştır. Hibrit modeller geliştirildikten sonra, rüzgâr hızı tahminini yapmak için her bir alan için örneklemden 48 veri kullanılmıştır ve sonuçlar ayrı çalışan ARIMA ve YSA modelleriyle karşılaştırılmıştır. Üç yöntemi karşılaştırmak için ortalama hata (ME), ortalama kare hatası (MSE) ve ortalama mutlak hata (MAE) gibi istatistiksel hata ölçümleri hesaplanmıştır. Sonuçlar, hibrit modellerin incelenen üç bölgedeki ARIMA ve YSA modellerinden daha yüksek bir doğrulukla rüzgâr hızlarını tahmin ettiğini göstermiştir.

Golestaneh ve ark. [32] yaptıkları çalışmada yenilenebilir enerji tahmininde yer alan doğal belirsizlik nedeniyle belirsizlik ölçümü, güç sistemi çalışmasında kabul edilebilir düzeyde güvenilirlik ve karlılık sağlamak için kilit bir girdi olduğu ve gökyüzü görüntüleme ve bulut geçişleri hakkında bilgi olmadan yalnızca güç ve meteorolojik ölçümlerin mevcut olduğu güneş enerjisi üretimi durumu için bir teklif formüle edilmiştir ve değerlendirilmiştir. Ampirik araştırma, tahmin hatalarının dağılımının yaygın parametrik yoğunlukların hiçbirini takip etmediğini ortaya koymuştur. Bu nedenle parametrik olmayan bir yaklaşımın, çok kısa vadeli tahmin yoğunlukları, yani hızlı frekans güncellemeleriyle birkaç dakika ila bir saat öncesindeki teslimat süreleri oluşturmak için önerisini motive etmiştir ve güneş enerjisi üretiminin hem nokta hem de nicelik tahminlerini elde etmek için çeşitli şekillerde eğitilmiş hızlı bir regresyon modeli olarak Aşırı Öğrenme Makinesine (ELM) güvenilmiştir. Kıyaslama olarak dört olasılıklı yöntem uygulanmıştır. Rakip yaklaşımlar, farklı iklim bölgelerindeki iki güneş enerjisi üretim tesisi için bir dizi test durumuna göre değerlendirilmiştir ve bu da yaklaşımımızın, hesaplama açısından verimli bir şekilde becerikli ve güvenilir olasılıklı tahminler üretmemizi sağladığını göstermemize olanak tanımıştır.

Jiang ve ark. [33] yaptıkları bu çalışmada geleneksel fosil enerjinin hakim olduğu enerji tüketimi modeli, küresel enerji kaynağı kısıtlamalarına ve ekolojik çevrenin bozulmasına yol açmıştır ve bu zorluklar tüm dünyada önemli bir sorun haline geldiğinden dolayı şu anda Çin hükümeti, terminal enerji tüketiminde fosil enerji tüketimi katkısını önemli ölçüde azaltmayı hedeflemiştir. Terminal enerji ve enerji dönüşüm bağlantılarında yenilenebilir enerjinin geliştirilmesi, temiz düşük karbonlu enerji oranını önemli ölçüde artırmıştır ve yenilenebilir enerji terminali güç tüketiminin oranını doğru bir şekilde elde etmek için ilk olarak bu çalışmada brüt GSYİH, yenilenebilir enerji endüstrisine sabit yatırım, iller arası ve bölgeler arası yüksek voltaj iletiminin toplam uzunluğu dahil olmak üzere ayrıca Çin'in yenilenebilir enerji tarafından üretilen elektrik tüketimi fraksiyonunu etkileyen birçok ana faktörleri seçmiştir. İkincisi, sinyal ayrıştırma perspektifinden bakıldığında, veriler kaçınılmaz olarak çok fazla parazit ve gürültüye sebep olmuştur. Bu çalışmada sinyal distorsiyonun derecesini azaltmak için ampirik mod ayrıştırma (EMD) algoritması kullanılmıştır. Sinyali, çeşitli içsel mod fonksiyonları ve bir artık terim dâhil olmak üzere doğal modlara ayırmıştır. Daha sonra, Ters Kare Kök Doğrusal Birimler (ISRLU) aktivasyon fonksiyonu ile optimize edilmiş yeni bir aşırı öğrenme makinesi (ELM) tahmin modeli önerilmiştir ve ISRLU fonksiyonu, orijinal ELM algoritmasındaki örtük katman aktivasyon fonksiyonunun yerini almak için kullanılmıştır. Ardından, optimize edilmiş ELM tahmin modelinin parametrelerini optimize etmek için yeni

bir bakteriyel arama algoritması (BFOA) uygulanmıştır. Birden fazla öğrenme ve eğitim işleminden sonra optimum parametreler elde edilmiştir. Son olarak, yenilenebilir enerji tarafından üretilen Çin'in enerji tüketimi miktarını elde etmek için her bir IMF ve Res eğitim görevinin çıktısı birleştirilmiştir. Birkaç akıllı makine tahmin algoritmasının doğruluğunu karşılaştırmak için kök ortalama kare hatası (RMSE) dahil olmak üzere bazı istatistiksel göstergeler uygulanmıştır. Önerilen tahmin modelinin daha yüksek tahmin doğruluğuna sahip olduğunu ve deneysel bir analizle daha hızlı eğitim hızı sağladığını kanıtlamıştır. Son olarak, önerilen birleşik tahmin algoritması, Çin'in 2018'den 2030'a kadar yenilenebilir enerji terminali güç tüketimini tahmin etmek için uygulanmıştır. Bu çalışmanın sonucuna göre Çin'in yenilenebilir enerji terminali elektrik tüketiminin kademeli bir büyüme trendi gösterdiği ve 2030'da 3300 milyar kWh'yi aşacağı ve bunun 2030'da yaklaşık olarak %38'lik bir yenilenebilir enerji terminal güç oranını temsil edeceği tahmin edilmiştir.

Wang ve ark. [34] yaptıkları çalışmada yenilenebilir enerji, küresel elektrik enerjisi şebekesinde giderek daha popüler hale geldikçe, yenilenebilir enerji tahminlerinin doğruluğunun iyileştirilmesi, güç sistemi planlaması, yönetimi ve operasyonları için kritik öneme sahip olduğu belirtilmiştir ancak yenilenebilir enerji verilerinin aralıklı ve kaotik doğası nedeniyle bu zorlu bir görev olmuştur. Bugüne kadar, yenilenebilir enerjinin tahmin doğruluğunu iyileştirmek için fiziksel modeller, istatistiksel yöntemler, yapay zekâ teknikleri ve bunların hibritleri dahil olmak üzere çeşitli yöntemler geliştirilmiştir. Bunlar arasında, verilerdeki doğal doğrusal olmayan özellikleri ve üst düzey değişmez yapıları keşfetme yeteneğine sahip, gelecek vaat eden bir makine öğrenimi türü olarak derin öğrenme, literatürde sık sık rapor edilmiştir. Bu belge, etkinliğini, verimliliğini ve uygulama potansiyelini keşfetmek için derin öğrenmeye dayalı yenilenebilir enerji tahmin yöntemlerinin kapsamlı incelemesini sunmuştur. Derin öğrenmeye dayalı mevcut deterministik ve olasılıksal tahmin yöntemlerini derin inanç ağı, yığın otomatik kodlayıcı, derin tekrarlayan sinir sistemi ve diğerleri olarak dört gruba ayırmıştır. Ayrıca, tahmin doğruluğunu iyileştirmek için, uygulanabilir veri ön işleme tekniklerini ve hata sonrası düzeltme yöntemlerini incelenmiştir. Çeşitli derin öğrenmeye dayalı tahmin yöntemlerinin kapsamlı analizi ve tartışması verilmiştir. Çalışmanın sonucunda ise mevcut araştırma faaliyetlerini, zorlukları ve gelecekteki potansiyel arama yönleri keşfedilmiştir.

Sweeney ve ark. [35] yaptıkları çalışmada rüzgâr ve güneş yenilenebilir enerji için tahmin, bu kaynaklardan üretilen enerji miktarı arttıkça daha da önem kazanmıştır. Tahmin

becerisi gelişmiştir ancak tahminlerin kullanım şekli de artmıştır. Çalışmada rüzgâr ve güneş enerjisini tahmin etmenin son teknolojisine kısa bir genel bakış sunulmuştur. Hem deterministik hem de olasılıklı tahmin için, dakikalardan günlere kadar zaman ölçekleri için istatistiksel ve fiziksel modellemedeki yaklaşımları açıklanmıştır. Odak noktamız daha sonra yenilenebilir enerji tahminlerinin geleceğini düşünmek için değişmiştir. Tahmin becerisinde büyük gelişme potansiyeli gösteren son gelişmeleri tartışılmıştır. Tahminin ötesinde, risk kısıtlamalarına tabi karar vermeye yardımcı olması için gerekli olacak yeni ürünleri değerlendirilmiştir. Geleceğe yönelik tahmin ürünlerinin olasılık bilgilerini içermesi, ancak bunları son kullanıcıya ve onların belirli karar verme sorunlarına göre uyarlanmış bir şekilde sunulması gerekmiştir. Bu sektörde faaliyet gösteren işletmeler, bu alanda daha fazla insan rekabet ettikçe iş modellerinde bir değişiklik görebilmiştir ve farklı ürünler için farklı beceri, veri ve modelleme kombinasyonları gerekli olabilmektedir. Verilerin işleminin kendisi, sağlayıcıların ve son kullanıcıların güvenilir, ancak merkezi olmayan bir şekilde etkileşime girmesine olanak tanıyan blockchain teknolojisinin benimsenmesiyle değişebilmiştir. Bu çalışmanın sonucuna göre yüksek miktarda yenilenebilir enerji içeren senaryolar için yeni endüstri gereksinimleri ve zorlukları tartışılmıştır. Yeni tahmin ürünleri, yenilenebilir enerjilerin güç sistemi üzerindeki etkisini modelleme ve sistem güvenliğini garanti etmede dağıtım araçlarına yardımcı olma potansiyeline sahip olmuştur.

Tharani ve ark. [36] yaptığı çalışmada dünya sanayilere, evlere, taşıtlara vb. güç sağlamak için geleneksel olmayan enerjiyi oybirliğiyle kullanmaya çalışırken, temel sınırlayıcı faktörlerden biri, geleneksel olmayan enerji kaynaklarının enerji mevcudiyetindeki dalgalanmalar olmuştur. Rüzgâr türbinleri veya fotovoltaik hücreler tarafından üretilen güç; rüzgâr hızı, nem, güneş ışınımı gibi kontrol edilemeyen faktörlere bağlı olmuştur. Böyle bir durumda yenilenebilir enerjinin şebekelere entegrasyonu zorlaşmıştır. Sonunda, enerji arzı ve talebi arasında bir denge sağlamak düzensiz olmuştur. Bu, yenilenebilir enerjiye geçişin elektrik şebekesini istikrarsızlaştırmadan son derece verimli bir şekilde yapılabilmesi için rüzgâr ve güneş gibi geleneksel olmayan enerji kaynakları aracılığıyla mevcut olan enerji miktarını tahmin etmeyi gerektirmiştir. Yenilenebilir enerji endüstrisi, yenilenebilir enerji tahminlerinde kullanılacak bol miktarda veriye sahip olduğundan, makine öğrenimi teknikleri yenilenebilir enerji ile başa çıkma şeklimizde devrim yaratmıştır. Çalışmanın sonucuna göre dünya yüzeyinde bulunan güneş ışınımını tahmin etmek için Doğrusal Regresyon, Sinir Ağları Regresyonu, Rastgele Orman Regresyonu ve Ekstra Ağaç Regresyon modellerinin etkinliği açıklanmıştır.

Severiano ve ark.[37] yaptıkları çalışmada yenilenebilir enerji sistemlerinde öngörü, girdileri veri dağıtımında bazı belirsizlikler ortaya çıkardığı için zorlu bir sorun olduğunu ve bu tür sistemler tarafından kaydedilen ve geliştirilmiş performans beklentisiyle bir tahmin modeli ile araştırılabilen artan bir bilgi hacmi var olduğu tespit edilmiştir. Çalışma, e-MVFTS' yi (evrimleşen çok değişkenli bulanık zaman serileri), bulanık zaman serilerine dayanan gelişen bir tahmin modelini ve mekansal-zamansal olarak çok değişkenli zaman serilerini kullanan TEDA (Tipiklik ve Eksantriklik Veri Analitiği) Çerçevesine dayalı gelişen bir kümeleme yöntemini tanıtmıştır. Model, veri dağıtımındaki veya veri akışlarındaki kavram sapmalarındaki değişikliklerle başa çıkmak için bir uyarlama mekanizmasına sahip olduğunu ve gelişen kümeleme yöntemi, veri noktaları ulaştıkça ve çevrimiçi bir şekilde işlendikçe ayarlanmıştır. Güneş ve rüzgâr enerjisi tahmin problemlerinin yanı sıra konsept kayması olaylarını uygulamada performansı değerlendirilmiştir. Model, pyFITS kütüphanesi kullanılarak python programlama dilinde geliştirilmiştir. Bu çalışma sonucunda farklı deneylerdeki iyi sonuçlar, e-MVFTS modelinin yenilenebilir enerji sistemlerinde veri akışı ile ilgili sorunları tahmin etmede kullanılması sağlanmıştır.

Gonçalves ve ark. [38] yaptıkları çalışmada birden fazla yenilenebilir enerji santrali sahibi arasındaki veri alışverişi, zaman serisi verilerindeki mekansal-zamansal bağımlılıklar sayesinde tahmin becerisinde bir iyileşmeye yol açabildiği ancak, ticari rekabet faktörleri sebebiyle bu farklı şehirler verilerini paylaşmak konusunda isteksiz olmuşlardır. Bu gizlilik sorununu çözmek için bu makale, veri dönüştürme tekniklerini çarpanların alternatif yön yöntemiyle birleştiren yeni bir gizliliği koruyan çerçeve formüle etmiştir ve bu yaklaşım, modeli yalnızca dağıtılmış şekilde tahmin etmeye değil, aynı zamanda veri gizliliğini, katsayıları ve kovaryans matrisini korumaya da izin vermiştir. Ayrıca, eşler arasındaki asenkron iletişim, model uydurmada ele alınmıştır. İki farklı iş birliği şeması merkezi ve eşler arası dikkate alınmıştır. Çalışmanın sonucuna göre bir güneş enerjisi veri kümesinin sonuçları, önerilen yöntemin gizlilik ihlallerine ve iletişim başarısızlıklarına karşı sağlam olduğunu ve gizlilik koruması olmayan bir modelle karşılaştırılabilir bir tahmin becerisi sunduğunu göstermiştir.

Çizelge 2.1. 2010-2019 Yılları Arasında Yapılan Önceki Çalışmalar

Yıl	Çalışma Adı	Yapanlar	Kullandıkları Tahmin Yöntemleri
2010	Sinir ağları kullanılarak önceden işlenmiş günlük güneş radyasyonu zaman serilerinin tahmini	Christophe Paoli, Cyril Voyant, Marc Muselli ve Marie-Laure Nivet	MLP, ARIMA, k-en yakın komşu algoritması, bayes çıkarımı ve markov zincirleri
2010	Melez bir ARIMA-ANN modeli kullanılarak Meksika'nın üç farklı bölgesinde rüzgar hızı tahmini	Eresmo Cadenas ve Wilfrido Rivera	ARIMA ve YSA
2013	Yenilenebilir Enerji Modellemesi için zaman serisi tabanlı bir yaklaşım	Fatih Onur Hocaoglu ve Fatih Karanfil	Granger nedensellik ve dürtüsel yanıt analizi
2016	Yenilenebilir enerji üretiminin güneş enerjisine uygulama ile çok kısa vadeli parametrik olmayan olasılıksal tahmini	Faranak Golestaneh, Pierre Pinson ve H. B. Gooi	ELM
2019	Çin'in yenilenebilir enerji terminali güç tüketiminin ampirik mod ayrıştırmasına ve bakteriyel toplayıcı algoritması ile optimize edilmiş gelişmiş aşırı öğrenme makinesine dayalı tahmini	Peng Jiang, Jun Dong ve Hui Hang	ELM, BFOA ve EMD
2019	Yenilenebilir enerji tahmini için derin öğrenmenin gözden geçirilmesi	Huaizhi Wang, Zhenxing Lei, Xian Zhang, Bin Zhou ve Jianchun Peng	Derin öğrenmeye dayalı mevcut deterministik ve olasılıksal tahmin yöntemleri, inanç ağı, yığın otomatik kodlayıcı ve derin tekrarlayan sinir sistemi
2019	Yenilenebilir enerji için tahminin geleceği	Conor Sweeney, Ricardo J. Bessa, Jethro Browell ve Pierre Pinson	Deterministik tahmin ve olasılıklı tahmin

Çizelge 2.2. 2020-2021 Yılları Arasında Yapılan Önceki Çalışmalar

Yıl	Çalışma Adı	Yapanlar	Kullandıkları Tahmin Yöntemleri
2020	Yenilenebilir enerji tahmini için makine öğrenimi modelleri	Kusum Tharani, Neeraj Kumar, Vishal Srivastava, Sakshi Mishra ve M. Pratyush Jayachandran	Doğrusal regresyon, sınır ağları regresyonu, rastgele orman regresyonu ve ekstra ağaç regresyon
2021	Yenilenebilir enerji tahmini için gizliliği koruyan dağıtılmış öğrenme	Carla Goncalves, Ricardo J. Bessa ve Pierre Pinson	Vektör otoregresyonu, gizliliği koruyan ve dağıtılmış öğrenme
2021	Yenilenebilir enerji sistemlerinde uzaysal-zamansal tahmin için gelişen bulanık zaman serileri	Carlos A. Severiano, Petronio Candido de Lima Silva, Miri Weiss Cohen ve Frederico Gadelha Guimaraes	TEDA ve e-MVFTS

3. MATERYAL VE YÖNTEM

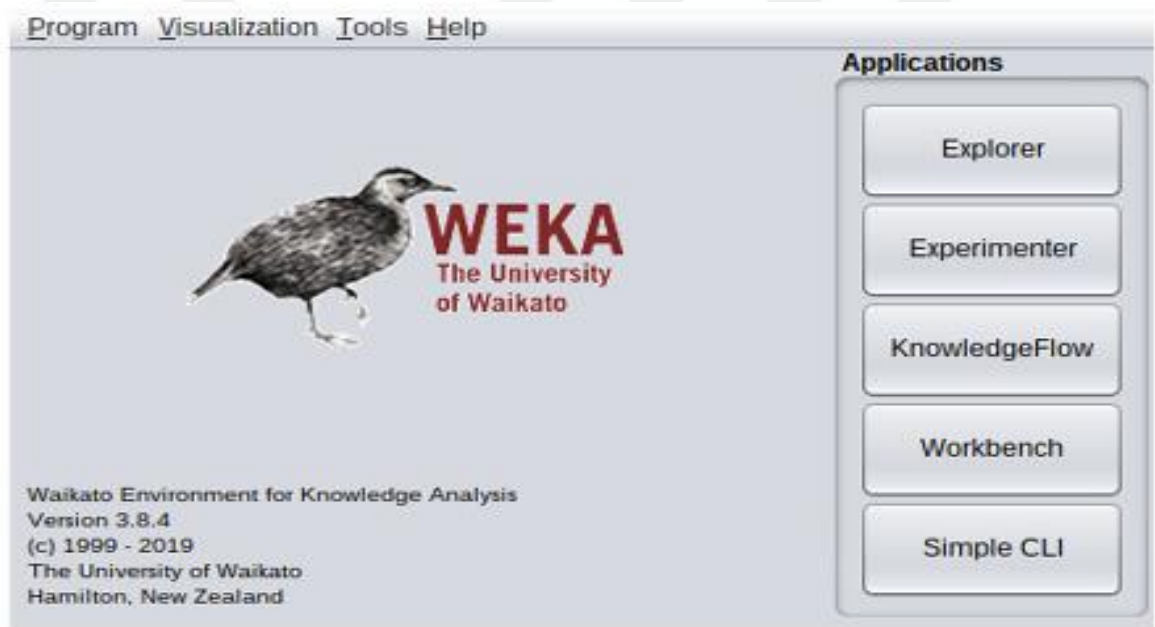
3.1. MATERYAL

Çalışmada kullanılan ana materyal olarak 1960-2019 yılları arasında Türkiye'nin hidroelektrik hariç yenilenebilir enerji kaynaklarından elektrik üretiminin toplam oranını gösteren veri seti kullanılmıştır. 1960-2015 yılları arasındaki veriler dünya bankası sitesinden ve 2015-2019 yılları arasındaki veriler enerji bakanlığından alınmıştır. Tahminleme yöntemleri Weka uygulaması ve R-Studio uygulaması üzerinde R programlama dili kullanılarak yapılması hedeflenmiştir. Weka ve R-Studio hakkında bilgi verilmiştir. Ayrıca R programlama dili hakkında da bilgi verilmiştir.

Weka, Java'da yazılan ve GPL (Genel Kamu Lisansı) altında yayınlanan veri madenciliği görevleri için kapsamlı bir makine öğrenimi algoritmaları koleksiyonu olup, veri ön işleme, sınıflandırma, regresyon, kümeleme, ilişkilendirme kuralları ve görselleştirme için araçlar içeren makine öğreniminde önde gelen açık kaynak uygulamasıdır [39]. Yeni

Zelanda'daki Waikato Üniversitesi'nde geliştirilmiştir [40]. Weka uygulamasında GUT'si aracılığıyla, kullanıcıların kod yazmasına gerek kalmadan derin öğrenmeyi erişilebilir kılar. GUI kullanıcıların nitelik-ilişki dosya formatında (ARFF) veri yükleme, bir sinir ağı mimarisini yapılandırma, bir deneysel protokol seçme ve denemeyi çalıştırma olan basit adımları kullanarak deneyler gerçekleştirmesini sağlar [41]. Ek olarak Java veritabanı bağlantısı sayesinde veri tabanları üzerinden veri alma desteği mevcuttur [42].

Uygulamayı açtığımızda kullanıcının kullanım için beş Weka uygulamasındaki arabiriminden olan explorer, knowledgeflow, experimenter, workbench ve simple CLI'den birini seçmesine izin vermesini sağlayan aynı zamanda bu pencereden paket yöneticisi ve görselleştirme araç takımları gibi çeşitli diğer araçlara erişmek de mümkündür [43].

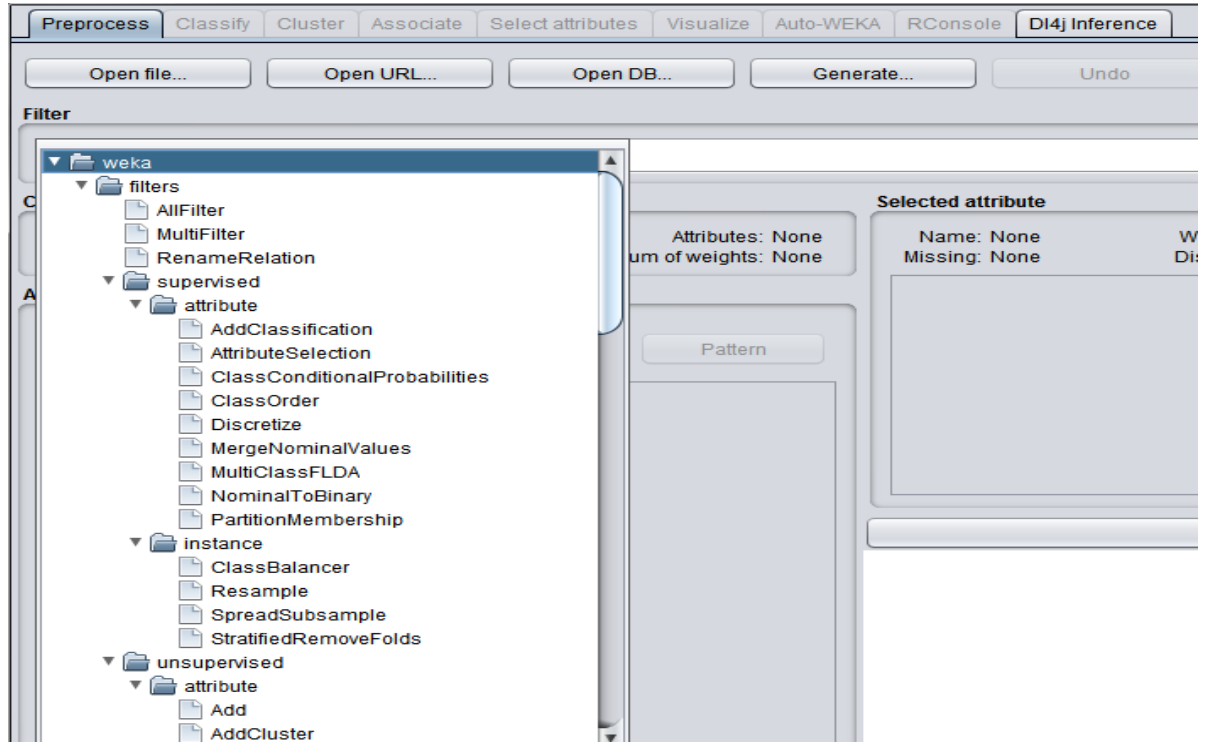


Şekil 3.1.Weka GUI Seçme Ekranı [43]

Explorer eldeki verileri keşfetmek ve çeşitli görselleştirmeler yoluyla verilerin nasıl olduğuna dair iç görüler elde etmek görevini yürütür [44]. Knowledgeflow kullanıcının akan veri işleme için yapılandırmalar tasarlamasına olanak tanır, uygulamanın üçüncü ara yüzü olan experimenter kullanıcının sınıflandırma ve regresyon yöntemlerini uygularken temel bir pratik soruyu yanıtlamasına yardımcı olmayı amaçlamakta olup verilen problem için hangi teknikler ve parametre değerleri daha iyi çalıştığını tespit eder, workbench olarak adlandırılan dördüncü ara yüz ise diğer üçünü tek bir uygulamada birleştiren birleşik bir grafik kullanıcı ara yüzüdür

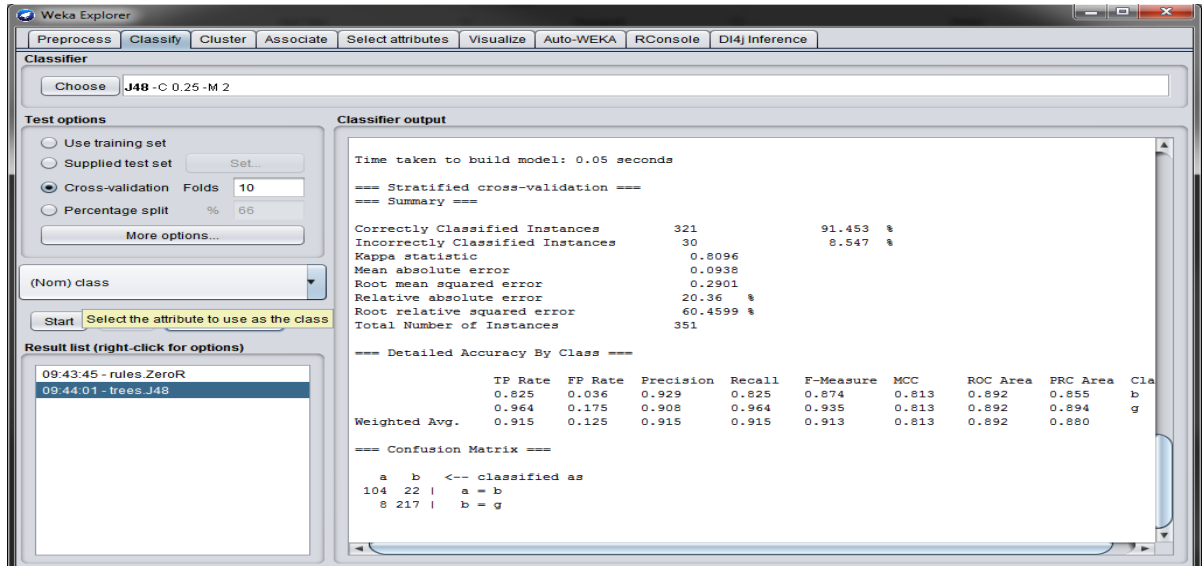
[45]. Simple CLI (Komut Satırı Ara Yüzü), sınıf yolunun ağırlaşmasına bakılmaksızın her Weka sınıfına yani kümeleyişler, filtreler vb. için tam giriş sağlar [46].

Weka uygulamasında ön işleme yeteneği, verilerin eş görünüm ve öznitelik düzeylerinde işlenmesini sağlayan filtreler adı verilen kapsamlı bir yönetim birimi kümesinde toplanmıştır [47].



Şekil 3.2. Weka Uygulamasında Filtre Algoritmaları

Weka uygulamasında kullanılan makine öğrenmesi modelinin performans ölçütleri, WEKA sınıflandırıcı çıktısında bulunan kök ortalama kare hatası (RMSE) ve ortalama mutlak hata (MAE) kullanılarak değerlendirilir [48].



Şekil 3.3. Weka Uygulamasında Kullanılan Modelin Performans Ölçütleri

R programlama dili, özellikle veri bilimciler, akademisyenler ve sağlık araştırmacıları tarafından istatistiksel analizler için kullanılan bir programlama dilidir [49]. 70'lerde Bella laboratuvarlarından ortaya çıkan S istatistiksel programlama dilinin bir uygulamasıdır [50]. R programlama dili bir bilgisayar dili olduğundan, kullanıcıların aşına olduğu programların çoğundan biraz farklı çalışıp yürütülmekten ziyade program tarafından değerlendirilen komutları yazmanız gerekmektedir [51]. R programlama dili etkileşimli ve toplu olmak üzere iki durumda çalışmakta olup tipik olarak kullanılan etkileşimli durumdur [52]. Bir veri çerçevesi, R programlama dilinde veri depolama ve istatistiksel analiz için normal birimdir, aynı uzunluktaki değişkenlerden oluşur, ancak aynı türden olma zorunluluğu yoktur ayrıca Şekil 3.4'te bir veri çerçevesi satır adlarına (burada: 1, 2, 3) ve değişken (sütun) adlarına (burada: ad, gelir) sahiptir [53].

```
> dat <- data.frame(name = c("Homer", "Flanders", "Skinner"),
                    income = c(1000, 23000, 85000))
> dat
      name income
1   Homer   1000
2 Flanders 23000
3  Skinner 85000
```

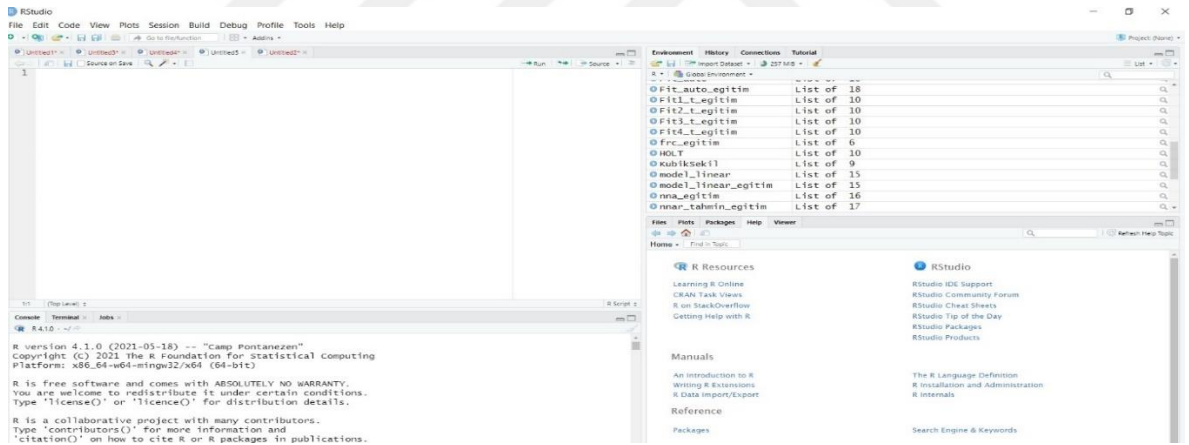
Şekil 3.4. R Programlama Dili Veri Çerçevesi Komutları [53]

Açık kaynak felsefesi ve çok yönlülüğü nedeniyle R programlama dili; esnekliği, kullanıcının C++ ve FORTRAN¹ gibi düşük seviyeli programlama dillerine, UNIX gibi betik

¹Sayısal hesaplama ve bilimsel hesaplama için uygun olan genel amaçlı, yordamsal, zorunlu programlama dilidir.

dillerine bağlanmasına, LaTeX² gibi kelime işlemcilerle harmanlanmasına ve paralel hesaplama yetenekleriyle büyük veri kümelerini işlemesine olanak tanır [54].

R programlama dili, tüm paketlerini ve diğer yazılımları ayrı makinelerde kurmaya bir alternatif, R-Studio uygulaması kullanmaktır [55]. R-Studio uygulaması, R programlama dilinin tümleşik geliştirme ortamıdır [56]. R-Studio uygulaması, R programlama dilinin çeşitli bileşenlerini (konsol, kaynak düzenleme, grafikler, geçmiş, yardım vb.) sorunsuz ve üretken tek bir çalışma ortamında birleştirmeyi amaçlamak hem yeni R kullanıcıları için öğrenme eğrisini kolaylaştırmak hem de daha ileri düzey kullanıcılar için yüksek üretkenlik araçları sağlamak üzere tasarlanmıştır [57]. R-Studio uygulaması, kullanıcıların işleme araçlarını ücretsiz olarak indirmelerine olanak tanır ve bu nedenle öğrenme kapsamı daha büyüktür [58]. R programlama dilindeki gereklilik, öğrencileri diğer istatistiksel yazılımlar gibi açılır menülere tıklamak yerine komutları girmeye zorlar [59]. R-Studio editörü, yaptığımız gibi etkileşimli olarak değil fakat eş zamanlı birden fazla kod satırının işletilmesi için gerekli durumlarla mücadele etmek için makuldür [60]. Şekil 3.5'te R-Studio uygulamasının kullanıcı ara yüzü gösterilmiştir.



Şekil 3.5. R-Studio Uygulamasının Grafiksel Kullanıcı Ara Yüzü

3.2. YÖNTEM

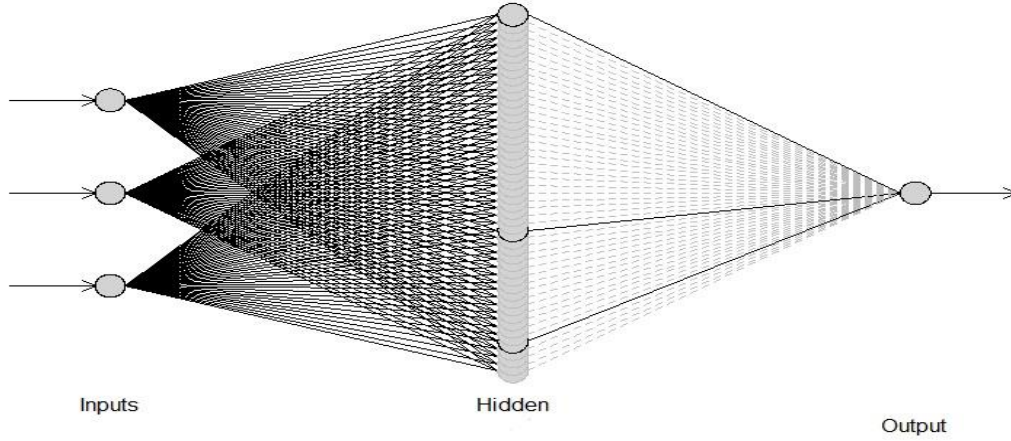
Yapılan çalışmada ana materyali işlemek için kullanılan yapay sinir ağları ve zaman serisi analizi yöntemleri ayrı başlıklar altında detaylı olarak incelenmiştir. Ayrıca zaman serisi analizi yöntemi ile ilgili bilgi verilmiş olup elimizde olan veri seti üzerinde kullanmış olduğumuz zaman serisi yöntemleri alt başlıklar altında ARIMA modelleme yöntemi, üstel

²TEX dizgi programı için belge düzenleme biçimidir.

düzeltilme yöntemi, doğrusal regresyon yöntemi, torbalama (bagging) yöntemi, sıralı minimum en uygun şekle sokma regresyon (SMOreg) algoritması ve yerel ağırlık öğrenme (LWL) yöntemi olarak incelenmiştir. Elimizde olan veri seti üzerinde kullanmış olduğumuz yapay sinir ağları yöntemde kendi içinde çok katmanlı algılayıcı (MLP), ileri beslemeli ağ (NNAR) ve aşırı öğrenme makinesi (ELM) yöntemlerinin her biri hakkında kısa bilgiler verilmiştir.

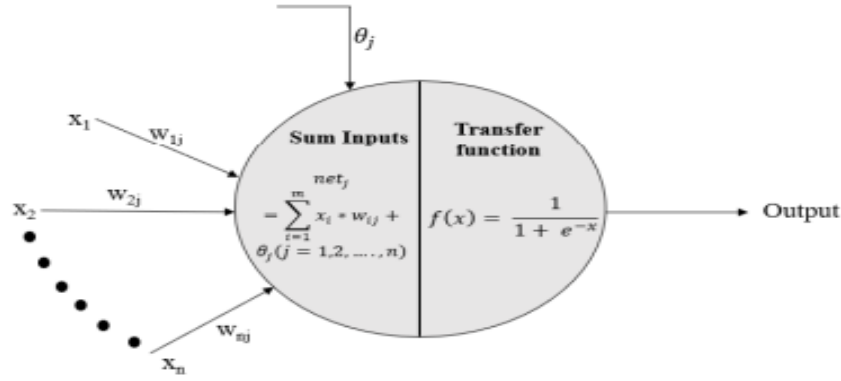
3.2.1. YAPAY SİNİR AĞLARI

Yapay sinir ağları (YSA), veri işleme tekniklerinden biridir [61]. YSA uygulamaları, zaman serilerinin geçmiş ve gelecek değerleri arasındaki işlevsel ilişkilere iyi bir yaklaşım oluşturma kabiliyetine dayanmaktadır [62]. Ek olarak verilerdeki dağınık ilişkileri öğrenme kabiliyeti vardır [63]. YSA geleneksel mimarisi, üç nöron katmanından oluşmaktadır bunlar giriş katmanı, gizli katman ve çıktı katmanıdır [64]. Yapay sinir ağları (YSA), bilgi setlerini görüntülemeye ve bilgi araştırmasında başlıklar vermede akıl almaz bir değerli şey olarak ortaya çıkabilir [65]. Giriş katmanı kullanıcı tarafından sağlanan birkaç farklı formattaki girdileri kabul etmekte olup gizli katmanı ise giriş ve çıkış katmanları arasında yer almakta olup gizli özellikleri ile kalıpları bulmak için tüm hesaplamaları yapmasıyla sonuç katmanında ise girdi, gizli katmanı kullanarak bir dizi dönüşümden geçer ve işlem sonunda çıkış katmanı kullanılarak iletilen çıktıyla sonuçlanmaktadır [66]. YSA modelleri, insan beyninin yapısından ve işlevinden esinlendiğinden nöronlar beynin en önemli bir bileşeni olup aralarında adens ağı oluşturan sinir hücreleridir ve tipik YSA modelleri, çok basitleştirilmiş nöron modellerini kullanır fakat yapay sinir ağları modeli yalnızca insan nöronlarının çok kaba özellikleri olabilir [67]. YSA modelinin şematik görünümü Şekil 3.6'da gösterilmiştir:



Şekil 3.6. YSA Modelinin Şematik Gösterimi

Güçlü örüntü tanıma yeteneklerine sahip olmasıyla mevcut gerçek yaşam problemine çok uygundur [68]. YSA uygulama alanımıza uygun, olası denetimli makine öğrenimi yaklaşımlarından yalnızca biridir [69]. Yapay sinir ağındaki bilgi ağı, ağıdaki nöronların sinapsları olarak bilinen ara bağlantıların gücü ile belirlenir [70]. Yapay bir nöronun modeli Şekil 3.7’de gösterilmiştir:



Şekil 3.7. Yapay Bir Nöron Modeli [71]

Bu şekilde $x_1, x_2, x_3, \dots, x_n$ nöronun girdileri olup nörona girdilerle birlikte bir önyargı da eklenmektedir, genellikle önyargı değeri 1'e ayarlanır, ağırlıkları $w_0, w_1, \dots, w_n$ olduğundan dolayı ağırlık, sinyale olan bağlantı olup ağırlık ve girdinin çarpımı, sinyalin gücünü verir. Bir nöron, farklı kaynaklardan çok sayıda girdi alır ve tek bir çıktıya sahiptir [72].

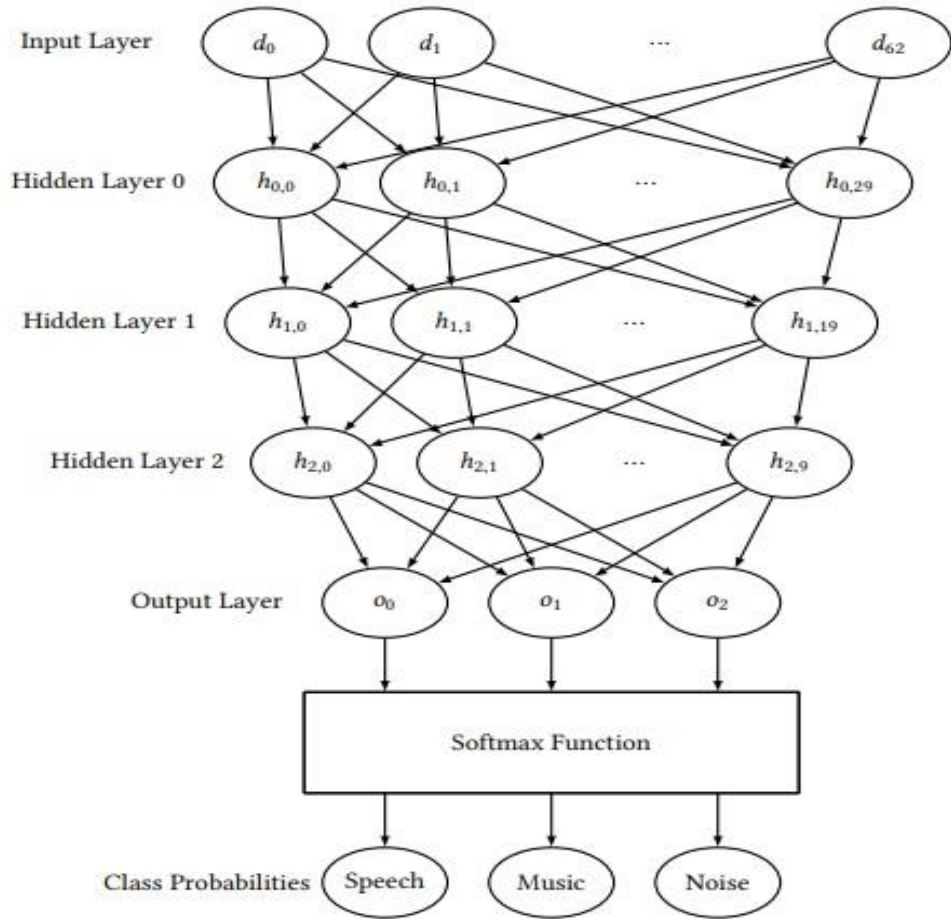
İşleme yeteneği, ağırlıklar adı verilen birimler arası bağlantı güçlerinde depolanır [73]. Giriş gücü, ağırlık değerine bağlı olup ağırlık değeri pozitif, negatif veya sıfır olabilmekte, negatif ağırlık sinyalin azaldığı veya engellendiği anlamına gelmekte, sıfır ağırlık ise iki nöron arasında bağlantı olmadığı anlamına gelmekte, ağırlıklar gerekli çıktı elde etmek üzere ayarlandığında gerekli çıktıyı elde etmek için YSA ağırlıklarını ayarlayan algoritmalar var olup bu ağırlık ayarlama sürecine öğrenme veya eğitim denmektedir[74]. En yaygın kullanılan aktivasyon fonksiyonlarından biri, $F(x)$ ile verilen sigmoid fonksiyon olup toplam (sum) kısmındadır [72]. $F(x)$ fonksiyonunu (3.1) denkleminde ve toplam (sum) değeri ise (3.2) denkleminde gösterilmiştir:

$$F(x) = \frac{1}{1 + e^{-sum}} \quad (3.1)$$

$$sum = \sum_{i=0}^n x_i W_i \quad (3.2)$$

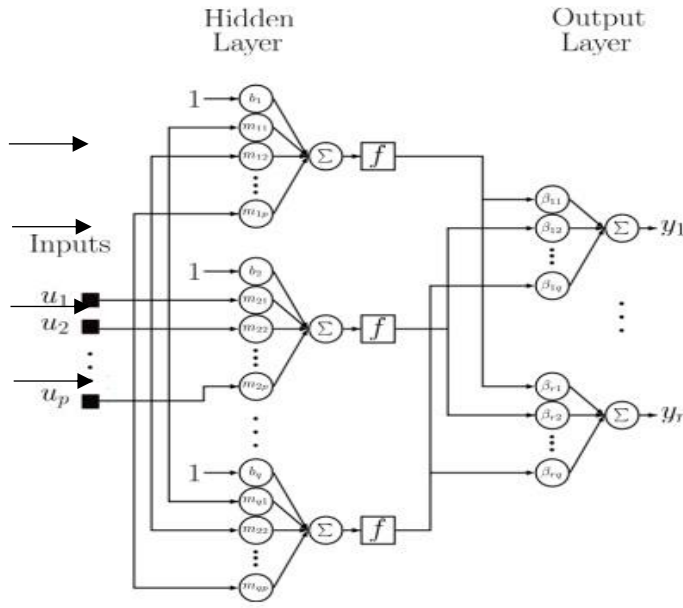
Çok katmanlı algılayıcı (MLP), bir giriş katmanı, bir çıkış katmanı ve bazı gizli katmanlara sahip ileri beslemeli katmanlı bir ağıdır [75]. Çok katmanlı algılayıcı (MLP) ağı, yapay sinir ağının en yaygın ve pratik mimarilerinden biridir [76]. MLP terimi belirsiz bir şekilde, ara sıra serbestçe herhangi birileri beslemeli yapay sinir ağı belirtmek için ara sıra dikkatli bir şekilde algılayıcının çeşitli katmanlarından oluşan sistemlere atıfta bulunmak için kullanılır[77]. MLP, genellikle karmaşık sorunlar için kullanılan bir algılayıcıdır. Çok katmanlı algılayıcının denklemi (3.3) denkleminde gösterilmiştir. (3.3) denkleminde w ağırlık vektörünü, x girdi vektörünü, b önyargı ve φ doğrusal olmayan aktivasyon fonksiyonudur [77]. Şekil 3.8’de Çok katmanlı algılayıcı (MLP) yönteminin şematik görünümü gösterilmiştir.

$$y = \varphi \left(\sum_{i=1}^n w_i x_i + b \right) = \varphi (w^T x + b) \quad (3.3)$$



Şekil 3.8. Çok Katmanlı Algılayıcı (MLP) Yönteminin Şematik Gösterimi [78]

Aşırı öğrenme makinesi (ELM), tek gizli katman ileri beslemeli sinir ağı (SLFN) için hızlı bir eğitim yöntemi olan tek gizli katman ileri beslemeli sinir ağının üç nöron katmanı var olmasına rağmen tek adı modeldeki doğrusal olmayan gizli katmanından gelmekte olup girdi katmanı veri özellikleri sağlar ve hiçbir hesaplama yapmazken, çıktı katmanı bir dönüştürme işlevi veya önyargı olmadan doğrusaldır [79]. Aşırı öğrenme makinesinin yapısı genel olarak Şekil 3.9’da gösterilmiştir:



Şekil 3.9.Aşırı Öğrenme Makinesi (ELM) Yönteminin Şematik Gösterimi [80]

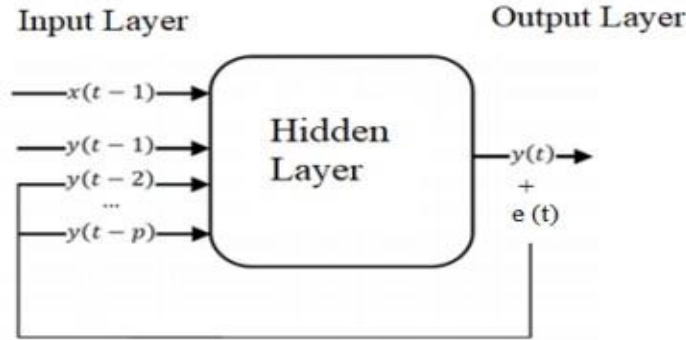
Girdi katmanının ağırlıkları rastgele belirlenirken çıktı ağırlıkları önceden tanımlanmış bir eğitim mekanizması aracılığıyla analitik olarak tahmin edilmesi nedeniyle ELM modelindeki eğitim prosedürü son derece hızlıdır ve genelleme kabiliyeti çok yüksektir [81]. Aşırı öğrenme makinesi (ELM) yönteminde L 'ye sahip tek katmanlı gizli katmanlı sinir ağı modeli gizli nöronlar ve bir aktivasyon fonksiyonu olan $g(x)$ (3.4) denkleminde gösterilmiştir:

$$\sum_{j=1}^L \beta_j \cdot g((w_j x_i) + b_j) = y_i, \quad i = 1, 2, \dots, N \quad (3.4)$$

(3.4) denkleminde w_j ve β_j sırasıyla girdilerden gizli katmana ve gizli katmandan çıktı katmanına ağırlık vektörlerini temsil eder, b_j sembolü j .gizli nöronun önyargısıdır ve $g((w_j x_i) + b_j)$ ise x_i girdi örneğine göre j .gizli nöronun çıktısıdır [82].

İleri beslemeli ağ, zaman serisi modelleme ve tahmin için en yaygın kullanılan sinir ağı modeli olan sinir ağı otoregresyonu (NNAR) modeli doğrusal bir kombinasyon fonksiyonu ve aktivasyon fonksiyonunu içeren üç katmanlı ileri beslemeli bir sinir ağıdır[83]. $NNAR(p, P, k)_m$ modelindeki sembolleri incelediğinde p giriş gecikmelerinin sayısıdır, P mevsimsel

gecikmelerdir, k gizli katmandaki düğüm sayısıdır ve m mevsimsel periyodun uzunluğudur [84]. NNAR modelinin şematik gösterimi Şekil 3.10’da gösterilmiştir:



Şekil 3.10. NNAR Modelinin Şematik Gösterimi [85]

3.2.2. ZAMAN SERİSİ ANALİZİ

Bir zaman serisi, zaman içinde sırayla üretilen bir gözlemler dizisi olup küme sürekli ise zaman serisinin sürekli olduğu söylenir, aksi takdirde ayrıktır [86]. Zaman serisi analizi, fizikte merkezi bir konudur ve biyoloji, tıp ve ekonomideki verileri karakterize etmek ve bunların altında yatan dinamik kökeni anlamak için güçlü bir yöntemdir [87]. Zaman serisi analizi yapmamızın ana nedeni tahmin olduğu için temel fikir, geleceği tahmin etmek için geçmiş gözlemleri denemek ve kullanmak olup verileri en iyi tanımlayan model daha sonra geçmiş kayıtlara dayalı olarak geleceği tahmin etmek için kullanılır [88]. Zaman serisinde gözlemler saatlik, günlük, yıllık olarak belirli zaman aralığında yapılır [89]. Belirli bir zaman serisi verildiğinde, zaman serisi analizinin birincil hedefleri şunlardır: Veriyi üreten mekanizmaya ilişkin iç görüler elde etmek için seriyi temsil edecek varsayım şeklinde bir istatistiksel model kurmak ve tatmin edici bir model olduğunda zaman serisinin gelecekteki değerlerini tahmin etmek için modelden tahminde bulunmak üzere formüle edilmiştir [90]. Zaman serisi analizi, bir diziyi bileşenlerine, eğilimlerin tanımlanmasına, tahminlerin yapılmasına olanak tanıyan açıklanabilir bölümlere ayıran yöntemleri veya süreçleri içerir [91]. Zaman serisi analizi yöntemi; birçok sürecin geçmişlerine ait bir hafızaya sahip olduğu gerçeğini yakalarlar, yani süreçlerin mevcut durumu bir dereceye kadar geçmiş durumlarına

bağılıdır [92]. Zaman serisi verileri söz konusu olduğunda, verilerin ayrı değerli mi yoksa gerçek değerli mi tek biçimli mi yoksa tek biçimli mi örneklenmiş tek değişkenli mi yoksa çok değişkenli mi olduğu ve veri serilerinin eşit mi yoksa eşit olmayan uzunlukta mı olduğu konusunda ayrımlar yapılır [93]. Zaman serisi analizinde uzun vadeli hareketleri temsil eden trend ve belirli bir zaman aralığında düzenli olarak tekrar eden mevsimsel model olup bu nedenlerinin bir modelinin bu özellikleri yakalaması gerekir [94]. Zaman serisi verilerinin doğal bir zamansal sıralaması var olup her veri noktasının öğrenilecek kavramın bağımsız bir örneği olduğu ve bir veri seti içindeki veri noktalarının sıralanmasının önemli olmadığı tipik veri madenciliği veya makine öğrenimi uygulamalarından farklıdır ve çeşitli algoritmalar kullanmayı önerir [95].

En temel zaman serisi tahmin yöntemi tahmini son zaman serisi gözlemine ayarlayan saf tahmin yöntemi olup başka bir basit yöntem ise tahminin, eski ve potansiyel olarak uygulanamaz gözlemleri atarak, zaman serisinin en son değerlerinin aritmetik ortalaması olduğu hareketli ortalama yöntemidir [96].

Kök ortalama kare hatası (RMSE) bir kesinlik ölçümüdür ve bir sistemden başka bir sisteme dönüşümün doğruluğunu belirlemekte kullanılır[97]. Ortalama mutlak hata (MAE) ise düzgün dağılmakta olan hataları tanımlamak için uygundur ve model hatalarının düzgün bir şekilde dağılımdan ziyade normal dağılıma sahip olması muhtemeldir [98]. Performans ölçütlerinin denklemleri MAE (3.5) denkleminde ve RMSE (3.6) denkleminde gösterilmiştir:

$$MAE = \frac{\sum_{n=1}^N |\hat{r}_n - r_n|}{N} \quad (3.5)$$

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (\hat{r}_n - r_n)^2}{N}} \quad (3.6)$$

(3.5) ve (3.6) denklemlerinde kullanılan sembollerden $\hat{r}_n$ tahmin derecelendirilmesi, r_n test veri setindeki gerçek değer anlamına, N ise test verileri ile tahmin sonucu arasındaki derecelendirme tahmin çiftlerinin sayısını temsil etmektedir [99].

Belirleme katsayısı (R^2) ağın genel tahmin performansını değerlendirmesi amaçlanırken, çağ sayısı ve yakınsama için geçen süre, farklı eğitim algoritmalarının göreceli

performansını değerlendirilmek için kullanılır [100]. Belirleme katsayısının denklemi (3.7) denkleminde gösterilmiştir:

$$R^2 = \left[\frac{\sum_{r=1}^R [(z_r - \bar{z}_r)(y_r - \bar{y}_r)]}{\sum_{r=1}^R (z_r - \bar{z}_r)^2 \sum_{r=1}^R (y_r - \bar{y}_r)^2} \right]^2 \quad (3.7)$$

(3.7) denkleminde z_r ve y_r sırasıyla gözlemlenen ve tahmin edilen değerler olduğunda; $\bar{z}_r$ ve $\bar{y}_r$ sırasıyla gözlemlenen ve tahmin edilen değerlerin ortalamasıdır ve R ise hatanın hesaplandığı toplam gözlem sayısıdır [100].

Dikkate alınan her zaman serisinin otomatik korelasyon fonksiyonunu karşılaştıran uygun bir benzerlik ölçüsü kullanarak zaman serilerini kümelenmekte olup otokorelasyon, aynı zaman serileri arasındaki korelasyondur ve zaman standardında değişkenlik gösterir [101].

Saf model, geçen yılın gözlemlenen değerleri cari yılın tahminleri olup örnek vermek gerekirse 1996 yılı tahmini, 1995 yılının gözlenen değeri, 1997 yılı tahmini, 1996 yılının gözlenen değeri şeklinde yapılır [102].

Ortalama model, ortak bir olasılık dağılımından bağımsız olarak çizilmiş gibi görünmektedir, bu gelecekteki gözlemlerin aynı dağılımdan alınacağını düşündürmesiyle tüm gelecekteki değerler için kullanılacak doğal tahmin, geçmiş verilerin örnek ortalamasıdır [103].

3.2.2.1. ARIMA MODELLEME YÖNTEMİ

Box ve Jenkins (1976) tarafından zaman serisi verilerini analiz etmek ve tahmin etmek için güçlü ve yaygın olarak kullanılan istatistiksel bir araç olarak kabul edilen otomatik gerilimli entegre hareketli ortalama (ARIMA) modelini geliştirmiştir [104]. ARIMA, otoregresif entegre hareketli ortalama modelinin kısaltması olup farklı otoregresif model ile hareketli ortalama modelin yüksek bir tahmin doğruluğu ile birleşimidir [105]. ARIMA modeli Box – Jenkins yaklaşımı izlenerek tahmin edilebilir [106]. Değişken bir ARIMA serisinin gelecekteki değeri, rastgele hata varsayımının doğrusal bir fonksiyonu ile belirlenen geçmiş değerlerdir [107]. Regresyon modelinin dezavantajları, yapısal değişikliklerin hızla ilerlediğini yeterince yansıtamaması ve değişkenlerin seçiminin tasarımcının özneliği tarafından belirlenmesi sonucunda ARIMA bunun üstesinden gelmenin bir yolu olarak sunulmuştur [108]. ARIMA prosedürü, otoregresif entegre hareketli ortalama (ARIMA) veya otoregresif hareketli

ortalama (ARMA) modelini kullanarak eşit aralıklı tek değişkenli zaman serisi verilerini, transfer fonksiyonu verilerini analiz etmeyi, tahmin etmeyi ve verilere müdahale etmeyi sağlar [109]. Modelin temel avantajları, mevsimsel değişiklikleri tespit edebilmesi ve zaman serileri içinde seri korelasyonu dikkate almasıdır [110]. ARIMA modelinin hareketli ortalama (MA) modeli, otoregresif (AR) modeli ve entegre (I) modeli olmak üzere üç temel türü vardır [111]. ARIMA modelleri, doğrusal zaman serisi verilerini çok iyi bir doğrulukla tahmin etmektedir [112].

Birçok ARIMA modelinden biri olan mevsimsel olmayanın genel modeli ise ARIMA (p, d, q) olup AR(p) modelin otoregresyon sırasını, I(d) modeli durağan hale getirmek için farklılaşma derecesini ve MA(q) ise modelin hareketli ortalama yönünün sırasını göstermektedir [113]. Otoregresif model olan AR(p) modelinin genel denklemi (3.8) denkleminde ve hareketli ortalama model olan MA(q) modelinin genel denklemi (3.9) denkleminde gösterilmiştir:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \varepsilon_t \quad (3.8)$$

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3.9)$$

Y_t sembolü t zamanında yanıt (bağımlı) değişkenini, $Y_{t-1}; Y_{t-2}; \dots; Y_{t-p}$ sembolleri sırasıyla $t-1; t-2; t-q$ sembolleri zamandaki gecikmelerindeki tepki değişkenlerini, $\phi_0; \phi_1; \phi_p$ ve $\theta_1; \theta_2; \theta_p$ tahmin edilecek katsayılarını, ε_t sembolü t zamanında hata terimini, $\varepsilon_{t-1}; \varepsilon_{t-2}; \varepsilon_{t-q}$ ise önceki dönemlerdeki hataları temsil etmektedir [114]. ARIMA (p, d, q) modelinin genel denklemi (3.10) denkleminde ve modelin kabaca denklemi (3.11) denkleminde gösterilmiştir:

$$U_t = \phi_1 U_{t-1} + \phi_2 U_{t-2} + \dots + \phi_p U_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (3.10)$$

$$U_t = X_t - X_{t-d} \quad (3.11)$$

Bu denklemlerde d sembolü farklılıkların sayısını, t sembolü ayırık zamanı, ϕ_p ise otoregresif parametreyi, ε_t sembolü artıktaki kalanları, θ_q sembolü hareketli ortalama

parametresini, X_t sembolü güvenilir değişkeni, U_t sembolü ise güvenilir değişkenin d. farkını göstermektedir [115].

ARIMA modelleme üç aşamadan oluşmaktadır. İlk adımda model araştırması, normal olarak örnek otokorelasyon fonksiyonunun (ACF) ve örnek kısmi otokorelasyon fonksiyonunun (PACF) çizimleriyle elde edilen uygun modeli ve model sıralarını belirlemektir, ikinci adımda tahmin, modelin katsayılarını tahmin etmektir ve son adımda ise seçilen modelin tanısıl bir kontrolüdür [116].

Kısmi otokorelasyonlar, diğer zaman gecikmelerinin etkileri 1,2,3,4, ,k-1 kaldırıldığında X_t ve X_{t-k} arasındaki ilişkinin derecesini ölçer[117].Gecikme k cinsinden ACF denklemi (3.12) denkleminde gösterilmiştir:

$$Pk = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X}) (X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad (3.12)$$

(3.12) denkleminde Pk sembolü k gecikmesinde otokorelasyon fonksiyonu değeri zaman serisini, X_{t-k} sembolü t ve t-k zamanlarında değişkenlerin veya zaman serisi verilerinin sırasıyla değerlerini, $\bar{X}$ ise değişkenlerin veya zaman serisi verilerinin ortalama değerini göstermektedir [117].

ARIMA modelini doğrulamak ve doğrulamak için çalışma, (3.13) denkleminde gösterilen ARIMA modelindeki uyum eksikliğini test etmek için bir tanısıl araç olarak Box Ljung testi kullanılır [118]. Kullanılan test (3.13) denkleminde gösterilmiştir:

$$Q = n(n + 2) \sum_{k=1}^m \frac{\hat{r}_k^2}{n - k} \quad (3.13)$$

(3.13) denkleminde Q 'nun asimptotik dağılımının, $h = m-p-q$ serbestlik derecesine sahip bir ki-kare dağılımına sahip olduğu durumlarda, p ve q işaretleri AR ve MA 'nın sıralarıdır, n örneklem büyüklüğüdür, $\hat{r}_k$ gecikmesindeki zaman serilerinin tahmini otokorelasyonu ve m ise test edilecek gecikme sayısıdır [118].

3.2.2.2. ÜSTEL DÜZELTME YÖNTEMİ

Üstel düzeltme yöntemi, tahmini veriler hakkında daha yeni bilgiler ışığında bir tahmini sürekli olarak revize etmek için geleneksel prosedürlerden oluşur [119]. Sadece gelecekteki talebin değişmeyeceğini tahmin etmek ve ardından hatayı düzeltmek yerine, tahmin edilen ve gerçek harici talep arasındaki farklardan kaynaklanan çarpmaları doğrusal olarak düzeltmektedir [120].

Üstel düzeltme yöntemi tahmini denklemi (3.14) denkleminde gösterilmiştir:

$$\widehat{X_{T|T+1}} = \alpha X_T + \alpha(1 - \alpha)X_{T-1} + \alpha(1 - \alpha)^2 X_{T-2} + \dots \quad (3.14)$$

Üstel düzeltme yöntemi tahmini denkleminde $0 \leq \alpha \leq 1$ düzeltme parametresini, (3.14) denkleminde görülebileceği gibi ağırlıklar üssel olarak azalır ve α hangi hızda azalacağını belirlenmesini, (3.14) denkleminde α 1'e yakın olduğunda, nispeten daha fazla ağırlık daha yeni gözlemlere atanır ve α 0'a yakın olduğunda, geçmişte daha fazla gözlemlere nispeten daha fazla ağırlık verilmektedir [121].

3.2.2.2.1. BASİT ÜSTEL DÜZELTME YÖNTEMİ

Basit üstel düzeltme (SES - Simple Exponential Smoothing) yöntemi genellikle, zaman serilerinin seviyesinin sabit bir seviyede dalgalanması veya zaman içinde yavaşça değişmesi gerektiği önermesine dayanmaktadır [122]. Eğilim ve mevsimsellik içermeyen en basit durum, basit üstel yumuşatma yöntemi olarak bilinmektedir [123]. Basit üstel düzeltme yönteminde $z_1, z_2, z_3, \dots, z_n$ bir zaman serisi olduğu kabul edilirse oluşan denklem (3.15) denkleminde gösterilmiştir ve (3.15) denkleminin (3.16) denklemindeki gibi gösterilmiştir. (3.15) ve (3.16) denklemlerindeki semboller incelendiğinde Z_t söz konusu dizi olduğunu, λ düzeltme parametresini, ξ_t hata değerini, F_t ile F_{t+1} sırasıyla t ve $t+1$ 'de öngörülen değerini, λ değeri ise 0 ve 1 arasındadır [123].

$$F_{t+1} = \lambda Z_t + (1 - \lambda)F_t \quad (3.15)$$

$$F_{t+1} = \lambda \xi_t + F_t \dots\dots\dots \quad (3.16)$$

3.2.2.2.2. HOLT WINTERS YÖNTEMİ

Birçok şirket, satış verileri bir eğilim ve mevsimsel bir model içerdiğinde kısa vadeli talep tahminleri üretmek için Holt Winters (HW) yöntemini kullanır[124]. Holt Winters yöntemi; seviye, trend ve mevsimsel faktörle ilişkili üç düzeltme parametrelerini tahmin etmektedir [125]. Holt (2004); eğilimi olmayan seriler, eklemeli (doğrusal), çarpan (doğrusal olmayan) eğilim ve mevsimsel olmayan seriler, eklemeli, çok yönlü modellerde mevsimsel değişim içeren seriler için üstel ağırlıklı hareketli ortalamaları genişletmiştir [126].

Holt Winters yönteminde çarpımsal modeli (3.17), (3.18) ve (3.19) denklemleri gösterilmiştir:

$$L_{t-1} = \alpha \frac{y_{t-1}}{c_{t-1-r}} + (1 - \alpha)(L_{t-2} + B_{t-2}) \quad (3.17)$$

$$S_{t-1} = \gamma \frac{y_{t-1}}{F_{t-1}} + (1 - \gamma)S_{t-1-r} \quad (3.18)$$

$$B_{t-1} = \beta (F_{t-1} - F_{t-2}) + (1 - \beta) B_{t-2} \quad (3.19)$$

(3.17), (3.18) ve (3.19) denklemlerin sembollerini incelediğimizde L_{t-1} sembolü basit bir üstel yumuşatma modelinden elde edilen düzleştirilmiş değer eşdeğerini(ağırlıklı aritmetik ortalama), B_{t-1} sembolü an veya dönem t-1 büyüme eğiliminin değerlendirilmesini, S_{t-1} sembolü mevsimsel göstergenin bir an veya dönem t-1 için değerlendirilmesini, r sembolü mevsimsel döngünün uzunluğu - aşamaların sayısını ve α, β, γ sembolleri ise $[0, 1]$ aralığındaki değerlerle modelin yumuşatma parametrelerini yukarıdaki üç denklemin sembollerinin açıklamasıdır [127].

Holt Winters yönteminde eklemeli modeli (3.20), (3.21) ve (3.22) denklemleri ve eklemeli modelin Holt Winters üstel yumuşatma tahmin değeri için (3.23) denkleminde gösterilmiştir:

$$S_t = \alpha (X_t - I_{t-L}) + (1 - \alpha)(S_{t-1} + b_{t-1}) \quad (3.20)$$

$$b_t = \beta (S_t - S_{t-1}) + (1 - \beta) b_{t-1} \quad (3.21)$$

$$I_t = \gamma(X_t - S_t) + (1 - \gamma) I_{t-L} \quad (3.22)$$

$$F_{t+m} = S_t + mb_t + I_{t-L+m} \quad (3.23)$$

S_t sembolü t yılında üstel düzeltmeyi, S_{t-1} sembolü mevsimler unsurlarına göre t-1 yılında üstel düzeltmeyi, b_t sembolü t yılında trend unsurlarını düzeltmeyi, b_{t-1} sembolü t-1 yılında trend unsurlarını düzeltmeyi, X_t sembolü t yılındaki veriyi, α sembolü veriler için üstel düzeltme parametreyi, I_t sembolü mevsimsel faktör düzeltmeyi, L sembolü mevsimsel uzunluğunu, β sembolü trend için üstel düzeltme parametresini ve γ sembolü mevsimsellik için üstel yumuşatma parametresini, F_{t+m} sembolü t+m süredeki tahmin sonuçlarını, F_t sembolü tahmin etmek istediğiniz değerini ve m ise tahmin edilecek dönemi gösterir [128].

3.2.2.3. DOĞRUSAL REGRESYON ANALİZİ

Regresyon analizi, istatistikte değişkenler arasındaki ilişkiyi araştırmak ve modellemek için kullanılan bir tekniktir [129]. Doğrusal model, modern istatistiksel analizde en çok kullanılan araç olup ayrıca daha karmaşık modeller oluşturmak için önemli bir kavramsal çerçeve sağlamaktadır [130]. Doğrusal regresyonun yararlı işlevi, bağımlı değişkenin değerlerini tahmin etmektir [131]. Doğrusal regresyon, birçok modern modelleme aracının yapı taşı olup özellikle, örnek boyutu küçük olduğunda veya sinyal nispeten zayıf olduğunda, lineer regresyon genellikle temeldeki regresyon fonksiyonuna tatmin edici bir yaklaşım sağlar[132]. Regresyona girmeden önce, mevcut projede kullanıldığı için korelasyondan bahsetmek gerekirse iki değişken arasındaki ilişkinin bir ölçüsüdür [133]. Değişken, diğer değişkende aynı yönde bir değişiklik ile ilişkilidir yani bir değişkendeki artış, diğer değişkendeki

artışla ilişkilidir ve bir değişkendeki azalma, diğer değişkendeki azalmayla ilişkilidir [134]. En basit tahmin yöntemlerinden biri, veri noktaları aracılığıyla en iyi çizgiyi bulmaya çalıştığımız doğrusal regresyondur [135].

Çoklu R kare (Multiple R squared), birden fazla tahmin değişkenine sahip modeller için basitçe bir R kare (R-squared) ölçüsü olduğundan bağımlı değişkenindeki bağımsız değişkenlerle açıklanabilecek varyasyon miktarını ölçer[136]. Düzeltilmiş R kare (Adjusted R-squared), bir regresyon modelinde anlamlı olmayan yordamları hesaba katan değiştirilmiş bir R-kare sürümü olup özetle ek öngörücüler eklemenin bir regresyon modelini iyileştirip iyileştirmediğini gösterir [137].

Breusch ve Godfrey test, otokorelasyonu kontrol etmek için en yaygın kullanılan LM testini geliştirdiler ve bu olasılığa dayalı LM tipi test, dinamik modellerde otokorelasyon varlığını kontrol etmek için en uygun testlerden biridir [138]. Alternatif hipotez $H_a: \rho = 0$ durumunda ise iki taraflı Breusch ve Godfrey testi tamamen uygun değildir [139].

Box - Cox dönüşümü, bir zaman serisinin değişkenliğini dengelemek için popüler bir dönüşümdür [140]. Box - Cox dönüşümü; bütün bir dönüşüm sınıfını kullanmak, araştırmacıya yanıt süresi verilerini analiz etmede daha fazla özgürlük sağlamak ve normal olarak dağıtılmış verileri elde etmek için uygun bir dönüşüm seçilmesine izin vermektedir[141].Box -Cox dönüşümünün denklemi (3.24) denkleminde gösterilmiştir:

$$\omega_t = \begin{cases} \log(y_t), & \lambda = 0 ; \\ (y_t^\lambda - 1)/\lambda, & \lambda \neq 0. \end{cases} \quad (3.24)$$

(3.24) denklemindeki λ sembolü neredeyse sonsuz sayıda değer alabildiğini düşünürsek, teorik olarak bir değişkeni normalliğe doğru hareket ettirmede en üst düzeyde etkili olacak şekilde negatif veya pozitif olarak çarpık olmasına bakılmaksızın ayarlayabilmekte olup ek olarak bu dönüşüm ailesi birçok geleneksel dönüşümü içerir [142].

Kübik düzeltme spline, bir tahminci arasındaki işlevsel ilişkileri tahmin etmek için güçlü bir yaklaşımdır [140]. Kübik düzeltme spline yönteminin temel avantajı, yumuşatma parametresinin filtreleme yöntemlerinde benimsenen öznel deneme ve hata ayarı yerine çapraz doğrulama yöntemi ile belirlenebilmesidir [143].

Tek deęişkenli regresyon analizi ile bir baęımlı ve bir baęımsız deęişken arasındaki ilişkiler analiz edilir ve baęımlı ve baęımsız deęişkenler arasındaki doğrusal ilişkileri temsil eden denklem formüle edilmekte olup bir baęımlı deęişkeni birden fazla baęımsız deęişkeni olan regresyon modelleri, çok deęişkenli regresyon analizi olarak bilinir [144].

3.2.2.3.1. BASİT DOĞRUSAL REGRESYON ANALİZİ

Basit doğrusal regresyon, iki deęişken arasındaki ilişkiyi modellemektedir [145]. Basit regresyon analizi denklemi (3.25) denkleminde gösterilmiştir:

$$Y = \alpha + \beta X + \varepsilon \quad (3.25)$$

(3.25) denkleminde Y sembolü baęımlı deęişkeni, X sembolü tahmin edici deęişkeni, ε sembolü artık veya tahmin hatasını son olarak α ve β sembolleri ise regresyon parametreleri veya katsayılarını göstermektedir [146]. (3.26), (3.27), (3.28) denklemlerinde α , β ve ε sembollerinin denklemleri gösterilmiştir:

$$\varepsilon_i = y_i - \hat{y}_i \quad (3.26)$$

$$\alpha = \sum_{i=1}^N y_i - \beta \sum_{i=1}^N x_i \quad (3.27)$$

$$\beta = \frac{\sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{\sum_{i=1}^N x_i^2 - \sum_{i=1}^N x_i \sum_{i=1}^N x_i} \quad (3.28)$$

Yukarıdaki denklemlerdeki sembolleri incelersek x_i sembolü son verilerdeki ortalama deęerini, y_i sembolü verinin gerçek deęerini, $\hat{y}_i$ sembolü y_i 'deki deęerlerin her birinin tahmin deęerini ve N sembolü ise örnek sayısını temsil eder [147].

3.2.2.3.2. ÇOKLU DOĞRUSAL REGRESYON ANALİZİ

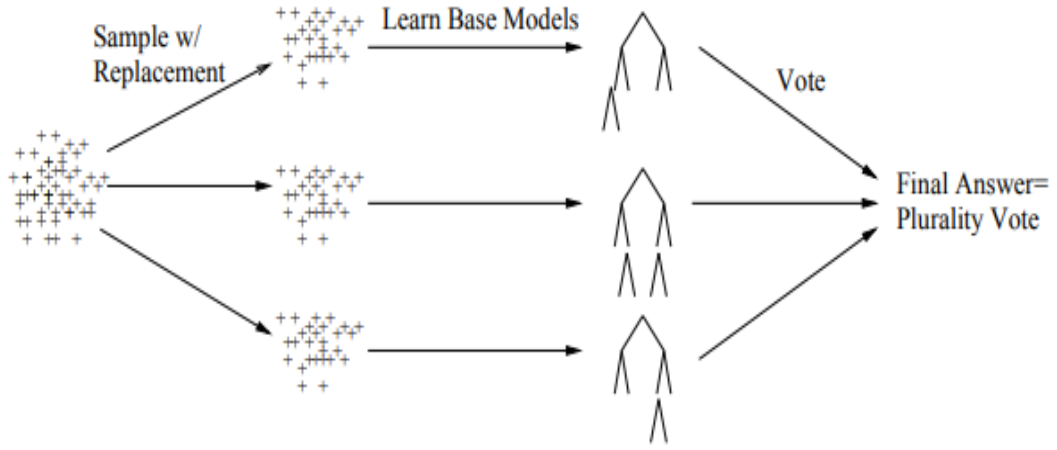
Çoklu doğrusal regresyon (MLR), bir veya daha fazla değişkenin $X_1, X_2, \dots, X_k$ (bağımsız veya açıklayıcı faktörler veya tahmin ediciler olarak adlandırılabilir) bir nicel faktör Y üzerindeki etkilerini değerlendirmek için popüler bir veri analiz yöntemidir (bağımlı veya tepki değişkeni olarak adlandırılabilir) [148]. Ayrıca çoklu doğrusal regresyon modeli, basit doğrusal regresyon ile aynı temel üzerine inşa edilmiştir [149]. Çoklu doğrusal regresyon modeli k değişkenleri için genel denklemi (3.29) denkleminde gösterilmiştir:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i \quad i = 0, 1, 2, 3, \dots, n \quad (3.29)$$

(3.29) denkleminde y_i sembolü bağımlı değişkeni, x_i sembolü bağımsız değişkeni, β_i sembolü regresyon katsayısını gösterir [150] ve e_i sembolü ise istatistiksel hatayı belirtir [151]. Çoklu doğrusal regresyon, birden fazla açıklayıcı değişkeni içerecek şekilde basit doğrusal regresyonu genişletmekte olup her iki durumda da doğrusal terimi kullanır çünkü yanıt değişkeninin doğrudan açıklayıcı değişkenlerin doğrusal bir kombinasyonu ile ilişkili olduğu varsayılmaktadır [152].

3.2.2.4. TORBALAMA YÖNTEMİ

Torbalama yöntemi, Breiman (1996) tarafından tanıtılmış olup öğrenme verilerinin önyükleme örnekleri temelinde sınıflandırma ağaçları oluşturur [153]. Buluşu sırasında, torbalamanın neden işe yarayacağına dair yalnızca buluş argümanları sunulmuş olup daha sonra torbalama yönteminin, regresyon veya sınıflandırma ağaçlarının tahmin performansını iyileştirmeyi amaçlarken avantajlı olduğu ortaya çıkan bir yumuşatma işlemi olduğu ortaya çıkmıştır [154]. Şekil 3.11’de torbalama yönteminin çalışma şeması gösterilmiştir:



Şekil 3.11. Torbalama Yönteminin Çalışma Şeması [155]

Şeklin sol tarafındaki noktalar torbalama algoritmasının çağrıldığı orijinal eğitim setini temsil etmektedir, eğitim kümesinden uzağa ve üç nokta kümesine işaret eden üç ok değiştirmeli örneklemeyi temsil etmektedir, temel model öğrenme algoritması bir temel model oluşturmak için bu örneklerin her birinde çağrılmaktadır ve son üç ok ise sınıflandırılacak yeni bir örnek geldiğinde ne olduğunu gösterir [156].

3.2.2.5. SIRALI MİNİMUM OPTİMİZASYON REGRESYON ALGORİTMASI

John Platt tarafından 1998'de sıralı minimal optimizasyon regresyonunu (SMOreg) olarak icat edilmiş olup destek vektör makinelerinin eğitimi sırasında ortaya çıkan ikinci dereceden programlama (QP) problemini çözmek için yaygın olarak kullanılan bir algoritmadır [157]. Temel mantığı, doğrusal olarak ayrılabilen veriler için en iyi ayırıcı hiper düzlemin belirlenmesi olan destek vektör makineleri, doğrusal olarak ayrılamayan verileri dönüştürme tekniği uygulayarak çözmektedir [158]. Sıralı minimum optimizasyon regresyon algoritması küresel olarak tüm eksik değerleri değiştirir ve nominal öznitelikleri ikili olana dönüştürür [159]. SMOreg algoritması, regresyon tahminini gerçekleştirme konusunda iyi bir yeteneğe sahip olan model olarak kısıtlamalar yapısal risk minimizasyonunu kullanmaktadır [160]. SMOreg, model olarak kısıtlamaları yapısal risk minimizasyonu kullanır; iyi bir regresyon tahmini ve doğrusal olmayan, küçük örnek verileri işleme yeteneğine sahip olmasıyla SMOreg, zaman serisi verilerini doğru bir şekilde tahmin edebilir [161].

3.2.2.6. YEREL AĞIRLIKLIL ÖĞRENME YÖNTEMİ

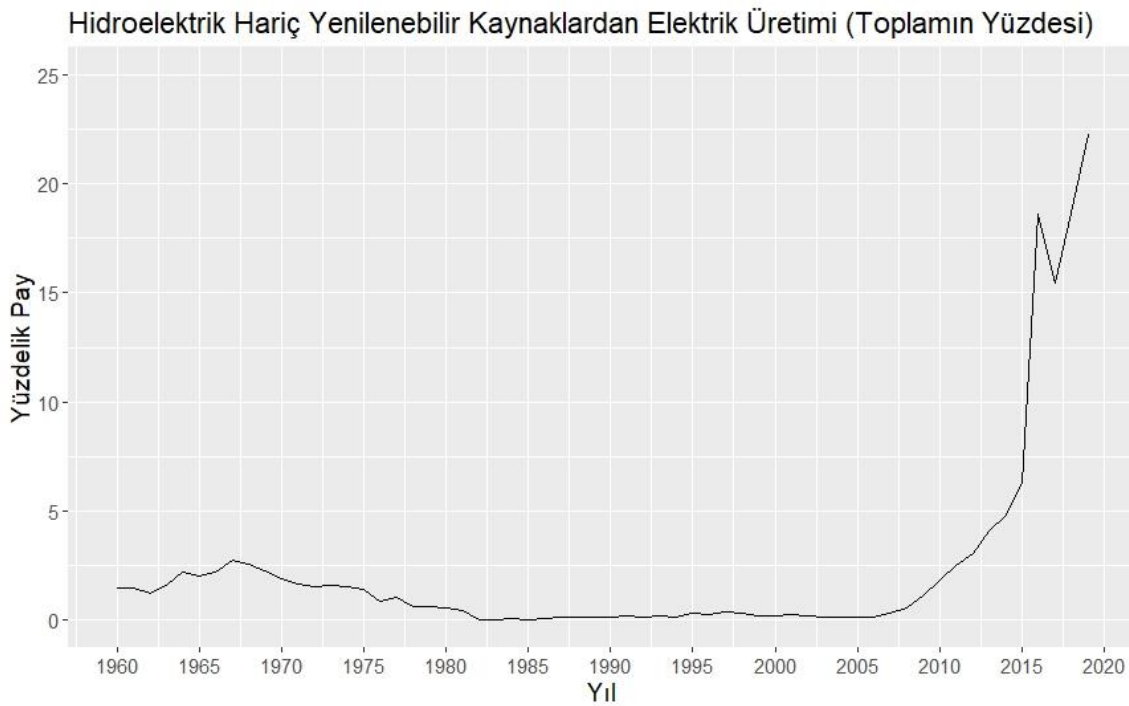
Yerel ağırlık öğrenme (LWL) yönteminin temel konsepti birinci dereceden Taylor serisi açılımına benzer şekilde, parçalı doğrusal modeller aracılığıyla doğrusal olmayan fonksiyonları yaklaşık olarak bulmaktır [162]. LWL yöntemi son derece uyarlanabilir ve uzun vadede kesin bir model sağlar [163]. LWL yönteminin genel denklemi (3.30) denkleminde gösterilmiştir:

$$F = \frac{1}{2N} \sum_{K=1}^N w_k \left(\alpha_{K0} + \sum_{n=1}^M \alpha_{kn} x_{kn} + \varepsilon_k - y_k \right)^2 \quad (3.30)$$

(3.30) denkleminde F sembolü LWL fonksiyonu, w sembolü ağırlık fonksiyon matrisini, M sembolü toplam değişken sayısını, ε_k sembolü rastgele hata değerini, $\alpha_{K0} \dots \alpha_{kn}$ sembolleri regresyon katsayılarını ve y_k ise rastgele değişkenin değerini göstermektedir [164].

4. ARAŞTIRMA BULGULARI

Çalışmada kullanılan ana materyal olarak 1960-2019 yılları arasında Türkiye'nin hidroelektrik hariç yenilenebilir enerji kaynaklarından elektrik üretiminin toplam oranını gösteren veri seti kullanılmıştır. 1960-2015 yılları arasındaki veriler dünya bankasının sitesinden ve 2015-2019 yılları arasındaki veriler enerji bakanlığından alınmıştır. Zaman serisi tahminleme yöntemi ve yapay sinir ağları tahminleme yöntemlerinin her birinin kendi içinde yöntemler kullanılarak 2020'den 2025 yılı dâhil olmak üzere tahminleme işlemleri yapılmıştır. Yapılan tahminler doğrultusunda elde edilen araştırma bulguları gösterilmiştir ve hangi yöntemin daha iyi olduğunu tespit etmek için performans ölçütleri olan ortalama mutlak hata (MAE) ve kök ortalama kare hatası (RMSE) değerleri karşılaştırılmıştır. Kullanılan teknikler ise Weka ve R-Studio uygulamaları üzerinde işlenmiştir. Kullanılan veri seti Şekil 4.1'de grafik üzerinde gösterilmiştir:



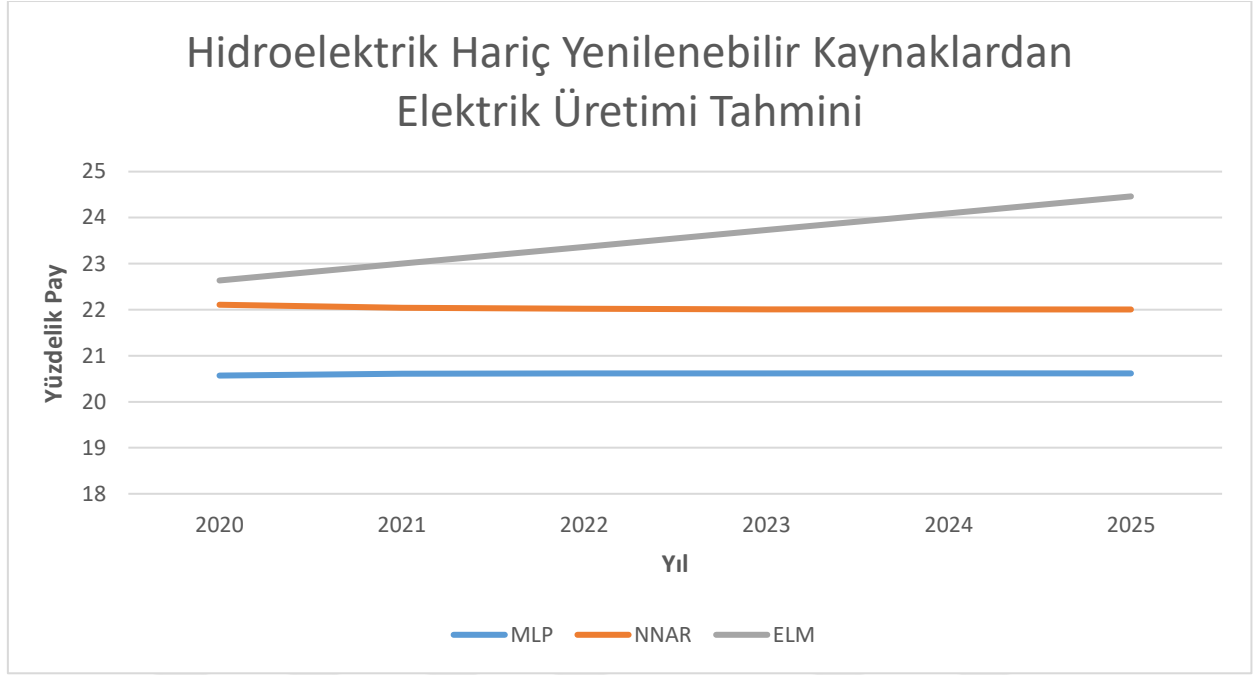
Şekil 4.1. Ana Materyaldeki Veri Setinin Gösterimi

Şekil 4.1'deki grafik incelendiğinde yüzdelik pay çizgileri belli bir artış ve azalışlar meydana gelmiştir fakat bu meydana gelen değişimler düzensiz bir şekilde olduğu görülmektedir. Grafikte gösterilen veri setinin trend bileşeni olduğu ortaya çıkmıştır. Mevsimsellik bileşenini içermemesinin nedeni ise grafikteki çizgilerin düzensiz bir şekilde dağılımı olduğu tespit edilmiştir.

Kullanılan ana materyalin tahminleme yöntemlerinden yapay sinir ağı kendi içinde çok katmanlı algılayıcı (MLP), aşırı öğrenme makinesi (ELM) ve sinir ağı otoregresyonu (NNAR) yöntemleri kullanılarak geleceğe yönelik tahmin değerleri oluşturulmuş, her bir yöntem için grafik gösterimi oluşturulmuş ve her bir yöntemin performans ölçütlerine göre hangi durumun daha iyi olup olmadığını tespit edilmesini sağlayan işlemler yapay sinir ağı için araştırma bulguları kısmında açıklanmıştır. Çok katmanlı algılayıcı (MLP) yöntemi için Weka uygulaması, aşırı öğrenme makinesi (ELM) ve sinir ağı otoregresyonu (NNAR) için R-Studio uygulaması üzerinde R programlama dili yazılarak kullanılan veri seti işlenmiştir. Diğer taraftan, zaman serisi analizi kendi içinde doğrusal regresyon modeli, ortalama model, saf model, kübik düzeltme spline, ARIMA modelleme, üstel düzeltme yöntemi, yerel ağırlık öğrenme (LWL) yöntemi, torbalama yöntemi ve sıralı minimum optimizasyon regresyonu (SMOreg) algoritması kullanılarak geleceğe yönelik tahmin değerleri oluşturulmuş, yöntemler için kendi içlerinde ayrı ayrı gruplar halinde grafik gösterimi oluşturulmuş ve kendi içlerinde ayrı ayrı gruplar halinde yöntemlerin performans ölçütlerine göre hangi durumun daha iyi olup olmadığını tespit edilmesini sağlayan işlemler yapay sinir ağı için araştırma bulguları kısmında açıklanmıştır. Yerel ağırlıklı öğrenme (LWL) yöntemi, torbalama yöntemi, sıralı minimum optimizasyon regresyon (SMOreg) algoritması için Weka algoritması kullanılmıştır. Doğrusal regresyon modeli, ortalama model. Saf model, kübik düzeltme spline, ARIMA modelleme ve üstel düzeltme yöntemi için R-Studio uygulaması üzerinde R programlama dili yazılarak kullanılan veri seti işlenmiştir. Kullanılan tahminleme yöntemlerinden yapay sinir ağı kapsamında kullanılan yöntemler ve diğer tahminleme yöntemi olan zaman serisi analizinde kullanılan yöntemler ise aşamalı olarak elde edilen yüzdelik pay değerleri grafikler üzerinde incelenmiştir.

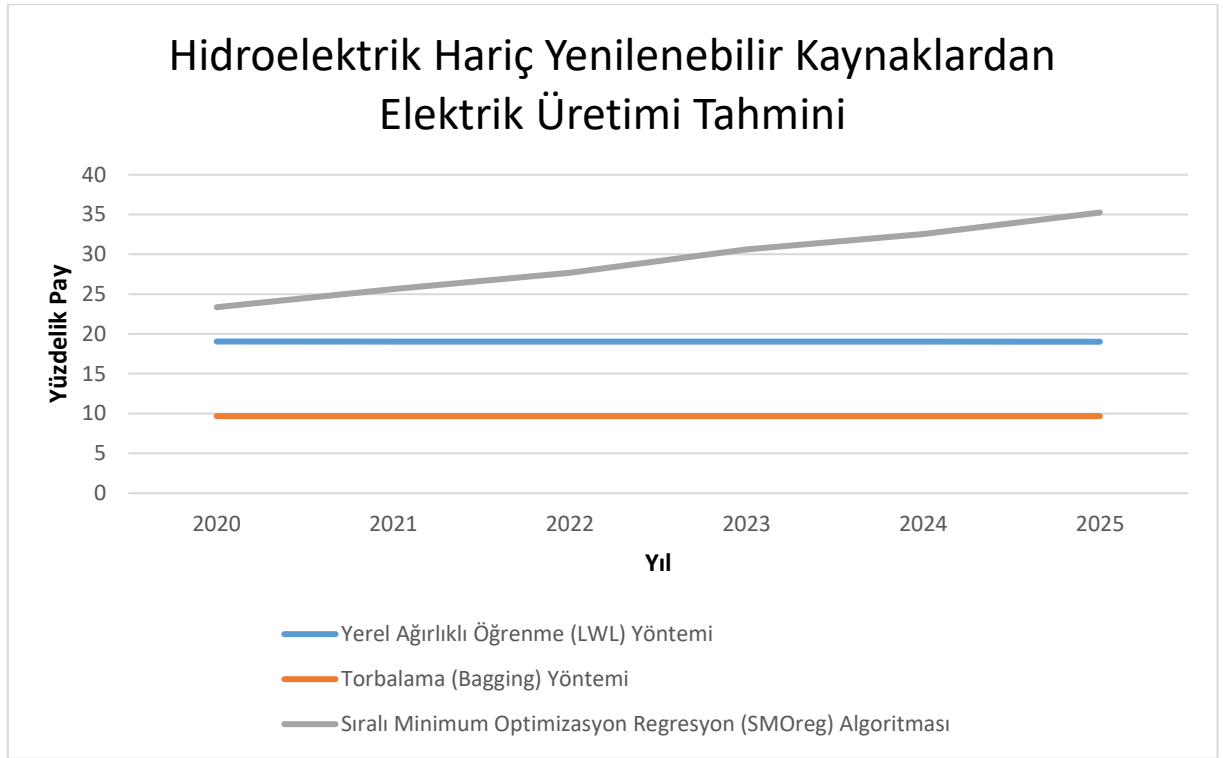
Yapay sinir ağı kendi içinde çok katmanlı algılayıcı (MLP), sinir ağı otoregresyonu (NNAR) ve aşırı öğrenme makinesi (ELM) yöntemleri kullanılarak 2020 yılından 2025 yılı dâhil olmak üzere tahmin edilen yüzdelik payları Şekil 6'daki grafikte karşılaştırılmıştır. 2020 yılından 2025 yılı dâhil edilerek tahmin edilen yüzdelik pay değerleri incelendiğinde çok katmanlı algılayıcı (MLP) yönteminde tahmin edilen verilerin çizgisinin düz olduğu, sinir ağı otoregresyonu (NNAR) yönteminde tahmin edilen yüzdelik pay değerlerinin her bir yıl geçtiğinde devasa bir artış söz konusu olduğu ve aşırı öğrenme makinesi (ELM) yönteminde tahmin edilen yüzdelik pay değerlerin aşırı olmasa da her bir yıl geçtikçe belli bir artış olduğu tespit edilmiştir. Yapay sinir ağı kendi içinde çok katmanlı algılayıcı (MLP), sinir ağı

otoregresyonu (NNAR) ve aşırı öğrenme makinesi (ELM) yöntemleri kullanılarak 2016 yılından 2025 yılı dâhil olmak üzere tahmin edilen yüzdeler payları Şekil 4.2’deki grafikte gösterilmiştir. Üç yöntemde belli bir noktaya kadar artıp grafik çizgisi sabit kalmıştır.



Şekil 4.2. Yapay Sinir Ağları Yöntemlerinin Tahminleri

Yerel ağırlıklı öğrenme (LWL) yöntemi, torbalama (bagging) yöntemi ve sıralı minimum optimizasyon regresyon (SMOreg) algoritması yöntemi kullanılarak 2020 yılından 2025 yılı dahil olmak üzere tahmin edilen yüzdeler pay değerleri Şekil 4.3’deki grafikte karşılaştırılmıştır. Sıralı minimum optimizasyon regresyon (SMOreg) algoritması yöntemi grafik çizgisi yukarıya doğru artma göstermiştir. Torbalama yöntemi ve yerel ağırlıklı öğrenme (LWL) yöntemi grafik çizgilerinin sabit olduğu görülmüştür.



Şekil 4.3. Birinci Aşamada Kullanılan Yöntemlerin Tahminleri

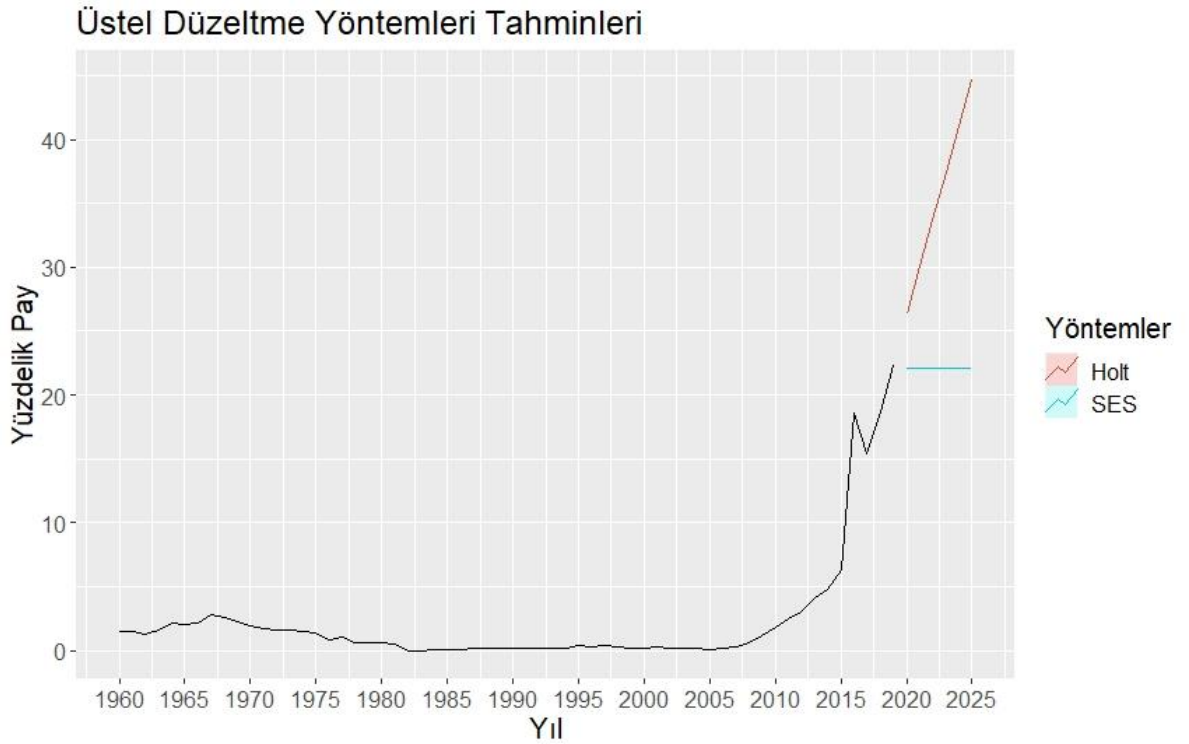
Üstel düzeltme yönteminin birden fazla alt yöntemlerden basit üstel düzeltme (SES) yöntemi ve Holt's Winters (HW) yöntemi başlıkları altında incelenmiştir. Şekil 4.1'deki grafikte kullanılan veri seti incelendiğinde verilerin trend bileşen içerdiği fakat verilerin düzensiz dağılımından dolayı mevsimsel bileşen içermediği belirlenmiştir. Böylece Holt's Winters yöntemi ise Holt's trend yöntemi başlığı altında incelenmiştir. Ek olarak gerçek serilerde ise trend bileşendeki verilerde ilerleyen süreçlerde durgunlaşma ve sönümlenme olabileceği için Holt's trend (bükülmüş) yöntemi olarak da incelenmiştir.

Kullanılan yöntemlerin kullanılan veri setini işlemeden önce performans ölçütleri olan ortalama mutlak hata (MAE) ve kök ortalama kare hatası (RMSE) değerleri karşılaştırılmıştır ve Çizelge 4.5'te gösterilmiştir. Tahminleme işleminden önce yapılan performans ölçütleri karşılaştırılmasında en iyi sonuç veren yöntemin basit üstel düzeltme yöntemi olduğu ortaya çıkmıştır. En kötü sonucu veren ise Holt's trend (bükülmüş) yöntemi olduğundan dolayı tahminleme işlemine dahil edilmemiştir.

Çizelge 4.1. İkinci Aşamada Kullanılan Yöntemlerin Tahminleme Öncesi Performans Ölçütleri Karşılaştırılması

	Basit Üstel Düzeltme	Holt's Trend Yöntemi	Holt's Trend (Bükülmüş) Yöntemi
MAE	1,326384	1,424873	1,408318
RMSE	3,464281	3,724091	3,748261

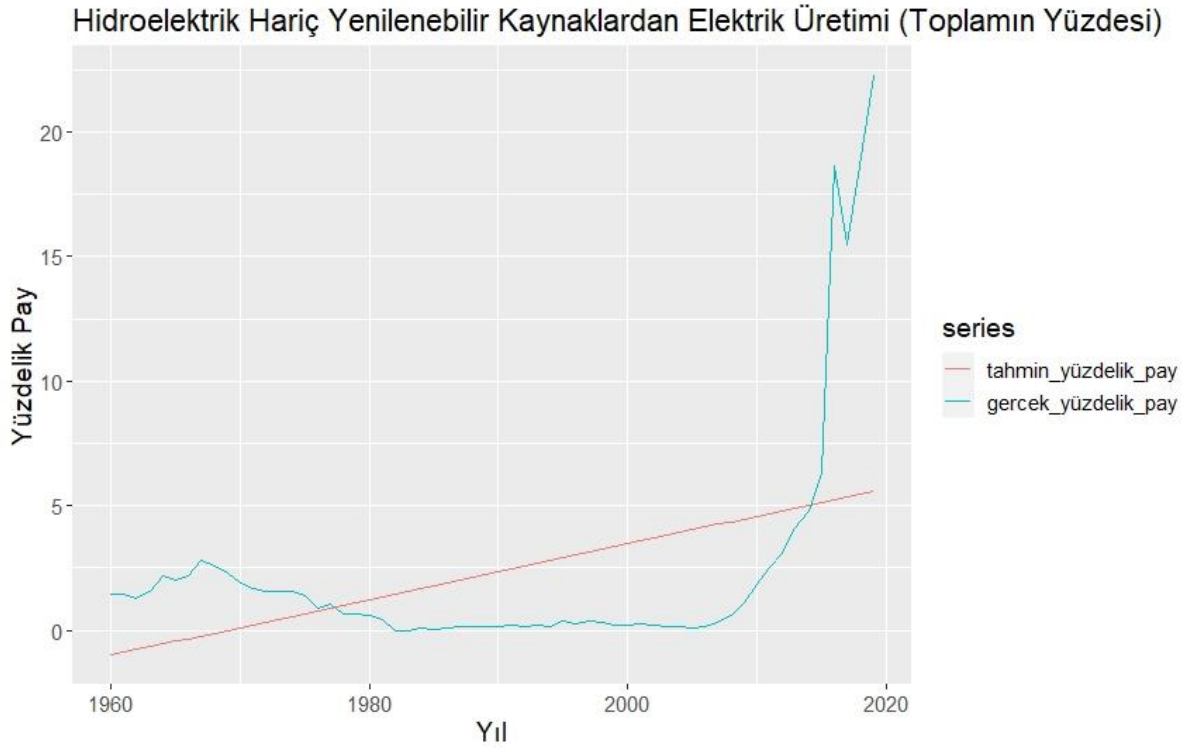
Basit üstel düzeltme yöntemi ve Holt's trend yöntemi sonucunda yapılan tahminleme işleminde elde edilen değerlerin Şekil 4.4'te grafik üzerinde gösterilmiştir. Şekil 4.4'te gösterilen grafikte 2019 yılındaki gerçek değeri 2020 yılında tahmin edilen yüzdelik pay değerinde basit üstel düzeltme yönteminde sabit ve Holt's trend yöntemi kullanılarak elde edilen yüzdelik pay değerleri ise grafik çizgisinin yukarıya doğru artmış bir şekilde arttığı görülmüştür. Holt's trend yöntemi ise grafik çizgisi artan bir eğilim göstermiştir.



Şekil 4.4. İkinci Aşamada Kullanılan Yöntemlerin Tahminleri

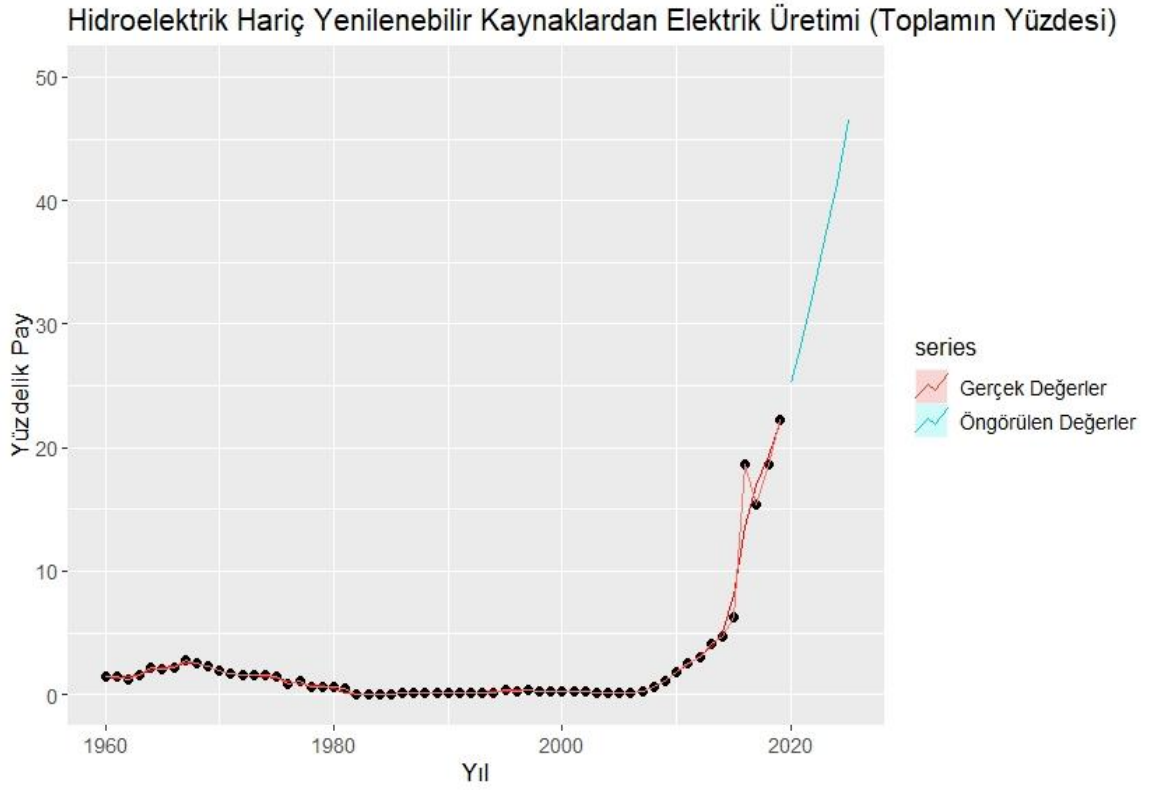
Doğrusal regresyon analizinde doğrusal regresyon model çizgisi değeri 1,1158 değeri elde edilmiş olup doğrusal regresyon model de yıldan yıla 0,1117 adetlik trend yükselme olmuştur.

Doğrusal regresyon modeline aldığımız değişkenler bağımlı değişkendeki değişkenlerin %17,62 civarında olduğu çoklu-R kare değerine bakarak tespit edilmiştir. Doğrusal regresyon analizinde Breusch – Godfrey testi yaptığımızda p değerinin 0,0008418 değeri elde edilmiş olup p değeri 0,05'ten az olduğundan sıfır hipotez olmadığı böylece Breusch – Godfrey testine uygun olduğu ortaya çıkmıştır. Elde edilen p değerinin 0,05'ten küçük olmasıyla kalıntılar arasındaki otokorelasyonun ortaya çıkan doğrusal regresyon modeli ile kapatamadığımız ortaya çıkmıştır. Şekil 4.5'deki grafikte doğrusal regresyon analizi yönteminde kullanılan verilerin gerçek ve tahmin edilen değerler karşılaştırıldığında çok başarısız kalmıştır. Sonuç olarak doğrusal regresyon yöntemi yerine doğrusal olmayan kübik düzeltme spline yöntemi kullanılmıştır.



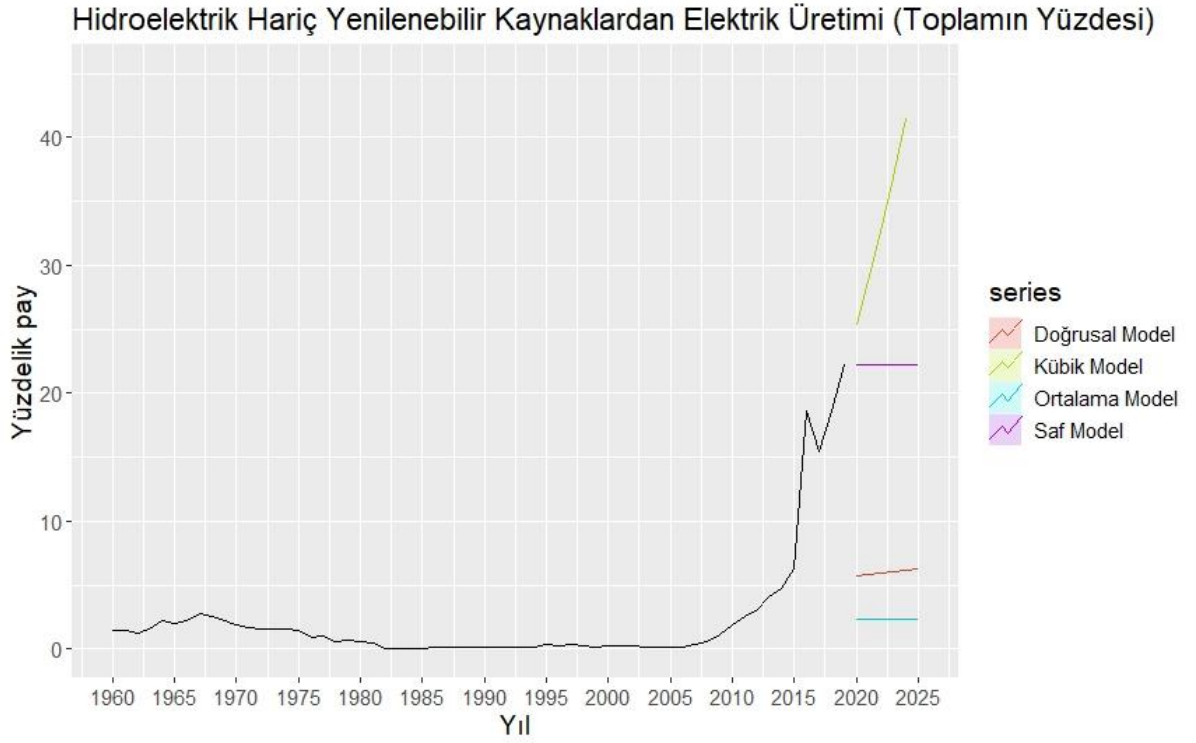
Şekil 4.5. Doğrusal Regresyon Yöntemindeki Verilerin Gerçek ve Tahmin Değerleri

Kübik düzeltme yöntemi işlemek için Box – Cox dönüşümü yapılmıştır. Yapılan dönüşümde 0,2642824 lambda değeri elde edilmiştir. Şekil 4.6'daki grafikte ortaya çıkan siyah çizgi gerçek yüzdelik pay değerlerini, kırmızı çizgi kübik düzeltme spline yöntemi kullanılarak elde edilen tahmin yüzdelik pay değerlerini ve mavi çizgi ise gerçek yüzdelik pay değerleri üzerinde gelecek için tahmin edilen yüzdelik pay değerlerini göstermektedir. Ayrıca grafik üzerinde kullanılan verilerin gerçek ve tahmin edilen değerler karşılaştırıldığında başarılı sonuçlar olduğu ortaya çıkmıştır.



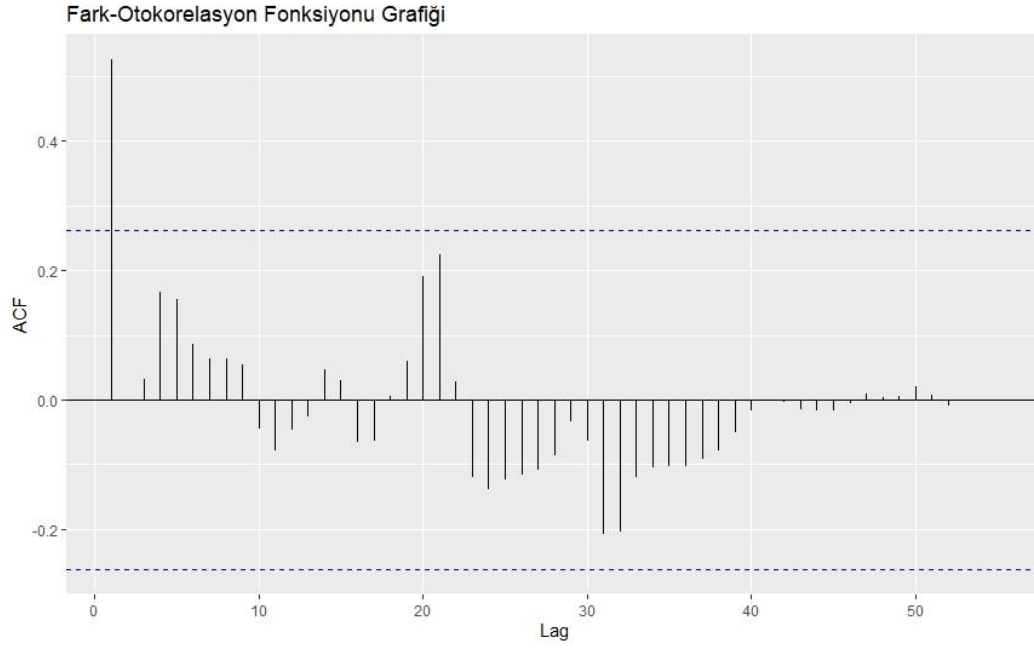
Şekil 4.6. Kübik Düzeltme Spline Yöntemindeki Verilerin Gerçek ve Tahmin Değerleri

Doğrusal regresyon analizi yöntemi, kübik düzeltme spline yöntemi, ortalama model yöntemi ve saf model yöntemi sonucunda yapılan tahminleme işleminde elde edilen yüzdelik pay değerlerin Şekil 4.7’de grafik üzerinde gösterilmiştir. Şekil 4.7’de gösterilen grafikte 2019 yılındaki gerçek değeri 2020 yılında tahmin edilen yüzdelik pay değerinde doğrusal regresyon analizi yöntemi ve ortalama model yöntemi kullanılarak elde edilen değerler azalmış bir şekilde birbirine yakın olmasına rağmen ilerleyen yıllarda tahmin edilen yüzdelik pay değerler grafikte incelendiğinde ortalama model yöntemi grafik çizgisi düz bir şekilde görünmüştür. Doğrusal regresyon analizi yönteminde ise ilerleyen yıllara tahmin edilen yüzdelik pay değerlerin az bir artma yapmasıyla düze yakın bir çizgi olduğu görünmektedir. 2019 yılındaki gerçek yüzdelik pay değeri 2020 yılında tahmin edilen yüzdelik pay değerinde saf model yöntemi kullanılarak elde edilen yüzdelik pay değeri artmıştır ama ileriki yıllarda grafik çizgisi düz bir şekilde görünmüştür ve kübik düzeltme spline yönteminde ise ileriki yıllarda grafik çizgisinin yukarıya doğru arttığı görülmüştür.

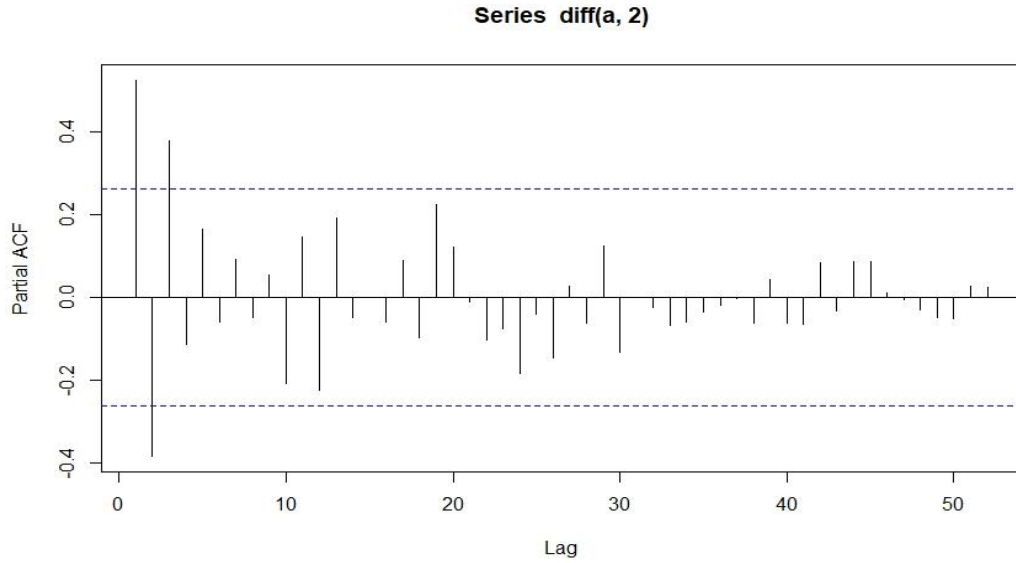


Şekil 4.7. Üçüncü Aşamada Kullanılan Yöntemlerin Tahminleri

ARIMA modelleme yöntemini incelersek kullanılan verinin zamandan etkilenmemesi için durağan hale getirmek amacıyla 2 normal fark alınması gerektiği ortaya çıkmıştır. Durağan hale getirilmiş verinin 52 gecikmeye kadar ACF grafiği Şekil 4.14'te ve PACF grafiği ise Şekil 4.15'te gösterilmiştir. ARIMA modellemede p ve q değerlerini bulabilmek için Şekil 4.8'deki ACF grafiği ile Şekil 4.9'daki PACF grafikleri incelenmiştir. Şekil 4.8'deki grafik incelendiğinde 1. gecikmede anlamlı çubuk olması MA(1) modelini yani q değerinin 1 olabileceği ortaya çıkmıştır. Şekil 4.9'daki grafik incelendiğinde 1. gecikme, 2. gecikme ve 3. gecikmede anlamlı çubuk olması AR(1), AR(3) ve AR(13) modelini yani p değerinin 1, 2 veya 3 olabileceği ortaya çıkmıştır. Veriyi durağan hale getirmek için 2 normal fark alınmasıyla d değerinin 2 olduğu ortaya çıkmıştır.



Şekil 4.8. Durađan Hale Getirilen Verinin ACF Grafiđi

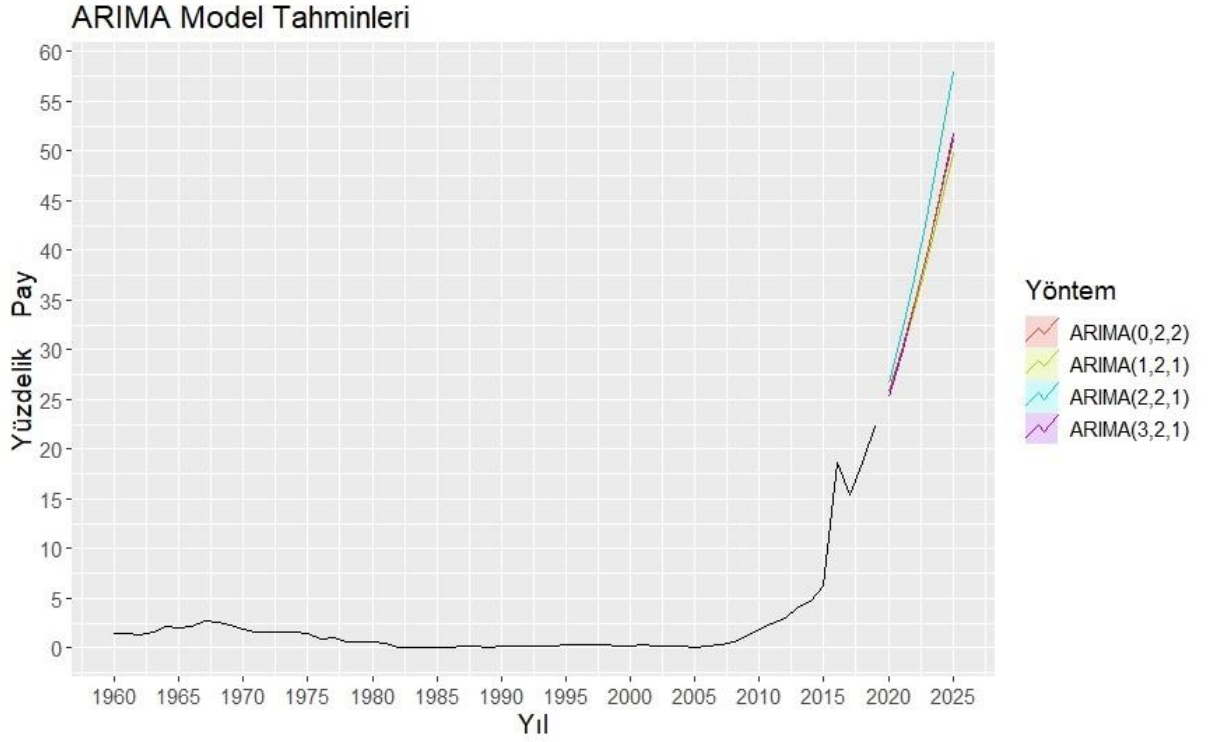


Şekil 4.9. Durađan Hale Getirilen Verinin PACF Grafiđi

ARIMA modelleme yöntemini kullanmak için ARIMA (1,2,1), ARIMA (2,2,1) ve ARIMA (3,2,1) modelleme yöntemleri kullanılmıştır. Ek olarak R-Studio uygulamasında otomatik olarak oluşturulan ARIMA (0,2,2) modelleme yöntemi de kullanılmıştır.

ARIMA (1,2,1), ARIMA (2,2,1), ARIMA (3,2,1) ve ARIMA (0,2,2) modelleme yöntemleri kullanılarak 2020 yılından 2025 yılı dâhil olmak üzere tahmin edilen yüzdelik pay

değerlerin Şekil 4.10'daki grafikte gösterilmiştir. Kullanılan yöntemlerin 2020 yılından itibaren çizgi grafiğinin yüksek bir trend doğrultusunda arttığı görülmüştür ve çizgilerin birbirine yakın olduğu tespit edilmiştir fakat ARIMA (1,2,1) modelinin grafik çizgisi diğer modellere göre daha yüksek bir trend göstermiştir.



Şekil 4.10. Dördüncü Aşamada Kullanılan Yöntemlerin Tahminleri

2020'den 2025 yılı dahil olmak üzere yapay sinir ağları ve zaman serisi analizi tahminleme işleminde kullanılan her bir yöntemin ortalama mutlak hata (MAE) ve kök ortalama kare hatası (RMSE) değerleri Çizelge 4.2'de gösterilmiştir. En iyi sonucu veren yöntem koyu renkle gösterilen sinir ağı otoregresyonu (NNAR) yöntemi, ikincisi Holt's trend yöntemi, üçüncüsü ARIMA (0,0,2) model yöntemi ve dördüncüsü ise ARIMA (1,2,1) model yöntemi olmuştur. Ek olarak, 2020'den 2025 yılı dahil olmak üzere yapay sinir ağları ve zaman serisi analizi tahminleme işleminde kullanılan her bir yöntemin tahmin edilen yüzdelik pay değerleri Çizelge 4.3'te gösterilmiştir. İki tabloda elde edilen değerler virgülden sonra üç rakam şeklinde gösterilmiştir.

Çizelge 4.2. Kullanılan Yöntemlerin Tahminleme İşlemi Sonrası Performans Karşılaştırılması

Kullanılan Tahminleme Yöntemleri	Ortalama Mutlak Hata (MAE)	Kök Ortalama Kare Hata (RMSE)
Çok Katmanlı Algılayıcı (MLP) Yöntemi	2,433	4,708
Sinir Ağı Otoregresyonu (NNAR) Yöntemi	0,476	1,464
Aşırı Öğrenme Makinesi (ELM) Yöntemi	0,779	1,835
Yerel Ağırlıklı Öğrenme (LWL) Yöntemi	2,229	5,054
Torbalama (Bagging) Yöntemi	1,328	2,726
Sıralı Minimum Optimizasyon Regresyon (SMOreg) Algoritması Yöntemi	1,742	3,812
Basit Üstel Yöntemi	0,577	1,789
Holt's Trend Yöntemi	0,438	1,548
Doğrusal Regresyon Yöntemi	2,895	4,184
Kübik Düzeltme Spline Yöntemi	0,704	2,623
Ortalama Model Yöntemi	2,529	4,611
Saf Model Yöntemi	0,592	1,809
ARIMA (1,2,1)	0,475	1,683
ARIMA (2,2,1)	0,526	1,695
ARIMA (3,2,1)	0,537	1,718
ARIMA (0,0,2)	0,462	1,660

Çizelge 4.3. Kullanılan Yöntemlerin Tahmin Değerleri

Kullanılan Tahminleme Yöntemleri	2020	2021	2022	2023	2024	2025
Çok Katmanlı Algılayıcı (MLP) Yöntemi	20,569	20,607	20,614	20,616	20,618	20,620
Sinir Ağı Otoregresyonu (NNAR) Yöntemi	22,108	22,044	22,020	22,010	22,006	22,005
Aşırı Öğrenme Makinesi (ELM) Yöntemi	22,635	23,001	23,366	23,731	24,096	24,461
Yerel Ağırlıklı Öğrenme (LWL) Yöntemi	19,005	19,046	19,035	19,031	19,026	19,023
Torbalama (Bagging) Yöntemi	9,682	9,682	9,682	9,682	9,682	9,682
Sıralı Minimum Optimizasyon Regresyon (SMOreg) Algoritması Yöntemi	23,361	25,638	27,674	30,639	32,532	35,251
Basit Üstel Yöntemi	22,053	22,053	22,053	22,053	22,053	22,053
Holt's Trend Yöntemi	22,410	30,071	33,733	37,395	41,057	44,719
Doğrusal Regresyon Yöntemi	5,701	5,813	5,924	6,036	6,148	6,259
Kübik Düzeltme Spline Yöntemi	25,253	28,773	32,647	36,878	41,507	46,549
Ortalama Model Yöntemi	2,292	2,292	2,292	2,292	2,292	2,292
Saf Model Yöntemi	22,270	22,270	22,270	22,270	22,270	22,270
ARIMA (1,2,1)	25,760	29,675	34,010	38,793	44,054	49,822
ARIMA (2,2,1)	26,632	31,564	37,082	43,292	50,249	57,989
ARIMA (3,2,1)	25,324	29,555	34,432	39,557	45,319	51,750
ARIMA (0,0,2)	25,897	29,985	34,530	39,562	45,114	51,219

4.1. TEST VE EĞİTİM VERİSİ AYRIMI İLE YAPILAN DENEYLER

Araştırma bulgularının alt kısmı olan test ve eğitim verisi ayrımı ile yapılan deneyler bölümünde kullanılan tahminleme yöntemleri arasında en iyi performansı belirlemek için ana materyaldeki veriler eğitim ve test verisi olmak üzere iki çeşide ayrılmıştır. Ana materyalde bulunan 1960-2014 yılları arasındaki veriler eğitim verisi ve 2015-2019 yılları arasındaki veriler ise test verisi olarak belirlenmiştir. Eğitim verisi kullanılarak 2015-2019 yılları arası tahmin işlemleri gerçekleştirilmiştir. Kullanılan yöntemlerin performansı tahminleme işlemleri sonuçlarında ortaya çıkan test verilerindeki ortalama mutlak hata (MAE) ve kök ortalama kare hata (RMSE) değerleri üzerinde değerlendirilmiştir. Çizelge 4.4'te en iyi sonucu veren üç tane tahminleme yöntemleri gösterilmiştir. Çizelge 4.4'te en iyi performansı veren tahminleme yönteminin test verisi kısmına bakılarak sinir ağı otoregresyonu (NNAR) yöntemi olmuştur. İkinci en iyi yöntem ARIMA (2,2,1) modeli yöntemi ve üçüncü en iyi tahmini veren yöntem ise kübik spline düzeltme yöntemi olmuştur. Çizelge 4.5'te en iyi sonucu veren yöntemlerin tahmin değerleri ve gerçek değerleri gösterilmiştir. Çizelge 4.5 incelendiğinde kübik düzeltme spline yöntemi ve ARIMA (2,2,1) modeli yönteminde elde edilen değerler birbirine yakın ve artma göstermesine rağmen gerçek değerlerle uyuşmadığı görülmüştür. Diğer taraftan ise sinir ağı otoregresyonu (NNAR) yönteminde elde edilen değerler ile gerçek değerlerin birbirine diğer yöntemlere göre daha uygun değerler olduğu ortaya çıkmıştır.

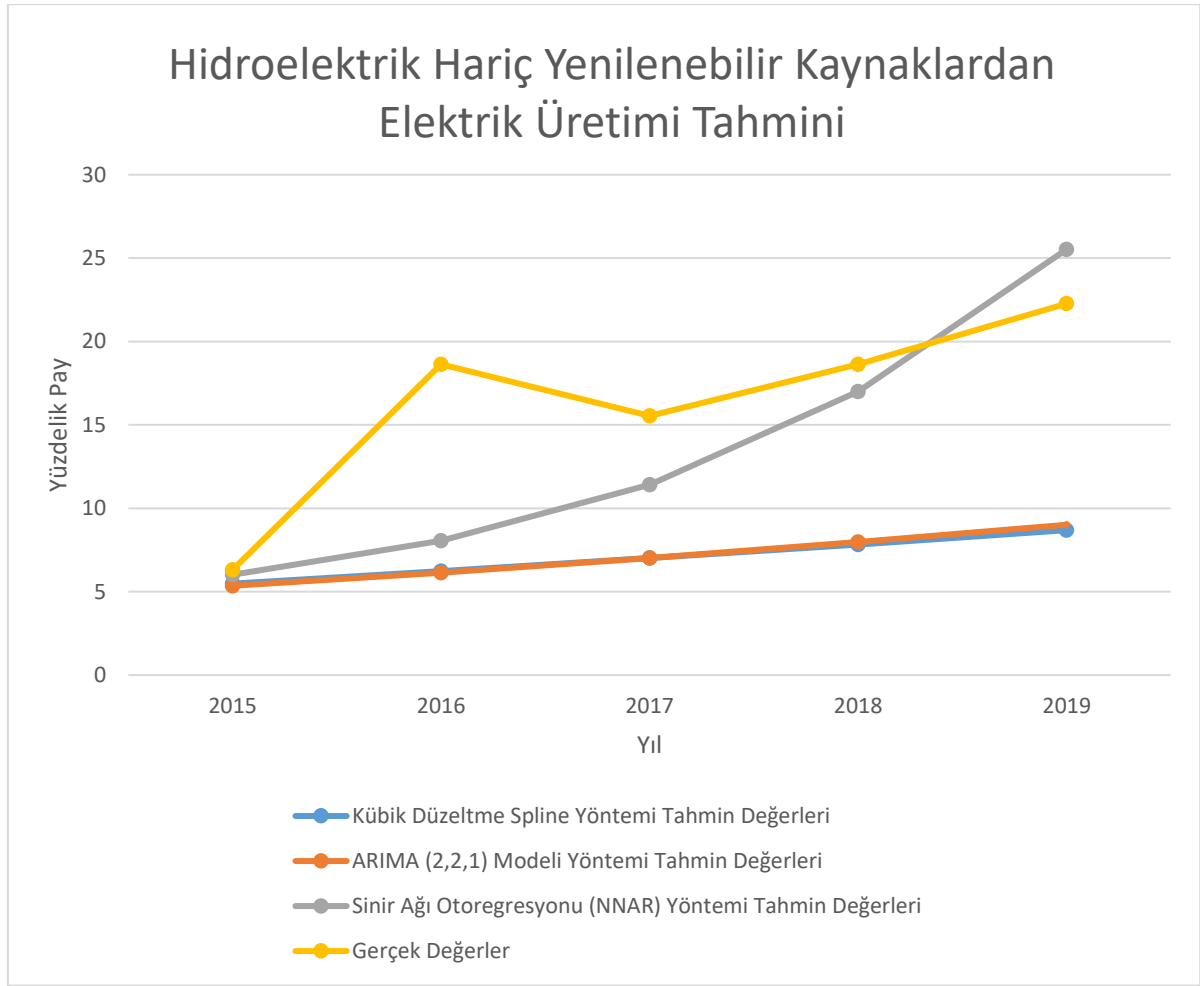
Çizelge 4.4. En İyi Sonucu Veren Üç Tane Yöntemin Eğitim ve Test Verilerinin Performans Karşılaştırılması

Tahminleme Yöntemleri	Veri Çeşitleri	Ortalama Mutlak Hata (MAE)	Kök Ortalama Kare Hata (RMSE)
ARIMA (2,2,1)	Eğitim Verisi	0,185	0,2678
	Test Verisi	9,161	10,176
Sinir Ağı Otoregresyonu (NNAR) Yöntemi	Eğitim Verisi	0,195	0,290
	Test Verisi	3,961	5,323
Kübik Düzeltme Spline Yöntemi	Eğitim Verisi	0,202	0,287
	Test Verisi	9,211	10,265

Çizelge 4.5. En İyi Sonucu Veren Üç Tane Yöntemin Tahmin Değerleri ve Gerçek Değerleri

Yıl	Kübik Düzeltme Spline Yöntemi Tahmin Değerleri	ARIMA (2,2,1) Modeli Yöntemi Tahmin Değerleri	Sinir Ağı Otoregresyonu (NNAR) Yöntemi Tahmin Değerleri	Gerçek Değerler
2015	5,482	5,353	6,016	6,308
2016	6,229	6,131	8,047	18,631
2017	7,013	7,022	11,420	15,453
2018	7,834	7,974	16,994	18,639
2019	8,689	9,011	25,525	22,270

Çizelge 4.5'teki kullanılan yöntemlerin tahmin değerleri ve gerçek değerler Şekil 4.11'de grafik üzerinde gösterilmiştir. Kübik düzeltme spline yöntemi tahmin değerleri ile ARIMA (2,2,1) modeli yöntemi tahmin değerleri birbirine yakın çizgide görülmüştür ama gerçek değerlerin grafik çizgisinden uzakta olduğu görülmüştür. Sinir ağı otoregresyonu (NNAR) yöntemi tahmin değerlerinin grafik çizgisinin diğer tahminleme yöntemlerine göre gerçek değerlerin grafik çizgisine daha yakın çizgide olduğu görülmüştür.



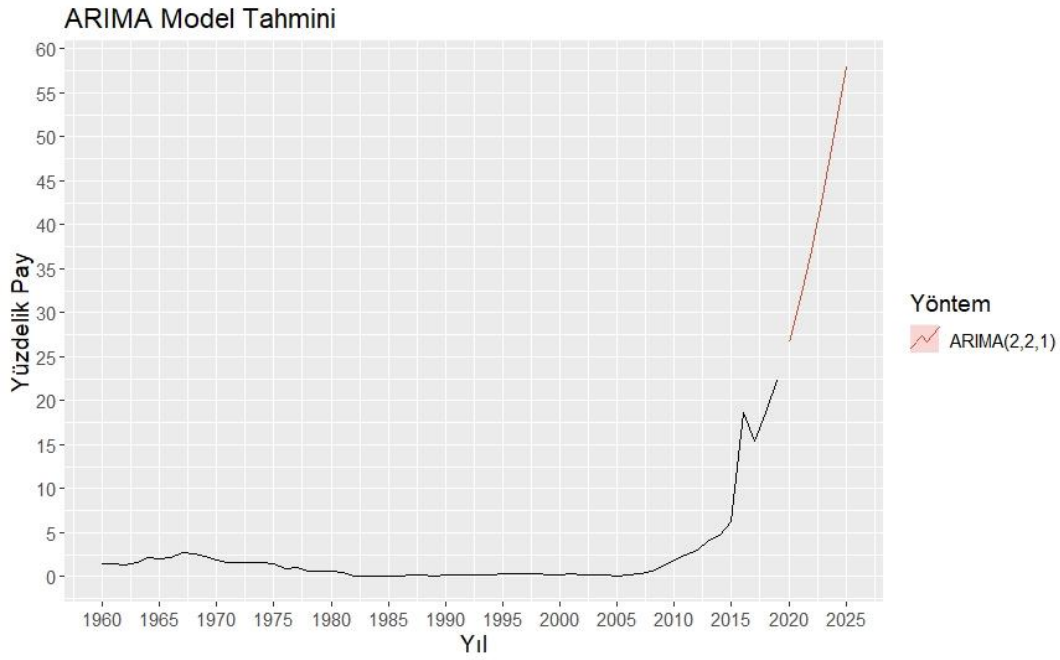
Şekil 4.11. En İyi Sonucu Veren Üç Tane Yöntemin Tahmin Değerleri ve Gerçek Değerleri

Ana materyaldeki verilerin hepsini eğitim verisi alınarak test verileri üzerinde en iyi performans veren sinir ağı otoregresyon (NNAR) yönteminin gelecekte Şekil 4.1'deki grafik gibi çizgi dağılımı gösterip göstermediği kontrol edilmiştir. Bu doğrultuda 2020'den 2025 yılına kadar verileri tahmin etme süreci izlenmiştir. Sinir ağı otoregresyonu (NNAR) yönteminin 2020'den 2025 yılına kadar tahmin edilen değerler incelendiğinde her bir yıl azalmıştır ve değerler sabit aralıkta ve sabit çizgide kalmıştır.

Test verileri üzerinde en iyi performansı veren ikinci yöntem olan ARIMA (2,2,1) modelleme yöntemi en iyi performansı veren yöntemde izlenen sürecin aynısı izlenmiştir. Tahmin edilen değerler Çizelge 4.6'da gösterilmiştir. Şekil 4.12'deki grafik incelendiğinde 2019 yılından sonra tahmin edilen değerler çizgisinin sabit olmadığı yukarıya doğru artmakta olduğu ortaya çıkmıştır.

Çizelge 4.6. ARIMA (2,2,1) Modelleme Yönteminin Tahmin Değerleri

Tahminleme Yöntemi	2020	2021	2022	2023	2024	2025
ARIMA (2,2,1)	26,632	31,564	37,082	43,292	50,249	57,989

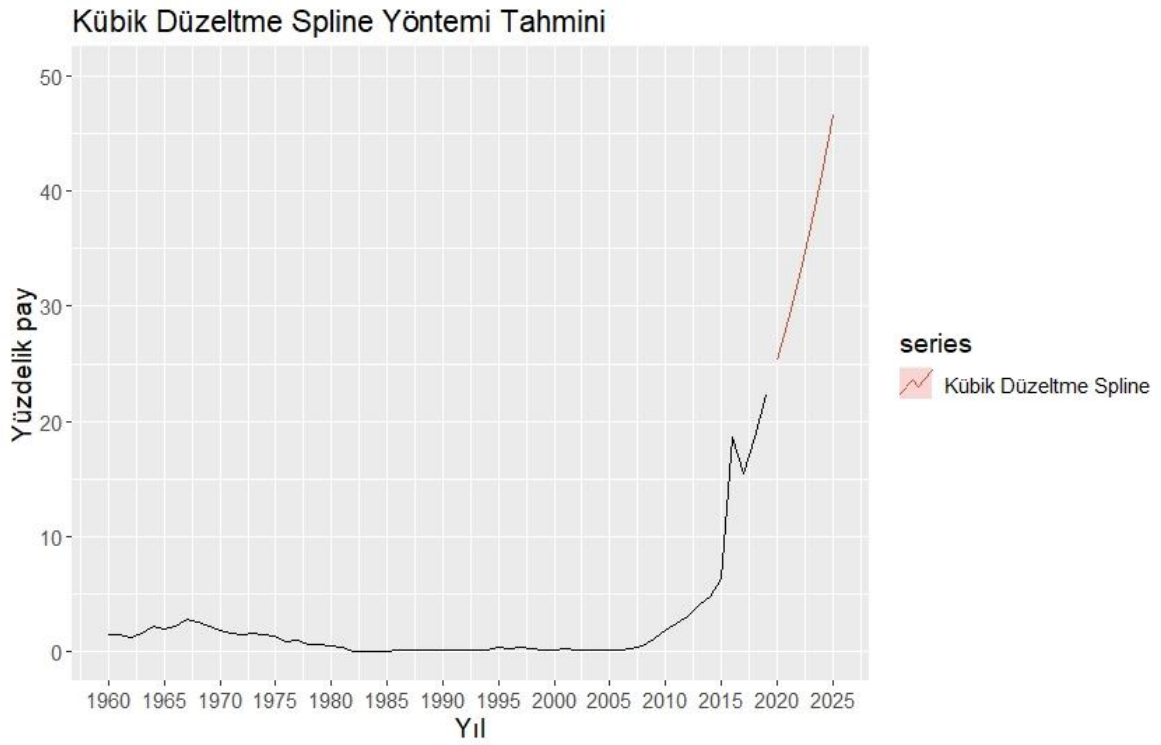


Şekil 4.12. ARIMA (2,2,1) Modelleme Yönteminin Tahminleri

Ayrıca test verileri üzerinde en iyi performansı veren üçüncü yöntem olan kübik düzeltme spline yönteminde de sürecin aynısı izlenmiştir. Tahmin edilen değerler Çizelge 4.7’de gösterilmiştir. Şekil 4.13’deki grafik incelendiğinde 2019 yılından sonra tahmin edilen değerler çizgisinin de sabit olmadığı yukarıya doğru artmakta olduğu ortaya çıkmıştır. Ama bu artma Şekil 4.12’deki grafikte görüldüğü kadar olmamıştır.

Çizelge 4.7. Kübik Düzeltme Spline Yöntemi Tahmin Değerleri

Tahminleme Yöntemi	2020	2021	2022	2023	2024	2025
Kübik Düzeltme Spline Yöntemi	25,253	28,773	32,647	36,878	41,507	46,549



Şekil 4.13. Kübik Düzeltme Spline Yönteminin Tahminleri

5. TARTIŞMA VE SONUÇ

Yapılan bu tez çalışmasında daha önce işlenmemiş veri kümesi kullanılmıştır. Yenilenebilir enerji kaynakları ile ilgili olarak birçok tahminleme çalışması yapılmıştır. Önceki yapılan çalışmalar incelendiğinde benzer olarak ARIMA, ELM ve MLP tahminleme yöntemlerinin kullanıldığı görülür. Bu çalışmada çok sayıda heuristik yöntem karşılaştırmalı ve geniş kapsamlı olarak incelenmiştir. Türkiye’de 1960-2019 yıllarına (60 yıl) ait yenilebilir enerji kaynaklarından elde edilen elektrik miktarı verisi kullanılmıştır. Veri kümesi içerisinde ilk 55 veri eğitim verisi ve son 5 veri ise test verisi olarak kullanılmıştır. Veri kümesi üzerinde zaman serisi tahminleme yöntemleri denenmiş ve en iyi performans sonucu sahip yöntemin yapay sinir ağları yöntemlerinden biri olan sinir ağı otoregresyonu (NNAR) yönteminin olduğu görülmüştür. Ancak geçmiş yıllardan süregelen verilerde yükseliş yönünde bir trend olduğundan sabit kalacağını öngören sinir ağı otoregresyonu (NNAR) yöntemi tercih edilmemiştir. Yükseliş trendi yönünde tahminlerde bulunan yöntemler incelendiğinde ise en iyi sonuçları ARIMA (2,2,1) ve kübik düzeltme spline yöntemleri ortaya koymuştur. Türkiye için ARIMA (2,2,1) ve kübik düzeltme spline yöntemleri kullanılarak 2021 ile 2025 yılı arasında artış oranları ise sırasıyla yaklaşık olarak %83 ve %62 olarak belirlenmiştir. Elde ettiğimiz bulgulara göre, Türkiye’de 2025 yılında yenilebilir enerjinin toplam elektrik üretimindeki payı %62-83 arasında öngörülebilir. Gelecekte araştırmacılar yapılacak olan çalışmalarda yenilenebilir enerji kaynaklarını önceden öngörerek bunun için, gerekli araç, gereç, personel temini gibi birtakım ihtiyaçlar daha verimli planlanabilir. Bu konuda regülasyonlar yayınlayan karar vericiler teşvik, artış sağlamak amaçlı birtakım planlamalar ve organizasyonlar yapılabilir.

Gelecekte veri kümesi içerisinde test ve eğitim amacıyla kullanılacak olan veri aralığı değişik oranlarla yeniden değerlendirilebilir. Kullanılan veri kümesi ve teknikler dışındaki yaklaşımlar kullanılarak çalışma tekrarlanabilir. Yapılan çalışmada kullanılan teknikler yeraltı ve yerüstü kaynakların sayısal tahminleme analizlerinde denenebilir. Araştırmacılar yapacakları benzer testlerde ARIMA (2,2,1) ve kübik düzeltme spline yöntemlerini tercih edebilirler.

KAYNAKLAR

- [1]E. B. Abusaada, “THE RENEWABLE ENERGY TRANSITION THE RENEWABLE ENERGY TRANSITION,” 2020.
- [2]H. Bölen Türkmen, “DETAILED RENEWABLE ENERGY SYSTEMS ANALYSIS-THE CASE OF TURKEY IN YEARS 2023 AND 2053 DETAILED RENEWABLE ENERGY SYSTEMS ANALYSIS-THE CASE OF TURKEY IN YEARS 2023 AND 2053,” *Master Thesis MASTER THESIS Dep. Mech. Eng. Thesis*, 2019.
- [3]O. KULAÇ and M. CİĞEROĞLU ÖZTEPE, “THE RENEWABLE ENERGY POLICY OF TURKEY UNDER THE IMPACT OF THE EUROPEAN UNION,” *Süleyman Demirel Üniversitesi Vizyoner Derg.*, pp. 886–897, 2020, doi: 10.21076/vizyoner.693835.
- [4]T. S. Uyar and D. Beşikci, “Integration of hydrogen energy systems into renewable energy systems for better design of 100% renewable energy communities,” *Int. J. Hydrogen Energy*, vol. 42, no. 4, pp. 2453–2456, 2017, doi: 10.1016/j.ijhydene.2016.09.086.
- [5]A. S. Şekerci, “RENEWABLE ENERGY PROMOTION STRATEGIES OF INDIVIDUAL STATES AND REGIONAL FRAMEWORKS: A COMPARATIVE APPROACH,” *Partial FULFILLMENT Requir. DEGREE MASTER Sci. Dep. Int. RELATIONS*, vol. 4, no. July, pp. 9–15, 2017.
- [6]M. A. Destek and A. Aslan, “Renewable and non-renewable energy consumption and economic growth in emerging economies: Evidence from bootstrap panel causality,” *Renew. Energy*, vol. 111, pp. 757–763, 2017, doi: 10.1016/j.renene.2017.05.008.
- [7]L. Olatomiwa, S. Mekhilef, M. S. Ismail, and M. Moghavvemi, “Energy management strategies in hybrid renewable energy systems: A review,” *Renew. Sustain. Energy Rev.*, vol. 62, pp. 821–835, 2016, doi: 10.1016/j.rser.2016.05.040.
- [8]H. C. Gils and S. Simon, “Carbon neutral archipelago – 100% renewable energy supply for the Canary Islands,” *Appl. Energy*, vol. 188, pp. 342–355, 2017, doi: 10.1016/j.apenergy.2016.12.023.
- [9]K. W. ABID, “DEVELOPMENT OF RENEWABLE ENERGY CIRCUIT COMPONENT,” *Univ. GAZIANTEP Grad. Sch. Nat. Appl. Sci. Sc. THESIS Electr. Electron. Eng.*, no. January, 2018.

- [10] X. Yuan, X. Wang, and J. Zuo, "Renewable energy in buildings in China-A review," *Renew. Sustain. Energy Rev.*, vol. 24, pp. 1–8, 2013, doi: 10.1016/j.rser.2013.03.022.
- [11] O. Ocal and A. Aslan, "Renewable energy consumption-economic growth nexus in Turkey," *Renew. Sustain. Energy Rev.*, vol. 28, pp. 494–499, 2013, doi: 10.1016/j.rser.2013.08.036.
- [12] U. Bulut and G. Muratoglu, "Renewable energy in Turkey: Great potential, low but increasing utilization, and an empirical analysis on renewable energy-growth nexus," *Energy Policy*, vol. 123, no. May, pp. 240–250, 2018, doi: 10.1016/j.enpol.2018.08.057.
- [13] Ö. Yüksek, "Reevaluation of Turkey's hydropower potential and electric energy demand," *Energy Policy*, vol. 36, no. 9, pp. 3374–3382, 2008, doi: 10.1016/j.enpol.2008.05.004.
- [14] I. Yüksel, "Hydropower for sustainable water and energy development," *Renew. Sustain. Energy Rev.*, vol. 14, no. 1, pp. 462–469, 2010, doi: 10.1016/j.rser.2009.07.025.
- [15] E. Toklu, "Overview of potential and utilization of renewable energy sources in Turkey," *Renew. Energy*, vol. 50, pp. 456–463, 2013, doi: 10.1016/j.renene.2012.06.035.
- [16] S. Keleş and S. Bilgen, "Renewable energy sources in Turkey for climate change mitigation and energy sustainability," *Renew. Sustain. Energy Rev.*, vol. 16, no. 7, pp. 5199–5206, 2012, doi: 10.1016/j.rser.2012.05.026.
- [17] M. Balat, "Use of biomass sources for energy in Turkey and a view to biomass potential," *Biomass and Bioenergy*, vol. 29, no. 1, pp. 32–41, 2005, doi: 10.1016/j.biombioe.2005.02.004.
- [18] K. Kaygusuz and S. Keles, "Use of biomass as a transitional strategy to a sustainable and clean energy system," *Energy Sources, Part A Recover. Util. Environ. Eff.*, vol. 31, no. 1, pp. 86–97, 2009, doi: 10.1080/15567030701468225.
- [19] N. Saracoglu, "The biomass potential of Turkey for energy production: Part i," *Energy Sources, Part B Econ. Plan. Policy*, vol. 5, no. 3, pp. 272–278, 2010, doi: 10.1080/15567240802532981.
- [20] A. Akpınar, M. I. Kömürcü, H. Önsoy, and K. Kaygusuz, "Status of geothermal energy amongst Turkey's energy sources," *Renew. Sustain. Energy Rev.*, vol. 12, no. 4, pp. 1148–1161, 2008, doi: 10.1016/j.rser.2006.10.016.

- [21] K. Kaygusuz and A. Kaygusuz, "Geothermal energy in Turkey: The sustainable future," *Renew. Sustain. Energy Rev.*, vol. 8, no. 6, pp. 545–563, 2004, doi: 10.1016/j.rser.2004.01.001.
- [22] I. Akkus, O. Aydogdu, H. Akıllı, O. Gokmenoglu, and S. Sarp, "Geothermal energy and its economic dimension in Turkey," *Geotherm. energy its Econ. Dimens. Turkey. Proc. world Geotherm. Congr. Antalya, Turkey*.
- [23] A. Hepbasli and L. Ozgener, "Development of geothermal energy utilization in Turkey: A review," *Renew. Sustain. Energy Rev.*, vol. 8, no. 5, pp. 433–460, 2004, doi: 10.1016/j.rser.2003.12.004.
- [24] Ö. Güler, "Wind energy status in electrical energy production of Turkey," *Renewable and Sustainable Energy Reviews*, vol. 13, no. 2. Pergamon, pp. 473–478, Feb. 01, 2009, doi: 10.1016/j.rser.2007.03.015.
- [25] H. Ergün, "The Effect of Problem Posing on Problem Solving in Introductory," vol. 6, no. 3, pp. 1–10, 2010.
- [26] A. D. Şahin, "Yenilebilir Enerji Kaynaklarının Potansiyeli, Kullanabilirliği ve Temel Sorunları," *İ.T.Ü. Vakfı Derg.*, vol. 60, 2013.
- [27] E. Ekinci and H. Okutan, "Küresel İklim Değişikliği: Türkiye Enerji Sektörü ve Temiz Teknolojiler,Küresel İklim Değişimi ve Su Sorunların Çözümünde Ormanlar," *Küresel İklim Değişimi ve Su Sorunlarının Çözümünde Ormanlar Sempozyumu, İstanbul Üniversitesi Orman Fakültesi, Bahçeköy*, 2007.
- [28] B. D. Çift and H. Okutan, "TURKEY'S ENERGY VIEW, CLEAN ENERGY TECHNOLOGIES AND DETERMINATION OF APPROPRIATE ENERGY POLICY," *J. Navaal Sci. Eng.*, vol. 9, no. 1, pp. 81–97, 2013.
- [29] C. Paoli, C. Voyant, M. Muselli, and M. L. Nivet, "Forecasting of preprocessed daily solar radiation time series using neural networks," *Sol. Energy*, vol. 84, no. 12, pp. 2146–2160, 2010, doi: 10.1016/j.solener.2010.08.011.
- [30] F. O. Hocaoglu and F. Karanfil, "A time series-based approach for renewable energy modeling," *Renew. Sustain. Energy Rev.*, vol. 28, pp. 204–214, 2013, doi: 10.1016/j.rser.2013.07.054.

- [31] E. Cadenas and W. Rivera, “Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model,” *Renew. Energy*, vol. 35, no. 12, pp. 2732–2738, 2010, doi: 10.1016/j.renene.2010.04.022.
- [32] F. Golestaneh, P. Pinson, and H. B. Gooi, “Very short-term nonparametric probabilistic forecasting of renewable energy generation - With application to solar energy,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 3850–3863, 2016, doi: 10.1109/TPWRS.2015.2502423.
- [33] P. Jiang, J. Dong, and H. Huang, “Forecasting China’s renewable energy terminal power consumption based on empirical mode decomposition and an improved extreme learning machine optimized by a bacterial foraging algorithm,” *Energies*, vol. 12, no. 7, 2019, doi: 10.3390/en12071331.
- [34] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, “A review of deep learning for renewable energy forecasting,” *Energy Convers. Manag.*, vol. 198, no. July, p. 111799, 2019, doi: 10.1016/j.enconman.2019.111799.
- [35] C. Sweeney, R. J. Bessa, J. Browell, and P. Pinson, “The future of forecasting for renewable energy,” *Wiley Interdiscip. Rev. Energy Environ.*, vol. 9, no. 2, pp. 1–18, 2019, doi: 10.1002/wene.365.
- [36] K. Tharani, N. Kumar, V. Srivastava, S. Mishra, and M. Pratyush Jayachandran, “Machine learning models for renewable energy forecasting,” *J. Stat. Manag. Syst.*, vol. 23, no. 1, pp. 171–180, 2020, doi: 10.1080/09720510.2020.1721636.
- [37] C. A. Severiano, P. C. de L. e. Silva, M. Weiss Cohen, and F. G. Guimarães, “Evolving fuzzy time series for spatio-temporal forecasting in renewable energy systems,” *Renew. Energy*, vol. 171, pp. 764–783, 2021, doi: 10.1016/j.renene.2021.02.117.
- [38] C. Goncalves, R. J. Bessa, and P. Pinson, “Privacy-preserving Distributed Learning for Renewable Energy Forecasting,” *IEEE Trans. Sustain. Energy*, vol. 3029, no. c, pp. 1–10, 2021, doi: 10.1109/TSTE.2021.3065117.
- [39] K. Hornik, C. Buchta, and A. Zeileis, “Open-source machine learning: R meets Weka,” *Comput. Stat.*, vol. 24, no. 2, pp. 225–232, 2009, doi: 10.1007/s00180-008-0119-7.
- [40] M. Varouqa and B. Hammo, “WIT: Weka interface translator,” *Int. J. Speech Technol.*, vol. 19, no. 2, pp. 359–371, 2016, doi: 10.1007/s10772-015-9283-7.

- [41] S. Lang, F. Bravo-Marquez, C. Beckham, M. Hall, and E. Frank, "WekaDeeplearning4j: A deep learning package for Weka based on Deeplearning4j," *Knowledge-Based Syst.*, vol. 178, pp. 48–50, 2019, doi: 10.1016/j.knosys.2019.04.013.
- [42] L. Gula and R. Norberg, "Information Data Management for the Future of Communication," *Master Sci. Eng. Technol. Comput. Sci. Eng. Luleå Univ. Technol. Dep. Comput. Sci. Electr. Sp. Eng.*
- [43] D. Merlini and M. Rossini, "Text categorization with WEKA: A survey," *Mach. Learn. with Appl.*, vol. 4, no. March, p. 100033, 2021, doi: 10.1016/j.mlwa.2021.100033.
- [44] P. Kotak and H. Modi, "Enhancing the data mining tool WEKA," *Proc. 2020 Int. Conf. Comput. Commun. Secur. ICCCS 2020*, pp. 4–9, 2020, doi: 10.1109/ICCCS49678.2020.9276870.
- [45] F. M. Nafie Ali and A. A. Mohamed Hamed, "Usage Apriori and clustering algorithms in WEKA tools to mining dataset of traffic accidents," *J. Inf. Telecommun.*, vol. 2, no. 3, pp. 231–245, 2018, doi: 10.1080/24751839.2018.1448205.
- [46] R. R. Bouckaert, E. Frank, R. Kirkby, P. Reutemann, A. Seewald, and D. Scuse, "WEKA Manual for Version 3-7-2," 2002, [Online]. Available: <http://netcologne.dl.sourceforge.net/project/weka/documentation/3.7.x/WekaManual-3-7-12.pdf>.
- [47] I. H. Witten, E. Frank, L. Trigg, M. Hall, G. Holmees, and S. J. Cunningham, "WEKA: Practical Machine Learning Tools and Techniques with Java Implementations," *Dep. Comput. Sci. Univ. Waikato*, 1994, [Online]. Available: <https://researchcommons.waikato.ac.nz/bitstream/handle/10289/1040/uow-cs-wp-1999-11.pdf?sequence=1&isAllowed=y>.
- [48] R. Meenal, A. Immanuel Selvakumar, K. Brighta, S. Christy Jeba Joice, and C. P. Richerd, "Solar radiation resource assessment using WEKA," *Proc. 2nd Int. Conf. Inven. Syst. Control. ICISC 2018*, no. Icisc, pp. 1038–1042, 2018, doi: 10.1109/ICISC.2018.8398960.
- [49] M. J. Lanovaz and B. Adams, "Comparing the Communication Tone and Responses of Users and Developers in Two R Mailing Lists: Measuring Positive and Negative Emails," *IEEE Softw.*, vol. 36, no. 5, pp. 46–50, 2019, doi: 10.1109/MS.2019.2922949.
- [50] J. Davis, *Practical Data Science Cookbook*. 2014.

- [51] M. Gardener, "Beginning R: The Statistical Programming Language - Introducing R: What It Is and How to Get It," pp. 1–24, 1995.
- [52] N. Matloff, *The art of R programming : tour of statistical software design*. 2011.
- [53] J. Maindonald and W. J. Braun, "A brief introduction to R," *Data Anal. Graph. Using R – an Example-Based Approach*, pp. 1–42, 2013, doi: 10.1017/cbo9781139194648.004.
- [54] D. L. Carlson, *Introduction to R*, vol. 32. 2017.
- [55] J. Stander and L. Dalla Valle, "On enthusing students about big data and social media visualization and analysis using R, RStudio, and RMarkdown," *J. Stat. Educ.*, vol. 25, no. 2, pp. 60–67, 2017, doi: 10.1080/10691898.2017.1322474.
- [56] M. Katamaneni, G. Guttikonda, and M. Suneetha, "Implementing of decision tree algorithm using R-studio and Java," *Proc. 4th IEEE Int. Conf. Adv. Electr. Electron. Information, Commun. Bio-Informatics, AEEICB 2018*, pp. 16–20, 2018, doi: 10.1109/AEEICB.2018.8481001.
- [57] J. . Allaire, "RStudio: Integrated Development Environment for R," *Kaleidosc. Ic, The R User Conf. 2011*, p. 14, 2011.
- [58] D. Arevalo and C. Padilla, "Medición de la Confiabilidad del Aprendizaje del Programa RStudio Mediante Alfa de Cronbach Measurement reliability RStudio Learning Program using Cronbach ' s alpha," *Rev. Politécnica*, vol. 2, no. January, pp. 1–9, 2016.
- [59] N. R. Salim, K. Gopal, and A. F. M. Ayub, "Experiential Statistics Learning with RStudio: Study on Students' Engagement," *J. Phys. Conf. Ser.*, vol. 1132, no. 1, 2018, doi: 10.1088/1742-6596/1132/1/012039.
- [60] K. B. Mustapha, *R for Finite Element Analyses of Size dependent Microscale Structures*. 2019.
- [61] S. Kim, Y. Matsumi, S. Pan, and H. Mase, "A real-time forecast model using artificial neural network for after-runner storm surges on the Tottori coast, Japan," *Ocean Eng.*, vol. 122, pp. 44–53, 2016, doi: 10.1016/j.oceaneng.2016.06.017.
- [62] S. Trajkovic, B. Todorovic, and M. Stankovic, "Forecasting of Reference Evapotranspiration by Artificial Neural Networks," *J. Irrig. Drain. Eng.*, vol. 129, no. 6, pp. 454–457, 2003, doi: 10.1061/(asce)0733-9437(2003)129:6(454).

- [63] A. Mossalam and M. Arafa, "Using artificial neural networks (ANN) in projects monitoring dashboards' formulation," *HBRC J.*, vol. 14, no. 3, pp. 385–392, 2018, doi: 10.1016/j.hbrcj.2017.11.002.
- [64] M. Dehghani, B. Saghafian, F. Nasiri Saleh, A. Farokhnia, and R. Noori, "Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation," *Int. J. Climatol.*, vol. 34, no. 4, pp. 1169–1180, 2014, doi: 10.1002/joc.3754.
- [65] Y. B. NEGEWO, "ANN PREDICTION OF UNCONFINED COMPRESSIVE STRENGTH DEVELOPMENT OF CEMENT MORTAR," 2019.
- [66] "Artificial Neural Network Tutorial - Javatpoint." <https://www.javatpoint.com/artificial-neural-network> (accessed May 08, 2021).
- [67] P. T. Nastos, K. P. Moustris, I. K. Larissi, and A. G. Paliatsos, "Rain intensity forecast using Artificial Neural Networks in Athens, Greece," *Atmos. Res.*, vol. 119, pp. 153–160, 2013, doi: 10.1016/j.atmosres.2011.07.020.
- [68] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," *Int. J. Forecast.*, vol. 14, no. 1, pp. 35–62, 1998, doi: 10.1016/S0169-2070(97)00044-7.
- [69] C. Gleue, D. Eilers, H. J. von Mettenheim, and M. H. Breitner, "Decision Support for the Automotive Industry: Forecasting Residual Values Using Artificial Neural Networks," *Bus. Inf. Syst. Eng.*, vol. 61, no. 4, pp. 385–397, 2019, doi: 10.1007/s12599-018-0527-3.
- [70] A. R. Pazikadin, D. Rifai, K. Ali, M. Z. Malik, A. N. Abdalla, and M. A. Faraj, "Solar irradiance measurement instrumentation and power solar generation forecasting based on Artificial Neural Networks (ANN): A review of five years research trend," *Sci. Total Environ.*, vol. 715, p. 136848, 2020, doi: 10.1016/j.scitotenv.2020.136848.
- [71] B. Sathelly, "An Artificial Neural Network Approach to Predict Liver Failure Likelihood," 2018.
- [72] S. Harvey and R. Harvey, "An introduction to artificial intelligence," *Appita J.*, vol. 51, no. 1, 1998, doi: 10.2514/6.1994-294.
- [73] S. Kuldeep and G. S. Anitha, "Neural Network Approach for Processing Substation Alarms," *Int. Journals Power Electron. Control. Convert.*, vol. 1, no. 1, pp. 21–28, 2017.

- [74] C. Gershenson, "Artificial Neural Networks for Beginners," 2003, [Online]. Available: <http://arxiv.org/abs/cs/0308031>.
- [75] A. Rahimi, "Short-term prediction of NO₂ and NO_x concentrations using multilayer perceptron neural network: a case study of Tabriz, Iran," *Ecol. Process.*, vol. 6, no. 1, 2017, doi: 10.1186/s13717-016-0069-x.
- [76] E. Heidari, M. A. Sobati, and S. Movahedirad, "Accurate prediction of nanofluid viscosity using a multilayer perceptron artificial neural network (MLP-ANN)," *Chemom. Intell. Lab. Syst.*, vol. 155, pp. 73–85, 2016, doi: 10.1016/j.chemolab.2016.03.031.
- [77] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for COVID-19 pandemic in India," *Stoch. Environ. Res. Risk Assess.*, vol. 34, no. 7, pp. 959–972, 2020, doi: 10.1007/s00477-020-01827-8.
- [78] M. Wetzel, M. Sperber, and A. Waibel, "Audio Segmentation for Robust Real-Time Speech Recognition Based on Neural Networks," *Int. Work. Spok. Lang. Transl.*, no. June, 2016.
- [79] A. Akusok, "Extreme Learning Machines : novel extensions and application to Big Data," *Univ. Iowa Iowa Res. Online, A thesis Submitt. Partial fulfillment Requir. Dr. Philos. degree Ind. Eng. Grad. Coll. Univ. Iowa*, 2016.
- [80] A. L. FREIRE, "ON THE EFFICIENT DESIGN OF EXTREME LEARNING MACHINES USING INTRINSIC PLASTICITY AND EVOLUTIONARY COMPUTATION APPROACHES," 2015.
- [81] H. Azimi and H. Shiri, "Sensitivity analysis of parameters influencing the ice–seabed interaction in sand by using extreme learning machine," *Nat. Hazards*, vol. 106, no. 3, pp. 2307–2335, 2021, doi: 10.1007/s11069-021-04544-9.
- [82] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, "E2LMs: Ensemble extreme learning machines for hyperspectral image classification," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 7, no. 4, pp. 1060–1069, 2014, doi: 10.1109/JSTARS.2014.2301775.
- [83] A. Maleki, S. Nasser, M. S. Aminabad, and M. Hadi, "Comparison of ARIMA and NNAR Models for Forecasting Water Treatment Plant's Influent Characteristics," *KSCE J. Civ. Eng.*, vol. 22, no. 9, pp. 3233–3245, 2018, doi: 10.1007/s12205-018-1195-z.

- [84] G. Yu, H. Feng, S. Feng, J. Zhao, and J. Xu, "Forecasting hand-foot-and-mouth disease cases using wavelet-based SARIMA-NNAR hybrid model," *PLoS One*, vol. 16, no. 2 February, pp. 1–13, 2021, doi: 10.1371/journal.pone.0246673.
- [85] K. Gibson, "THE APPLICATION OF MACHINE LEARNING FOR GROUNDWATER LEVEL PREDICTION IN THE STEENKOPPIES COMPARTMENT OF THE GAUTENG AND NORTH WEST DOLOMITE AQUIFER , SOUTH AFRICA," 2020.
- [86] S. Başkan, "EFFECT OF LIGAND BINDING ON PROTEIN DYNAMICS : A TIME SERIES ANALYSIS," *Boğaziçi Univ.*, p. 77, 2008.
- [87] L. Lacasa, V. Nicosia, and V. Latora, "Network structure of multivariate time series," *Sci. Rep.*, vol. 5, pp. 1–10, 2015, doi: 10.1038/srep15508.
- [88] Y. Ben Baccar, "Comparative Study on Time Series Forecasting Models," *Master Sci. (Data Sci. Bertrand LAMY, Jacques DOAN HUU*, no. October, pp. 1–92, 2019, doi: 10.13140/RG.2.2.32241.02408.
- [89] R. Strandberg, "Using Time Series Analysis to Forecast Daily Municipal Water Demand," *Mat. institutionen, Math. Stat. Stock. Univ. Bachelor Thesis 201524*, p. 24, 2015.
- [90] D. Preve, "Essays on Time Series Analysis With Applications to Financial Econometrics," *Digit. Compr. Summ. Uppsala Diss. from Fac. Soc. Sci.* 39, 2008.
- [91] C. A. K. TODOKO, "TIME SERIES ANALYSIS OF WATER CONSUMPTION IN THE HOHOE MUNICIPALITY OF THE VOLTA REGION OF GHANA," *A Thesis Submitt. to Dep. Math. Kwame Nkrumah Univ. Sci. Technol. Partial fulfillment Requir. degree MASTER Sci. (Industrial Math., 2013.*
- [92] X. Huang, "Topics in Multivariate Time Series Analysis : Statistical Control , Dimension Reduction Visualization and Their Business Applications," *Submitt. to Grad. Sch. Univ. Massachusetts Amherst Partial fulfillment Requir. degree Dr. Philos.*, p. 148, 2010.
- [93] T. Warren Liao, "Clustering of time series data - A survey," *Pattern Recognit.*, vol. 38, no. 11, pp. 1857–1874, 2005, doi: 10.1016/j.patcog.2005.01.025.
- [94] N. Uçar, "COMPARISON OF THE FORECAST PERFORMANCES OF LINEAR TIME SERIES AND ARTIFICIAL NEURAL NETWORK MODELS WITHIN THE

CONTEXT OF TURKISH INFLATION,” *Partial Fulfilment Requir. Degree MASTER ARTS*, p. 40, 2001.

- [95] D. Mehrotra, “Performance Comparison of Time Series Data Using Predictive Data Mining Techniques,” *Adv. Inf. Min.*, vol. 4, no. 1, pp. 57–66, 2012.
- [96] C. Lemke, “Combinations of time series forecasts: when and why are they beneficial?,” *Bournemouth Univ.*, no. January, 2010, [Online]. Available: <http://eprints.bournemouth.ac.uk/14609/4/licence.txt>.
- [97] K. P. Moses and M. D. Devadas, “An Approach to Reduce Root Mean Square Error in Toposheets,” *Eur. J. Sci. Res.*, vol. 91, no. 2, pp. 268–274, 2012.
- [98] T. Chai and R. R. Draxler, “Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature,” *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.
- [99] W. Wang and Y. Lu, “Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model,” in *IOP Conference Series: Materials Science and Engineering*, 2018, vol. 324, no. 1, doi: 10.1088/1757-899X/324/1/012049.
- [100] M. R. Mustafa, R. B. Rezaur, S. Saiedi, and M. H. Isa, “River suspended sediment prediction using various multilayer perceptron neural network training algorithms-A case study in Malaysia,” *Water Resour. Manag.*, vol. 26, no. 7, pp. 1879–1897, 2012, doi: 10.1007/s11269-012-9992-5.
- [101] P. D’Urso and E. A. Maharaj, “Autocorrelation-based fuzzy clustering of time series,” *Fuzzy Sets Syst.*, vol. 160, no. 24, pp. 3565–3589, 2009, doi: 10.1016/j.fss.2009.04.013.
- [102] C. P. Dhakal, “A Naïve Approach for Comparing a Forecast Model,” *Int. J. Thesis Proj. Diss.*, vol. 5, no. 1, pp. 1–3, 2017.
- [103] “Mean (constant) model,” *Fuqua Duke University*. https://faculty.fuqua.duke.edu/~rnau/Decision411_2007/411mean.htm (accessed Jun. 13, 2021).
- [104] M. A. Rahman, L. Yunsheng, and N. Sultana, “Analysis and prediction of rainfall trends over Bangladesh using Mann–Kendall, Spearman’s rho tests and ARIMA model,” *Meteorol. Atmos. Phys.*, vol. 129, no. 4, pp. 409–424, 2017, doi: 10.1007/s00703-016-0479-4.

- [105] F. Jiang, X. Yang, and S. Li, "Comparison of forecasting India's energy demand using an MGM, ARIMA model, MGM-ARIMA model, and BP neural network model," *Sustain.*, vol. 10, no. 7, 2018, doi: 10.3390/su10072225.
- [106] L. Zhang *et al.*, "Trend analysis and forecast of PM2.5 in Fuzhou, China using the ARIMA model," *Ecol. Indic.*, vol. 95, no. August, pp. 702–710, 2018, doi: 10.1016/j.ecolind.2018.08.032.
- [107] M. A. Ekmiş, M. Hekimoğlu, and B. Atak Bülbül, "REVENUE FORECASTING USING A FEED-FORWARD NEURAL NETWORK AND ARIMA MODEL," *OBASE A.Ş. Res. Dev. Center, ISTANBUL*, vol. 8, no. 2, pp. 129–134, 2017.
- [108] S. Kang, "A Comparative Study on the Prediction of the Final Settlement Using Preexistence Method and ARIMA Method," *J. Korean Geo-Environmental Soc.*, vol. 20, no. 10, pp. 29–38, 2019.
- [109] S. Chaudhuri and D. Dutta, "Mann-Kendall trend of pollutants, temperature and humidity over an urban station of India with forecast verification using different ARIMA models," *Environ. Monit. Assess.*, vol. 186, no. 8, pp. 4719–4742, 2014, doi: 10.1007/s10661-014-3733-6.
- [110] K. Yürekli, H. Simsek, B. Cemek, and S. Karaman, "Simulating climatic variables by using stochastic approach," *Build. Environ.*, vol. 42, no. 10, pp. 3493–3499, 2007, doi: 10.1016/j.buildenv.2006.10.046.
- [111] Q. Yang, J. Wang, H. Ma, and X. Wang, "Research on COVID-19 based on ARIMA model—Taking Hubei, China as an example to see the epidemic in Italy," *J. Infect. Public Health*, vol. 13, no. 10, pp. 1415–1418, 2020, doi: 10.1016/j.jiph.2020.06.019.
- [112] C. N. Babu and B. E. Reddy, "A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data," *Appl. Soft Comput. J.*, vol. 23, pp. 27–38, 2014, doi: 10.1016/j.asoc.2014.05.028.
- [113] E. Cadenas, W. Rivera, R. Campos-Amezcu, and C. Heard, "Wind speed prediction using a univariate ARIMA model and a multivariate NARX model," *Energies*, vol. 9, no. 2, pp. 1–15, 2016, doi: 10.3390/en9020109.
- [114] R. Nochai and T. Nochai, "Arima Model for Forecasting Oil," *Proc. 2nd IMT-GT Reg. Conf. Math. Statistics Appl.*, no. May 2005, pp. 1–7, 2006.

- [115] M. Nyatuame and S. K. Agodzo, "Stochastic ARIMA model for annual rainfall and maximum temperature forecasting over Tordzie watershed in Ghana," *J. Water L. Dev.*, vol. 37, no. 1, pp. 127–140, 2018, doi: 10.2478/jwld-2018-0032.
- [116] F. Cui, "ARIMA models for bank failures: Prediction and comparison," 2011, [Online]. Available: <http://oatd.org/oatd/record?record=oai%5C:digitalscholarship.unlv.edu%5C:thesisdissertations-2028&q=bank risk management data in nigeria>.
- [117] G. Shiri, B. Salahi, R. Samadzadeh, and M. Shiri, "The investigation and forecasting of relative humidity variation of Pars Abad-e-Moghan, North-West of Iran, by ARIMA model," *Res. J. Appl. Sci.*, vol. 6, no. 2, pp. 81–87, 2011, doi: 10.3923/rjasci.2011.81.87.
- [118] S. A. Sarkodie, "Estimating Ghana's electricity consumption by 2030: An ARIMA forecast," *Energy Sources, Part B Econ. Plan. Policy*, vol. 12, no. 10, pp. 936–944, 2017, doi: 10.1080/15567249.2017.1327993.
- [119] E. M. de Oliveira and F. L. Cyrino Oliveira, "Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods," *Energy*, vol. 144, pp. 776–788, 2018, doi: 10.1016/j.energy.2017.12.049.
- [120] L. Ferbar, D. Čreslovník, B. Mojškerč, and M. Rajgelj, "Demand forecasting methods in a supply chain: Smoothing and denoising," *Int. J. Prod. Econ.*, vol. 118, no. 1, pp. 49–54, 2009, doi: 10.1016/j.ijpe.2008.08.042.
- [121] M. De Ruiter, "Using Exponential Smoothing Methods for Modelling and Forecasting Short-Term Electricity Demand By," no. July, 2017.
- [122] E. Ostertagová and O. Ostertag, "Forecasting using simple exponential smoothing method," *Acta Electrotech. Inform.*, vol. 12, no. 3, p. 10198, 2013, doi: 10.2478/v10198-012-0034-2.
- [123] V. de A. Silva, C. C. M. Cordeiro, M. L. de B. Baltar, J. E. C. Bender, and J. Bortolini, "FORECASTING STORAGE CAPACITY USING EXPONENTIAL SMOOTHING METHOD," *Rev. Produção e Desenvol.*, pp. 1–19, 2020.
- [124] P. Goodwin, "The Holt-Winters Approach to Exponential Smoothing : 50 Years Old and Going Strong," *Foresight*, no. April, pp. 30–33, 2010.
- [125] L. Ferbar Tratar, "Forecasting method for noisy demand," *Int. J. Prod. Econ.*, vol. 161, pp. 64–73, 2015, doi: 10.1016/j.ijpe.2014.11.019.

- [126] Y. J. Puah, Y. F. Huang, K. C. Chua, and T. S. Lee, "River catchment rainfall series analysis using additive holt–Winters method," *J. Earth Syst. Sci.*, vol. 125, no. 2, pp. 269–283, 2016, doi: 10.1007/s12040-016-0661-6.
- [127] D. Dąbrowska, M. Sołtysiak, and J. Waligóra, "Short-term forecasting of the chloride content in the mineral waters of the Ustroń Health Resort using SARIMA and Holt-Winters models," *Environ. Socio-economic Stud.*, vol. 3, no. 4, pp. 57–65, 2015, doi: 10.1515/environ-2015-0074.
- [128] N. Nurhamidah, N. Nusyirwan, and A. Faisol, "Forecasting Seasonal Time Series Data Using the Holt-Winters Exponential Smoothing Method of Additive Models," *J. Mat. Integr.*, vol. 16, no. 2, p. 151, 2020, doi: 10.24198/jmi.v16.n2.29293.151-157.
- [129] S. Sunthornjittanon, "Linear Regression Analysis on Net Income of an Agrochemical Company in Thailand," *Sunthornjittanon, Supichaya, "Linear Regres. Anal. Net Income an Agrochem. Co. Thailand" (2015). Univ. Honor. Theses*, p. 131, 2015.
- [130] E. N. Brown, "Simple Regression Model I, II & III," *MIT OpenCourseWare, 9.07 Stat. Brain Cogn. Sci.*, no. 2016, pp. 1–15, 2016.
- [131] S. Rosenthal, "Regression Analysis, linear," *Nanyang Technol. Univ.*, p. 17, 2013.
- [132] X. Su, X. Yan, and C.-L. Tsai, "Linear Regression," *Wiley Interdiscip. Rev. Comput. Stat.*, vol. 4, no. 3, pp. 275–294, 2012.
- [133] A. Mantzaris, "Predicting the amount of breast cap in broilers Predicting the Amount of Breast Cap in Broilers," *Eindhoven Univ. Technol.*, 2016.
- [134] K. I. Soufya, "Regression and Correlation," *Univ. Belgrade Virtual Libr. Fac. Math. - Univ. Belgrade Fac. Math. Master thesis, candidate Karima Ibrahim Soufya Superv. Vesna Jevremović June 2014 Virtual Libr. Fac.*, no. June, p. 48, 2014.
- [135] N. Altman and M. Krzywinski, "Points of Significance: Simple linear regression," *Nat. Methods*, vol. 12, no. 11, pp. 999–1000, 2015, doi: 10.1038/nmeth.3627.
- [136] "regression - What's the difference between multiple R and R squared?," *Cross Validated*. <https://stats.stackexchange.com/questions/241283/what-is-the-main-difference-between-multiple-r-squared-and-adjusted-r-squared/241298> (accessed May 23, 2021).

- [137] “Adjusted R-squared,” *CFI*.
<https://corporatefinanceinstitute.com/resources/knowledge/other/adjusted-r-squared/>
 (accessed May 23, 2021).
- [138] E. Toor and T. U. Islam, “Power Comparison of Autocorrelation Tests in Dynamic Models,” pp. 58–69, 2019.
- [139] R. Rois, T. Basak, M. M. Rahman, and A. K. Majumder, “Modified Breusch-Godfrey Test for Restricted Higher Order Autocorrelation in Dynamic Linear Model – A Distance Based Approach,” *Int. J. Bus. Manag.*, vol. 7, no. 17, pp. 88–97, 2012, doi: 10.5539/ijbm.v7n17p88.
- [140] N. E. Helwing, “Smoothing Spline Regression,” *Department of Psychology & School of Statistics University of Minnesota*, 2021. <http://users.stat.umn.edu/~helwig/notes/smooth-spline-notes.html> (accessed May 24, 2021).
- [141] R. H. Klein Entink, W. J. van der Linden, and J. P. Fox, “A Box-Cox normal model for response times,” *Br. J. Math. Stat. Psychol.*, vol. 62, no. 3, pp. 621–640, 2009, doi: 10.1348/000711008X374126.
- [142] J. W. Osborne, “Improving your data transformations: Applying the Box-Cox transformation,” *Pract. Assessment, Res. Eval.*, vol. 15, no. 12, 2010.
- [143] C. P. Lin, J. Cabrera, D. Y. W. Yu, F. Yang, and K. L. Tsui, “SOH Estimation and SOC Recalibration of Lithium-Ion Battery with Incremental Capacity Analysis & Cubic Smoothing Spline,” *J. Electrochem. Soc.*, vol. 167, no. 9, p. 090537, 2020, doi: 10.1149/1945-7111/ab8f56.
- [144] G. K. Uyanık and N. Güler, “A Study on Multiple Linear Regression Analysis,” *Procedia - Soc. Behav. Sci.*, vol. 106, pp. 234–240, 2013, doi: 10.1016/j.sbspro.2013.12.027.
- [145] K. A. Marill, “Advanced Statistics: Linear Regression, Part I: Simple Linear Regression,” *Acad. Emerg. Med.*, vol. 11, no. 1, pp. 87–93, 2004, doi: 10.1197/j.aem.2003.09.005.
- [146] C. Cheng, “SYMMETRIC AND TRIMMED SOLUTIONS OF SIMPLE LINEAR REGRESSION,” *A Diss. Present. to Fac. Grad. Sch. Univ. South. Calif. Partial Fulfillment Requir. Degree MASTER Sci.*, no. December, p. 55, 2006.

- [147] J. J. Liaw, Y. F. Huang, C. H. Hsieh, D. C. Lin, and C. H. Luo, "PM2.5 concentration estimation based on image processing schemes and simple linear regression," *Sensors (Switzerland)*, vol. 20, no. 8, pp. 1–14, 2020, doi: 10.3390/s20082423.
- [148] B. Beigzadeh, M. Bahrami, M. J. Amiri, and M. R. Mahmoudi, "A new approach in adsorption modeling using random forest regression, Bayesian multiple linear regression, and multiple linear regression: 2,4-D adsorption by a green adsorbent," *Water Sci. Technol.*, vol. 82, no. 8, pp. 1586–1602, 2020, doi: 10.2166/wst.2020.440.
- [149] K. A. Marill, "Advanced Statistics: Linear Regression, Part II: Multiple Linear Regression," *Acad. Emerg. Med.*, vol. 11, no. 1, pp. 94–102, 2004, doi: 10.1197/j.aem.2003.09.006.
- [150] S. H. Brown, "Multiple Linear Regression Analysis: A Matrix Approach with MATLAB," *Alabama J. Math.*, no. Spring/Fall, pp. 1–3, 2009, [Online]. Available: <http://ajmonline.org/2009/brown.pdf>.
- [151] S. J. Kim, C. H. Kim, S. Y. Jung, and Y. J. Kim, "Shape Optimization of a Hybrid Magnetic Torque Converter Using the Multiple Linear Regression Analysis," *IEEE Trans. Magn.*, vol. 52, no. 3, pp. 3–6, 2016, doi: 10.1109/TMAG.2015.2482964.
- [152] M. Tranmer, J. Murphy, M. Elliot, and M. Pampaka, "Multiple Linear Regression (2nd Edition)," *Cathie Marsh Inst. Work. Pap.*, no. 01, p. 59, 2020, [Online]. Available: <https://hummedia.manchester.ac.uk/institutes/cmist/archive-publications/working-papers/2020/2020-1-multiple-linear-regression.pdf>.
- [153] P. Fink, "Ensemble methods for classification trees under imprecise probabilities," *Ludwig-Maximilians Univ.*, vol. 5, pp. 1–62, 2012.
- [154] J. E. Gentle, W. K. Härdle, and Y. Mori, *Handbook of computational statistics: Concepts and methods: Second Edition*, no. January 2012. 2012.
- [155] A. A. Kostin, A. V. Kopylov, V. V. Mottl, and I. Muchnik, "Dynamic Programming Procedures in Nonstationary Signal Analysis," *Pattern Recognit. Image Anal.*, vol. 11, no. 1, pp. 205–208, 2001, [Online]. Available: <http://nauka.maik.rssi.ru/cgi-perl/search.pl?type=abstract&name=patrec&number=1&year=1&page=205>.
- [156] A. A. Kostin, A. V. Kopylov, V. V. Mottl, and I. Muchnik, "Dynamic Programming Procedures in Nonstationary Signal Analysis," *Pattern Recognit. Image Anal.*, vol. 11, no. 1, pp. 205–208, 2001.

- [157] P. Kumar, "Slope Movement Prediction and Early Warnings via Machine Learning Algorithms and IoT Technologies," *Sch. Comput. Electr. Eng. Indian Inst. Technol. Mandi*, *This Diss. is Submitt. degree Master Sci. (by Res.*, no. September, p. 24, 2020.
- [158] M. Akyol and E. Uçar, "Carbon footprint forecasting using time series data mining methods: the case of Turkey," *Environ. Sci. Pollut. Res.*, 2021, doi: 10.1007/s11356-021-13431-6.
- [159] P. Mohammadi, A. Hatamlou, and M. Masdari, "A Comparative Study on Remote Tracking of Parkinson's Disease Progression Using Data Mining Methods," *Int. J. Found. Comput. Sci. Technol.*, vol. 3, no. 6, pp. 71–83, 2013, doi: 10.5121/ijfcst.2013.3607.
- [160] P. Singh and S. Agrawal, "Node localization in wireless sensor networks using the M5P tree and SMOreg algorithms," *Proc. - 5th Int. Conf. Comput. Intell. Commun. Networks, CICN 2013*, pp. 104–108, 2013, doi: 10.1109/CICN.2013.32.
- [161] L. Wang, L. Tan, C. Yu, and Z. Wu, "Study and application of non-linear time series prediction in ground source heat pump system," *2012 2nd Int. Conf. Consum. Electron. Commun. Networks, CECNet 2012 - Proc.*, no. 19005114009, pp. 3522–3525, 2012, doi: 10.1109/CECNet.2012.6201751.
- [162] S. Schaal, C. G. Atkeson, and S. Vijayakumar, "Real-time robot learning with locally weighted statistical learning," *Proc. - IEEE Int. Conf. Robot. Autom.*, vol. 1, no. April, pp. 288–293, 2000, doi: 10.1109/robot.2000.844072.
- [163] A. Chellam, L. Ramanathan, and S. Ramani, "Intrusion Detection in Computer Networks using Lazy Learning Algorithm," *Procedia Comput. Sci.*, vol. 132, pp. 928–936, 2018, doi: 10.1016/j.procs.2018.05.108.
- [164] R. M. Adnan, A. Jaafari, A. Mohanavelu, O. Kisi, and A. Elbeltagi, "Novel Ensemble Forecasting of Streamflow Using Locally Weighted Learning Algorithm," 2021.

EKLER

