



T. C.

ORDU ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

**PYTHON İLE MAKİNE ÖĞRENME Sİ ALGORİTMALARINI
KULLANARAK RÜZGAR SANTRALİNİN ELEKTRİK
ÜRETİM TAHMİNİ**

MUHSİN BAKDEMİR

YÜKSEK LİSANS

YENİLENEBİLİR ENERJİ

ORDU 2024

TEZ BİLDİRİMİ

Tez yazım kurallarına uygun olarak hazırlanan ve kullanılan intihal tespit programının sonuçlarına göre; bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, tezin içerdiği yenilik ve sonuçların başka bir yerden alınmadığını, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

MUHSİN BAKDEMİR

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı, 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

ÖZET

PYTHON İLE MAKİNE ÖĞRENMESİ ALGORİTMALARINI KULLANARAK RÜZGAR SANTRALİNİN ELEKTRİK ÜRETİM TAHMİNİ

MUHSİN BAKDEMİR

ORDU ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

YENİLENEBİLİR ENERJİ

YÜKSEK LİSANS TEZİ, 61 SAYFA

(TEZ DANIŞMANI: DOÇ. DR. SİBEL AKKAYA OY)

Dünyadaki enerji ihtiyacının hızlı artışı ile birlikte tükenen kaynakların da enerji talebini karşılamada yetersiz kalması, ülkelerin özellikle yenilenebilir enerji kaynaklarına ilgilerinin artmasına yol açmaktadır. Yenilenebilir enerji kaynakları içerisinde rüzgâr enerjisi bu aşamada ön plana çıkmaktadır. Rüzgâr santralleri kurulum alanı, bakım ve süreklilik açısından kolaylık sunması, bölgesel olarak rüzgâra elverişli olmasıyla birleştiğinde elektrik üretimi için yeni fırsatlar sunmaktadır. Son yıllarda Türkiye’de gerek coğrafi olarak gerek teknolojik gelişmeler açısından kurulumundaki hızlı artış ile beraber rüzgâr santralleri öne çıkmaktadır.

Bu tezde, Türkiye’deki Ordu ili Akkuş ilçesindeki Rüzgâr Enerji Santrali’ne ait 392712 satır ve 27 sütundan oluşan, 16.10.2020 ile 16.04.2023 tarihleri arasındaki 10 ‘ar dakika aralıklarla ölçülen değerlerin yer aldığı ve üretilen aktif güç, reaktif güç, rüzgâr hızı, nem, kanat pozisyonları, hava basıncı vb. sütun değerlerini bulunduran bir veri seti kullanılmıştır. Veri seti, veri madenciliği metotları kullanılarak veri ön işleme işlemi ve aykırı (uç) veri analizi yapılarak makine öğrenmesi için hazır hale getirilmiştir. Makine öğrenmesi modelleri uygulanmadan önce model çerçevesi belirlenmiş ve veri setindeki değerler normalizasyon yöntemiyle 0 ile 1 arasında değerlere ölçeklendirilmiştir. Bu sayede python programlayıcının makine öğrenmesi modellerini hızlı ve daha kolay uygulaması sağlanmıştır. Verilerin %80’i train (öğrenme) ve %20’si ise test verisi olarak ayrılmıştır. Veri seti lineer olmayan karmaşık ve sayıca çok verilerden oluştuğu için makine öğrenmesi modellerinden Simple RNN, GRU ve LSTM modelleri uygulanarak sonuçlar karşılaştırılmıştır. Tahmin verileri gerçek verilerle karşılaştırılmış ve hata oranları RMSE ve MAE hata oranı belirleme yöntemleri kullanılarak hesaplanmıştır. Sonuçlar karşılaştırıldığında Simple RNN modelinin daha gerçeğe yakın sonuç verdiği saptanmıştır.

Bu tez çalışmasıyla, bir rüzgâr enerji santralinin elektrik üretiminin Python programlayıcı ve kütüphaneleri yardımıyla makine öğrenmesi metotları kullanılarak gerçeğe yakın olarak tahmin edilebildiği gösterilmiştir. Bu sayede günlük, aylık ve yıllık elektrik üretim planlamasının yapılmasında kolaylık sağlanabilmektedir.

Anahtar Kelimeler: Makine öğrenmesi, Python Programlama, Rüzgâr Enerji Santrali, Yenilenebilir Enerji, Veri Madenciliği, Veri Ön işleme.

ABSTRACT

ELECTRICITY GENERATION PREDICTION OF WIND FARM USING MACHINE LEARNING ALGORITHMS WITH PYTHON

MUHSİN BAKDEMİR

ORDU UNIVERSITY INSTITUTE OF NATURAL AND APPLIED
SCIENCES

RENEWABLE ENERGY DEPARTMENT

MASTER THESIS, 61 PAGES

(SUPERVISOR: DOÇ. DR. SİBEL AKKAYA OY)

The rapid increase in the world's energy needs and the inadequacy of depleted resources to meet the energy demand cause countries to increase their interest in renewable energy sources. Among renewable energy sources, wind energy comes to the fore at this stage. Its ease of installation area, maintenance and continuity, combined with its regional suitability for wind, offers new opportunities for electricity generation. In recent years, wind power plants have come to the fore in Turkey with the rapid increase in their installation, both geographically and in terms of technological developments.

In this thesis, the values measured at 10-minute intervals between 16.10.2020 and 16.04.2023, consisting of 392712 rows and 27 columns, belonging to the Wind Power Plant in Akkuş district of Ordu province in Turkey, are included and the produced active power, reactive power A data set containing column values such as, wind speed, humidity, wing positions, air pressure, etc. was used. The data set was made ready for machine learning by performing data preprocessing and outlier data analysis using data mining methods. Before applying machine learning models, the model framework was determined and the values in the data set were scaled to values between 0 and 1 using the normalization. In this way, the Python programmer is enabled to implement machine learning models faster and easier. 80% of the data is divided as train (learning) and 20% as test data. Since the data set consists of non-linear, complex and numerous data, the results were compared by applying Simple RNN, GRU and LSTM models, which are machine learning models. Prediction data were compared with real data and error rates were calculated using RMSE and MAE error methods. When the results were compared, it was determined that the Simple RNN model gave more realistic results.

With this thesis study, it has been shown that the electricity production of a Wind Power Plant can be predicted realistically by using machine learning methods with the help of Python programmer and libraries. In this way, daily, monthly and annual electricity production planning can be made easier.

Keywords: Machine learning, Python Programming, Wind Power Plant, Renewable Energy, Data Mining, Data Preprocessing.

TEŐEKKÖR

Bu yűksek lisans tez alıŐmasında tez konunun belirlenmesi, alıŐmanın yűrűtűlmesi ve yazımı esnasında bilgi ve tecrűbesini paylaŐarak yol gűsteren baŐta danıŐman hocam Sayın Do. Dr. Sibel AKKAYA OY'a ve AkkuŐ Rűzgâr Enerji Santrali (RES) verilerin kullanımı esirgemeyen Sayın Cemalettin AYDOĐDU'ya teŐekkűr ederim.

Aynı zamanda, manevi desteklerini her an űzerimde hissettiĐim ailem ve űzellikle eŐim Pakize BAKDEMİR'e teŐekkűrű bir bor bilirim.

İÇİNDEKİLER

	<u>Sayfa</u>
TEZ BİLDİRİMİ	I
ÖZET	II
ABSTRACT	III
TEŞEKKÜR	IV
İÇİNDEKİLER	V
ŞEKİL LİSTESİ	VI
ÇİZELGE LİSTESİ	VII
SİMGELER ve KISALTMALAR LİSTESİ	VIII
1. GİRİŞ	1
2. GENEL BİLGİLER	3
2.1 Yenilenebilir Enerji.....	3
2.1.1 Türkiye'nin Enerji Potansiyeli.....	8
2.1.2 Türkiye'nin Rüzgâr Enerjisi Potansiyeli	9
2.2 Literatür Özeti.....	10
2.2.1 Dünyadaki Çalışmalar	10
2.2.2 Türkiye'deki Çalışmalar	16
3. MATERYAL ve YÖNTEM	21
3.1 Makine Öğrenmesi.....	21
3.2 Makine Öğrenmesi Yöntemleri	22
3.2.1 Denetimli Öğrenme	22
3.2.2 Regresyon Analizi	27
3.2.3 Denetimsiz Öğrenme.....	32
3.2.4 Yarı Denetimli Öğrenme	34
3.2.5 Takviyeli Öğrenme.....	34
3.2.6 Pekiştirmeli Öğrenme.....	35
3.2.7 Özellik Öğrenme	35
3.3 Python Programlama ile Veri Analizi	36
3.3.1 Python Programlama ve Kütüphaneleri	36
3.3.2 Veri Analizi	39
3.4 Akkuş Rüzgâr Enerji Santrali Verilerinin Modellenmesi	41
3.4.1 Veri Ön işleme	42
3.4.2 Aykırı (Uç) Veri Analizi	43
3.4.3 Veri Tahmin Modellemesi.....	48
4. ARAŞTIRMA BULGULARI	51
5. TARTIŞMA ve SONUÇ	56
6. KAYNAKLAR	58
ÖZGEÇMİŞ	61

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 Türkiye’de 2023 Yılı Elektrik Üretim Dağılımı (Anonim, 2023a).	9
Şekil 2.2 Türkiye Rüzgâr Enerjisine Dayalı Kurulu Güç Grafiği (Anonim, 2023a)..	10
Şekil 3.1 Makine Öğrenmesi Döngüsü	21
Şekil 3.2 Makine Öğrenmesi Yöntemleri	22
Şekil 3.3 Destek Vektör Makineleri Karar Sınırlarının Belirlenmesi	25
Şekil 3.4 Bir LSTM Modelinin Yapısı	29
Şekil 3.5 Bir GRU Modelinin Yapısı	30
Şekil 3.6 Bir SimpleRNN Modelinin Yapısı.....	31
Şekil 3.7 Düzenlenmiş Akkuş Res Veri Seti.....	42
Şekil 3.8 IQR Yöntemi Görseli	43
Şekil 3.9 Wind (Rüzgâr) Sütunu Box-Plot Çizimi	44
Şekil 3.10 Airhumidity (Nem) Sütunu Box-Plot Çizimi	45
Şekil 3.11 Airpressure (Basınç) Sütunu Box-Plot Çizimi	45
Şekil 3.12 Illumination (Aydınlık) Sütunu Box-Plot Çizimi	46
Şekil 3.13 Akkuş Res Veri Setindeki Verilerin Grafiği.....	47
Şekil 3.14 Wind (Rüzgâr) - Power (Güç) Karşılaştırma Grafiği.....	47
Şekil 3.15 Airhumidity (Nem) - Power (Güç) Karşılaştırma Grafiği.....	48
Şekil 3.16 Airpressure (Basınç) – Power (Güç) Karşılaştırma Grafiği	48
Şekil 3.17 Normalize Edilmiş Akkuş Res Veri Seti Görseli.....	49
Şekil 3.18 Çerçevenilmiş Akkuş Res Veri Seti Görseli	49
Şekil 4.1 Simple RNN Train ve Test Veri Grafiği	51
Şekil 4.2 GRU-LSTM Train ve Test Veri Grafiği	52
Şekil 4.3 Simple RNN Model Özeti Görseli.....	52
Şekil 4.4 GRU-LSTM Model Özeti Görseli	53
Şekil 4.5 Simple RNN Modeline Ait Gerçek Veri-Tahmin Verisi Grafiği	54
Şekil 4.6 GRU-LSTM Modellerine Ait Gerçek Veri-Tahmin Verisi Grafiği.....	54
Şekil 4.7 Simple RNN Modeline Ait Gerçek Veri-Tahmin Verisi Regresyon Grafiği	55
Şekil 4.8 GRU-LSTM Modellerine Ait Gerçek Veri-Tahmin Verisi Regresyon Grafiği	55

ÇİZELGE LİSTESİ

Sayfa

Çizelge 4.1 Modellerin RMSE ve MAE Hata Oranlarının Karşılaştırması.....	53
--	----

SİMGELER ve KISALTMALAR LİSTESİ

%	: Yüzde
2D	: İki Boyut
A3C	: Asynchronous Actor-Critic Agents
ABD	: Amerika Birleşik Devletleri
ANN	: Yapay Sinir Ağları
ARIMA	: Otoregresif Entegre Hareketli Ortalama
BiLSTM	: Çift Yönlü Uzun Kısa Süreli Bellek
CNN	: Evrişimli Sinir Ağları
Csv	: Virgülle Ayrılmış Veri Seti Dosya Uzantısı
CV	: Çapraz Doğrulama
DBSCAN	: Gürültülü Uygulamada Yoğunluğa Dayalı Konumla Kümeleme
DELM	: Derin Aşırı Öğrenme Makinesi
DT	: Karar Ağaçları
DVM	: Destek Vektör Makinası
EM	: Beklenti Maksimizasyonu
GBT	: Gradyan Artırma Ağaçları
GMM	: Gauss Karışım Modelleri
GP	: Gauss Süreci
GPR	: Gauss Süreç Regresyonu
GRU	: Kapılı Tekrarlayan Birim
GW	: Gigawatt
HAC	: Hiyerarşik Kümeleme Analizi
HO	: Hold-Out
IEA	: Uluslararası Enerji Ajansı
IQR	: Aykırı Veri Analizi Yöntemi
ISO	: Kaliforniya Bağımsız Sistem Operatörü
IT	: İnternet Teknolojileri
kNN	: K-En Yakın Komşu
kWh	: Kilowattsaat
LASSO	: Least Absolute Shrink age and Selection Operator
LSTM	: Uzun Kısa Süreli Bellek
M	: Metre
MAE	: Ortalama Mutlak Hata
MCTS	: Monte-Carlo Tree Search
MDP	: Markov Karar Süreci
MKS	: Markov Karar Süreci
ML	: Makine Öğrenmesi
MOCSO	: Çapraz Optimizasyon
MSE	: Ortalama Karesel Hata
MVMD	: Çok Değişkenli Değişken Mod Ayrıştırma
MW	: Megawatt
NARNET	: Zaman Serileri Uygulama Yazılımı
NCA	: Komşuluk Bileşen Analizi
NMAE	: Normalize Ortalama Mutlak Hata
NN	: Sinir Ağları
NOAA	: Ulusal Okyanus ve Atmosfer İdaresi

NRMSE	: Normalize Ortalama Karesel Hata
NumPy	: Python Programlama Kütüphanesi
O2	: Oksijen
PCA	: Principal Component Analysis
Q1	: Aykırı Veri Analizinde 1.Çeyreklik
Q3	: Aykırı Veri Analizinde 3.Çeyreklik
R2	: Determinasyon Katsayısı
REPA	: Rüzgâr Enerji Potansiyel Atlası
RES	: Rüzgâr Enerji Santrali
RF	: Rastgele Orman
RMSE	: Ortalama Karesel Hata
RNN	: Tekrarlayan Sinir Ağları
RT	: Regresyon Ağaçları
SCADA	: Gözetleyici Kontrol ve Veri Toplama Sistemi
SciKit-Learn	: Python Programlama Kütüphanesi
SciPy	: Python Programlama Kütüphanesi
SVR	: Destek Vektör Makinesi
TD	: Temporal Difference
t-SNE	: t-Distributed Stochastic Neighbor Embedding
TÜBİTAK	: Türkiye Bilimsel ve Teknik Araştırma Kurumu
TWh	: Terawattsaat
WSPC	: Rüzgâr ve Güneş Enerji Kesintileri
xGBoost	: Extreme Gradient Boosting
YSA	: Yapay Sinir Ağları

1. GİRİŞ

20'inci yüzyıl itibariyle büyük bir gelişme gösteren sanayi sektörü ve bunun sonucunda ihtiyaç duyulan makinalaşma hergün artan enerji talebini ortaya çıkarmıştır. 1970 yılından sonraki on yılda petrolün varil fiyatı 2,5 dolardan 30 dolara kadar çıkmış ve enerji açığını karşılamada baskı unsuru olarak kullanılmaya başlanan petrolün değeri artmıştır. 1970'li yıllardaki bu petrol kriziyle tüm ülkeler enerjide bağımsızlık politikalarını geliştirmeye başlamış, böylece enerji kaynaklarındaki çeşitliliği artırmanın ve özellikle yenilenebilir enerji kaynaklarına önem vermenin gerekliliği ortaya çıkmıştır (Anonim, 2023b).

Yenilenebilir enerjiye olan ilginin artmasına karşın bu alandaki dezavantaj oluşturan hususların da giderilmesi ihtiyacı ortaya çıkmıştır. Yenilenebilir enerji kaynakları, hava durumu, iklim şartları ve mevsimsel değişikliklere bağlı olarak dalgalanmalar gösterebilmekte, sürekli ve sabit bir enerji üretimi sağlamamakta ve elde edilen enerjinin depolanması zorlu bir süreç olduğundan anlık olarak üretilen enerjinin kullanıma sunulması gerekmektedir. Bu bağlamda, yenilenebilir enerjiyle birlikte makine öğrenmesi tahmin modelleri ve python yazılımı modern dünyanın ihtiyaçlarına uygun olarak ön plana çıkmaktadır.

Makine öğrenmesi tahmin modelleri bu tez çalışmasında yenilenebilir enerji çeşitlerinden olan rüzgâr enerjisi tahmini için önerilmiştir. Rüzgâr enerjisi, rüzgârın yoğun olduğu bölgelerde kinetik enerjiyi elektrik enerjisine dönüştürerek, çevreye zarar vermeden enerji üretme potansiyeline sahiptir. Rüzgâr enerjisinin çevre dostu olmasının yanı sıra, sürdürülebilir bir enerji kaynağı olması da önemli bir avantajdır. Rüzgâr enerjisi, tükenmez bir kaynak olan rüzgârın sürekli olarak var olması sayesinde uzun vadeli enerji üretimini destekler. Ayrıca, enerji güvenliğini artırarak enerji arzının çeşitlenmesine katkı sağlar. Gelişen teknoloji ile birlikte rüzgâr enerjisi sektörü de sürekli olarak yenilikçi çözümlerle kendini güçlendirmektedir. Rüzgâr türbinlerinin verimliliği artırırken maliyetleri düşürmesiyle birlikte rüzgâr enerjisi, ekonomik bir seçenek haline gelmiştir. Ayrıca, rüzgâr enerjisi projeleri dünya genelinde birçok ülkede yaygınlaşarak enerji ihtiyacına sürdürülebilir çözümler sunmaktadır.

Günümüzde enerji sektöründe hızla ilerleyen teknoloji, rüzgâr enerjisi üretiminde de önemli değişikliklere yol açmaktadır. Bu değişimlerin merkezinde,

Python programlama dili ve makine öğrenmesi algoritmalarıyla desteklenen gelişmiş tahmin modelleri yer almaktadır. Rüzgâr gücü tahmini, enerji üreticilerine, dağıtım şirketlerine ve enerji piyasasındaki diğer paydaşlara daha kesin ve etkili bir enerji yönetimi sağlama potansiyeli taşımaktadır.

Python, kolay öğrenilebilir yapısı, geniş kütüphane desteği ve açık kaynak olması nedeniyle birçok veri bilimi ve makine öğrenmesi uygulamasında tercih edilen bir dil haline gelmiştir. Rüzgâr enerjisi tahmin modellerinin geliştirilmesinde, Python'un esnekliği ve veri işleme yetenekleri büyük avantaj sağlamaktadır.

Makine öğrenmesi algoritmaları, rüzgâr hızı ve yönü, sıcaklık ve diğer meteorolojik verileri kullanarak gelecekteki rüzgâr gücünü tahmin etme konusunda önemli bir kolaylık sağlamaktadır. Bu algoritmalar, büyük veri setleri üzerinde eğitilerek, rüzgâr enerjisi tesislerinin üretim kapasitelerini optimize etmek ve enerji şebekelerindeki dalgalanmaları minimize etmek amacıyla kullanılır.

Rüzgâr gücü tahmini, enerji üreticilerine daha etkili bir enerji planlaması yapma, bakım süreçlerini optimize etme ve enerji verimliliğini artırma fırsatı sunar. Aynı zamanda, enerji dağıtım şirketleri için de güç kaynaklarını daha etkili bir şekilde yönetme imkanı sağlar. Böylece enerji şebekelerinin kararlılığını artırarak enerji tüketici ve üreticileri arasındaki denge korunabilmektedir.

Bu tezde, Python ve makine öğrenmesi tekniklerinin rüzgâr gücü tahmini alanındaki uygulaması Ordu ili Akkuş ilçesindeki RES'e ait veri seti ile modellenmiştir. Veri setinin modele uygun hale getirilmesinde veri analiz yöntemleri ve veri madenciliği uygulamaları yapılmıştır. Tahmin modelinin uygulanmasında ve zaman serisi analizlerinde ise makine öğrenmesi yöntemlerinden Simple RNN, LSTM ve GRU metotları kullanılmıştır. Modelleme sonucunda etkili bir tahmin uygulaması yapılmıştır. Rüzgâr enerjisi sektöründeki bu tahmin çalışmasıyla, daha öngörülebilir, daha sürdürülebilir ve verimli enerji üretiminin uygulanabilirliği gösterilmiştir.

2. GENEL BİLGİLER

2.1 Yenilenebilir Enerji

Yenilenebilir enerji, doğada kendiliğinden bulunan ve tükenmeyen kaynaklardan elde edilen enerji olarak tanımlanabilir. Bu enerji, diğer fosil kaynaklı enerji türleriyle karşılaştırıldığında sürekli olarak kendini doğal yollarla yenileyebilmesi ve hiç tükenmemesi bakımından öne çıkmaktadır. Ayrıca çevre kirliliği açısından oldukça az olumsuz etkisinin olması, karbon salınımını minimum seviyelere indirmesi, ülkeler açısından ithalata gereksinim duyulmaması, enerjide tam bir bağımsızlık sağlamaya katkı sunması gibi avantajlarından dolayı enerji yapılanmasında ilk sıralarda yer almaktadır.

Yenilenebilir enerji kaynaklarına bakıldığında özellikle hidrolik, güneş, rüzgâr, biyokütle, jeotermal, dalga vb. kaynaklar ön plana çıkmaktadır. Tüm bu kaynakların temelinde ise güneş enerjisi bulunmaktadır. Tükenen kaynaklar olarak bilinen fosil yakıtlar da güneş enerjisi sayesinde meydana gelmektedirler. Böylece güneş enerjisini tüm enerjiler bakımından en önemli enerji kaynağı olarak tanımlayabiliriz.

Hidrolik Enerji: Su kaynaklarının kullanılarak enerjiye dönüştürülmesindeki çalışmalar hidrolik enerjinin konusunu oluşturmaktadır. Su baraj rezervuarlarında depolanarak türbinler vasıtasıyla generatörlere hareket enerjisi sağlamakta ve generatörler bu sayede hareket enerjisini elektrik enerjisine dönüştürmektedir. Hidrolik enerji kurulum maliyetleri açısından avantajlı olmasından dolayı tüm ülkeler tarafından yaygın olarak kullanılmaktadır. Kaynak bakımından bol olan ülkelerde daha çok hidrolik santral bulunmaktadır. Hidrolik enerji kaynakları bulunduğu bölgenin iklimsel özelliklerine etki etse de özellikle çevre kirliliği bakımından diğer fosil yakıtlı kaynaklara göre daha tercih edilebilir kaynaklar olarak öne çıkmaktadır.

Özellikle 2015 yılında dünyadaki hidrolik enerji kapasitesinde meydana gelen artış ile birlikte, dünyadaki hidrolik enerji kullanım seviyesi yaklaşık 1.000 GW'ı geçmiştir. Böylelikle dünyadaki elektrik enerjisinin yaklaşık olarak %16'sı hidrolik enerjiden karşılanmıştır. Ülkeler bazında bakıldığında ise özellikle Çin Halk Cumhuriyeti toplam enerji kapasitesi ve hidrolik enerjideki artış açısından öne çıkmaktadır (Karagöl & Kavaz, 2017).

Hidrolik enerji kaynağı olarak su kaynakları kullanıldığından bu enerjiyi yoğun olarak kullanan ülkeler, coğrafi bakımdan su kaynaklarına en çok hakim olan ülkelerdir. Bu durum ise coğrafi konumun önemini ortaya çıkarmaktadır. Coğrafi konumu bakımından su kaynaklarına ve engebeli arazi koşullarına en çok sahip olan ülkeler; Brezilya, Türkiye, Hindistan, Vietnam ve Malezya gibi ülkelerdir. Bu ülkeler hidrolik enerji bakımından uygun koşullara sahiptir.

Güneş Enerjisi: Diğer enerji kaynaklarının da temeli olan güneş enerjisi dünyada en çok kullanılabilen enerji kaynağıdır. Güneş enerjisinden enerji üretiminde en çok fotovoltaik paneller öne kullanılmaktadır. Fotovoltaik paneller güneş enerjisini elektrik enerjisine dönüştüren Silisyum ve Germanyum gibi yarı iletken malzemelerden üretilen malzemelerdir. Yeni geliştirilmekte olan bu malzemeler, maliyetinin yüksek olmasıyla dezavantaj oluşturmakta ve güneş enerjisinin kullanımını maliyetli kılmaktadır. Ancak ilerleyen teknolojik gelişmelerle birlikte bu dezavantajın da ortadan kalkacağı söylenebilir. Şuanki teknolojik ürünlerle güneş enerjisinin sadece %4'ü enerji olarak kullanılabilmektedir. Yatırımların ve gelişmiş teknolojik ürünlerin artmasıyla bu oranında da artacağı değerlendirilmektedir. Dünyada güneş enerjisi kullanımı özellikle 2014 yılından sonra ciddi bir büyüme göstermiştir. Yaklaşık %25'lik bir büyüme meydana gelmiştir. 2015 yılından sonra ise güneş enerjisinde tüm dünyada toplamda 227 GW'lık bir potansiyele ulaşılmıştır. Dünyada güneş enerjisi kurulu güç bakımından en çok Avrupa kıtasında, daha sonra ise Asya ve Kuzey Amerika ülkelerinde kullanılmıştır. Teknolojik ürün bakımından ise fotovoltaik sistem kapasitesi yüksek olan ülkeler sırasıyla Çin, Almanya, Japonya, ABD ve İtalya olmuştur. Fotovoltaik sistem kapasitesi nüfusa oranlandığında ise 2015 yılında Almanya ilk sırada yer almıştır. Tüm bunlardan anlaşıldığı üzere özellikle Çin Halk Cumhuriyeti güneş enerjisinden elektrik üretimi konusunda ilk sıralarda yer almıştır (Karagöl & Kavaz, 2017).

Güneş enerjisi kullanımı, yüksek maliyetine bakılmaksızın teknolojik gelişmeler ışığında yatırımların ve teşviklerin de artmasıyla hem dünyada hem de Türkiye'de giderek yaygınlaşmaktadır. Teknolojik gelişmelerin ilerlemesi ve maliyetlerin zamanla düşmesiyle birlikte güneş enerjisinin büyük enerji santrallerinde ve son kullanıcı düzeyinde kullanımını artıracak projelerle desteklenmelidir.

Rüzgâr Enerjisi: Rüzgâr enerjisi, rüzgârın kinetik enerjisini kullanarak elektrik üreten bir yenilenebilir enerji kaynağıdır. Rüzgâr türbinleri, rüzgârın döndürdüğü kanatlarıyla mekanik enerjiyi elektrik enerjisine dönüştürür. Rüzgâr türbinleri genellikle açık arazilere veya deniz sahillerine kurulur. Bu alanlarda sürekli olarak rüzgâr olması ve rüzgâr hızının yüksek olması tercih edilir. Rüzgâr türbinlerinin kanatları, rüzgârın etkisiyle dönmeye başlar. Kanatlar döndüğünde, türbinlerin jeneratörleri de harekete geçer ve mekanik enerji elektrik enerjisine dönüştürülür. Rüzgâr enerjisiyle elektrik üretimi çevreye zarar vermez. Rüzgâr enerjisi, fosil yakıtların kullanıldığı güç santrallerine kıyasla daha az karbondioksit emisyonu üretir. Rüzgâr enerjisi potansiyeli, rüzgârın sıklığı, hızı ve yoğunluğuna bağlıdır. Rüzgâr türbinlerinin optimal şekilde çalışabilmesi için uygun rüzgâr koşullarının olduğu bölgeler tercih edilir. Rüzgâr enerjisi, elektrik şebekesine entegre edilebilir. Rüzgâr enerjisi santralleri, ürettikleri elektriği şebekeye vererek bölge halkının elektrik ihtiyacını karşılayabilmektedir. Rüzgâr enerjisi, yenilenebilir ve sürdürülebilir bir enerji kaynağıdır. Daha fazla rüzgâr türbini kurularak ve teknolojik yeniliklerle birlikte kullanımı artmaktadır. Elektrik üretiminde çevreye duyarlı bir seçenek olarak rüzgâr enerjisi tercih edilmektedir.

Dünya ülkelerine bakıldığında, özellikle 2015 yılından sonra Danimarka ve Almanya rüzgâr enerjisi kullanımını en çok artıran ülkeler olarak öne çıkmaktadır. Uruguay, Portekiz, İrlanda ve İspanya'da ise onlardan sonra gelen ülkeler olmuşlardır. Rüzgâr enerjisini dünya pazarına sokan ülkeler olarak ise ABD ve Çin başta gelmektedir. Bu bakımdan rüzgâr enerjisi kullanımı giderek artmakta ancak rüzgâr bakımından coğrafi konum da oldukça öne çıkmaktadır. IEA verilerine göre, 2050 yılında dünyada kullanılan elektriğin yaklaşık %18'i rüzgâr enerjisinden temin edileceği öngörülmektedir (Karagöl & Kavaz, 2017).

Rüzgâr türbinleri, rüzgâr santrallerinin temel elemanını oluşturmaktadır. Hareketli havanın türbin kanatlarına çarpmasıyla kanatlarda meydana gelen mekanik enerji rotor ve dişliler yardımıyla generatöre ulaştırılır ve generatörde bu hareket enerjisi elektrik enerjisine dönüştürülür. Güneşin yeryüzüne ulaşan ışınları sayesinde, yeryüzünün eşit olmayan bir şekilde ısınması ve soğumasıyla meydana gelen basınç farkları sonucu hava hareketleri ortaya çıkarak rüzgâr meydana gelir. Böylece rüzgâr enerjisinin güneş enerjisinden dönüşmüş olduğu söylenebilir. Yeryüzünün coğrafi

faklılıkları ve farklı iklim özelliklerine baęlı olarak rüzgârın özellikleri de deęişmektedir. Rüzgâr enerjisi ifade edilirken hız ve yön olmak üzere iki özellik öne çıkar. Rüzgâr hızı, yükseklikle artar ve rüzgâr gücü ise hızın küpü ile orantılıdır. Yeryüzüne gelen güneş ışınlarının yaklaşık %2'si rüzgâr enerjisini meydana getirir.

Rüzgâr türbinleri çeşitli tiplerde bulunabilir. En yaygın olanları yatay eksenli ve dikey eksenli rüzgâr türbinleridir. Yatay eksenli rüzgâr türbinleri, döner kanatların yatay bir mil etrafında döndüğü türbinlerdir. Dikey eksenli rüzgâr türbinleri ise kanatların dikey bir mil etrafında döndüğü şekildedir. Türbinlerin eksenine, gücüne, devir ve kanat sayısına, dişli olup olmamasına ve kurulum yerlerine göre de gruplandırılabilirler. Tarihte kullanılan ilk türbinler, boyut olarak büyük ve gürültülü çalışırken günümüzdeki türbinler boyutları daha küçük ve daha sessiz çalışmaktadırlar. Rüzgâr enerjisinin daha yüksek potansiyelde kullanılabilmesi için türbin teknolojilerine yatırım yapılması gerekmektedir.

Rüzgâr türbininin parçaları; rotor, gövde, sensor, kule, kanat ve generatörden meydana gelmektedir. Rotor, rüzgâr enerjisini elektrik enerjisine dönüştürmek için kullanılan kanatlarla birlikte dönen elemandır. Rotor, rüzgârın enerjisini türbinin mekanik parçalarına iletmek için hava akımını yakalar. Göbek, rotoru türbin kulesine bağlar ve dönme hareketini ileterek türbinin diğer bileşenlerine enerji sağlar. Jeneratör rüzgâr türbininin mekanik enerjisini elektrik enerjisine dönüştürmeyi sağlar. Rotorun dönmesiyle birlikte manyetik alan oluşur ve elektrik üretimi gerçekleşir. Rüzgâr yönü sensörü, rüzgârın yönünü tespit eden bir sensördür. Bu sensördeki bilgi, türbinin rotorunun optimum açıda konumlandırılmasına yardımcı olmaktadır. Aero-dinamik kanatlar, rüzgârın enerjisini yakalamak için tasarlanmış kanatlardır. Aerodinamik şekilleri sayesinde rüzgâr kuvvetini maksimum düzeyde kullanarak türbini hareket ettirirler. Jeneratör kulesi, rüzgâr türbininin yüksek bir noktaya yerleştirildiği yapıdır. Kule, türbinin jeneratörünü ve rotorunu destekler ve rüzgâr türbininin yükseklik kazanmasını sağlar. Kontrol sistemi, rüzgâr türbininin çalışmasını izler ve düzenler. Rüzgârın hızına ve yönüne baęlı olarak türbinin rotorunun hızını ve kanat açısını ayarlar. Aynı zamanda güvenlik için türbinin durmaya geçme mekanizmasını da yönetir. Bu karmaşık sistem ve bileşenler bir araya gelerek rüzgâr enerjisini elektrik enerjisine dönüştüren rüzgâr türbininin çalışmasını sağlar.

Rüzgâr enerjisi, elektrik üretiminin yanında tarımsal üretim amacıyla da kullanılmaktadır. Özellikle tarım arazilerinin sulamasında kullanılan pompa sulama sistemleri için rüzgâr enerjisi avantaj sunmaktadır. Elektriğin ulaştırılmadığı tarım arazilerine kurulacak rüzgâr türbinleri vasıtasıyla ihtiyaç duyulan elektrik enerjisi sağlanabilmektedir. Türkiye’de Doğu ve Güneydoğu bölgeleriyle Ege bölgesinde birçok rüzgar tesisi bulunmaktadır. Sanayi tesislerindeki endüstriyel işletmelerin elektrik ihtiyacı da şebeke dışında yine rüzgâr türbinleri ile karşılanabilir. Rüzgâr enerjisi konutlarda sınırlı olarak kullanılsa da gelişen teknoloji ile birlikte bir hanenin ihtiyacını karşılayacak kadar enerji üretilebilecektir. Bahçe ve sokak aydınlatmalarında, soğutma sistemlerinde, lojistikte, batarya ve şarj sistemlerinde rüzgâr enerjisinden faydalanılmaktadır. Rüzgâr enerjisinin en önemli avantajı; tükenmeyen, yerli, temiz ve çevre dostu bir enerji kaynağı olmasıdır. Enerjide tam bağımsızlık adına yerli imkanların kullanılması açısından avantaj sağlar. İstihdama ve ekonomiye katkı sağlar. Kurulum ve bakım onarımı kolaydır. Rüzgâr enerjisinden elde edilen elektrik enerjisinde sürekliliğin sağlanabilmesi için diğer hidrolik ya da termik santraller ile birlikte entegre çalışması daha verimli olacaktır. Doğal yaşamdaki canlılar açısından özellikle kuşların kanatlara çarpmasıyla hayvan ölümleri meydana gelmesi ise rüzgâr türbinlerinin dezavantajlarından (Anonim, 2022).

Jeotermal Enerji: Yeryüzünün alt katmanlarında bulunan termal kaynaklar ve gaz yoğunluklarından kaynaklı ortaya çıkan ısı enerjisinden faydalanarak elde edilen enerjidir. Jeotermal enerji düşük maliyetli, çevreye dostu bir enerjidir. Coğrafi olarak yoğun olduğu yerlerde kurulumu daha uygundur. İklim ve hava koşullarından etkilenmediğinden diğer yenilenebilir kaynaklara göre daha avantajlıdır.

Jeotermal enerji kullanılarak elektrik üretimi için genellikle jeotermal kaynaklardan çıkan sıcak su veya buhar kullanılır. Bu su veya buhar, yüksek basınçlı buhar türbinleri kullanılarak elektrik enerjisine dönüştürülür. Elde edilen elektrik enerjisi, şebekeye verilebilir veya yerel alanlarda kullanılabilir. Isıtma amaçlı jeotermal enerji kullanımında ise sıcak su doğrudan kullanılır veya bir ısı eşanjörü aracılığıyla ısıtma sistemlerine transfer edilir. Bu yöntem ile binaların ısıtılması, sera tarımı, endüstriyel işlemler ve termal turizm gibi alanlarda enerji elde edilebilir. Jeotermal enerjinin yenilenebilir enerji kaynaklarından biri olarak değerlendirilmesinin sebebi sürekli olarak yenilenen doğal kaynaklara dayanmasıdır.

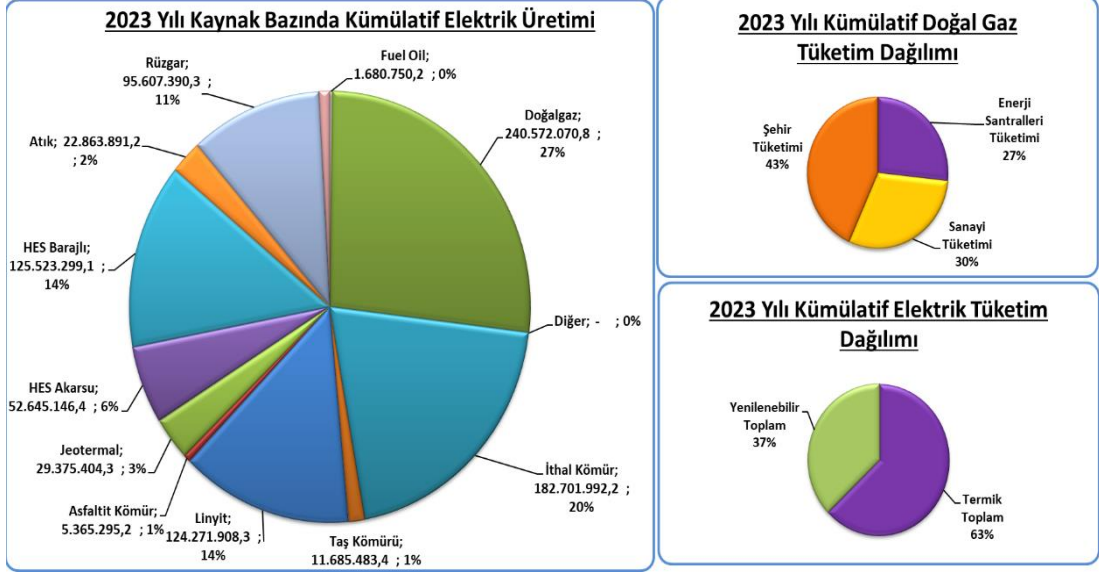
Bununla birlikte, jeotermal enerjinin kullanımını özellikle jeotermal kaynakların sınırlı olduğu bölgelerde yaygın değildir. Jeotermal enerjinin bazı çevresel etkileri olabildiğinden dikkatli bir şekilde kullanılmalısı gerekmektedir.

Dünyada jeotermal enerjiden elektrik üretimi, yaklaşık %0,04 seviyelerindedir. IEA'nın yaptığı tahminlere göre, 2050 yılında üretilecek olan toplam elektrik enerjisinin yaklaşık %3,5'inin jeotermal enerjiden üretileceği tahmin edilmektedir. Türkiye ise 2015 yılında ciddi bir kapasite artırımı yapmasına karşın yeterli seviyede değildir.

Biyokütle Enerjisi: Biyoenerji veya biyokütle enerjisi, organik malzemelerin (bitkiler, hayvan atıkları, atık malzemeler vb.) yanması veya fermantasyonu yoluyla elde edilen enerjiyi ifade eder. Bu enerji, elektrik üretimi, ısıtma, soğutma veya yakıt olarak kullanılabilir. Biyokütle enerjisi çeşitli şekillerde elde edilebilir. Örneğin, tarımsal ürünlerden, odun, saman veya bitki atıklarından elde edilen biyokütle, yakıt için kullanılabilir. Bunun yanı sıra, biyogaz üretimi de yaygın bir biyoenerji yöntemidir. Organik atıkların mikroorganizmalar tarafından ayrıştırılmasıyla meydana gelen biyogaz, ısınma ve elektrik üretimi için kullanılabilir. Biyokütle enerjisi, yenilenebilir bir enerji kaynağı olarak kabul edilir çünkü organik malzemeler doğal olarak yenilenebilir. Ayrıca, biyokütle enerjisi kullanılarak elde edilen yanma gazları, fosil yakıtların yanmasından kaynaklanan sera gazı emisyonlarının azaltılmasına da yardımcı olabilir. Bu da çevresel sürdürülebilirlik açısından önemli bir avantajdır. Verimi yüksek olup maliyeti düşüktür.

2.1.1 Türkiye'nin Enerji Potansiyeli

Enerji ve Tabii Kaynaklar Bakanlığı verilerine göre, 2022 yılında Türkiye'deki elektrik enerjisi tüketimi 2021 yılına oranla %1.25 azalarak 328.7 TWh olarak, elektrik üretimi ise %2.60 azalarak 326 TWh olarak gerçekleşmiştir. Geleceğe yönelik olarak 2025 yılında elektrik tüketiminin 380.2 TWh, 2030 yılında 455.3 TWh, 2035 yılında ise 510.5 TWh seviyelerine yükselmesi beklenmektedir. 2022 yılındaki bu elektrik üretiminin, %34.6'sı kömürden, %22.2'si doğal gazdan, %20.6'sı hidrolik enerjiden, %10.8'i rüzgâr enerjisinden, %4.7'si güneş enerjisinden, %3.3'ü jeotermal enerjiden ve %3.7'si diğer kaynaklardan elde edilmiştir (Anonim, 2023a).



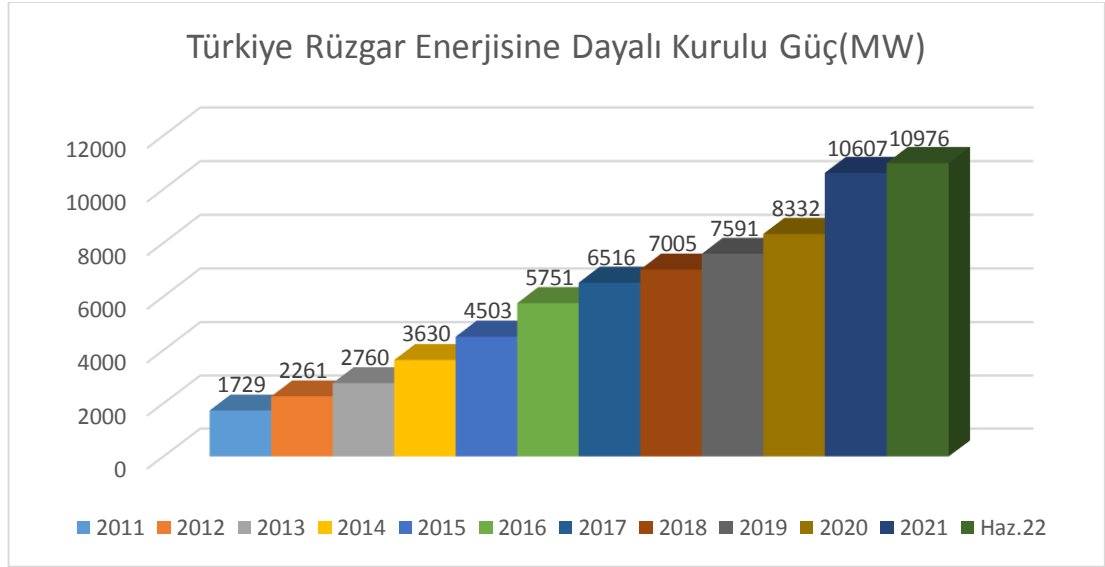
Şekil 2.1 Türkiye’de 2023 Yılı Elektrik Üretim Dağılımı (Anonim, 2023a).

Enerji ve Tabii Kaynaklar Bakanlığı verilerine dayalı olarak Şekil 2.1’de Türkiye’de 2023 Yılı Elektrik Üretim Dağılımı gösterilmiştir. Grafığe göre, Ekim 2023 sonu itibarıyla Türkiye elektrik kurulu gücü 105.964 MW olarak gerçekleşmiştir. Kurulu gücün kaynaklara göre dağılımı ise yaklaşık %20’si hidrolik enerji, %27’si doğal gaz, yaklaşık %36’sı kömür, %11’i rüzgâr, %10.5’i güneş, %3’ü jeotermal ve %2.5’i ise diğer kaynaklardan sağlanmıştır. Türkiye’deki elektrik enerjisi üretim santrali sayısına bakıldığında, Ekim 2023 sonu itibarıyla 12.868’e ulaşmıştır. Bu santrallerin 754’ü hidroelektrik, 68’i kömür, 363’ü rüzgâr, 63’ü jeotermal, 344’ü doğal gaz, 10.784’ü güneş, 492’si ise diğer enerji kaynaklarından üretim yapan santrallerdir (Anonim, 2023a).

2.1.2 Türkiye’nin Rüzgâr Enerjisi Potansiyeli

Türkiye’de kurulan ilk rüzgâr santrali 1998 yılında İzmir’de kurulmuştur. Ekim 2023 sonu itibarıyla ise RES sayısı 363’e ulaşmıştır. Türkiye’deki rüzgâr enerji santralleri kaynaklı toplam kurulu güç 10.930 MW’dır. Tüketilen enerjinin yaklaşık %7.1’i rüzgâr santrallerinden karşılanmaktadır. 363 adet RES’in bir kısmı henüz işaat halinde olduğundan lisans kurulu gücü kadar kurulu güce erişmemiştir. İnşaatı devam eden RES’lerin de devreye alınmasıyla mevcut kurulu güce 1.267 MW’lık kapasite ilave edilerek rüzgâr kurulu gücü 12.197 MW kapasiteye ulaşacaktır. Diğer ilavelerle birlikte Türkiye rüzgâr santrali kurulu gücünün toplamda 12.328 MW düzeyine çıkacağı görülmektedir (Anonim, 2023).

Enerji ve Tabii Kaynaklar Bakanlığı verilerine göre 2022 yılı Haziran ayı sonu itibariyle Türkiye'nin rüzgâr enerjisi kaynaklı elektrik kurulu gücü 10.976 MW, rüzgâr kurulu gücünün toplam kurulu güç içerisindeki oranı ise % 10.81 olarak gerçekleşmiş olup yıllara göre kurulu güç değişimi ve toplam kurulu güç içerisindeki oranı Şekil 2.2'de Türkiye Rüzgâr Enerjisine Dayalı Kurulu Güç grafiğinde yer almaktadır. Grafiğe göre özellikle 2021 yılında önemli bir kurulu güç artışı olduğu görülmektedir.



Şekil 2.2 Türkiye Rüzgâr Enerjisine Dayalı Kurulu Güç Grafiği (Anonim, 2023a).

2.2 Literatür Özeti

2.2.1 Dünyadaki Çalışmalar

Ma & Mei (2022) yaptıkları çalışmada, yapay sinir ağlarının birbirinden farklı şekillerini kullanarak hibrit bir derin öğrenme modeli üzerinde çalışmışlardır. Bu çalışmada hibrit model içerisindeki her bir sinir ağının tahmin yeteneğinin en ileri avantajından yararlanmak amaçlanmıştır. Çalışmada önerilen modelle, diğer tahmin yöntemlerine kıyasla zamanın etkilerini en aza indirerek daha verimli bir tahmin sunulmakta ve Evrişimli Sinir Ağı, Bi LSTM birleşiminden faydalanarak daha iyi bir tahmin sonucu elde edilmektedir. Çalışmada Türkiye'deki ve dünyadaki rüzgâr enerjisi veri setleri karşılaştırılarak daha gerçeğe uygun çıkarımlar yapılmıştır.

Meng ve ark., (2022) tahmin doğruluğundan önce tahmin kararlılığına vurgu yapmışlardır. Bu amaçla çapraz optimizasyona (MOCSO) dayalı bir hibrit model önermişlerdir. Özellikle rüzgâr hızındaki farklı frekans değerlerini minimize etmek için verileri düzenlerken rüzgâr gücünü, meridyen rüzgâr hızını ve bölgesel rüzgâr

hızını ayırt etmek için Çok Değişkenli Değişken Mod Ayırıştırma (MVMD) kullanılmıştır. Optimizasyon aşamasında, tahmin doğruluğunu ve kararlılığını sağlamak için MOCSO, Derin Aşırı Öğrenme Makinesi (DELM) optimizasyon için uygulanmıştır. Deneysel sonuçlar, MOCSO'nun son teknoloji ürünü çok amaçlı optimizasyon algoritmalarından daha iyi performans gösterdiğini ve önerilen hibrit modelin bu çalışmada yer alan diğer modellere göre önemli avantajlara sahip olduğunu göstermiştir. Önerilen MOCSO'nun etkinliği daha sonra doğrulama yapılarak diğer üç optimizasyon algoritmasıyla (yani MOALO, MOEA/D, MOPSO) karşılaştırılmış ve MOCSO'nun daha iyi arama yeteneğine sahip olduğu değerlendirilmiştir. Önerilen tahmin modeliyle gerçek rüzgâr gücü değerlerine yaklaşık bir tahmin yapılmıştır.

Datta (2018) yaptığı çalışmada, az sayıda veri ile etkin ve hızlı bir çalışma için bir yapay sinir ağı (YSA) ile saatlik rüzgâr hızını tahmin etmek için bir model geliştirmiştir. Çalışmada ana hatlarıyla verilen YSA modeli ve Kalıcılık Tahmin Modeli performansları karşılaştırılmıştır. Bu çalışmada Ulusal Okyanus ve Atmosfer İdaresi (NOAA) web sitesindeki meteorolojik veriler kullanılmıştır. Tek değişkenli zaman serileri (saatlik rüzgâr hızı) metodu kullanılarak NARNET yardımı ile gerekli tahmin modeli geliştirilmiştir. NARNET modelinin performansı, Ortalama Kareysel Hata (MSE) yöntemi ile analiz edilmiş ve sonuçlar Kalıcılık Modeli ile karşılaştırılmıştır. Karşılaştırmada, YSA modelinin saatlik rüzgâr hızını tahmin etmede Kalıcılık Modeli'nden daha iyi performans gösterdiği gözlemlenmiştir. Herhangi bir saatteki rüzgâr hızını tahmin etmek için yapılan analizlerde, geliştirilen model iyi sonuç göstermiştir.

Shams ve ark., (2021) çalışmada, rüzgâr ve güneş enerjisi kesintilerini (WSPC'ler) doğru bir şekilde tahmin etmek için yapay zekâ tabanlı makine öğrenme modelleri kullanmışlardır. Bu bağlamda, farklı türde makine öğrenimi yöntemleri kullanılmış ve hem Hold-Out (HO) hem de Çapraz Doğrulama (CV) yaklaşımlarına dayalı olarak sonuçlar değerlendirilmiştir. Regresyon Ağaçları (RT), Gradyan Artırma Ağaçları (GBT), Rastgele Orman (RF), İleri Beslemeli Yapay Sinir Ağları (YSA), Uzun Kısa Süreli Bellek (LSTM) ve Destek Vektör Makineleri (SVR) gibi makine öğrenme yöntemleri kullanılmıştır. Bu tahmin modelleri, yük talepleri, santrallerin çıkış gücü, güneş tarlaları, rüzgâr türbinleri, ve enerji kesintileri WSPC'den oluşan veri girdileri eğitilmiştir. Bu veri seti, California bağımsız sistem operatörünün (ISO)

saatlik kayıtlarına dayanarak oluşturulan verilerdir. Bu verilere dayanılarak modellerin doğrulaması yapılmış ve en iyi sonuçlar için Bayes optimizasyonu kullanılmıştır. Tüm modeller arasında, RF modeli, en az tahmin hata oranı ve dolayısıyla en iyi performansı sağlamıştır. Sonuçlar, önerilen modellerin WSPC'lerin tahmininde etkili olduğunu ortaya koymuştur. Model, yaygın olarak kullanılan HO yaklaşımıyla karşılaştırıldığında, tahmin performansını “güneş enerjisi azaltma” için %46 ve “rüzgâr azaltma” için %66'ya kadar iyileştirmiştir. HO yaklaşımını kullanarak, RF yöntemi, SVR ve ANN yöntemlerine kıyasla sırasıyla %6 ve %27 oranında iyi performans göstermiştir. LSTM yöntemi, RF yaklaşımından %7 daha az RMSE'ye sahip olmasına rağmen, yılın bazı saatlerinde negative WSPC olduğunu tahmin etmiştir. Giriş özellikleri ile yüksek korelasyonu nedeniyle, hedef değişken güneş kesintileri, rüzgâr kesintilerinden daha etkili bir şekilde tahmin edilmiş ve rüzgâr kısıtlaması, tüm yıl için değil, yılın çoğu için düzenli bir model izleyen güneş kesintisine kıyasla oldukça değişken olması nedeniyle zayıf bir tahmin ortaya koymuştur.

Lee ve ark., (2022) yaptıkları çalışmada, rüzgâr enerjisi tahmini yapmak için rüzgâr hızı, rüzgâr gücü vb. veri kümelerinden oluşan veri setini ve Python programını kullanmışlardır. Çalışma, topluluk araçlarının zaman serisi tahmininde etkili olduğunu, ancak genellikle daha düşük bir performansa sahip olduğunu göstermiştir. Veri kümesindeki veriler medyan mutlak hata ve etkili bir değerlendirme puanı örneği olan Peirce başarı puanı ile analiz edilmiştir. Çok organizasyonlu topluluk ortalamasının zaman serisi tahmininde yeterli başarıya sahip olduğunu ve topluluk üyelerine kıyasla güç tahmininde daha düşük performans gösterdiği bulunmuştur. Spektral analiz, topluluk ortalamasının birçok frekansta üyelerinin çoğundan daha iyi performans gösterdiğini ve böylece tüm tahmin periyodu için düşük zaman serisi tahmin hataları ürettiğini ileri sürmüştür. Bu çalışmanın bir sonraki aşamasında, olasılıksal tahminleri değerlendirmek ve sistematik bir prosedürle tahmin belirsizliğini ölçmek için WE-Validate'e yetenekler eklenmiştir. Rüzgâr tahmini iyileştirmek için WE-Validate'e topluluk katkısını ele almıştır.

Tubulekas (2022) çalışmasında, kısa süreli rüzgâr için çeşitli makine öğrenimi algoritmalarını kullanarak güç tahmini yapmıştır. Çalışmada kullanılan algoritmalar, Otoregresyon, LASSO yöntemiyle Regresyon (LR), Random Forrest Regresyon ve

Gradient Boost algoritmalarıdır. Çalışmada, Gradient Boosting modelinin incelenen algoritmalar arasında en iyi performansı gösteren model olduğu saptanmıştır. Bu çalışmada incelenen tüm makine öğrenmesi modelleri, yani LASSO yöntemiyle AR, LR, RFR ve GB, Referans modelden daha iyi performans gösterirken bu girdi-çıkı ilişkisini gerçekleminin ve daha iyi bir tahmin değeri üretmenin mümkün olduğunu göstermiştir. Birden çok giriş özelliğine sahip olan LASSO, RFR ve GB gibi giriş özelliklerini de hesaba katan modeller, yalnızca geçmiş rüzgâr gücü çıkış değerleri olan AR modelinden daha iyi performans göstermiştir. İki grup modeli, RFR ve GB, GB modelinin biraz daha iyi olduğu test veri setinde en iyi genel performansa sahip olmuştur. Her iki model de eğitim verilerini tahmin ederken test verilerine kıyasla RMSE açısından daha düşük hata oranına sahip olmuştur. Güç çıktısını tahmin etme söz konusu olduğunda en büyük etkiye sahip en önemli girdi özellikleri, LASSO ve GB modeli LR'den elde edilen özellik önem sonuçlarına göre rüzgâr hızı, rüzgâr yönü ve günün saati olmuştur.

Lee ve Baldick (2014) yaptıkları çalışmada, modern literatürdeki yeni yaklaşımlara ve gelişmiş tahmin tekniklere dayanan kısa süreli bir topluluk rüzgâr enerjisi tahmin modeli önermiştir. Modelin doğrulaması, bir buçuk yıl boyunca yedi rüzgâr çiftliğinde 48 saate kadar rüzgâr gücü tahmin edilerek yapılmıştır. Tahmin modelinde 52 Sinir Ağı (NN) alt modeli ve beş Gauss Süreci (GP) alt modeli kullanılmıştır. 48 saat boyunca, NN alt modelleri, geçmiş rüzgâr gücü verilerine ve tahmini rüzgâr bilgilerine dayalı olarak gelecekteki rüzgâr gücü tahmin edilmiştir. Paralel olarak, ilk beş saat için, NN alt modellerine doğru rüzgâr gücü tahminleri sağlamak için yalnızca geçmiş rüzgâr gücünü kullanarak rüzgâr gücünü tahmin etmek için beş GP alt modeli kullanılmıştır. Rüzgâr bilgisini eğitmek için NN alt modelleri kullanılmış ve NN alt modellerine başlangıç güç seviyesini vermek için GP alt modelleri kullanılmıştır. Bu çalışmadaki WPF modeli, rüzgâr gücünün faz kaymasını modelleyemez. Gerçek rüzgâr tahmin edilenden daha yavaş veya daha hızlı esiyorsa, tahmin edilen değerler kaydırılır. Bu nedenle tahmin hataları büyük olur. Rüzgâr gücünün ayırt edici özelliği, çeşitli örnekleme süreleri için rüzgâr gücü farkının Laplace dağılımını takip etmesidir. Bu özelliği GP'ye dahil etmek için, parametrelerden önce gelen, hatalar üzerindeki olasılık veya kovaryans fonksiyonu Laplace dağılımına dönüştürülebildiği saptanmıştır.

Farah ve ark., (2022) yaptıkları çalışmada, Almanya’da bir rüzgâr enerjisi tahmininde önemli ölçüde daha iyi tahmin doğruluğunu göz önünde bulunduran Kapılı Tekrarlayan Birim (GRU) ve Uzun Kısa Süreli Bellek (LSTM) modelleri olmak üzere iki tekrarlayan sinir ağları (RNN) modeli sunmaktadır. Eğitim için veri dört eğitim bölümüne ayrılarak sonuçları literatürdeki ARIMA ve SVR yaklaşımları ile karşılaştırmıştır. Sonuçlar, RNN-GRU modelinin yalnızca daha yüksek tahmin doğruluğu elde etmekle kalmayıp, aynı zamanda uzun diziler üzerinde daha hızlı bir öğrenme hızına sahip olduğunu ortaya koymuştur. İki yeni tekrarlayan sinir ağı RNN-GRU ile RNN-LSTM derin öğrenme modelleri kısa vadeli tahmin sağlamak için geliştirilmiştir. Her iki model de doğrusal olmayan ortamda iyi öğrenme ve genelleme yetenekleri ile hesaplama verimliliğini göstermiştir. Rüzgâr enerjisi zaman serisi veri kümelerine ek olarak, Almanya için saatlik veri kümesine dayalı olarak üç özellik alt kümesi ve dört vaka yürütülmüştür. Bu modeller, istatistiksel doğruluk ölçümleri MAE, RMSE ve belirleme katsayısı (R2) açısından verimli, güvenilir ve sağlam tahminler üretmiştir. Aynı veri seti için kısa vadeli rüzgâr enerjisi tahminleri için yayınlanan verilerle karşılaştırıldığında elde edilen sonuçlar, Doğrusal veya RBF çekirdekleri ile yapılandırılmış destek vektör makinesi (SVR) olan makine öğrenme yöntemiyle karşılaştırıldığında, geliştirilen RNN modelleri daha iyi tahminler üretmiştir. %70 eğitim %30 testten oluşan eğitim ve test alt kümeleri arasında dört farklı veri tahsisi ile en düşük hataları üretmiştir. RNN-GRU modeli, saatler öncesinden rüzgâr enerjisi tahmininde RNN-LSTM, SVR-RBF ve SVR-Linear modellerinden önemli ölçüde daha iyi performans göstermiştir. RNN-GRU modeli tarafından elde edilen sonuçlar, güneş, jeotermal ve diğer yenilenebilir enerji ile ilişkili olanlar gibi kısa vadeli döngüler boyunca bağımlı değişkende yüksek oynaklık gösteren diğer zaman serisi veri setleri ile değerlendirmeye değer olduğunu göstermektedir. Sunulan sonuçlar, RNN-GRU modelinin güneş, jeotermal ve diğer enerji üretimi ile ilişkili olanlar gibi kısa vadeli döngüler boyunca bağımlı değişkende yüksek oynaklık gösteren diğer zaman serisi veri setlerine uygulanma potansiyeline sahip olduğunu göstermiştir.

Wang ve ark., (2021) yaptıkları çalışmada, rüzgâr gücünün tahmini için, veri temizleme ve ön işleme dayalı kısa vadeli bir rüzgâr enerjisi tahmin yöntemi önermişlerdir. Rüzgâr yönü, rüzgâr hızı ve rüzgâr gücünden oluşan çok sayıda tarihsel

veri analiz edilmiştir. Her örneğin yerel yoğunluğunu hesaplayarak aykırı değerler etkin bir şekilde tespit edilmiştir. Orijinal verilerin özellikleri, tahmin modeli için yüksek kaliteli bir eğitim seti sağlayarak büyük ölçüde zenginleştirilmiştir. Kısa vadeli rüzgâr gücünü tahmin etmek için yeniden tasarlanmış bir evrimsel sinir ağı kullanılmış ve önerilen yöntemler, Çin'deki gerçek bir rüzgâr çiftliğinin veri kümesine dayalı olarak eğitilmiş ve test edilmiştir. Çok sayıda deneysel sonuç, önerilen yöntemlerin iyi performans gösterdiğini ve kısa vadeli rüzgâr enerjisi tahmininin doğruluğunu etkili bir şekilde geliştirdiğini göstermiştir. Kısa vadeli rüzgâr enerjisi tahmini için havuz katmanlarının kaldırılması modelin performansını iyileştirmiştir. Ayrıca, kısa süreli atalet nedeniyle, her andaki rüzgâr koşulları, bir önceki saatin geçmiş verileriyle güçlü bir korelasyona sahip olmuştur. Özetlemek gerekirse, önerilen model kısa vadeli rüzgâr enerjisi tahmininde, enerji arzı ve talep dağıtımına elverişli olan rüzgâr enerjisi ve derin öğrenme kombinasyonu için teorik bir temel sağlamaktadır.

Sæther (2021) yaptığı çalışmada, Kuzey Norveç'teki beş farklı rüzgâr çiftliği için istatistik ve makine öğrenimi yöntemlerini kullanarak rüzgâr gücü tahminini ortaya koymuştur. Rüzgâr gücünün kısa ve orta vadeli, çok adımlı tahmini için geleneksel istatistiksel modellerden dört farklı tahmin modeli geliştirilmiştir. Bu çalışmada, dört farklı makine öğrenimi modeli, kuzeydeki beş farklı rüzgâr santralinden elde edilen güç çıkışının tahmin edilmesi için çok kademeli zaman serisi uygulanmış ve karşılaştırılmıştır. Kuzey Norveç'teki beş farklı rüzgâr santralinden toplanan veriler ve Norveç Meteoroloji Enstitüsü'nden alınan hava tahmini verileri işlenmiş ve çok adımlı ve çok değişkenli geleceğin rüzgâr tahmini amacıyla yapılandırılmıştır. Uygulanan modeller, performansları, ilk olarak rüzgâr çiftliği konumlarının her biri için ve tahmin ufukları ve daha sonra tüm rüzgâr çiftliği konumları karşılaştırılarak değerlendirilmiştir. Kalıcılık modelinin, daha karmaşık makine öğrenme algoritmeleri ile karşılaştırıldığında, kısa tahmin ufklarında iyi performans göstermiştir. Ancak, tahmin ufku arttığında, RF gibi daha gelişmiş modeller ve SVR modelleri, basit kalıcılık modelinden ve arimax modelinden daha iyi sonuçlar göstermiştir. Arimax modeli, kalıcılık modeline, diğer uygulanan modellerin hiçbirine ilişkin olarak iyileştirmeler göstermemiştir. Proje boyunca yoldaki engeller, uygulanan sinir ağı modelleri ile deneylerin azaltılmasıyla sonuçlanmış ve lstm modeli

yalnızca konumlardan biri ve üç farklı tahmin ufku için test edilmiştir. Makine öğrenme modellerinden ikisi, yani RF ve SVR modelleri, yinelemeli olarak uygulandığında veya doğrudan hedeflenen tahmin ufku için tahminler yapıldığında performanslarına göre değerlendirilmiştir. Bu sonuçlardan, özyinelemeli yaklaşımın, gelecekteki güç çıkışını tahmin etmede doğrudan yaklaşımdan daha iyi olduğu görülmüştür. Bununla birlikte, özyinelemeli yaklaşımdan gelen tahminler, güç çıkışındaki olayların gerçekte önceden tahmin edildiği, güç çıkışının kendisinin değişen serileri gibi görülmüştür.

2.2.2 Türkiye’deki Çalışmalar

Çift (2023) Osmaniye Korkut Ata Üniversitesinde yaptığı tez çalışmasında, Osmaniye ili Bahçe ilçesinde bulunan Gökçedağ Rüzgâr Enerji Santrali’ne ait ölçüm verilerinden oluşan bir veri seti kullanmıştır. Bu veri setindeki verileri 8 farklı veri grubuna ayırarak veri analizini gerçekleştirmiştir. Daha sonra yapay sinir ağları(YSA) kullanarak rüzgâr santralının ürettiği güç miktarını tahmin etmeye çalışmıştır. Sonuçların karşılaştırmasını yaparken R-Kare, MAE, RMSE hata oranı tespit yöntemlerini kullanarak tahmin doğruluğunu belirlemiştir. Yapılan çalışma farklı lokasyona ait veri seti kullanması ve LSTM, GRU gibi farklı modeller kullanmamasından dolayı bu tez çalışmasıyla farklılık göstermektedir.

Karaali (2023) İzmir Dokuz Eylül Üniversitesinde yaptığı tez çalışmasında, İç Anadolu bölgesinde bir RES’e ait veri seti kullanmıştır. Veri seti Ocak 2019- Aralık 2020 arasındaki 10’ar dk arayla ölçülen SCADA verilerinden oluşmaktadır. Çalışmada, makine öğrenmesi yöntemlerinden k En Yakın Komşu (k-NN), Destek Vektör Makineleri (SVR), Rasgele Orman (RF), GBM ve xGBoost algoritmaları kullanılarak sonuçlar karşılaştırılmıştır. Farklı lokasyon verileri ve farklı makine öğrenmesi algoritması kullanması ile bu tez çalışmasından ayrılmaktadır.

Tekin (2022) Çukurova Üniversitesinde yaptığı tez çalışmasında, Çukurova bölgesine ait 100.000 veri içeren bir rüzgâr gücü veri seti kullanmıştır. Tahmin modelini uygularken Yapay Sinir Ağları (YSA), Anfis ve Bulanık Mantık (Fuzzy) tabanlı uygulamalar kullanmıştır. Modellerin uygulaması MATLAB programı ile yapılmış ve tahmin sonuçlarındaki hata oranları karşılaştırılarak modeller incelenmiştir. Çalışma sonucunda Bulanık Mantık (Fuzzy) tabanlı modelin en iyi

performası gösterdiği tespit edilmiştir. Farklı lokasyon ve farklı makine öğrenmesi modelleri kullanılmasıyla bu tez çalışmasından ayrılmaktadır. Ayrıca çalışmada kullanılan yazılım programı da farklılık göstermektedir.

Yoldaş (2022) İzmir Teknoloji Üniversitesinde yaptığı çalışmada, EPIAŞ sisteminden alınan bir veri seti kullanılarak kısa süreli rüzgâr hızı ve gücü tahmin edilmiştir. Çalışmada tahmin yapılırken ARIMA, LSTM ve GRU modelleri kullanılarak sonuçlar karşılaştırılmıştır. Hata oranları karşılaştırılırken R-Kare, MSE, RMSE ve MAE hata tespit yöntemleri kullanılmıştır. Tahmin işlemi Python yazılımı kullanılarak yapılmıştır. Zaman serisi analizi için ise Facebook Prophet yazılımı kullanılmıştır. Çalışma kullanılan yazılım ve LSTM, GRU algoritmaları kullanılmasıyla bu tez ile benzerlik göstermesine karşın lokasyon ve bazı diğer tahmin algoritmalarıyla farklılık göstermektedir.

Yağmur ve Yağmur (2022) yaptıkları çalışmada, Türkiye’de bir rüzgâr türbinine ait 2018 yılı boyunca ölçülen SCADA verilerinden oluşan bir veri seti kullanmışlardır. Veri seti aynı zamanda aynı konuma ait NASA sisteminden alınan meteorolojik verileri de içermektedir. Verilerin %75’i öğrenme, %25’i test için kullanılmıştır. Öznitelik seçiminde k-NN, Çoklu Doğrusal Regresyon, Rassal Orman, Karar Ağaçları gibi farklı makine öğrenmesi algoritmaları kullanılmıştır. Hata oranı karşılaştırmasında R-Kare değeri kullanılmıştır. Farklı makine öğrenmesi algoritmaları ve farklı lokasyon verileri girdi olarak kullanılmasıyla bu tez çalışmasından ayrılmaktadır.

Tekinay (2022) Ankara Üniversitesinde yaptığı tez çalışmasında, TEİAŞ’tan alınan 10 Ocak 2015- 23 Ocak 2022 tarihleri arasındaki ölçümleri içeren bir veri seti kullanılmıştır. Tahmin işleminde LSTM, BiLSTM, GRU ve Evrişimli Sinir Ağları (CNN) algoritmaları kullanılmıştır. Hata oranlarının belirlenmesinde R-Kare, MSE, RMSE ve MAE yöntemleri kullanılmıştır. Çalışma kullanılan makine öğrenmesi yöntemleriyle bu tez ile benzerlik göstermesine karşın farklı lokasyon verilerini girdi olarak kullanmıştır.

İnal (2022) Ordu Üniversitesinde yaptığı tez çalışmasında, Ordu ilindeki Darıca-2 hidroelektrik enerji santraline ait 2018 ile 2020 yılları arasındaki 3 yıllık verileri kullanılmıştır. Çalışmada rezervuara gelen giriş akımlarını, mevcut su

seviyesini ve hava tahminini giriş verisi olarak kullanılmış ve bir sonraki gün için enerji üretim kapasitesi tahmin edilmiştir. Çalışmada, her biri 10 nörona sahip iki gizli katmanlı bir sinir ağı algoritması kullanılmıştır. Daha sonra ise gün öncesi elektrik fiyat tahmini için TEİAŞ'ın yayınladığı Yük Tahmin Planlarını ve EPIAŞ'ın yayınladığı Piyasa Takas Fiyatı verileri kullanılmıştır. Piyasa fiyat tahmininde her biri 20 nörona sahip bir gizli katmanlı bir sinir ağı kullanılarak tahmin yapılmıştır. Çalışmada HES verilerinin girdi olarak kullanılması, MATLAB yazılımının kullanılması ve sadece yapay sinir ağı algoritması kullanılmasıyla bu tezle farklılık göstermektedir.

Özkay (2021) Bilecik Şeyh Edebali Üniversitesinde yaptığı tez çalışmasında, Bilecik Meteoroloji Müdürlüğüne ait saatlik rüzgâr hızı verilerini kullanmıştır. 10 yıl boyunca ölçülen saatlik veri seti kullanılmıştır. Verilerin %67'si öğrenme, %33'ü test için ayrılmıştır. Tahmin işleminde 3 katmanlı LSTM modeli kullanılmış ve LSTM'nin iki farklı türü karşılaştırılmıştır. Tahmin modellerinin uygulaması Python yazılımı kullanılarak yapılmıştır. Çalışma LSTM algoritması ve Python yazılı kullanılmasıyla bu tez ile benzer olsa da diğer öğrenme algoritmaları ve farklı lokasyon verilerini girdi olarak sunmasıyla ayrılmaktadır.

Yazıcı (2021) yaptığı tez çalışmasında, statik ve dinamik olmak üzere 2 model kullanmıştır. Statik modelde 1 yıla ait tüm mevsimleri içeren sabit bir veri seti kullanılırken dinamik modelde zamana bağlı olarak değişken yapıda bir veri seti kullanılmıştır. Statik modelde Destek Vektör Makinaları, Karar Ağaçları, Gauss Süreç Regresyonu ile Bayes optimizasyonu algoritmaları kullanılmıştır. Dinamik modelde ise LSTM algoritması kullanılmıştır. Sonuçların değerlendirilmesinde hata oranlarının tespiti için RMSE, MAE ve R-Kare yöntemleri kullanılmıştır.

Çolaker ve ark., (2020) yaptığı çalışmada, Afyonkarahisar ili Dinar ilçesindeki rüzgâr enerji santrali verilerini içeren veri seti kullanmıştır. Veri setindeki verilerin %70'i öğrenme, %15'i doğrulama ve %15'i test için ayrılmıştır. Tahmin modeli olarak 12 nörondan oluşan gizli katmanlı bir Yapay Sinir Ağı (YSA) kullanılmıştır. Çalışmada modellerin uygulaması MATLAB yazılımı kullanılarak yapılmıştır. Çalışma kullanılan yazılım, öğrenme algoritması ve farklı lokasyon verisi kullanması yönüyle bu tez çalışmasından ayrılmaktadır.

Görgel ve Kavlak (2020) yaptıkları çalışmada, İzmir ili Urla ilçesindeki bir rüzgâr santraline ait 1 yıllık veri seti kullanılmıştır. Veri setinde 01.05.2017-31.05.2018 tarihleri arasındaki 10'ar dk arayla ölçülen rüzgâr gücü değerleri bulunmaktadır. Veri setindeki 300 veri öğrenme kalanı ise test için ayrılmıştır. Tahmin modelinde bu tez çalışmasıyla benzer olarak sinir ağları, LSTM ve GRU algoritmaları kullanılmıştır. LSTM modelinde 3 adet gizli katman kullanılmış ve 100 iterasyon işlemi 5 kez tekrarlanmıştır. Hata oranları MAE mutlak hata yöntemiyle belirlenmiştir. Farklı lokasyon verilerini girdi olarak kullanmasıyla bu tez çalışmasından ayrılmaktadır.

Demolli (2020) yaptığı çalışmada, Türkiye'deki 5 farklı lokasyon için meteorolojik ölçüm istasyonlarından alınan 5 yıl boyunca saatlik olarak ölçülen verilerden oluşan bir veri seti kullanılmıştır. Veri setindeki rüzgâr hızı ve güneş radyasyonu gibi veriler kullanılarak çeşitli makine öğrenmesi metodlarıyla kısa ve uzun dönem rüzgâr ve güneş güç tahminleri yapılmıştır. Ayrıca farklı konumlar için muhtemel yük talebine göre kurulması planlanan hibrit santraller için maliyet değeri göz önüne alınarak güç üretimi ve optimal boyut tahminleri yapılmıştır. Tahmin modellerinin uygulanmasında makine öğrenmesi algoritmalarından LASSO, k-NN, xGBoost, Rastgele Orman ve Destek Vektör Makineleri algoritmaları kullanılmıştır. Çalışmada kullanılan farklı lokasyon verileri ve makine öğrenmesi algoritmaları bu tezle ayrılmaktadır.

Varol (2019) yaptığı çalışmada, anlık değişken enerji talebinin hangi yenilenebilir enerji kaynağı tarafından ne zaman karşılanacağını belirlemek için önceden enerji üretim tahmin edilmesine yönelik atmosferik ölçümleri kullanarak santrale özgü ve saat bazlı enerji üretim tahmini yapılmıştır. Çalışmada Çoklu Doğrusal regresyon, Powell Optimizasyonu ve Markov Chain Monte Carlo algoritmaları kullanılarak tahmin sonuçları karşılaştırılmıştır. Çalışmayla yenilenebilir enerji sistemlerinin mevcut şebekeye entegrasyonunu kolaylaştıracağı ve akıllı şebekeyi daha yaygın hale getireceği değerlendirilmiştir. Çalışma kullanılan farklı makine öğrenmesi yöntemleri ile bu tezle farklılık göstermektedir.

Aksoy (2018) yaptığı çalışmada, Bursa ili Harmanlık ilçesindeki 52.8 MW'lık bir rüzgâr santraline ait verilerden oluşan bir veri seti kullanılmıştır. Çalışmada kısa

sürekli rüzgâr gücü üretim tahmini için Çoklu Doğrusal Regresyonu, Destek Vektör Regresyon, K-En Yakın Komşu Regresyonu ve Karar Ağacı Regresyonu gibi farklı makine öğrenmesi algoritmaları kullanılmıştır. Modelde kullanılan tahmin algoritmaları, Python yazılımı kullanılarak yapılmıştır. Veri setindeki veriler 01.09.2016 ve 31.08.2018 tarihleri arasında SCADA sisteminden ve rüzgâr santralindeki türbinlerdeki sensörlerden 10'ar dk arayla ölçülen değerlerden oluşmaktadır. Bu veriler saatlik değerlere dönüştürüldükten sonra modellerde kullanılmıştır. Tahmin modelinde bir türbin verisinde 2 analiz uygulanmıştır. Birinci analizde iki farklı girdiye sahip tahmin modellerinden elde edilen sonuçlar karşılaştırılmıştır. İlk analizde hem rüzgâr hızı hem de sıcaklık girdi olarak kullanılarak tahmin sonucu elde edilmiştir. İkinci analizde ise, sadece rüzgâr hızı girdi olarak kullanılarak sonuçlar elde edilmiştir. Hata oranlarının belirlenmesinde MSE, MAE, RMSE doğruluk oranları kullanılmıştır.

Cantürk (2018) yaptığı tez çalışmasında, Türkiye'deki bir rüzgâr santralinden elde edilen veri setini kullanmıştır. Tahmin modelinde makine öğrenmesi algoritmalarından Naïve Bayes algoritması kullanılmıştır. Modelde statik ve dinamik olmak üzere iki farklı sinir ağı kullanılmıştır. Bu iki farklı modelde hata oranları tespit edilerek karşılaştırma yapılmıştır. Modellemede farklı aktivasyon fonksiyonları için tahminler yapıldığında sonuçlar arasında benzerlik olduğu görülmüştür. Daha sonra modelin girdileri ile öğrenme ve test verilerinin oranları değiştirilerek sonuçlar gözlemlenmiştir. Çalışmada kullanılan öğrenme algoritmaları ve farklı lokasyon verileri kullanılmasıyla bu tezden ayrılmaktadır.

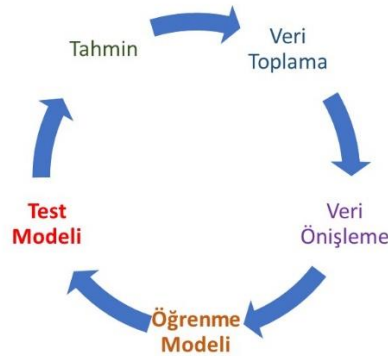
Kırbaş (2018) yaptığı çalışmada, TÜBİTAK T60 teleskopuna ait ortalama rüzgâr hızı verilerine ait veri setini kullanılarak rüzgâr hızı değerlerini geçmişe göre tahmin etmeye çalışılmıştır. Çalışmada, zaman serisi analizi için ARIMA modelleri ve NAR sinir ağı modelleri kullanılmıştır. 12 farklı model ile elde edilen sonuçlar istatistiksel ve grafiksel olarak karşılaştırılmıştır. Hata oranlarının belirlenmesi için MAE, MSE, RMSE yöntemleri kullanılmıştır. Çalışma farklı veri girdileri ve farklı öğrenme algoritmaları kullanması yönüyle bu tezle farklılık göstermektedir.

3. MATERYAL ve YÖNTEM

3.1 Makine Öğrenmesi

Makine Öğrenmesi, bir probleme yönelik olarak edinilen veri setlerinin düzenlenerek, istatistiksel metotlarla o problemle ilgili tahmin yapmaya yönelik kendi kendine davranış geliştiren algoritmalar ve matematiksel modellerin meydana getirilmesidir. Makine öğrenmesi, veri setleri ve algoritmalarından oluşan yapay zekanın bir alt kümesidir. Makine öğrenmesi, bilgisayar sistemlerinin veri analitiği yaparak kendi kendine öğrenme yeteneği kazanmasını sağlayan bir yapay zeka alt alanıdır. Bir makine öğrenmesi algoritması, belirli bir görevi veya problemi çözmek için veriye dayalı bir model oluşturur. Bu model, verilerdeki desenleri ve ilişkileri analiz ederek gelecekteki verileri tahmin etme veya kararlar verme yeteneği kazanır. Makine öğrenmesi algoritmaları, istatistiksel analiz metotlarıyla geçmiş veri setlerinin düzenlenmesi yoluyla öğrenmeyi ve sonuçları tahmin etmeyi sağlar. Makine öğrenmesi algoritmaları, gözetimli, gözetimsiz veya takviyeli öğrenme gibi farklı yaklaşımlar kullanabilir. Bu teknoloji, otomatik araçlar, öneri sistemleri, görüntü tanıma, doğal dil işleme gibi birçok alanda kullanılmaktadır.

Veri setlerinin eğitimi için kullanılan algoritmalar genel olarak ortak özelliklerine göre çeşitli öğrenme yöntemlerini kullanmaktadır. Bu öğrenme algoritmaları Hebb, Delta, Hopfield, Kohonen olmak üzere dört grupta toplanabilir.



Makine Öğrenmesi(ML) Döngüsü

Şekil 3.1 Makine Öğrenmesi Döngüsü

Bir makine öğrenmesi Şekil 3.1'de verilen döngüye göre öğrenmeyi gerçekleştirir. Buna göre, ilk olarak olaya ait veri toplanır, toplanan veri seti önleme

yapılarak öğrenmeye hazır hale getirilir, sonra model uygulanır ve uygulanan model test edilerek tahmin gerçekleştirilir.

3.2 Makine Öğrenmesi Yöntemleri



Makine Öğrenmesi(ML)

Şekil 3.2 Makine Öğrenmesi Yöntemleri

Makine öğrenmesi yöntemleri Şekil 3.2’de verilen yöntemlere göre çeşitlendirilir. Buna göre Makine Öğrenmesi; Denetimli Öğrenme, Denetimsiz Öğrenme, Yarı Denetimli Öğrenme, Takviyeli Öğrenme ve Özellik Öğrenme olarak çeşitlenmektedir.

3.2.1 Denetimli Öğrenme

Denetimli öğrenme, makine öğrenmesinin bir alt dalıdır. Bu yöntem, bir modelin belirli bir girdiye ilişkin doğru çıktıyı tahmin etmesine yardımcı olmak için etiketlenmiş eğitim verilerinin kullanıldığı bir öğrenme sürecini ifade eder. Denetimli öğrenme, veri setindeki örnekleri kullanarak bir modelin desenleri ve ilişkileri anlamasını sağlar ve daha sonra bu modelin yeni girişleri doğru bir şekilde tahmin etmesini bekler. Örneğin, bir evin maliyetini tahmin etmek için denetimli öğrenme kullanılabilir. Eğitim verilerinde evlerin özellikleri (odaların sayısı, konum gibi) ve gerçek maliyetleri etiket olarak bulunur. Bu verileri kullanarak bir model eğitilebilir ve daha sonra yeni bir evin özelliklerine dayanarak tahmini maliyeti hesaplaması istenebilir.

Denetimli öğrenme, makine öğrenmesindeki bir tekniktir ve bir modelin belirli bir girdiye dayalı olarak doğru çıktıyı tahmin etmesi için eğitildiği bir süreci ifade eder. Genellikle, veri setleri üzerinde çalışılır ve bu veri setlerinde girdi-veri çiftleri yani bir girdiye karşılık gelen doğru çıktı değeri bulunur.

Denetimli öğrenmede ilk adımda, kullanılacak olan veri seti toplanır. Bu veri seti, örneklerin girdi ve çıktı değerlerini içerir. Veri ön işleme aşamasında veri setindeki gürültüleri temizlemek, eksik verileri tamamlamak veya ölçeklemek gibi adımlarla veri ön işleme gerçekleştirilir. Model seçimi aşamasında denetimli bir öğrenme algoritması seçilir. Karar ağaçları, destek vektör makineleri (SVM), rastgele ormanlar, yapay sinir ağları gibi çeşitli modeller bulunmaktadır. Seçilen model, eğitim veri setindeki girdi ve çıktı örnekleri üzerinde eğitilir. Bu adım, modelin girişlere nasıl tepki vermesi gerektiğini öğrenmesini içerir. Eğitim veri setini kullanarak modelin performansı değerlendirilir. Bu aşamada, hata metrikleri kullanılarak modelin ne kadar doğru tahminler yaptığı belirlenir. Eğitilen model, gerçek dünya verilerinde kullanılarak girdilere dayalı olarak tahminler yapar. Bu tahminler, keşif, sınıflandırma, regresyon gibi çeşitli uygulamalarda kullanılabilir. Bu şekilde, denetimli bir öğrenme modeli, veri setlerindeki desenleri tanıyarak ve eğitim sürecinde öğrendiklerini kullanarak girdi değerlerine dayalı olarak çıktıları tahmin edebilir.

Denetimli öğrenme algoritmaları, veri setindeki özellikleri analiz eder ve bu özelliklerin hedef çıktıyla ilişkisini belirlemeye çalışır. Ardından, yeni bir giriş veri verildiğinde, algoritma bu ilişkiyi kullanarak çıktıyı tahmin eder. Bazı popüler denetimli öğrenme algoritmaları Lineer Regresyon, Lojistik Regresyon, Karar Ağaçları, K-En Yakın Komşu (KNN), Destek Vektör Makineleri (SVM), Yapay Sinir Ağlarıdır (YSA).

Karar Ağaçları: Karar ağaçları, makine öğrenmesi alanında sınıflandırma ve regresyon problemleri için yaygın olarak kullanılan bir algoritmadır. Bu algoritma, bir veri kümesine dayalı olarak bir karar ağacı oluşturur ve bu ağaç üzerinden tahminler yapar. Karar ağaçları, veri kümesindeki öznitelikleri değerlendirerek sınıflandırma veya regresyon için bir dizi karar kuralı üretir. Özniteliklerin değerleri kullanılarak veri kümesi içindeki örnekler birbirinden ayrılır ve tahminler yapılır. Karar ağaçlarının çıktısı, sınıflandırma problemleri için genellikle bir sınıf etiketi veya regresyon

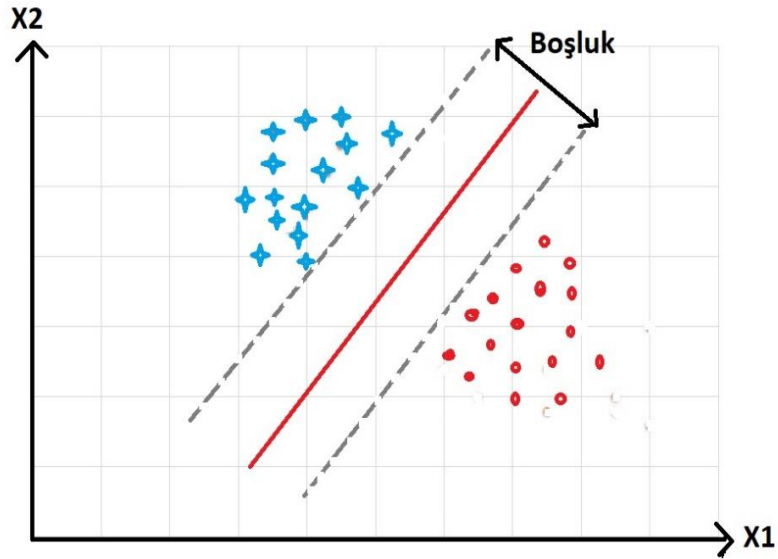
problemleri için bir sayısal değerdir. Karar ağaçları aynı zamanda sınıflandırma, regresyon ve özellik seçimi gibi işlemler için de kullanılabilir. Bu algoritma, sınıflandırma veya regresyon problemlerini çözmek için genellikle tercih edilen bir seçenektir. Çünkü; anlaşılması kolay, kullanımı basit ve yüksek doğruluk elde edebilen bir algoritmadır. Ayrıca, anlamlı özniteliklerle bağlantılar kurarak veri analizi yapma yeteneği sağlar.

Karar ağaçları, bir veri setini sınıflandırmak veya tahmin yapmak için kullanılabilir. Her bir düğüm, bir özellik değeriyle karşılaştırılarak verileri alt düğümlere böler. Bölme işlemi, verileri homojen alt gruplara ayırmak için en iyi özelliği ve eşik değeri seçmeye çalışır. Bu işlem, enformasyon kazancı veya hata azaltma gibi farklı ölçütler kullanılarak yapılabilir. Karar ağacı oluşturulduktan sonra, yeni bir veri noktası ağaçta ilerleyerek sonuç elde eder. Her bir düğüm, karar ağacında bir karar kriterini temsil eder ve alt düğümlere geçişi belirler. Bu şekilde, karar ağacı modele dayalı bir tahmin yapabilir veya sınıflandırma gerçekleştirebilir. Karar ağaçları algoritması, kolay yorumlanabilir olmaları ve çalışma prensiplerinin anlaşılabilir olması sebebiyle tercih edilir. Aynı zamanda çok yönlüdür ve birçok veri tipi ile kullanılabilir. Karar ağaçları genellikle toplu veri analizi, desizyon destek sistemleri, veri madenciliği ve doğal dil işleme gibi alanlarda kullanılır.

K-NN En Yakın Komşu: K-NN En Yakın Komşu (k-NN), bir makine öğrenimi algoritmasıdır. K-NN algoritması bir sınıflandırma veya regresyon problemi çözmek için kullanılır. Kullanıcının veri noktalarının belirli bir etiket veya değerle ilişkilendirildiği bir veri seti, algoritmanın temelini oluşturur. Algoritma, yeni bir veri noktasının sınıflandırılması veya tahmin edilmesi gerektiğinde kullanılır. Bu durumda, algoritma, k en yakın komşularının etiketlerini veya değerlerini analiz eder ve yeni veri noktasına en yakın olan k komşusunun etiketini veya değerini tahmin eder. K-NN algoritmasında "k" yakın komşular sayısını temsil eder ve kullanıcı tarafından belirlenen bir parametredir. K-NN algoritması, basit ve anlaşılır bir yapısı olan bir tahminleme yöntemidir. Ancak, algoritmanın veriye odaklanma ve hesaplama maliyeti yüksek olabilir. Ayrıca, uygun k değerini belirlemek de hata oranını etkileyebilir. Kısacası, k-NN algoritması, yeni bir veri noktasını belirli bir etiket veya değerle sınıflandırmak veya tahmin etmek için kullanılan basit bir makine öğrenimi algoritmasıdır. K-NN'nin diğer bir avantajı, modelin gerçek zamanlı olarak

eđitilebilmesidir. Veri setindeki yeni verileri anında kullanılarak tahmin yapılabilir. Ancak, k-NN algoritmasının belirli bir veri setinde yüksek hesaplama maliyetlerine sahip olabildiđi gibi veri noktası sayısı arttıkça performansı da azalabilmektedir.

Destek Vektör Makineleri: Destek vektör makineleri (SVM), makine öğrenmesi alanında sınıflandırma ve regresyon problemleri için kullanılan bir yöntemdir. SVM, veri noktalarını çok boyutlu uzayda temsil ederek farklı sınıflara ait verileri birbirinden ayırmayı hedefler. SVM, sınıflandırma problemleri için kullanıldığında, veri noktalarını farklı sınıflara bölmeyi amaçlayan bir "karar sınırı" veya "hiper düzlem" oluşturur. Bu hiper düzlem, farklı sınıflara ait veri noktalarını en iyi şekilde ayıran bir marjinal oluşturmayı hedefler. SVM'nin temel amacı, her iki sınıf arasında maksimum marjinal sınıflandırma yapmaktır. SVM'nin avantajları arasında veri boyutundan bağımsız olarak etkili çalışabilmesi, yüksek boyutlu verilerle başa çıkabilmesi ve sınıflandırma performansının iyi olabilmesi yer alır.



Şekil 3.3 Destek Vektör Makineleri Karar Sınırlarının Belirlenmesi

Şekil 3.3'de Destek Vektör Makineleri karar sınırlarının belirlenmesine ait grafik yer almaktadır. Grafikte yer alan x_1 ve x_2 sınıflarını karşılaştırılırken bu iki sınıfın özelliklerini ayırt edebilmek için en az hatayla bu özelliklerin ayrıldığı bölge tespit edilir. SVM ile bir düzlem üzerinde belirli bir mesafede tek bir doğru ile bu iki

özellik için ayırım yapılır. Grafikte, kırmızı çizgi verilerin yoğunlaştığı bölgenin ortasından geçen doğrudur. Mavi yıldızlar ve kırmızı daireler x_1 ve x_2 sınıflarını ayıran karar sınırları olarak belirlenir. Tireli çizgiler ise x_1 ve x_2 sınıflarının destek vektörleri arasında veri topluluğunu ayıran bölge sınır hattıdır. Veri bölgeleri sınırı üzerindeki verilerin kırmızı çizgiye uzaklığı hesaplanır. Veri olmayan boş kısımlar ne kadar iyi olursa, SVM o kadar iyi sonuç verir.

Naive Bayes: Naive Bayes algoritması, makine öğrenmesi ve istatistiksel sınıflandırma problemleri için kullanılan bir algoritmadır. Bu algoritma, Bayes teoremini temel alır ve sınıflandırma yaparken özelliklerin birbirinden bağımsız olduğunu varsayar. Naive Bayes algoritması, sınıflandırma problemlerinde sıklıkla kullanılır. Özellikle doğal dil işleme, spam filtreleme, duygu analizi gibi problemlerde tercih edilen bir algoritmadır. Algoritmanın çalışma prensibi, veri kümesinde bulunan özelliklerin sınıf etiketini tahmin etmek için kullanılmasıdır. Veri kümesi, etiketlenmiş örneklerden oluşur. Algoritma, bu örnekleri kullanarak her bir özelliğin sınıf etiketi üzerindeki etkisini tahmin eder. Naive Bayes algoritması, özellikleri ve sınıf etiketini belirlemek için bir olasılık modeli kullanır. Bu modeli oluşturmak için veri kümesindeki örneklerin özelliklerinin ve sınıf etiketlerinin dağılımlarını kullanır. Ardından, sınıf etiketini tahmin etmek için Bayes teoremini kullanır. Naive Bayes algoritması, çalışma prensibi ve basitliği nedeniyle tercih edilen bir sınıflandırma algoritmasıdır. Ancak, özellikler arasındaki bağımsızlık varsayımı gerçek dünyadaki bazı problemler için geçerli olmayabilir. Bu nedenle, veri setinin özellikleri bağımsız olmalıdır.

Rassal Orman Modeli: Rassal orman algoritması (Random Forest Algorithm), makine öğrenmesi ve veri analitiği alanında sıkça kullanılan bir algoritmadır. Bu algoritma, birçok karar ağacının bir araya getirilerek oluşturduğu bir algoritmadır. Rassal orman algoritması, sınıflandırma ve regresyon problemlerinin çözümünde etkili bir şekilde kullanılır. Algoritma, farklı özelliklere sahip bir veri setini kullanarak birden çok karar ağacı oluşturur. Her ağacın sonuçları toplanarak anlamlı bir sonuç üretilmesi hedeflenir. Her bir karar ağacı, rastgele örnekler ve rastgele özellikler kullanarak eğitilmektedir. Bu sayede her bir ağacın bağımsız bir öngörü yeteneği bulunur. Rassal orman algoritması, özelleştirme yeteneği sayesinde başarılı sonuçlar üretirken, aynı zamanda aşırı uydurma (overfitting) problemini de azaltır. Algoritma,

veri setinin özelliklerini analiz ederek hangi özelliklerin model için daha önemli olduğunu belirlemektedir. Bu özellik önemleri kullanılarak daha iyi bir tahmin yapabilen bir rassal orman modeli oluşturulur. Rassal orman algoritması genellikle sınıflandırma ve regresyon problemlerinde kullanılırken, özellikle büyük veri kümeleri ve yüksek boyutlu verilerle çalışmak için uygundur. Esnekliği, doğruluk oranı ve performansı sayesinde birçok uygulama alanında tercih edilen bir algoritmadır.

3.2.2 Regresyon Analizi

Regresyon analizi veri setindeki sayısal değişkenler arasındaki ilişkilerin belirlendiği bir algoritmadır. Veri kümesindeki verilerde ayrık bir özellik belirlenerek hedef çıktılar tahmin yapılmasına dayanır. Doğrusal ve Lojistik regresyon analizi çeşitleri mevcuttur (Diri, 2014).

Doğrusal Regresyon: Bir veri setindeki veri değişkenlerinin birbiriyle olan ilişkilerini inceleyerek benzerlik veya farklılıklarına göre ayıran istatistiksel bir yöntemdir. Veri setindeki verilerin durumuna göre bir ayırım çizgisi çizer. Bu sayede veri setindeki özelliklere göre çıkış verisi tahmin edilir (Kaysal, Akarslan, & Hocaoglu, 2022). Doğrusal regresyon, ekonomik büyüme, enflasyon, işsizlik gibi ekonomik göstergeler arasındaki ilişkileri anlamak ve tahmin etmek için kullanılır. Ayrıca finansal verilerde varlık fiyatları, faiz oranları ve benzeri faktörlerle ilgili analizleri yapmak, reklam harcamaları ile satışlar arasındaki ilişkileri anlamak, fiyatlandırma stratejilerini belirlemek ve müşteri davranışlarını tahmin etmek, hastalıkların risk faktörlerini belirlemek, tedavi yanıtlarını tahmin etmek ve sağlık sonuçlarını analiz etmek, öğrenci başarıları ile faktörler arasındaki ilişkileri anlamak, eğitim programlarının etkilerini değerlendirmek ve öğrenci performansını tahmin etmek, mühendislik alanında süreç kontrolü, kalite kontrolü yapmak, biyolojik verilerde gen ifadesi, ilaç etkileri ve diğer biyolojik süreçler arasındaki ilişkileri anlamak için kullanılır.

Lojistik Regresyon: Lojistik regresyon, istatistik ve makine öğrenimi alanlarında sıkça kullanılan bir sınıflandırma tekniğidir. Temelde, bir bağımlı değişkenin iki kategoriden birine ait olma olasılığını tahmin etmek için kullanılır. Bu iki kategori genellikle "0" ve "1" olarak kodlanır, yani bir olayın gerçekleşip gerçekleşmemesi durumlarını temsil eder. Lojistik regresyon, doğrusal regresyonun bir genişlemesi olarak düşünülebilir. Ancak, lojistik regresyon, doğrusal regresyondan farklı olarak,

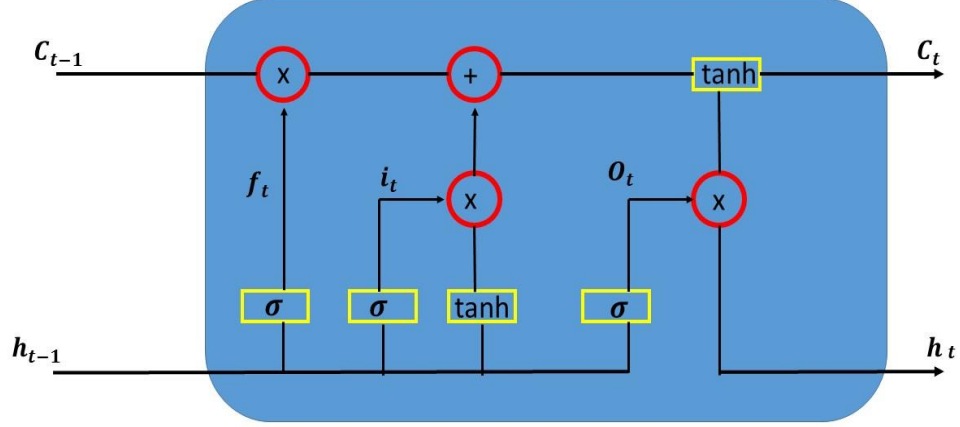
çıkıyı bir olasılık değeri (0 ile 1 arasında) ve ardından bu olasılık değerini belirli bir eşik değeriyle karşılaştırarak sınıflandırma yapar.

Lojistik regresyonun tıp, ekonomi, pazarlama ve birçok diğer alanlarda kullanımı yaygındır. Hastalıkların risk faktörlerini belirleme, hastalık teşhisi, tedavi yanıtlarını tahmin etme ve sağlık sonuçlarını değerlendirme gibi tıbbi uygulamalarda, müşteri davranışlarını analiz etme, müşteri segmentasyonu, kampanya etkileşimlerini tahmin etme gibi pazarlama stratejilerini belirleme amacıyla, işe alım süreçlerini değerlendirme, personel performansını tahmin etme ve iş gücü planlaması gibi konularda, kredi riski değerlendirmesi, finansal dolandırıcılık tespiti, varlık fiyat tahminleri ve ekonomik göstergelerin analizi gibi finansal uygulamalarda kullanılır.

Uzun Kısa Süreli Bellek (LSTM)

Uzun Kısa Süreli Bellek (LSTM), özellikle uzun süreli bağımlılıkları başarıyla ele alabilen bir tür tekrarlayan sinir ağı çeşididir. Geleneksel RNN'ler, uzun zaman aralıklarında bağımlılıkları modellemekte zorlanır ve bu sorunu çözmek için LSTM geliştirilmiştir. LSTM modeli kısa süreli öğrenilmiş bilgileri hafıza birimlerinde tutarak uzun süreli faydalanmayı sağlar (Erden, 2023).

LSTM modeli veriyi işlerken kapılar kullanır. Giriş kapısı, yeni bilgileri hücre durumuna ekler. Unutma kapısı hücre durumundan gereksiz bilgileri siler. Çıkış kapısı hücre durumundan çıktı alır ve yeni hücre durumunu günceller. Zaman serileri analizi, geçmiş verilerden gelecekteki değerleri tahmin etme amacı taşır. LSTM, uzun vadeli bağımlılıkları başarıyla modelleyerek bu tür problemlere çözüm sunar. LSTM modelleri, eğitim verileri üzerinde öğrenir ve ardından doğrulama seti üzerinde performansını değerlendirir. Aşırı uydurma gibi sorunları önlemek için uygun düzenleme teknikleri kullanılabilir. Metin verilerindeki uzun vadeli bağımlılıkları modelleme, hisse senedi fiyatları ve ekonomik göstergelerin tahmini, hastalık prognozları ve tedavi yanıtlarının tahmini alanlarında kullanılır.



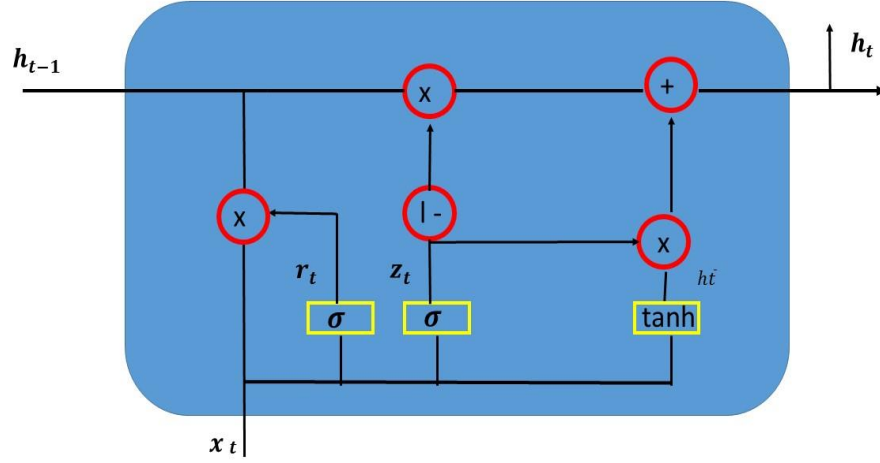
Şekil 3.4 Bir LSTM Modelinin Yapısı

Yukarıda Şekil 3.4'te LSTM yapısındaki kapılar gösterilmektedir. Buna göre, i_t girdi kapısı, f_t unutmaya kapısı, O_t çıktı kapısı, c_t hücre durumu, h_t gizli durum, x_t t zamanındaki girdi, σ sigmoid aktivasyon fonksiyonudur.

Tekrarlayan Kapılı(Geçitli) Birim(GRU)

GRU, "Gated Recurrent Unit" (Kapılı Tekrarlayan Ünite) kısaltmasıyla bilinen bir tür tekrarlayan sinir ağı modelidir. GRU, özellikle uzun vadeli bağımlılıkları başarılı bir şekilde ele almak amacıyla tasarlanmış bir sinir ağı türüdür ve özellikle zaman serileri analizi ve doğal dil işleme gibi alanlarda kullanılır.

GRU, LSTM (Long Short-Term Memory) modeli gibi uzun vadeli bağımlılıkları modeller, ancak daha basitleştirilmiş bir yapıya sahiptir. Tekrarlayan Kapılı Birim (GRU) modelinde LSTM'den farklı olarak iki kapı yer almaktadır. Bu kapılar güncelleme ve sıfırlama kapılarıdır. Güncelleme kapısı, hangi bilgilerin güncellenip saklanacağını kontrol eder. Sıfırlama kapısı, mevcut girdi ile önceki hücre durumunu karıştırarak yeni bir bilgi oluşturur. Bu sayede, GRU önceki bilgileri hafızasında tutabilir veya unutabilir.



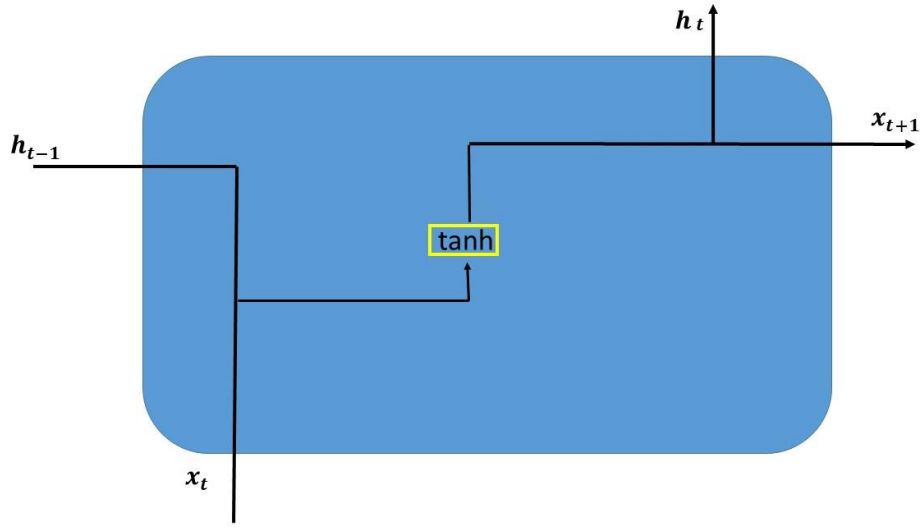
Şekil 3.5 Bir GRU Modelinin Yapısı

Şekil 3.5'te GRU modelinin yapısı gösterilmiştir. Buna göre, LSTM modelinden farklı olarak h_{t-1} gizli durum, r_t sıfırlama birimi ve z_t güncelleme birimini yapısında bulundurmaktadır. GRU'nun avantajlarından biri, LSTM'den daha az parametreye sahip olması ve daha hızlı eğitilebilir olmasıdır. Ayrıca, GRU'nun daha basit bir yapıya sahip olması, bazı durumlarda aşırı uyumaya (overfitting) karşı daha dirençli olmasını sağlayabilir. GRU, özellikle dil modelleri, metin sınıflandırma, zaman serileri tahminleri ve benzeri uygulamalarda LSTM ile birlikte veya alternatif olarak kullanılabilir. Her iki model de uzun vadeli bağımlılıkları daha iyi modelleme yeteneğine sahiptir, ancak hangi modelin tercih edileceği, uygulama ve veri setine bağlı olarak değişebilir. Tahmini yapılacak olaya ait veri setinin karakterine göre model de değişiklik gösterir.

Basit Tekrarlayan Sinir Ağları (SimpleRNN)

Simple RNN (Recurrent Neural Network), tekrarlayan sinir ağı (RNN) ailesine ait bir modeldir ve özellikle zaman serileri analizi, doğal dil işleme ve diğer sıralı veri problemlerinde kullanılmaktadır. Simple RNN, temelde bir hücre yapısına sahiptir ve geçmiş bilgiyi gelecekteki tahminlerde kullanabilme yeteneği ile bilinir. Simple RNN, bir döngü (loop) aracılığıyla geçmiş zamandaki bilgiyi mevcut zamandaki girdilerle birleştirerek gelecekteki tahminlerde kullanır. Bu yapı, modelin zamansal bağımlılıkları anlama kabiliyetini temsil eder. Simple RNN'nin hücre yapısı, her

adımında gelen girdileri ve önceki adımdaki çıktıyı kullanarak bir çıktı üretir. Ancak, Simple RNN'nin temel bir sorunu olan "gradient kaybı" problemi vardır. Bu problem, uzun vadeli bağımlılıkları başarılı bir şekilde modelleme konusunda zorluklar yaratır. Simple RNN'nin parametreleri, modelin öğrenme sürecindeki ağırlık ve bias değerlerini temsil eder. Bu parametreler, veri setine uyarlanarak öğrenilir (Özsağır, Erden, Bol, Sert, & Özocak, 2022).



Şekil 3.6 Bir SimpleRNN Modelinin Yapısı

Şekil 3.6'da Tekrarlayan Sinir Ağları RNN modeli yapısı verilmiştir. Buna göre, iki RNN katmanı ile iki yoğun katmandan oluşmaktadır. İlk adımda girdi katmanında, modele giriş verileri eklenir. Bu girdi genellikle bir zaman serisi verisi veya sıralı bir veri seti olabilir. Her bir zaman adımındaki girdiler, modelin öğrenmesini ve zaman içindeki bağımlılıkları anlamasını sağlar. Simple RNN'nin ana bileşeni olan tekrarlayan katman, geçmiş zaman adımlarındaki bilgiyi mevcut zamandaki girdilerle birleştirir. Her bir zaman adımında bir hücrenin çıktısı, bir sonraki adımda tekrar kullanılır. Bu katman sayesinde model, zaman içindeki bağımlılıkları öğrenir. Çıktı katmanı, tekrarlayan katmandan gelen bilgileri kullanarak bir çıktı üretir. Bu çıktı, genellikle bir zaman serisi değeri veya sıralı bir veri setindeki bir özellik olabilir. Simple RNN modeli, öğrenme süreci boyunca güncellenen ağırlık ve bias parametrelerine sahiptir. Bu parametreler, modelin giriş verilerini ve geçmiş zaman adımlarındaki bilgileri nasıl birleştireceğini ve çıktıyı nasıl oluşturacağını belirler.

Simple RNN'nin dezavantajlarından biri, uzun vadeli bağımlılıkları başarılı bir şekilde ele alamamasıdır. Bu, uzun süreli sıralı verilerde performans sorunlarına yol açabilir. Bu nedenle, daha gelişmiş RNN türleri, özellikle LSTM (Long Short-Term Memory) ve GRU (Gated Recurrent Unit) gibi modeller, bu dezavantajları aşmak için geliştirilmiştir. Simple RNN'nin kullanım alanları, dil modellemesi, zaman serileri tahminleri ve benzeri sıralı veri problemleridir. Ancak, daha karmaşık ve gelişmiş RNN türleri genellikle uzun vadeli bağımlılıkları daha iyi modelleyebildiği için bu alanlarda daha yaygın olarak tercih edilmektedir.

3.2.3 Denetimsiz Öğrenme

Denetimsiz makine öğrenmesi, veri setindeki örneklerin etiketlenmeden kullanıldığı bir makine öğrenmesi yaklaşımıdır. Bu yöntemde, algoritma, veri setindeki yapıları ve desenleri anlamaya çalışırken, veri örnekleri hakkında önceden bilgi veya yönlendirme sunulmaz. Denetimsiz makine öğrenmesiyle, algoritma, veri setindeki benzerlikleri, grupları veya dağılımları keşfedebilir, veri setindeki anlamlı öznitelikleri tanımlayabilir veya yeni verileri kümeler halinde gruplandırabilir. Bu tür bir öğrenme, bir veri setinin içinde gizli trendleri veya yapıları ortaya çıkarabilir ve anlam kazandırabilir. Örneğin, pazarlama analizi için denetimsiz öğrenme teknikleri kullanarak, benzer tüketici profillerini belirlemek veya satış verilerindeki segmentleri keşfetmek mümkün olabilir. Denetimsiz makine öğrenmesi, verilerdeki desenleri keşfetmek ve anlamak için çok yönlü bir yaklaşımdır. Ancak etiketlenmiş verilere dayalı olarak doğruluğu doğrulamak yerine, veri setindeki yapı ve desenleri belirlemeye odaklanmaktadır.

Denetimsiz öğrenme, bir makine veya yazılımın doğrudan insan müdahalesi olmadan, veri analizi ve makine öğrenimi yöntemleri kullanarak bilgi edinmesi ve sonuç çıkarabilmesidir. Denetimsiz öğrenme algoritmaları, veri setlerindeki yapıları, desenlere veya ilişkilere dayanarak bilinmeyen veya gizli bilgileri keşfetmek için kullanılır. Bu yöntem genellikle veri madenciliği, gruplandırma, yoğunluk tahmini ve veri sıkıştırma gibi alanlarda kullanılır. Denetimsiz öğrenme, veri setindeki kural ve yapıları kavramak ve tahminlerde bulunmak için kullanılan bir yaklaşımdır.

Denetimsiz makine öğrenmesi algoritmaları, verilerdeki yapıyı analiz etmek ve ilişkileri ortaya çıkarmak için kullanılan bir grup algoritmadır. Kümeleme

algoritmaları, verileri benzer özelliklere sahip gruplara bölen algoritmalarlardır. K-means, hiyerarşik kümeleme ve yoğunluk tabanlı kümeleme gibi algoritmalar bu kategoriye girer. Boyut indirgeme algoritmaları, veri setindeki değişken sayısını azaltmak için kullanılan algoritmalarlardır. Veri setinin gizli özetini çıkarırken boyut indirgeme ve anomalileri tespit etme gibi görevleri gerçekleştirirler. Bu algoritmalarından her biri, verilerin yoğunluğu, dağılımı ve problem alanı gibi faktörlere bağlı olarak tercih edilir ve denetimsiz öğrenme yöntemleri kullanılarak çeşitli sorunlar çözülür. Ancak, bu algoritmaların başarısı veri seti kalitesi ve doğru parametre ayarlarına da bağlıdır.

Kümeleme (Clustering): Kümeleme algoritması, bir veri setinin benzerlik ölçütlerine göre bir veya daha fazla kümeye ayrılmasını sağlayan bir yöntemdir. Bu algoritma, veri noktalarının özelliklerini analiz ederek benzer olanları bir araya getirir ve farklı olanları birbirinden ayırır. Kümeleme algoritmaları genellikle veri madenciliği, istatistiksel analiz ve makine öğrenmesi gibi alanlarda kullanılır. Veri setlerini anlamak, örüntüleri çıkarmak ve verileri daha anlamlı bir şekilde sınıflandırmak için kümeleme algoritmalarından faydalanılır.

K-Means Algoritması: K-means kümeleme (k-means clustering), bir veri setinin benzer örneklerden oluşan gruplara ayrılması işlemidir. Bu algoritma, veri noktalarını birbirine yakın olan gruplara ayırarak, noktalar arasındaki benzerlikleri ölçer. Küme merkezlerini belirlemek için iteratif bir yaklaşım kullanır ve her noktanın en yakın olduğu merkeze atanmasını sağlar. K-means, pazar segmentasyonu, görüntü işleme, genetik analiz ve örüntü tanıma gibi çeşitli uygulamalarda kullanılabilir. Kullanıcının belirlemesi gereken iki önemli parametresi küme sayısı (k) ve başlangıç küme merkezleridir. Algoritmada başlangıçta k tane rastgele küme merkezi seçilir. Sonra her nokta en yakın küme merkezine atanır. Her küme merkezi, atanan noktalardan yeni bir merkez konumu hesaplayarak güncellenir. Bu adımlar, küme merkezleri sabit olana kadar tekrarlanır. Sonuçta, veri seti kümeler halinde gruplandırılır ve her bir küme belirli bir örüntüyü temsil eder. K-means, veri keşfi, veri düzenleme, veri sıkıştırma, veri azaltma gibi birçok alanda kullanılan popüler kümeleme algoritmalarından birisidir.

3.2.4 Yarı Denetimli Öğrenme

Yarı denetimli öğrenme, hem etiketli veri hem de etiketsiz veri içeren bir makine öğrenme yaklaşımıdır. Bu yöntemde, bazı veriler etiketlenirken, diğer veriler etiketlenmez. Yarı denetimli öğrenme, karmaşık bir problemde etiketleme maliyetini azaltmayı amaçlar. Etiketli veri sayısı genellikle daha azdır ve bu nedenle etiketleme süreci zaman alıcı ve maliyetlidir. Ancak, etiketsiz veri genellikle daha fazla miktarda bulunur ve bunların kullanılmasıyla modele daha fazla çeşitlilik ve genelleyebilme yetisi kazandırılabilir. Bir yarı denetimli öğrenme algoritması, etiketli verileri kullanarak modele bir başlangıç noktası sağlar ve ardından etiketsiz verilerin kullanılmasıyla modelin genişletilmesi ve iyileştirilmesi hedeflenir. Bu yaklaşım, etiketsiz verileri kullanarak belirsiz etiketleri bulma ve öğrenme sürecini iyileştirme potansiyeline sahiptir. Ancak, yarı denetimli öğrenme veri setinin doğru bir şekilde oluşturulması ve algoritmanın optimal şekilde ayarlanması gibi bazı zorlukları da beraberinde getirebilir. Yarı denetimli öğrenmede, öğrenme etiketleri gürültülü, sınırlı veya kesin olmamakla beraber, bu etiketlerin belirlenmesi genellikle daha kolaydır. Bu da daha büyük ve etkili veri setleriyle sonuç verir (Karakuş, 2023).

3.2.5 Takviyeli Öğrenme

Takviyeli öğrenme, makine öğrenmesinin bir dalıdır. Bu yöntemde bir sistem, bir çevreyle etkileşime geçer ve bu etkileşimler sonucunda deneyimler edinir. Amacı, belirli bir görevi en iyi şekilde yerine getirebilecek stratejileri öğrenmektir. Bu süreçte, sistem bir davranışı gerçekleştirir, ardından çevreden gelen geribildirimlerle bu davranışın ne kadar iyi veya kötü olduğunu değerlendirir. Daha sonra, bu geribildirimleri kullanarak gelecekteki davranışları geliştirmek için öğrenme algoritmasıyla kendisini günceller. Takviyeli öğrenme genellikle oyun teorisi, kontrol sistemleri ve robotik gibi alanlarda kullanılır. Bu yöntem, bir yapay zekanın belirli bir durumu nasıl algılayacağını, nasıl hareket edeceğini ve bir görevi nasıl yerine getireceğini öğrenmesine olanak tanır. Q learning, Temporal Difference, Monte Carlo Tree Search, Asynchronous Actor Critic Agents algoritmaları örnek verilebilir (Karakuş, 2023).

Takviyeli öğrenme algoritmaları, çevrenin durumlarını ve bu durumlara göre alınması gereken aksiyonları öğrenmek için kullanılır. Bir algoritma, durumu gözlemleyerek aksiyon seçer, bu aksiyonu çevre üzerinde gerçekleştirir ve çevreden

geribildirim alır. Geribildirimler, seçilen aksiyonun ne kadar iyi olduğunu ifade eder ve algoritma, daha iyi sonuçlar elde etmek için aksiyon seçim stratejisini günceller. Takviyeli öğrenme algoritmaları, oyunlarda, robot kontrolünde, otomatik araçlarda ve diğer pek çok uygulamada kullanılabilir. Takviyeli öğrenme algoritmalarının bazı popüler örnekleri arasında Q-learning, SARSA, Deep Q-Networks ve Proximal Policy Optimization sayılabilir. Her bir algoritma farklı özelliklere sahip olup, belirli bir uygulama için en uygun olanını seçmek önemlidir.

3.2.6 Pekiştirmeli Öğrenme

Pekiştirmeli öğrenme algoritmaları, makine öğrenmesi alanında kullanılan bir tür algoritmadır. Pekiştirmeli öğrenmede, bir sistem belirli bir ortamda bulunur ve çeşitli eylemler gerçekleştirir. Bu eylemler, sistem için birer "karar" olarak değerlendirilir ve her eylem sonucunda bir "ödül" veya "ceza" alınır. Bu algoritmalarda amaç, sistem için en fazla ödülü elde edecek eylemleri öğrenmektir. Başlangıçta sistem rastgele eylemler gerçekleştirerek çevreyi keşfeder ve zamanla daha fazlasını öğrenerek daha iyi kararlar vermeyi öğrenir. Pekiştirmeli öğrenme algoritmaları, çeşitli uygulamalarda kullanılabilir. Örneğin, oyunlarda veya robot kontrolünde en iyi hareketleri belirlemek için kullanılabilirler. Ayrıca, reklam stratejileri oluşturmak veya tedarik zinciri yönetimi gibi alanlarda optimize edilmiş kararlar vermek için de kullanılabilirler.

3.2.7 Özellik Öğrenme

Özellik öğrenme algoritmaları, veri analizi ve makine öğrenmesinde kullanılan bir dizi teknik ve yöntemdir. Bu algoritmalar, girdi verilerinden anlamlı özellikleri (öznitelikleri) otomatik olarak çıkarmak veya seçmek için kullanılır. Özellik öğrenme algoritmaları, denetimli veya denetimsiz öğrenme yaklaşımlarına dayanabilir. Denetimli öğrenme durumunda, önceden etiketlenmiş bir veri kümesi kullanılarak bir model eğitilir. Bu model, girdi verilerinin etiketlerine bağlı olarak en iyi öznitelikleri öğrenmeye çalışır. Denetimsiz öğrenme durumunda ise, veri kümesinde etiketler bulunmaz. Bu durumda, algoritmalar verilerdeki yapıları ve desenleri tanımlamaya çalışır. Kümeleme ve boyut indirgeme gibi teknikler kullanılarak verilerdeki benzerlikler ve farklılıklar keşfedilir. Örneğin, özellik öğrenme algoritmaları arasında en yaygın olanları temel bileşen analizi, lineer ve doğrusal olmayan diskriminant analizi, destek vektör makineleri, veri madenciliği ve derin öğrenme yöntemleri

bulunmaktadır. Bu algoritmalar, çeşitli uygulama alanlarında verilerin daha iyi anlaşılmasını ve modelleme performansının artmasını sağlar.

3.3 Python Programlama ile Veri Analizi

Bir makine öğrenmesi probleminde modelin yanlılığını veya tutarlılığını değerlendirmek ve doğru bir tahmin yapmak için matematik istatistik gibi teorik bilgilerin yanında uygulamaya geçebilmek için bir değerlendirme stratejisine de ihtiyaç olacaktır. Belirli bir durum için en iyi makine öğrenimi algoritmasını seçmek için farklı makine öğrenimi algoritmalarının nasıl çalıştığını anlamak çok önemlidir. Bu nedenle, makine öğrenmesi algoritmalarının yanında bu algoritmaları uygulayabilecek bir Java, R, Python ve C++ gibi dillerde programlamayı bilmek gerekmektedir.

3.3.1 Python Programlama ve Kütüphaneleri

Python programlama, genel amaçlı, yüksek seviyeli bir programlama dilidir. 1990'ların sonunda Guido van Rossum tarafından geliştirilmiştir. Python, basit ve anlaşılır bir sözdizimiyle bilinir ve çeşitli uygulamalar için kullanılabilir. Python ile çeşitli programlar geliştirilebilir. Web uygulamaları, veritabanı yönetimi, bilimsel hesaplamalar, oyunlar, yapay zeka ve daha fazlası gibi farklı alanlarda kullanılabilir. Python programlama dilinin birçok avantajı vardır. Basit ve anlaşılır sözdizimi sayesinde okunabilir ve kolayca öğrenilebilir. Aynı zamanda çok sayıda kütüphane ve modül içermesi, geliştirme sürecini hızlandırır. Python topluluğu da oldukça geniş olduğu için, dökümantasyonlara ve yardıma kolaylıkla ulaşılabilir. Diğer karmaşık program dillerini öğrenmek çok zaman alır ve kullanım alanları büyük olmasından dolayı kullanımını öğrenmek çok zordur. Ancak Python programının sözdizimi hem okunabilir hem de ileriye dönüktür. Öğrenim, kararlı programlama dili sayesinde basittir. Yeni başlayanlar için de ideal bir seçim olarak ön plana gelir. Sonuç olarak, Python kullanarak program geliştirmeye yeni bir adım atmış olan herkes hızlı ve basit şekilde ilerleme kaydedebilir. Diğer karmaşık dillere göre basitlik söz konusu olduğunda, en önde yer almaktadır.

Veri analizleri günümüzde internet teknolojileri (IT) alanında önemli bir yer oluşturmaktadır. Python ise bu durumlar için en elverişli programlama dili olmuştur. Python arayüzündeki kütüphanelerin birçoğu makine öğrenimi ve veri bilimi üzerine elverişlidir. Bu alanlardaki kütüphanelerde ki yüksek kaliteli komutları, makine

öğrenimi kütüphanelerinin ve diğer nümerik algoritma kütüphanelerinin sürekli gelişmesine çok yardımcı olmuştur. Python programının basit ve anlaşılır sözdizimi, diğer birçok programlama diline göre daha az karmaşıktır ve yazılım geliştirmeyi kolaylaştırır. Bu nedenle, Python genellikle başlangıç seviyesindeki programcılar için tercih edilen bir dil haline gelmiştir. Python, yorumsal programlama yapmayı destekler. Programcılar, kodun anlaşılması için açıklayıcı yorumlar ekleyebilirler. Bu, başkalarının veya gelecekteki kendinizin yazdığınız kodu daha rahat anlamasını sağlar. Python, nesne yönelimli programlamanın avantajlarını sunar. Nesne yönelimli programlama, kodu daha organize hale getirir ve tekrar kullanılabilir, modüler ve sürdürülebilir yazılım geliştirmeyi sağlar. Python, dinamik olarak tip atama yapma yeteneğine sahiptir. Bu, değişkenlerin türlerini önceden belirtmek zorunda bırakmaz. Ayrıca, Python, düşük seviyeli işlemleri otomatik olarak yönetir ve yüksek seviyeli bir dil olarak karmaşıklığı gizler. Python, zengin bir standart kütüphane ve birçok üçüncü taraf kütüphane sunar. Bu kütüphaneler, çeşitli görevleri gerçekleştirmek için kullanılabilir ve yazılım geliştirme sürecini hızlandırır. Örneğin, veri analizi, yapay zeka, web geliştirme gibi birçok alanda popüler kütüphaneler bulunmaktadır. Python programında aynı kod farklı işletim sistemlerinde çalıştırılabilir. Python kodunu yazdıktan sonra, Linux, Windows, macOS gibi farklı işletim sistemlerinde çalışabilir. Python programı NumPy, SciPy, Matplotlib, Pandas, Scikitlearn, Keras vb. birçok veri analizine yardımcı kütüphaneleri barındırmaktadır.

NumPy Kütüphanesi: NumPy, Python programlama dili için bir veri analizi kütüphanesidir. Bilimsel hesaplamalar yapmak için kullanılır ve çok boyutlu dizilerin oluşturulmasını, manipüle edilmesini ve işlenmesini kolaylaştırır. Veri analitiği, veri işleme, matematiksel hesaplamalar, lineer cebir, istatistik ve diğer birçok alanda yaygın olarak kullanılır. NumPy'nin temel yapısı, hızlı ve verimli bir şekilde büyük veri dizileri üzerinde çalışabilen çok boyutlu ndarray (N-boyutlu dizi) nesnesidir. Bu nesne üzerinde bir dizi işlem ve operasyon yapılabilir. NumPy ayrıca matematiksel fonksiyonlar, rastgele sayı üretimi, lineer cebir işlemleri ve Fourier dönüşümleri gibi bir dizi özellik sağlar. Performansı önemli ölçüde artıran diziler ve matrisler üzerinde matematiksel işlemlerin vektörleştirilmesini sağlar. Ayrıca diğer birçok python kitaplığı NumPy üzerine kuruludur (Varol D. , 2021).

SciPy Kütüphanesi: Scipy kütüphanesi, bilimsel hesaplama ve veri analizi için kullanılan bir Python kütüphanesidir. Scipy, çeşitli alt modülleri aracılığıyla matematiksel fonksiyonları, optimizasyon algoritmalarını, sinyal ve görüntü işleme araçlarını, istatistiksel işlevleri ve daha fazlasını içerir. Scipy kütüphanesi, NumPy kütüphanesine dayalı olarak çalışır ve genellikle birlikte kullanılır. NumPy sayısal hesaplamalar için temel araçları sağlarken, Scipy daha yüksek seviye fonksiyonlar ve özellikler sunar. Scipy kütüphanesi, bilimsel ve mühendislik uygulamalarında yaygın olarak kullanılan birçok alanda kullanılabilir. Örneğin, lineer cebir, türev alma ve integral hesaplama, optimize edici algoritmalar, sinyal işleme, görüntü işleme ve istatistiksel analiz gibi alanlarda kullanılan fonksiyonları içerir.

Pandas Kütüphanesi: Pandas, Python programlama dilinde veri analizi ve manipülasyonu için kullanılan güçlü bir kütüphanedir. Veri işleme, temizleme, dönüştürme, keşif ve analizde kullanılan bir dizi veri yapıları ve işlevler sunar. Pandas, tabloları (DataFrame), sıralı verileri (Series) ve matrisler gibi yapıları destekler. Bu kütüphane, veri setleri üzerinde filtreleme, sıralama, gruplama, birleştirme ve daha birçok işlemi kolayca gerçekleştirmenizi sağlar. Pandas, veri analitiği ve veri bilimi projelerinde sıkça tercih edilen bir araçtır.

SciKit-Learn Kütüphanesi: Scikit-learn, Python programlama dili için açık kaynaklı bir makine öğrenimi kütüphanesidir. Scikit-learn, bir dizi makine öğrenimi algoritması, veri işleme işlevleri ve veri ön işleme araçları sağlar. Sınıflandırma, regresyon, kümeleme, boyut indirgeme ve model seçimi gibi birçok popüler makine öğrenimi işlemi için kullanılabilir. Scikit-learn'ün avantajlarından bazıları, kullanımının kolay olması, geniş bir algoritma ve fonksiyon koleksiyonuna sahip olması ve Python ekosistemiyle uyumlu olmasıdır.

Matplotlib Kütüphanesi: Matplotlib, veri görselleştirme için yaygın olarak kullanılan bir Python kütüphanesidir. Bilimsel ve sayısal hesaplamalarla ilgili olarak çizimler, histogramlar, çizgi grafikleri, dağılım grafikleri vb. gibi çeşitli grafikler oluşturmak için kullanılır. Matplotlib'in temel yapısı, veriyi temsil eden bir dizi noktanın koordinat düzleminde gösterilmesine dayanır. Bunun için çeşitli grafik nesnelere kullanılır ve bu nesnelere istenen özelliklerin ayarlanmasıyla grafikler oluşturulur. Matplotlib'i kullanırken, önce grafiğin çizimi için bir veri kümesi

oluşturmanız ve ardından bu verileri uygun grafik türüyle eşleştirmek gerekir. Daha sonra grafik üzerindeki detayları özelleştirerek istenilen sonuç elde edilebilir.

Seaborn Kütüphanesi: Seaborn, Python programlama dilinde veri görselleştirmesi için kullanılan bir kütüphanedir. Veri analizi ve görselleştirme işlemlerini kolaylaştırmak amacıyla Matplotlib kütüphanesinin üzerine inşa edilmiştir. Seaborn ile yapabilecek bazı görselleştirme işlemleri, ilişkiyi analiz etmek için scatter plotlar ve regresyon modelleri oluşturmak, kategorik değişkenler için bar plotlar, count plotlar ve box plotlar oluşturmak, dağılımı görmek için histogramlar, yoğunluk grafiği ve çizgi plotları oluşturmak, gruplar arasındaki farkları analiz etmek için violin plotlar ve point plotlar oluşturmaktır.

3.3.2 Veri Analizi

Veri analizinde ilk olarak veri toplama işlemi yapılmaktadır. Makineler öğrenme işlemine başlarken ilk olarak onlara verdiğiniz veri setlerinden öğrenmeyi gerçekleştirirler. Makine öğrenimi modelinizin doğru kalıpları bulabilmesi için güvenilir verilerin toplanması son derece önemlidir. Makineye beslediğiniz verilerin kalitesi, modelinizin ne kadar doğru olduğunu belirleyecektir. Yanlış veya güncel olmayan veriler varsa, konuyla ilgili olmayan yanlış sonuçlara veya tahminler elde edilir. İyi veriler çok az eksik ve tekrarlanan değer içerir ve mevcut çeşitli alt kategorileri/sınıfları iyi bir şekilde yansıtır. Veri Toplama, makine öğrenimi yaşam döngüsünün ilk adımıdır. Bu adımın amacı, verilerle ilgili tüm sorunları belirlemek ve elde etmektir. Bu adımda, veriler dosyalar, veritabanı, internet veya mobil cihazlar gibi çeşitli kaynaklardan toplanabileceğinden farklı veri kaynaklarını tanımlamak gerekir. Bu yaşam döngüsünün en önemli adımlarından biridir. Toplanan verilerin niceliği ve kalitesi, çıktının verimliliğini belirleyecektir. Veriler ne kadar fazla olursa, tahmin o kadar doğru olur (Karakuş, 2023).

Verileri topladıktan sonra, modelleme için veri hazır hale getirilmelidir. Veri hazırlama, verilerimizi uygun bir yere yerleştirip makine öğrenimi eğitimimizde kullanmak üzere hazırladığımız bir adımdır. Bu adımda, önce tüm veriler bir araya getirilir ve ardından verilerin sıralaması rastgele yapılır. Verilerin özelliklerini, biçimini ve kalitesini anlamak gerekir. Verilerin daha iyi anlaşılması, etkili bir sonuca yol açar. Bunun için de korelasyonlar, genel eğilimler ve aykırı değerler bulunması

gerekmektedir. Veriler alındıktan sonra tüm verileri bir araya getirerek ve rastgele hale getirilmelidir. Bu, verilerin eşit olarak dağıtıldığından ve sıralamanın öğrenme sürecini etkilemediğinden emin olmaya yardımcı olur. İstenmeyen verileri, eksik değerleri, satırları ve sütunları, yinelenen değerleri, veri türü dönüştürmeyi vb. kaldırmak için veriler temizlenmelidir. Hatta veri kümesi yeniden yapılandırılmalı ve satırları ve sütunları veya satır ve sütunların dizinin değiştirilmesi bile gerekebilir. Nasıl yapılandırıldığını anlamak ve çeşitli değişkenler ve mevcut sınıflar arasındaki ilişkiyi anlamak için veriler görselleştirilmelidir. Temizlenen veriler eğitim (train) ve test olmak üzere iki kümeye bölünmelidir. Eğitim seti, modelin öğrendiği settir. Eğitimden sonra modelin doğruluğunu kontrol etmek için bir test seti kullanılır.

Ham verileri temizleme ve kullanılabilir bir biçime dönüştürme işlemi ile verilerin temizlenmesi, kullanılacak değişkenin seçilmesi ve bir sonraki adımda analize daha uygun hale getirilmesi için verinin uygun formata dönüştürülmesi gerekmektedir. Bu tüm sürecin en önemli adımlarından biridir. Bu aşamada gerçek dünya uygulamalarında toplanan verilerde eksik değer, tekrarlayan değer, geçersiz veri, gürültü gibi istenmeyen durumdaki verilerin kullanılabilir hale getirilmesi oldukça önem taşımaktadır. Artık temizlenen ve hazırlanan veriler analiz aşamasına geçilir. Bu aşamada çeşitli analitik teknikleri kullanarak verileri analiz etmek ve sonucu gözden geçirmek için bir makine öğrenimi modeli oluşturmaktır. Sınıflandırma, regresyon, kümeleme analizi, ilişkilendirme vb. gibi makine öğrenmesi algoritmaları belirlenerek model oluşturulmaktadır.

Veri kümesinin yapısına ve problemin türüne göre algoritmalar belirlenir. Problemin türüne göre hangi makine öğrenmesi algoritmasının seçileceğine karar verilmelidir. Bir makine öğrenimi modeli, toplanan veriler üzerinde bir makine öğrenimi algoritması çalıştırıldıktan sonra elde ettiğiniz çıktıyı belirler. Problemin kapsamıyla ilgili bir model seçmek önemlidir. Yıllar içinde konuşma tanıma, görüntü tanıma, tahmin vb. gibi farklı görevlere uygun çeşitli modeller geliştirilmiştir. Bunun dışında modelin sayısal mı yoksa kategorik veriler için mi uygun olduğuna karar verilmeli ve buna göre seçim yapılmalıdır.

Modelin eğitilmesi aşamasında problemin daha iyi sonuçlanması için performansını geliştirmek üzere modeli çeşitli makine öğrenimi algoritmaları

kullanarak eğitmek için veri kümeleri kullanılır. Çeşitli kalıpları, kuralları ve özellikleri anlayabilmesi için bir modelin eğitimi gereklidir.

Makine öğrenimi modeli belirli bir veri kümesinde eğitildikten sonra model test edilir. Bu adımda, modele bir test veri seti sağlayarak modelin doğruluğu kontrol edilir. Modelin test edilmesi, proje veya problemin ihtiyacına göre modelin yüzde doğruluğunu belirler.

Eğitim, makine öğreniminde en önemli adımdır. Eğitimde, kalıpları bulmak ve tahminler yapmak için hazırlanan verileri makine öğrenme modeline iletilir. Görev setini gerçekleştirebilmesi için modelin verilerden öğrenmesiyle sonuçlanır. Zamanla, eğitimle model tahmin etmede daha iyi hale gelir.

Model eğitildikten sonra, nasıl performans gösterdiğini kontrol etmek gerekir. Bu, modelin performansını daha önce görülmemiş veriler üzerinde test ederek yapılır. Kullanılan görünmeyen veriler, verileri daha önce bölünen test kümesidir. Eğitim için kullanılan aynı veriler üzerinde test yapılırsa, model verilere zaten alıştığından ve daha önce yaptığı gibi aynı kalıpları bulduğundan doğru bir ölçüm elde edemez. Bu orantısız bir şekilde yüksek doğruluk sağlayacaktır. Test verileri üzerinde kullanıldığında, modelin nasıl performans göstereceğine ve hızına ilişkin doğru bir ölçüm elde edilir. Hazırlanan model, kabul edilebilir bir hızda ihtiyacımıza göre doğru bir sonuç üretiyorsa, modeli gerçek sisteme yerleştiririz. Ancak projeyi dağıtmadan önce, mevcut verileri kullanarak performansını iyileştirip iyileştirmediğini kontrol etmek gerekir.

Modeli oluşturup değerlendirdikten sonra, doğruluğunun herhangi bir şekilde geliştirilip geliştirilemeyeceğine bakılmalıdır. Bu, modelde bulunan parametreleri ayarlayarak yapılır. Parametreler, programcının genel olarak karar verdiği modeldeki değişkenlerdir. Parametrelerin belirli bir değerinde doğruluk maksimum olacaktır. Parametre ayarlama, bu değerleri bulmayı ifade eder. En iyileme yapılacak parametrelerin aralığı belirlenir. Sonunda, doğru tahminler yapmak için modeli öngörülemeyen, beklenen veriler üzerinde işlem yapılabilir.

3.4 Akkuş Rüzgâr Enerji Santrali Verilerinin Modellenmesi

Ordu ili Akkuş ilçesinde 1750 rakımlı Taz yaylasına 2020 yılı sonunda kurulumu yapılan rüzgâr enerji santralinde kule yüksekliği yaklaşık 116 metre ve her biri

yaklaşık 18 ton ağırlığında 3 adet türbinden elektrik üretilmektedir. Tam kapasiteyle çalışan üç türbinden, saatte 12 MWh elektrik üretilmektedir. Üretilen elektrik, ulusal elektrik şebekesine aktarılmakta ve Akkuş ilçesindeki ev ve tesislerin elektrik ihtiyacını karşılamaktadır.

3.4.1 Veri Önleme

Akkuş Rüzgâr Enerji Santrali (Akkuş RES) verileri makina öğrenmesi metodları ile modellenebilmesi için öncelikle ön işlem uygulanarak modelin eğitilmesi için uygun hale getirilmiştir.

Akkuş RES veri seti, 392712 satır ve 27 sütundan oluşan içeriğinde boş değerler ve 0 değerlerini bulduran bir excel veri setidir. 16.10.2020 ile 16.04.2023 tarihleri arasındaki 10 ‘ar dakika aralıklarla ölçülen değerlerin yer aldığı, üretilen aktif reaktif güç, rüzgâr hızı, nem değerleri, kanat pozisyonları, hava basınç değerleri vb sütun değerlerini bulduran bir veri setidir. Veri ön işleme aşamasında, sütunlarda kaç adet 0 verisi olduğuna bakılmış ve bunun sonucunda eksik olan çoğu verinin yerine 0 kaydedildiği görülmüştür. Bu durum, modelin doğru tahmin yapmasını oldukça olumsuz etkileyeceği için tüm 0 değerleri eksik veri (NaN) olarak kaydedilmiş ve bu eksik değerler yerine her sütunun ortalaması alınarak bu ortalama değerler yazdırılmıştır. Daha sonra ise üretilen güç değerini en az etkileyen ve çok sayıda eksik değer bulduran sütunlar, veri setinden düşürülerek daha doğru tahmin yapılması hedeflenmiştir. Yeni veri seti kaydedilerek aykırı veri analizine geçilmiştir. Düzenlenmiş Akkuş Res veri seti Şekil 3.7’de gösterilmiştir.

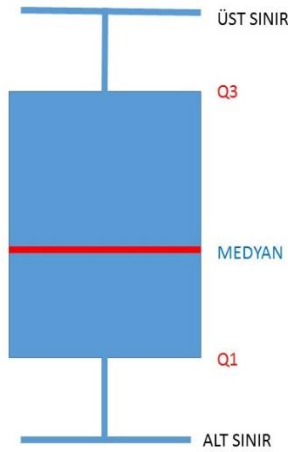
Time	Wind	Windmax	Windmin	Rotationspeed	Rotationspeedmax	Rotationspeedmin	Power	Powermax	Powermin	Power_Avail_wind	...	Reactivepowerm
2020-10-16 23:20:00	6.6	7.7	4.7	8.36	8.69	7.31	702.0	807.0	434.0	710.0	...	11:
2020-10-16 23:30:00	7.0	8.0	6.1	8.60	8.86	8.23	774.0	856.0	676.0	776.0	...	11:
2020-10-16 23:40:00	7.0	8.0	6.0	8.65	8.91	8.24	789.0	880.0	679.0	793.0	...	11:
2020-10-16 23:50:00	6.7	7.7	5.8	8.36	8.87	7.94	704.0	862.0	600.0	712.0	...	11:
2020-10-17 00:00:00	7.0	7.9	6.0	8.68	8.99	8.38	798.0	902.0	708.0	800.0	...	12:
2020-10-17 00:10:00	6.6	8.1	5.4	8.60	9.04	7.81	774.0	914.0	585.0	784.0	...	12:
2020-10-17 00:20:00	5.8	7.1	4.5	7.69	8.21	7.20	517.0	657.0	387.0	537.0	...	10:
2020-10-17 00:30:00	5.2	6.3	4.2	7.27	7.48	7.00	376.0	441.0	315.0	392.0	...	8:

Şekil 3.7 Düzenlenmiş Akkuş Res Veri Seti

3.4.2 Aykırı (Uç) Veri Analizi

Öncelikle veri setindeki tarih sütunu indeks yapılarak tarihe göre sıralama yapılmış ve tarih indeks değerleri, Python programlama dilinin veriyi tarih olarak tanıyabileceği data tipine çevrilmiştir. Veri setindeki sütunların box-plot görselleştirmeleri yapılarak sütun değerlerinin belirli bir aralıkta olup olmadığına bakılmış ve veri setinin karakteristiğini yansıtmayan aykırı (uç) değer olup olmadığı incelenmiştir. Aykırı (uç) veri tespit edilirken IQR yöntemi uygulanmıştır. IQR yönteminde verilerin yoğunlaştığı alan Q1 ve Q3 denilen çeyrekliklere ayrıldıktan sonra bu iki çeyreklik arasındaki mesafe IQR olarak tespit edilir. Q1 ve Q3 çeyrekliklere bu IQR mesafesinin 1.5 katı eklenerek veya çıkarılarak alt ve üst sınır değerler hesaplanmaktadır. Bu yöntem Python programlama dilinde box-plot yöntemi olarak yer almaktadır.

Box plot, veri dağılımını ve istatistiksel özelliklerini görselleştirmek için kullanılan bir grafik yöntemidir. Box plot, veri setinin özet istatistik değerini gösterir. Bu değer veri setinin genel özelliklerini anlamak için kullanılır. Box plot, veri setinin dağılımını, merkezi eğilimini ve aykırı değerleri hızlı bir şekilde görselleştirmek için kullanışlıdır. Ayrıca, farklı gruplar veya kategoriler arasındaki karşılaştırmaları yapmak için de kullanılabilir. Box plot, istatistiksel analizlerde ve veri keşfi aşamalarında sıkça tercih edilen bir yöntemdir.

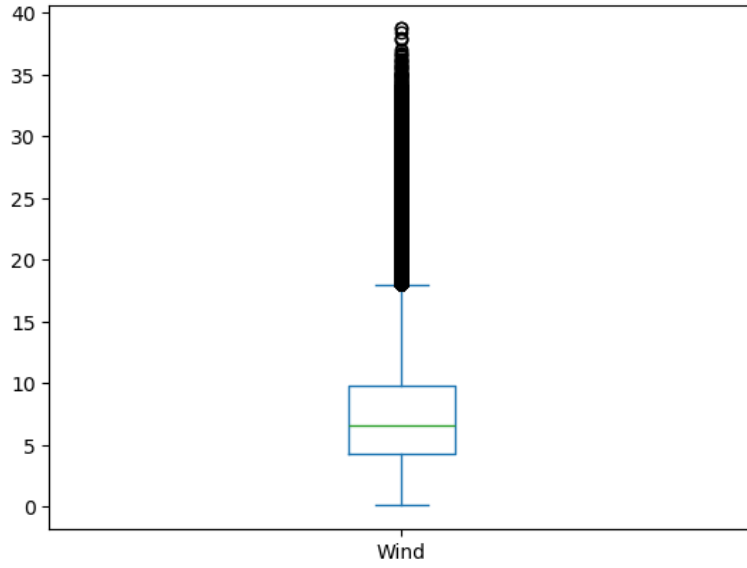


Şekil 3.8 IQR Yöntemi Görseli

Box-plot ya da IQR yönteminin yapısı Şekil 3.8’de gösterilmiştir. Bu çalışmada 1. Çeyreklik (Q1) %10 olarak, 3. Çeyreklik (Q3) %90 olarak uygulanmıştır. Çeyrekler arası uzaklık (IQR) hesaplanırken Q3 ve Q1 arasındaki fark alınmıştır. Üst sınır değer belirlenmesi aşamasında Q3 değerine 1,5xIQR eklenmiş, alt sınır değer belirlenmesinde ise Q1 değerinden 1,5xIQR çıkarılmıştır. Hesaplanan bu alt sınırın altındaki ve üst sınırın üstündeki değerler aykırı veriler olarak tespit edilmiştir.

$$\text{üst sınır} = Q_3 + 1,5xIQR \quad (3.1)$$

$$\text{alt sınır} = Q_1 - 1,5xIQR \quad (3.2)$$



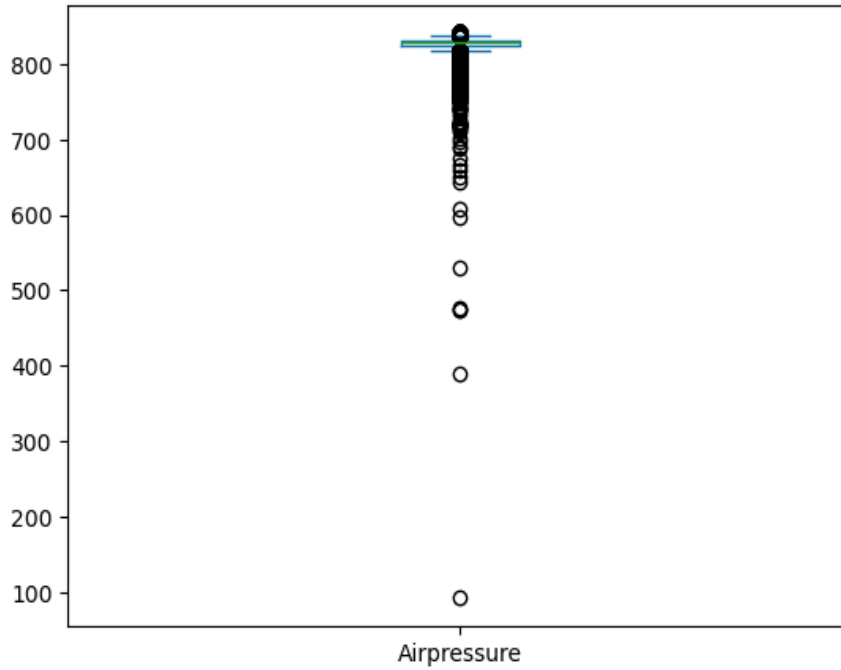
Şekil 3.9 Wind (Rüzgâr) Sütunu Box-Plot Çizimi

Şekil 3.9’da veri setindeki Wind (Rüzgâr) sütunundaki box-plot çizimine göre, değerlerde aykırı değer olduğu görülmüş ancak ihmal edilebilir olarak kabul edilmiştir. Şekil 3.10’da Airhumidity (hava nem değeri) sütunundaki box-plot çizimine göre, değerlerde aykırı veri olduğu tespit edilmiştir. Ortalama 100 birim seviyelerinde seyreden değerlerden farklı olarak 40000, 50000, 60000, ... gibi aykırı veriler tespit edilmiştir. Denklem (3.1) kullanılarak nem sütunu için 137.5 değeri üst değer olarak belirlenmiş ve bu değer üzerindeki veriler 137.5 olarak değiştirilmiştir.



Şekil 3.10 Airhumidity (Nem) Sütunu Box-Plot Çizimi

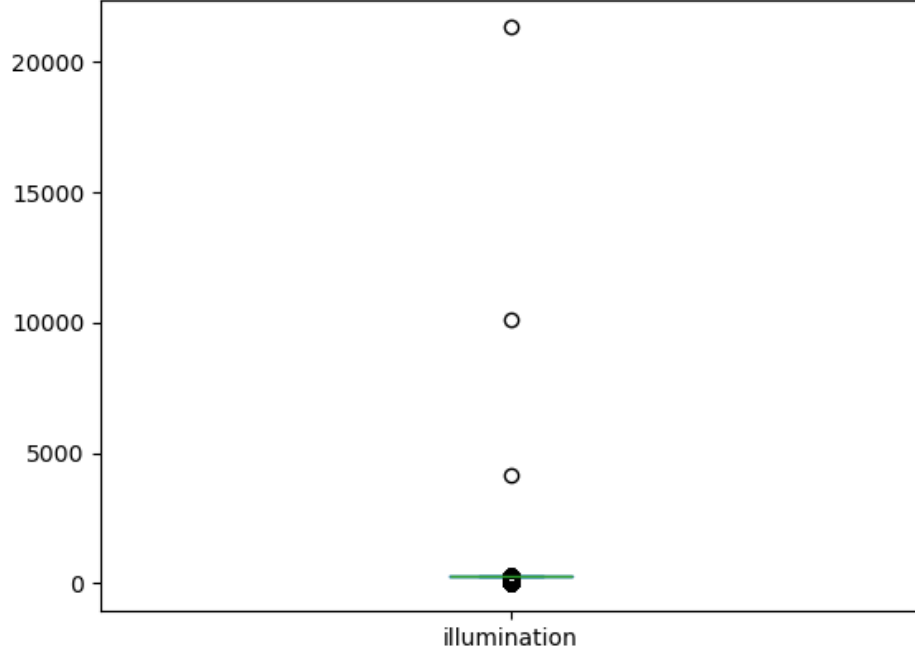
Şekil 3.11’de veri setindeki Airpressure (hava basınç) sütununa ait box-plot çiziminde, ortalama değerler 1000 bar seviyelerinde olmasına karşın 400 bar gibi aykırı veriler olduğu saptanmıştır. Denkelem (3.2) kullanılarak basınç sütunu için alt sınır değeri 807 olarak belirlenmiş ve bu değerin altındaki değerler 807 olarak kaydedilmiştir. Böylece



Şekil 3.11 Airpressure (Basınç) Sütunu Box-Plot Çizimi

Şekil 3.11’de veri setindeki illumination (aydınlık) sütununa ait box-plot çiziminde, ortalama değerler 500 lüks seviyelerinde olmasına karşın 2000 lüks gibi

aykırı veriler olduğu saptanmıştır. Denklem (3.1) kullanılarak aydınlık sütunu için üst sınır değeri 847 olarak belirlenmiş ve bu değerin üstündeki değerler 847 olarak kaydedilmiştir.



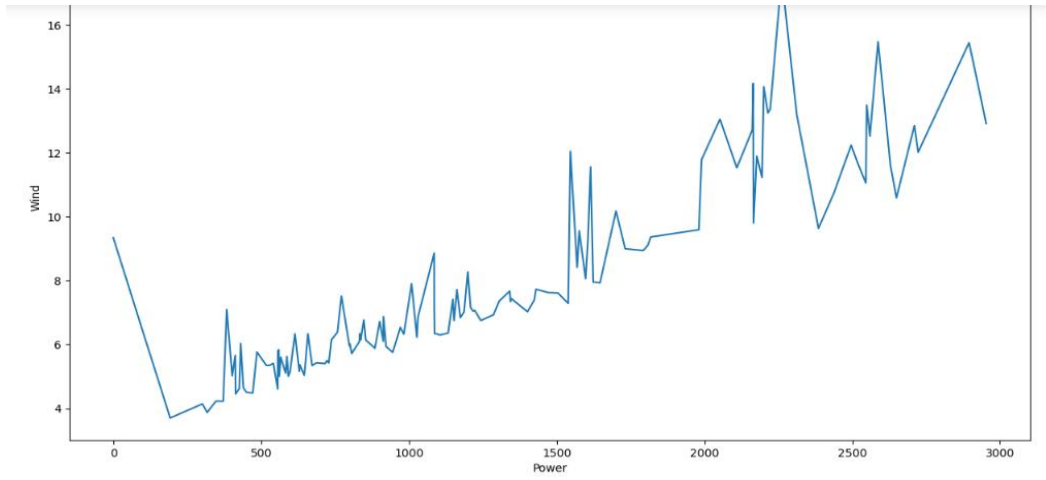
Şekil 3.12 Illumination (Aydınlık) Sütunu Box-Plot Çizimi

Son olarak Şekil 3.13'te Akkuş Res veri setindeki Wind (Rüzgâr hızı), Airpressure (hava basıncı), Airhumidity (hava nem değeri), Reactivepower (reaktif güç), Rotationspeed (kanat dönüş hızı) ve illumination (aydınlık) sütun değerleri hedef sütunu olan Power (üretilen elektrik gücü) verilerine daha çok etki ettiği için bu sütunların veri görselleştirmeleri yapılarak verilerde aykırı bir değer olup olmadığı tekrar gözlemlenmiştir. Ayrıca bu sütun değerlerinin hedef sütunu olan "Power" sütunu ile ilişkisi grafiklerle görselleştirilmiştir. Son düzenlemeleri yapılan veri seti ".csv" uzantılı olarak makina öğrenmesi modelimizde kullanılmak üzere kaydedilmiştir.

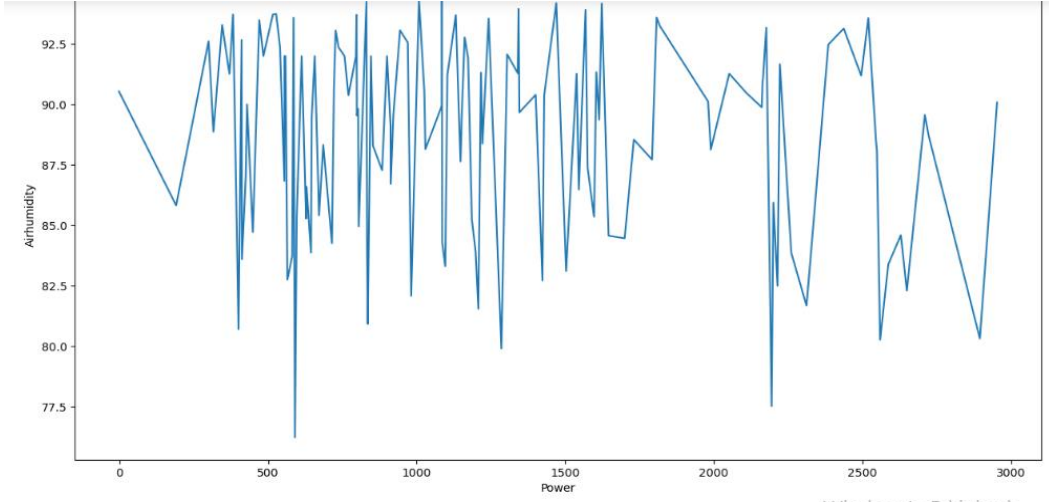


Şekil 3.13 Akkuş Res Veri Setindeki Verilerin Grafiği

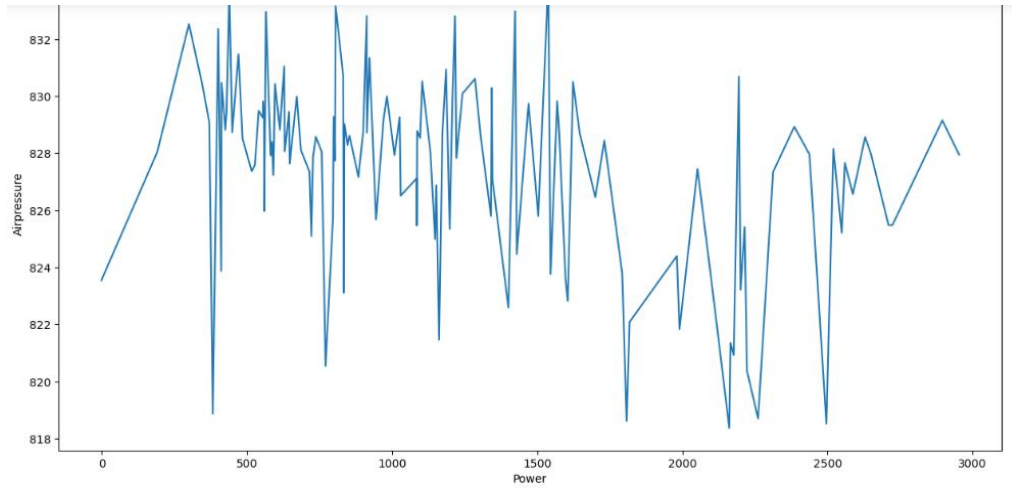
Akkuş RES veri setindeki bazı sütun verileri grafik ile görselleştirilerek aykırı veri veya verilerde herhangi olumsuzluk olup olmadığı gözlenmiştir. Böylece model için uygun hale getirilen veri setinin kullanılabilir olduğuna karar verilmiştir. Modelde kullanılacak olan önemli bazı sütun değerlerinin (Wind, Airhumidity, Airpressure), hedef değer olan Power (güç) değerleriyle olan ilişkisini incelemek amacıyla Şekil 3.14’de, Şekil 3.15’de ve Şekil 3.16’da karşılaştırma grafikleri görselleştirilmiştir. Özellikle Wind (Rüzgâr) değerlerinin Power (Güç) değerleriyle doğrusal olarak değiştiği gözlemlenmektedir. Böylelikle en çok, rüzgâr hızının üretilen güç değerini etkilediği gözlemlenmiştir.



Şekil 3.14 Wind (Rüzgâr) - Power (Güç) Karşılaştırma Grafiği



Şekil 3.15 Airhümidity (Nem) - Power (Güç) Karşılaştırma Grafiği



Şekil 3.16 Airpressure (Basınç) – Power (Güç) Karşılaştırma Grafiği

3.4.3 Veri Tahmin Modellemesi

Veri modellemede ilk olarak verinin belirli bir çerçevesi yapılmıştır. Burada 24 saatlik veriyi alarak 25. Saatteki veri tahmin edilmeye çalışılmıştır. Modelde kullanılan aktivasyon fonksiyonu ve diğer fonksiyonlar bu aşamada belirlenmiştir. Verinin %80'i eğitim için, %20'si ise test için ayrılmıştır. Öğrenme 10 dk arayla 10 aşamada yapılmıştır. Çalışmada "Power", "Wind", "Airpressure", "Airhumidity", "Reactivepower", "Rotationspeed", "illumination" sütunlarındaki veriler kullanılarak öğrenme sağlanmıştır.

İkinci aşamada, Python programlayıcının veri setindeki verilerle yapacağı işlemin kolay olması için veriler ölçeklendirilmiştir. Bu normalizasyon işlemi Python kütüphanelerinde bulunan MinMaxScaler ile yapılmıştır. Veriler 0 ile 1 arasında değerlere normalize edilen bu değerler Şekil 3.17’de gösterilmiştir.

```
scaled_res_veri
```

Time	Power	Wind	Airpressure	Airhumidity	Reactivepower	Rotationspeed	illumination
2020-10-16 23:20:00	0.207414	0.168394	0.705882	0.656604	0.560825	0.670683	0.353428
2020-10-16 23:30:00	0.224308	0.178756	0.705882	0.656604	0.564948	0.689960	0.353428
2020-10-16 23:40:00	0.227827	0.178756	0.705882	0.656604	0.564948	0.693976	0.353428
2020-10-16 23:50:00	0.207884	0.170984	0.705882	0.656604	0.564948	0.670683	0.353428
2020-10-17 00:00:00	0.229939	0.178756	0.705882	0.656604	0.571134	0.696386	0.353428
...
2023-04-16 23:10:00	0.859456	0.290155	0.676471	0.716981	0.781443	0.967068	0.353428
2023-04-16 23:20:00	0.958001	0.318653	0.647059	0.716981	0.872165	0.976707	0.353428
2023-04-16 23:30:00	0.750352	0.261658	0.676471	0.716981	0.637113	0.940562	0.353428
2023-04-16 23:40:00	0.666588	0.248705	0.676471	0.716981	0.593814	0.927711	0.353428
2023-04-16 23:50:00	0.734397	0.251295	0.676471	0.716981	0.674227	0.946988	0.353428

392712 rows x 7 columns

Şekil 3.17 Normalize Edilmiş Akkuş Res Veri Seti Görseli

Normalize edilen bu veriler ise Python programlayıcının işlem yapabilmesi için supervised learning denilen işleme tabii tutulmuştur. Supervised learning aşamasında kaç tane verinin giriş verisi kaç tane verinin çıkış verisi olacağı tanımlanmış ve Şekil 3.18’de veri seti görselinden görüleceği üzere, 24 saatlik veriler döngüyle öğrenme işlemine ilave edilmiştir. Herbir sütundaki 24 saat önceki veriler alınmış ve her sütunun 25’inci saatteki verisi ve enson sütunda ise ‘Power’ değerinin tahmin edilmesi sağlanmıştır.

```
reframed_res_veri.head()
```

Time	var1(t-24)	var2(t-24)	var3(t-24)	var4(t-24)	var5(t-24)	var6(t-24)	var7(t-24)	var1(t-23)	var2(t-23)	var3(t-23)	...	var6(t-2)	var7(t-2)	var1(t-1)	var2(t-1)	var3(t-1)
2020-10-17 03:20:00	0.207414	0.168394	0.705882	0.656604	0.560825	0.670683	0.353428	0.224308	0.178756	0.705882	...	0.512450	0.353428	0.056546	0.090674	0.705882
2020-10-17 03:30:00	0.224308	0.178756	0.705882	0.656604	0.564948	0.689960	0.353428	0.227827	0.178756	0.705882	...	0.512450	0.353428	0.043641	0.069948	0.676471
2020-10-17 03:40:00	0.227827	0.178756	0.705882	0.656604	0.564948	0.693976	0.353428	0.207884	0.170984	0.705882	...	0.467470	0.353428	0.041999	0.056995	0.676471
2020-10-17 03:50:00	0.207884	0.170984	0.705882	0.656604	0.564948	0.670683	0.353428	0.229939	0.178756	0.705882	...	0.384739	0.353428	0.042234	0.046632	0.676471
2020-10-17 04:00:00	0.229939	0.178756	0.705882	0.656604	0.571134	0.696386	0.353428	0.224308	0.168394	0.705882	...	0.336546	0.353428	0.041999	0.056995	0.676471

5 rows x 169 columns

Şekil 3.18 Çerçevenilmiş Akkuş Res Veri Seti Görseli

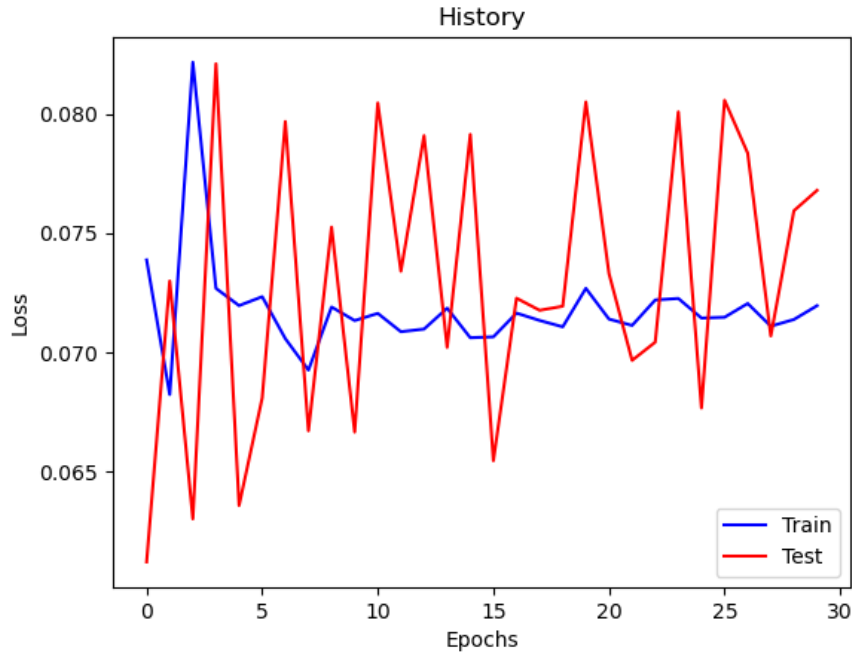
Daha sonra girdi ve çıktı deęerleri belirlenmiřtir. Son stn yani ‘‘Power’’ stn hari tm stnlar girdi deęiřkeni, son stn olan ‘‘Power’’ stn ise ıktı deęiřkeni olarak tanımlanmıřtır.

nc ařamada Modelleme yapılmıřtır. Modelleme ařamasında Simple RNN, LSTM ve GRU makine ęrenmesi modelleri uygulanmıřtır.

4. ARAŞTIRMA BULGULARI

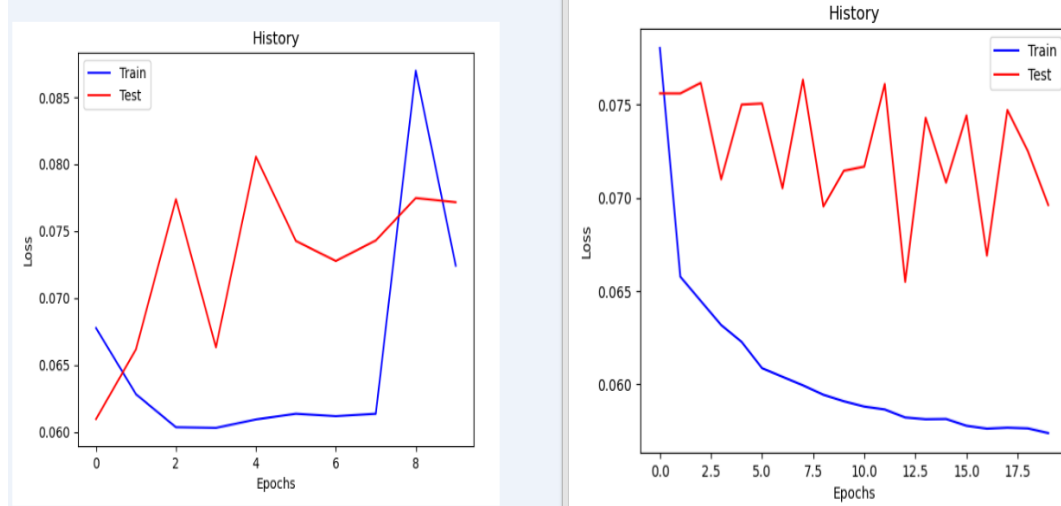
Bu tez çalışmasında, Akkuş rüzgâr enerji santralindeki elektrik üretim veri seti kullanılarak bir veri madenciliği uygulaması yapılmıştır. Veri madenciliği uygulamasıyla yapılan veri analizi sayesinde veri setindeki eksik ve yanlış (anlamsız) değerler ortalama değerlerle doldurularak daha doğru tahmin yapılabilmesi için düzenli hale getirilmiştir. Özellikle aykırı (uç) veri analizinde bir çok sütun değerinde sonucu olumsuz etkileyen, anlamlı sayılabilecek değer aralığının çok uzağında değerler olduğu saptanmış ve bu değerler uygun sınır değerlerine çekilmiştir. Makine öğrenmesi metotları kullanılırken tez çalışmasına uygun, uygulanabilir olan metotların tespiti yapılmıştır. Veri setimiz çok değişkenli giriş verilerinden ve değişken hedef (çıkış) verilerinden oluştuğu için hedef değişkenin sabit bir değer olmasına dayanan Regresyon ve Kümeleme uygulamalarının uygun olmadığı saptanmıştır. Bunun yerine RNN, LSTM, GRU vb. gibi katmanlı ve çok kapılı denetimli sinir ağları modellerinin daha uygulanabilir olduğuna karar verilmiştir. Girişteki her bir farklı değere karşılık farklı bir hedef değer oluşmaktadır.

Simple RNN, LSTM ve GRU makine öğrenmesi modelleri uygulamasında tanımlanan giriş ve çıkış değerleri $x_{train}(314149, 24, 7)$, $y_{train}(314149)$, $x_{test}(78538, 24, 7)$, $y_{test}(78538)$ olmak üzere train ve test verilerine ayrılmıştır.



Şekil 4.1 Simple RNN Train ve Test Veri Grafiği

Şekil 4.1’de ve Şekil 4.2’de train ve test verilerinin grafiği görselleştirilerek özellikle Simple RNN modelinde daha uyumlu varyasyon olduğu gözlenmiştir. GRU ve LSTM modellerinde bu uyumun daha az olduğu görülmektedir.



Şekil 4.2 GRU-LSTM Train ve Test Veri Grafiği

Model aşamaları (epochs) Simple RNN modelinde 30, LSTM modelinde 20 ve GRU modelinde 10 aşamalı olarak uygulandığında en verimli sonucun elde edildiği görülmüştür. Model çerçevesi aynı tutularak sadece Simple RNN, LSTM ve GRU farklı modeller uygulanarak sonuçlar gözlemlenmiştir. Farklı modeller uygulandığında parameter sayılarında değişiklik olduğu görülmektedir.

```
model.summary()
Model: "sequential"
-----
```

Layer (type)	Output Shape	Param #
simple_rnn (SimpleRNN)	(None, 24, 30)	1140
dropout (Dropout)	(None, 24, 30)	0
simple_rnn_1 (SimpleRNN)	(None, 30)	1830
dropout_1 (Dropout)	(None, 30)	0
dense (Dense)	(None, 1)	31

```
-----
Total params: 3,001
Trainable params: 3,001
Non-trainable params: 0
-----
```

Şekil 4.3 Simple RNN Model Özeti Görseli

Şekil 4.3’de ve Şekil 4.4’de uygulanan modellerin özetleri görülmektedir. Buna göre, modeller iki katmanlı olarak uygulanmıştır. Öğrenme verilerin yarısı ilk katmanada kalan yarısı ikinci katmanda uygulanmıştır. Train (öğrenme) ve test verileri görselleştirme yapılarak grafik üzerinde karşılaştırılmış ve öğrenme ve test verilerinin uyumlu olduğu gözlenmiştir.

Layer (type)	Output Shape	Param #
gru (GRU)	(None, 24, 30)	3510
dropout (Dropout)	(None, 24, 30)	0
gru_1 (GRU)	(None, 30)	5580
dropout_1 (Dropout)	(None, 30)	0
dense (Dense)	(None, 1)	31
Total params: 9,121		
Trainable params: 9,121		
Non-trainable params: 0		

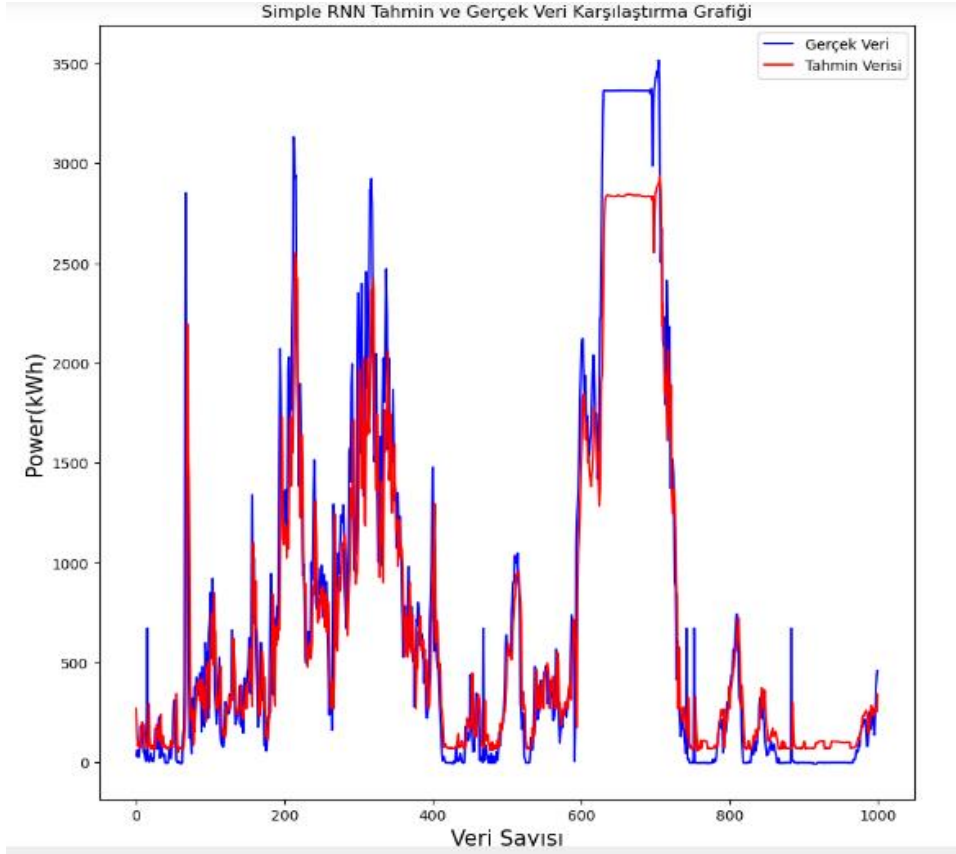
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 24, 30)	4560
dropout (Dropout)	(None, 24, 30)	0
lstm_1 (LSTM)	(None, 30)	7320
dropout_1 (Dropout)	(None, 30)	0
dense (Dense)	(None, 1)	31
Total params: 11,911		
Trainable params: 11,911		
Non-trainable params: 0		

Şekil 4.4 GRU-LSTM Model Özeti Görseli

Normalize edilerek tahmin uygulaması için daha kolay hale getirilen veri seti gerekli çerçeveler yapılarak modellenmiştir. Tahmin uygulamasında her modelin uygulanması sırasında farklı aşama (epochs) sayıları uygulanmıştır. Simple RNN modelinde epochs 30 için en uygun sonucu vermiştir. 40, 50 gibi daha yüksek epochs uygulandığında tahmin değerlerinde çok değişim olmadığı ancak programın sonuç verme süresinin çok arttığı gözlenmiştir. 10, 15 gibi daha az epochs uygulandığında ise tahminden sapmaların arttığı gözlenmiştir. LSTM ve GRU modellerinde ise epochs 10 olarak uygulandığında en uygun sonuçların elde edildiği gözlenmiştir. Son aşamada ise tahmin edilen veriler hangi doğrulukla tahmin yapıldığının tespiti için hata(error) oranı hesaplanması yapılmıştır. Bu hesaplamaların yapılabilmesi için tahmin verileri tekrar MinMaxScaler ile normalize edilerek 0 ile 1 arasındaki değerlere ölçeklendirilmiştir. Hata oranı hesaplanırken Mean Square Error (Ortalama Kareysel Hata) ve Mean Absolute Error (Ortalama Mutlak Hata) metotları kullanılmıştır. Çizelge 4.1’de bu hata oranları modeller için hesaplanmış ve her bir Simple RNN, LSTM, GRU modelleri için ayrı ayrı karşılaştırılarak hangi modelde daha doğru tahmin yapıldığı saptanmıştır.

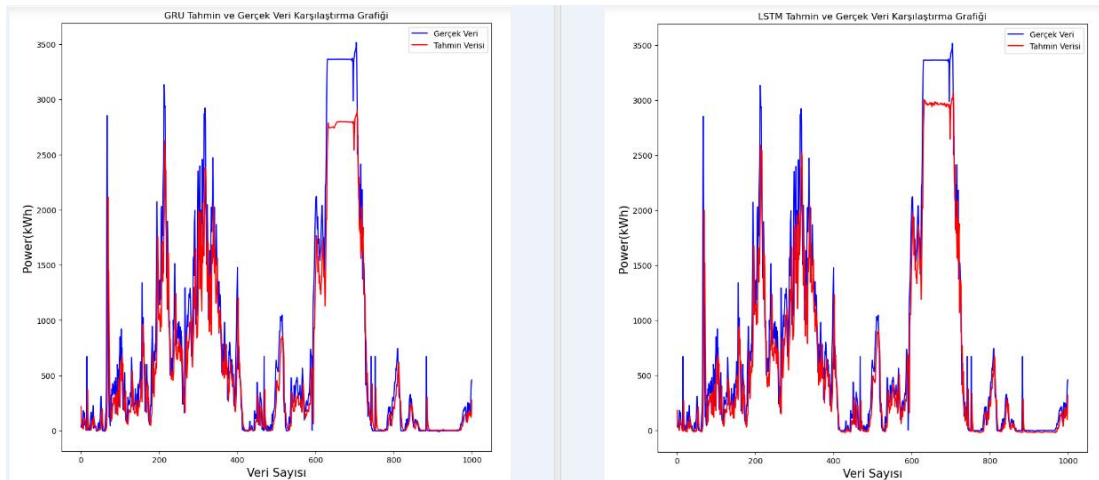
Çizelge 4.1 Modellerin RMSE ve MAE Hata Oranlarının Karşılaştırması

Hata Türü	SimpleRNN(%)	LSTM(%)	GRU(%)
RMSE	4.68	4.38	4.75
MAE	3.27	2.96	3.28

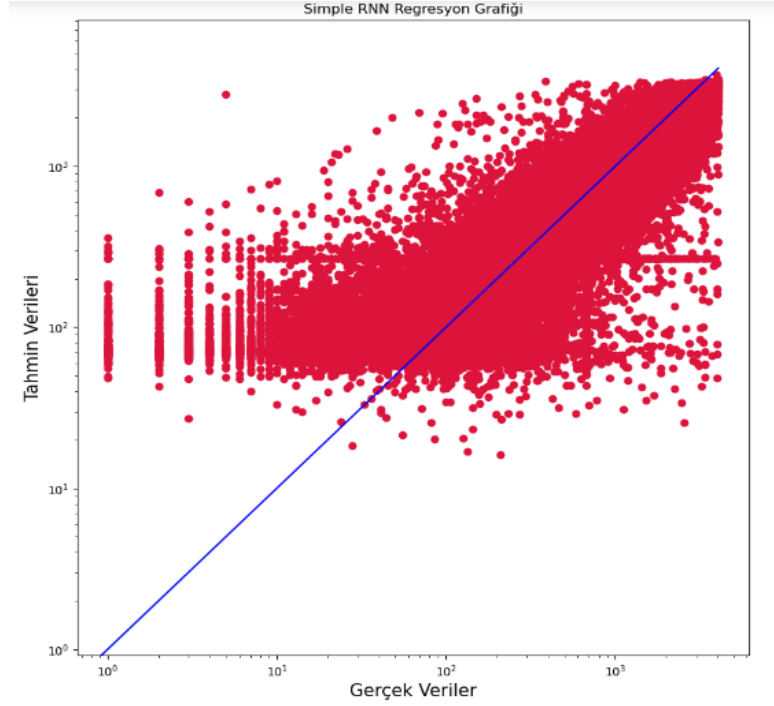


Şekil 4.5 Simple RNN Modeline Ait Gerçek Veri-Tahmin Verisi Grafiği

Şekil 4.5’de Simple RNN modeli uygulanarak elde edilen tahmin verileriyle gerçek verilerin grafiği görselleştirilmiştir. Grafikte ilk 1000 veri gösterilmiştir. Böylelikle tahmin değerlerinin gerçek değerlere oldukça yakın olduğu görülmüş ve doğruya yakın bir tahmin yapıldığı görülmüştür. Şekil 4.6’da ise GRU-LSTM modellerine ait tahmin ile gerçek veri grafiği karşılaştırılmış ve yine gerçeğe yakın doğru tahminler yapıldığı görülmüştür.

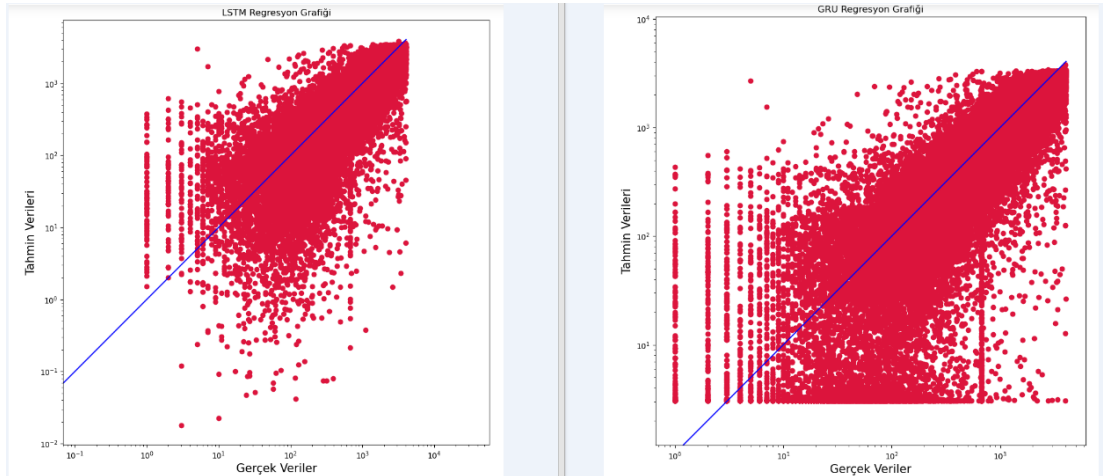


Şekil 4.6 GRU-LSTM Modellerine Ait Gerçek Veri-Tahmin Verisi Grafiği



Şekil 4.7 Simple RNN Modeline Ait Gerçek Veri-Tahmin Verisi Regresyon Grafiği

Tahmin verileriyle gerçek veriler karşılaştırıldığında her üç modelinde iyi sonuç verdiği görülmüştür. Şekil 4.7’de modellere ait regresyon grafiği görselleştirilmiştir. Regresyon grafiğinde, hangi modelin daha iyi sonuç verdiğini saptamak için tahmin ve gerçek veriler matris olarak işaretlenerek 45° eğimli bir $y=x$ doğrusu üzerinde yoğunlaşması ve bu doğrunun da orijin noktasına oldukça yakın geçmesi istenmektedir. Böylelikle tahmin değerlerinin gerçek değerlere yaklaştığı söylenebilecektir. Şekil 4.7’de ve Şekil 4.8’de regresyon grafikleri incelendiğinde, en iyi sonucun SimpleRNN modelinde ortaya çıktığı saptanmıştır.



Şekil 4.8 GRU-LSTM Modellerine Ait Gerçek Veri-Tahmin Verisi Regresyon Grafiği

5. TARTIŞMA ve SONUÇ

Bu tezde, Türkiye'deki Ordu ili Akkuş ilçesindeki Rüzgâr Enerji Santrali(RES)'e ait 392712 satır ve 27 sütundan oluşan, 16.10.2020 ile 16.04.2023 tarihleri arasındaki 10 'ar dakika aralıklarla ölçülen değerlerin yer aldığı ve üretilen aktif güç, reaktif güç, rüzgâr hızı, nem, kanat pozisyonları, hava basıncı vb. sütun değerlerini bulunduran bir veri seti kullanılmıştır. Veri seti, veri madenciliği metotları kullanılarak veri ön işleme ve aykırı (uç) veri analizi yapılarak makine öğrenmesi için hazır hale getirilmiştir. Makine öğrenmesi modelleri uygulanmadan önce model çerçevesi belirlenmiş ve veri setindeki değerler normalizasyon yöntemiyle 0 ile 1 arasında değerlere ölçeklendirilmiştir. Bu sayede, python programlayıcının makine öğrenmesi modellerini hızlı ve daha kolay uygulaması sağlanmıştır. Verilerin %80'i train (öğrenme) ve %20'si test verisi olarak ayrılmıştır. Veri seti lineer olmayan, karmaşık ve sayıca çok verilerden oluştuğu için makine öğrenmesi modellerinden Simple RNN, GRU ve LSTM modelleri uygulanarak sonuçlar karşılaştırılmıştır. Tahmin verileri gerçek verilerle karşılaştırılmış, hata oranları Ortalama Karesel Hata (RMSE) ve Ortalama Mutlak Hata (MAE) yöntemleri kullanılarak hesaplanmıştır. Sonuçlar karşılaştırıldığında, Simple RNN modelinin daha gerçeğe yakın sonuç verdiği saptanmıştır. Profesyonel yazılımlar ve makine öğrenmesi modellerinin daha da geliştirilmesiyle tahminlerdeki bu doğruluk oranı çok daha yüksek olacaktır.

Çalışmanın amacı, gelecek bir zamandaki rüzgâr hızı değerlerini geçmişe göre oldukça doğru tahmin etmeye çalışmaktır. Rüzgâr hızı zaman serileri, doğrusal ilişkinin yanı sıra doğrusal olmayan ilişkiler de içermektedir. Bu sebeple, sadece geçmiş değerlere bakılarak ileriye yönelik tahminlerde bulunmak, girdiler arasındaki karmaşık ilişkiler açısından oldukça zordur. Çalışmada, ileriye yönelik tahmini oldukça zor olan, sadece belli sınırlar içerisinde öngörülebilir ve rüzgâr yönü, basınç, nem, aydınlık gibi meteorolojik bir çok diğer değişkenden doğrudan etkilenen rüzgâr hızına ilişkin geçmiş değerlere göre tahmin yapılmıştır. Rüzgâr enerjisi güç tahmini için makine öğrenmesi yöntemlerinden LSTM, GRU ve SimpleRNN modelleri kullanılarak modellerinin performansları analiz edilmiş ve rüzgâr enerjisi güç tahmini için kullanılabilirlikleri incelenmiştir.

Tez çalışmasında, gerçeğe yakın güç üretim tahmini yapılmış ancak, veri setinde boş değerlerin ve aykırı verilerin sayısının fazla olması gerçek girdi verilerini sağlamadığından bu değerlerin sonuçtaki hata oranını artırdığı görülmüştür. Çünkü veri analizi yapılırken bu boş ve aykırı değerlerin yerine yine tahmine dayalı değerler girdi olarak verilmektedir. Veri setinin daha doğru girdi sağlaması durumunda sonucun da daha gerçeğe yakın olacağı değerlendirilmektedir.

Bu çalışmayla, bir rüzgâr enerji santralının elektrik üretiminin planlanmasında Python programlayıcı ve kütüphaneleri yardımıyla, makine öğrenmesi metotları kullanılarak gerçeğe yakın olarak tahmin edilebildiği gösterilmiştir. Bu sayede günlük, aylık ve yıllık elektrik üretim planlamasının yapılmasında kolaylık sağlanabilmektedir. Ayrıca sistemin ürettiği reaktif güç de tahmin edilebileceğinden şebeke üzerindeki fiderlerde yapılması muhtemel reaktif güç kompanzasyonu için de önceden tahmin verisi sağlanabileceği değerlendirilmektedir. Ulusal şebeke ağına anlık satılacak güç miktarı veya şebekeden satın alınması gereken güç miktarı için önceden tahmin verisi sağlanabilecektir. Bu sayede, ulusal elektrik güç ağına smart grid (akıllı şebeke) entegrasyonunda çok önemli bir kolaylık sunulabilecektir.

6. KAYNAKLAR

- Aksoy, F. (2018). Short-term forecasting of wind power production using machine learning and deep learning methods. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik Mühendisliği, İstanbul.
- Anonim, (2022). Rüzgâr enerjisi nedir? Rüzgâr türbinleri nasıl enerji üretir?. Aydem Perakende. [https://www.aydemperakende.com.tr/blog/ruzgar-enerjisi-nedir-ruzgar-turbinleri-enerji-uretimi-\(Erisim-Tarihi:24.12.2023\)](https://www.aydemperakende.com.tr/blog/ruzgar-enerjisi-nedir-ruzgar-turbinleri-enerji-uretimi-(Erisim-Tarihi:24.12.2023)).
- Anonim, (2023). Türkiye’de elektrik enerjisi üretimi ve yenilenebilir enerji kaynaklarının mevcut durumu. Kilis Üniversitesi. [http://mmf.kilis.edu.tr/-\(Erisim-Tarihi:20.12.2023\)](http://mmf.kilis.edu.tr/-(Erisim-Tarihi:20.12.2023)).
- Anonim, (2023a). Bilgi merkezi(enerji). Enerji ve Tabii Kaynaklar Bakanlığı(ETBK). [https://enerji.gov.tr/bilgi-merkezi-enerji-elektrik-\(Erisim-Tarihi:24.12.2023\)](https://enerji.gov.tr/bilgi-merkezi-enerji-elektrik-(Erisim-Tarihi:24.12.2023)).
- Anonim, (2023b). Yenilenebilir enerji kaynakları. Makina Mühendisleri Odası(MMO).[http://www1.mmo.org.tr/resimler/dosya_ekler/9514e888b8f2aca_ek.pdf-\(Erisim-Tarihi:20.06.2023\)](http://www1.mmo.org.tr/resimler/dosya_ekler/9514e888b8f2aca_ek.pdf-(Erisim-Tarihi:20.06.2023)).
- Cantürk, S. (2018). Bir rüzgâr çiftliğinden yapay sinir ağlarıyla kısa süreli elektrik üretim tahmini. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Temiz Tükenmez Enerjiler, Ankara.
- Çift, M. (2023). Rüzgâr türbininde üretilen gücün yapay sinir ağı kullanılarak modellenmesi. Yüksek Lisans Tezi, Osmaniye Korkut Ata Üniversitesi, Fen Bilimleri Enstitüsü, Enerji Sistemleri Mühendisliği Ana Bilim Dalı, Osmaniye.
- Çolaker, AH., Özyön, S. & Gün, A. (2020). Afyonkarahisar ili Dinar bölgesi için rüzgâr enerjisinden elektrik üretiminde yapay sinir ağları modelinin uygulanması. *Journal of Scientific Reports-B*, 2,29-38.
- Datta, PK. (2018). An artificial neural network approach for short-term wind speed forecast. Kansas State University, Department of Electrical and Computer Engineering, Manhattan, Kansas.
- Demolli, H. (2020). Makine öğrenmesi teknikleri kullanarak yenilenebilir enerji güç üretim tahmini ve optimum hibrit güç sistemi tasarımı. Doktora Tezi, Niğde Ömer Halisdemir Üniversitesi, Fen Bilimleri Enstitüsü, Makine Mühendisliği, Niğde.
- Diri, B. (2014). Makine öğrenmesine giriş. Bilgisayar Mühendisliği, Yıldız Teknik Üniversitesi, İstanbul.
- Duehee Lee, RB. (2014). Short-term wind power ensemble prediction based on gaussian processes and neural networks. *IEEE Trans Smart Grid*, 5(1), 501-510.
- Erden, C. (2023). Derin öğrenme ve ARIMA yöntemlerinin tahmin performanslarının kıyaslanması: bir borsa İstanbul hissesi örneği. *Yönetim ve Ekonomi*, 419-438.
- Farah, S., A, WD., Humaira, N., Aneela, Z. & Steffen, E. (2022). Short-term multi-hour ahead country-wide wind power prediction for Germany using gated recurrent unit deep learning. *Renewable and Sustainable Energy Reviews*, 167, 1-14.

- Görgel, P. & Kavlak, E. (2020). Uzun kısa süreli hafıza ve evrimsel sinir ağları ile rüzgâr enerjisi üretim tahmini. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 11(1),69-80.
- İnal, S. (2022). Yapay zeka ile hidroelektrik enerji santralleri baraj rezervuarlarının işletilmesi. Yüksek Lisans Tezi, Ordu Üniversitesi, Fen Bilimleri Enstitüsü, Yenilenebilir Enerji Bölümü, Ordu.
- Karaali, B. (2023). Makine öğrenmesi algoritmaları ile rüzgâr türbinlerinin güç tahmini. Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Ana Bilim Dalı, İzmir.
- Karagöl, ET. & Kavaz, İ. (2017). Dünyada ve Türkiye'de yenilenebilir enerji. Siyaset, Ekonomi ve Toplum Araştırmaları Vakfı(SETA). İstanbul.
- Karakuş, C. (2023). Makine öğrenmesine giriş. Ckk Ders Notları. <https://ckk.com.tr/ders.html>-(Erişim Tarihi: 02.03.2023).
- Kaysal, K., Akarşlan, E. & Hocaoğlu, FO. (2022). Türkiye kısa dönem elektrik yük talep tahmininde makine öğrenmesi yöntemlerinin karşılaştırılması. *BŞEÜ Fen Bilimleri Dergisi*.
- Keesstra, SD. (2016). The significance of soils and science towards rezlization of the united nations sutainable development goals. . *Soil*, 21(2), 133-139.
- Kırbaş, İ. (2018). İstatistiksel metotlar ve yapay sinir ağları kullanarak kısa dönem çok adımlı rüzgâr hızı tahmini. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22(1), 24-38.
- Lee, JC., Draxl, C. & Berg, LK. (2022). Evaluating wind speed and ower forecasts for wind energy applications using an open-source and systematic validation framework. *Renewable Energy*, 200, 457-475.
- Ma, Z. & Mei, G. (2022). A hybrid attention-based deep learning approach for wind power prediction. *Applied Energy*, 323(1), 1-33.
- Meng, A., Zhu, Z., Deng, W., Ou, Z., Lin, S., Wang, C. & Xu, X. (2022). A novel wind power prediction approach using multivariate variational mode decomposition and multi-objective crisscross optimization based deep extreme learning machine. *Energy*, 260, 1-21.
- Özkay, B. (2021). Derin öğrenme kullanılarak bilecik ili rüzgâr hızı tahmini. Yüksek Lisans Tezi, Bilecik Şeyh Edebali Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Bölümü, Bilecik.
- Özsağır, M., Erden, C., Bol, E., Sert, S. & Özocak, A. (2022). Machine learning approaches for prediction of fine-grained soils liquefaction. *Computers and Geotechnics*.
- Sæther, BB. (2021). Wind power prediction with machine learning methods in complex terrain areas. UiT The Arctic University of Norway, Faculty of Science and Technology, Department of Physics and Technology, Norway.
- Shams, MH., Niaz, H., Hashemi, B., Liu, J. J., Siano, P. & Anvari-Moghaddam, A. (2021). Artificial intelligence-based prediction and analysis of the oversupply

- of wind and solar energy in power systems. *Energy Conversion and Management*, 250, 1-13.
- Tekin, P. (2022). Çukurova bölgesi için kısa vadeli yapay zeka tabanlı rüzgâr güç tahmini. *Çukurova Üniversitesi Mühendislik Fakültesi Dergisi*, 37(4),1143-1153.
- Tekinay, Ç. (2022). Türkiye'nin günlük rüzgâr enerjisi üretiminin derin öğrenme ile modellenmesi. Yüksek Lisans Tezi, Ankara Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Ana Bilim Dalı, Ankara.
- Tubulekas, A. (2022). Exploring machine learning techniques for short-term wind power forecasting of multiple wind parks. Uppsala University, Department of Information Technology, Uppsala.
- Türkoğlu, NŞ. (2016). Effects of climate changes on phonological periods of apple cherry and hheat in Turkey. *Journal of Human Sciences*, 13(1), 1036-1057.
- Varol, D. (2021). Python ile kodlama II python bilim paketi numpy (numeric/sayısal python) scipy (scientific/bilimsel python) matplotlib (grafik kütüphanesi) pandas (veri bilimi modülü). *Research Gate*.
- Varol, NA. (2019). Olasılıksal programlama ile enerji üretiminin güneş enerji santrallerinde (GES) tahmini. Yüksek Lisans Tezi, Gazi Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik Elektronik Mühendisliği, Ankara.
- Wang, S., Li, B., Li, G., Yao, B., & Wu, J. (2021). Short-term wind power prediction based on multidimensional data cleaning and feature reconfiguration. *Applied Energy*, 292, 1-12.
- Yağmur, E., & Yağmur, S. (2022). Rüzgâr gücü tahmininde genetik algoritma ile öznitelik seçimi. *Afyon Kocatepe University Journal of Science and Engineering*, 055103,1028-1040.
- Yavuzaslanoğlu, EG. (2016). Distribution of the entamopathogenic nematods in apple growing areas of Karaman Turkey,. *Pakistan Journal of Nematology*, 34(1), 53-62.
- Yazıcı, K. (2021). Makine öğrenmesi yöntemleri kullanılarak kısa dönem rüzgâr gücü tahmini. Yüksek Lisans Tezi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği, Sakarya.
- Yoldaş, İ. S. (2022). Short-term wind speed and power forecasting: a comprehensive case study for three operational wind farms. Yüksek Lisans Tezi, İzmir Teknoloji Enstitüsü, Mühendislik Bölümü, İzmir.

ÖZGEÇMİŞ

Kişisel Bilgiler	
Adı Soyadı	MUHSİN BAKDEMİR
Doğum Yeri	
Doğum Tarihi	
Uyruğu	<input checked="" type="checkbox"/> T.C. <input type="checkbox"/> Diğer:
Telefon	
E-Posta Adresi	
Eğitim Bilgileri	
Lisans	
Üniversite	Karadeniz Teknik Üniversitesi(KTÜ)
Fakülte	Mühendislik Fakültesi
Bölümü	Elektrik Elektronik Mühendisliği
Mezuniyet Yılı	2012
Yüksek Lisans	
Üniversite	Ordu Üniversitesi
Enstitü	Fen Bilimleri Enstitüsü
Anabilim Dalı	Yenilenebilir Enerji
Mezuniyet Yılı	2024
Yayınlar	
Bakdemir, M., Kırca, G., Turan, S., Dabak, G.,(2012). Lisans Tezi, “Asenkron generatörlü rüzgâr türbini”, Elektrik Elektronik Mühendisliği, Karadeniz Teknik Üniversitesi, Trabzon.	