T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

TÜRKİYE ELEKTRİK TÜKETİMİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE TAHMİNİ

Deniz HERSEK

YÜKSEK LİSANS TEZİ

İstatistik Anabilim Dalı

İstatistik Programı

Danışman

Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK

Şubat, 2024

T.C.

YILDIZ TEKNİK ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

TÜRKİYE ELEKTRİK TÜKETİMİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE TAHMİNİ

Deniz HERSEK tarafından hazırlanan tez çalışması 09.02.2024 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, İstatistik Programı **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK, Danışman
Yıldız Teknik Üniversitesi _____

Prof. Dr. Fatma NOYAN TEKELİ, Asıl Üye
Yıldız Teknik Üniversitesi _____

Doç. Dr. Seher ARIKAN TEZERGİL, Asıl Üye
Marmara Üniversitesi _____

Danışmanım Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK sorumluluğunda tarafımca hazırlanan "Türkiye Elektrik Tüketiminin Makine Öğrenmesi Yöntemleri ile Tahmini" başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Deniz HERSEK

İmza

Canım aileme,

TEŞEKKÜR

Lisans ve yüksek lisans eğitimim boyunca tüm akademik bilgilerini ve deneyimlerini bana aktaran, tez çalışmam boyunca desteklerini hiç esirgemeyen, birlikte çalışmaktan mutluluk duyduğum saygı değer danışman hocam Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK'e teşekkürlerimi sunarım.

Eğitim öğretim hayatım boyunca en büyük fedakârlığı gösteren annem İdil Hersek, babam Mahmut Hersek, anneannem Pamuk Avcı, canım ablam İrem Selen'e, tez sürecimde kaybettiğim beni sevgiyle büyüten babaannem Sevgi Hersek'e, her daim sabır ve anlayışıyla yanımda olan erkek arkadaşım Melik Özan'a ve beni destekleyen arkadaşlarıma teşekkür ederim.

Deniz HERSEK

İÇİNDEKİLER

TEŞEKKÜR iv
İÇİNDEKİLER v
SİMGE LİSTESİ vii
KISALTMA LİSTESİ viii
ŞEKİL LİSTESİ ix
TABLO LİSTESİ x
ÖZET xi
ABSTRACT xiii
1 GİRİŞ 15
1.1 Literatür Araştırması 15
1.2 Tezin Amacı 16
1.3 Hipotez 16
2 ELEKTRİK TÜKETİMİ 18
2.1 Elektrik Tüketimi Nedir? 18
2.1.1 Dünya Genelinde Elektrik Tüketimi 18
2.1.2 Türkiye'de Elektrik Tüketimi 19
3 KURAMSAL TEMELLER 20
3.1 Yapay Zekâ 20
3.2 Makine Öğrenmesi 20
3.2.1 Makine Öğrenim Türleri 21
3.2.1.1 Denetimli Öğrenim 21
3.2.1.2 Denetimsiz Öğrenim 21
3.2.1.3 Yarı Denetimli Öğrenim 22
3.2.1.4 Pekiştirmeli Öğrenim 22
3.3 Derin Öğrenme 22
3.4 Makine Öğrenmesi ile Derin Öğrenme Arasındaki Temel Farklar 23
3.5 LSTM, XGBoost, RNN, GRU ve ANN Yöntemleri 24
3.5.1 Uzun Kısa Vadeli Bellek (LSTM) 25
3.5.2 Ekstrem Gradyan Arttırma (XGBoost) 25

3.5.3 Tekrarlayan Sinir Ağları (RNN) 26
3.5.4 Kapılı Tekrarlayan Birimler (GRU) 27
3.5.5 Yapay Sinir Ağları (ANN) 27
3.6 Model Performanslarının Karşılaştırılması 29
4 UYGULAMA **30**
4.1 Türkiye'de Elektrik Tüketimi Üzerine Bir Uygulama 30
4.1.1 Veri Manipülasyonu ve Görselleştirmesi 30
4.1.2 Regresyon Analizi 35
4.1.3 Tüketim Miktarı (MWh) Tahmini 40
4.1.4 Hızlı Sonuç Alma (Train, Dev, Test Seti Oluşturma) 40
4.1.5 Model Eğitimi ve Tahmini 41
4.1.6 Yöntemlerin Model Eğitimindeki (Train Verisi) Hiperparametre Başarıları 44
4.1.7 Yöntemlerin En Doğru Sonucu Veren Hiperparametler Seçilerek Test Verisi Tahmin Başarıları 46
4.1.8 Gelecek Değer Tahmini 49
5 SONUÇ **52**
KAYNAKÇA **53**
TEZDEN ÜRETİLMİŞ YAYINLAR **55**

SİMGE LİSTESİ

Y_i	Gerçek Değerler
$\bar{Y}_i$	Tahmin Değerleri
λ, γ	XGBoost modeli katsayıları

KISALTMA LİSTESİ

ADF	Augmented Dickey-Fuller
AI	Artifical Intelligence
ANN	Artifical Neural Network
BOİ	Biyolojik Oksijen İhtiyacı
CART	Classification and Regression Tree
EPDK	Enerji Piyasası Düzenleme Kurumu
GRU	Gated Recurrent Unit
JB	Jarque-Bera Testi
KOİ	Kimyasal Oksijen İhtiyacı
kWH	Kilowatt Saat
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MSE	Mean Squared Error
MWh	Megawatt Saat
OLS	Ordinary Least Squares
RMSE	Root Mean Square Error
RNN	Recurrent Neural Networks
SL	Denetimli Öğrenim
SVR	Support Vector Regression
TUIK	Türkiye İstatistik Kurumu
UL	Denetimsiz Öğrenim
VIF	Variance Inflation Factor
YSA	Yapay Sinir Ağları

ŞEKİL LİSTESİ

Şekil 3. 1 Makine öğrenimi türleri....21
Şekil 3. 2 Derin öğrenme modeli....23
Şekil 3. 3 LSTM mimari yapısı....25
Şekil 3. 4 RNN mimari yapısı....26
Şekil 3. 5 Genel bir yapay sinir ağı modeli....28
Şekil 4. 1 Elektrik tüketim miktarı veri seti....30
Şekil 4. 2 Yeni sütunların ekli olduğu veri seti....31
Şekil 4. 3 Günün saati zaman serisi grafiği....32
Şekil 4. 4 Haftanın günü zaman serisi grafiği....32
Şekil 4. 5 Ayın günü zaman serisi grafiği....33
Şekil 4. 6 Yılın ayı zaman serisi grafiği....33
Şekil 4. 7 Tüketim miktarı zaman serisi....34
Şekil 4. 8 Yıllık tüketim ortalamaları....34
Şekil 4. 9 Günün saati boxplot grafiği....35
Şekil 4. 10 Regresyon modelinin özet istatistikleri....36
Şekil 4. 11 Augmented dickey-fuller (ADF) testi....37
Şekil 4. 12 Variance inflation factor (VIF) testi....38
Şekil 4. 13 Tüketim miktarı fark grafikleri....38
Şekil 4. 14 Veri kümesi trend, mevsimsellik, gürültü grafikleri....39
Şekil 4. 15 Logaritması alınan tüketim miktarı ve ölçeklendirilen tüm veri....41
Şekil 4. 16 Modelin eğitim aşaması gözlem sayıları....42
Şekil 4. 17 LSTM modeli units değerleri....42
Şekil 4. 18 LSTM modeli hiperparametrelerin performansı....44
Şekil 4. 19 XGBoost modeli hiperparametrelerin performansı....44
Şekil 4. 20 RNN modeli hiperparametrelerin performansı....45
Şekil 4. 21 GRU modeli hiperparametrelerin performansı....45
Şekil 4. 22 ANN modeli hiperparametrelerin performansı....46
Şekil 4. 23 Tüm yöntemlerin test verisi tahmin performansları....46
Şekil 4. 24 XGBoost model tahminleri ve orjinal veri....48
Şekil 4. 25 Türkiye elektrik tüketim tahmini zaman serisi....49
Şekil 4. 26 Gelecek 1000 saat için tahmin sonuçları....49
Şekil 4. 27 Gelecek 1000 saat için tahmin verisi....50
Şekil 4. 28 24Nisan 2022 tahmin değeri....50

TABLO LİSTESİ

Tablo 3. 1 Makine öğrenmesi ve derin öğrenme farklılıkları **24**
Tablo 3. 2 Biyolojik sinir yapısı ile yapay sinir ağı **28**
Tablo 4. 1 Yöntemlerin hata metrikleri **47**
Tablo 4. 2 Gerçek zamanlı tüketim miktarı **51**

ÖZET

TÜRKİYE ELEKTRİK TÜKETİMİNİN MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE TAHMİNİ

Deniz HERSEK

İstatistik Anabilim Dalı

İstatistik Programı

Yüksek Lisans Tezi

Danışman: Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK

Tarih boyunca süregelen insanoğlunun günlük yaşantısında en çok ihtiyaç duyduğu enerji türü elektrik enerjisi olmuştur. Teknoloji, sanayi, üretim, tıp, ulaşım ve haberleşme gibi pek çok alanda elektrik enerjisinden faydalanılmaktadır. Elektrik tüketimi teknolojik ilerlemler, ülkelerin ekonomik büyümesi, endüstrileşme ve nüfus artışı gibi etkenlere bağlı olarak kesintisiz bir artış göstermektedir. Enerji talebinin büyük bir bölümünü oluşturan elektrik tüketimi, modern yaşamın da birçok yönünü desteklemektedir. Elektrik enerjisinin bir ülkenin ekonomi büyümesi üzerinde durulması gereken önemli bir nokta olduğunu ihmal etmemeliyiz. Türkiye, ekonomik yapısı ile dünyanın en hızlı büyüyen ülkelerinden biri olarak değerlendirilmektedir. Ülkemizin varlık içinde yaşama seviyesi, teknoloji, üretim ve sanayi gibi önemli unsurların yükselişi için elektrik tüketimi hakkında fikir sahibi olunması ve elektrik tüketim tahminin yapılması oldukça önem arz etmektedir. Bu çalışmalar enerji politikalarının oluşturulması açısından büyük önem taşımaktadır. Makine öğrenimi algoritmaları bizlere bu yolda büyük deneyimler sağlamaktadır. Literatürde yer alan çalışmalar kapsamında elektrik tüketim tahmini için çeşitli makine öğrenimi algoritmaları kullanılmıştır. Bu çalışma, Türkiye'de 31.12.2015 ile 22.03.2022 tarihleri arasındaki Elektrik

Tüketim Miktarı veri seti üzerinde makine öğrenmesi, derin öğrenme ve istatistiksel analiz tekniklerinin kapsamlı bir analiz incelemesini sunmaktadır. Uygulamada bu teknikler arasında LSTM, XGBoost, RNN, GRU ve ANN kullanılmıştır. Çalışmanın amacı, regresyon analizinde karşılaşılan otokorelasyon veya artıkların normal dağılmaması gibi problemlere çözüm getirebilecek, güçlü teorik bir yapıya sahip metodlar kullanarak başarılı tahminler üretmektir. Modellerin değerlendirilmesi aşamasında, 10 katlı çapraz doğrulama yöntemi kullanılmış ve çeşitli hiperparametrelerle denenmiştir. Çalışmada en düşük MAE (ortalama mutlak hata), MSE (ortalama kare hatası) ve RMSE (kök ortalama kare hatası) hata metriklerine sahip modeler seçilerek, test verisi üzerindeki performansları karşılaştırılmıştır. Çalışma sonucunda her bir modelin kendine özgür başarıları olmuştur. XGBoost ve ANN yöntemleri diğerlerine göre daha etkileyici sonuçlar göstermiştir. XGBoost modeli, MAE, MSE, ve RMSE için sırasıyla %10.8, %1,6 ve %12,8 olarak hesaplanmıştır. Çıktı değerlerine bakıldığında Türkiye'nin elektrik tüketimi tahmini için XGBoost modelinin daha iyi performans verdiği tespit edilmiştir.

Anahtar Kelimeler: Elektrik Tüketim Tahmini, Makine Öğrenme, Zaman Serileri

ABSTRACT

Prediction of Electricity Consumption in Turkiye with Machine Learning Methods

Deniz HERSEK

Department of Statistics

Master of Science

Supervisor: Prof. Dr. Gülhayat GÖLBAŞI ŞİMŞEK

Throughout history, the type of energy that human beings need most in their daily life has been electrical energy. Electrical energy is used in many areas such as technology, industry, production, medicine, transportation and communication. Electricity consumption increases due to factors such as technological advances, economic growth of countries, industrialization and population growth. Electricity consumption, which constitutes a large part of energy demand, also supports many aspects of modern life. We should not neglect that electrical energy is an important point that should be emphasized on the economic growth of a country. Turkey is considered one of the fastest growing countries in the world with its economic structure.It is very important to have an idea about electricity consumption and to estimate electricity consumption for the rise of important elements such as the level of living in our country, technology, production and industry.These studies are of great importance in terms of the creation of energy policies. Machine learning algorithms provide us with great experiences on this path. Within the scope of the studies in the literature, various machine learning algorithms have been used to estimate electricity consumption.

This study presents a comprehensive analysis review of machine learning, deep learning and statistical analysis techniques on the data set of Electricity Consumption between 31.12.2015 and 22.03.2022 in Turkey. In practice, LSTM, XGBoost, RNN, GRU and ANN were used among these techniques. The aim of the study is to produce successful predictions by using methods with a strong theoretical structure that can provide solutions such as autocorrelation or non-normal distribution of residues encountered in regression analysis. During the evaluation of the models, the 10-fold cross-validation method was used and tested with various hyperparameters. In the study, the models with the lowest MAE (mean absolute error), MSE (mean square error) and RMSE (root mean square error) error metrics were selected and their performance on the test data was compared. As a result of the study, each model has had its own successes. XGBoost, and ANN methods have shown more impressive results than others. The XGBoost model was calculated to be 10.8%, 1.6%, and 12.8% for MAE, MSE, and RMSE, respectively.Looking at the output values, it has been determi ned that the XGBoost model gives better performance for Turkey's electricity consumption estimation.

Keywords: Forecasting Electricity Consumption, Machine Learning, Time Series

1
GİRİŞ

1.1 Literatür Araştırması

Literatürde elektrik tüketimini tahmin etmek için makine öğrenimi algoritmaları ve derin öğrenme tekniklerinin uygulandığı, farklı veri ve çeşitli yöntemlerin kullanıma ilişkin çok sayıda çalışmalar yapılmıştır. Elektrik tüketim tahmini ve zamana göre değişim gösteren olayların analizlerine ilişkin literatür taramaları aşağıdaki gibi özetlenmiştir.

Hamzaçebi ve Kutay, Türkiye'nin 2010 yılına kadar olan elektrik enerjisi tüketimini yapay sinir ağları ile tahmin etmişlerdir. Çalışmada yapay sinir ağları, Box-Jenkins modelleri ve regresyon tekniği kullanılmıştır. Elde edilen bulgular sonucunda yapay sinir ağları tekniği elektrik enerjisi tüketiminde daha iyi bir tahmin aracı olduğu sonucuna ulaşılmıştır. [1]

Yavuz ve Deveci, istatistiksel normalizasyon tekniklerinden faydalanarak yapay sinir ağlarının tahmin performansı üzerine etkisi konusunda araştırma çalışması yapmışlardır. D_Min_Max tekniğinin kullanılması sonucu yapay sinir ağları ile elde edilen tahmin sonuçlarının, ham veri setinden elde edilen tahmin değerlerinden daha etkili performans gösterdiğini kabul etmişlerdir. [2]

Ribeiro ve diğ., enerji tüketimini tahmin etmek amacıyla araştırdıkları çalışmada, üç derin öğrenme modeli (LSTM, RNN ve GRU) ve iki makine öğrenme modeli (SVR ve Random Forest)'ni kullanarak GRU modelinin diğer modellerden daha iyi sonuçlar verdiğini göstermişlerdir. [3]

Oğcu ve diğ., çalışmalarında 2010 ve 2011 yıllarındaki elektrik üretimini tahmin etmek amacıyla mevsimsel bir SVR modeli ve YSA modellerini karşılaştırmışlardır. Performans karşılaştırılması yapıldığında mevsimsel SVR modelinin YSA modelini geride bırakarak daha iyi sonuçlar verdiğini tespit etmişlerdir. [4]

Balli, yüksek lisans tez çalışmasında atık su içerisindeki kimyasal oksijen ihtiyacı (KOİ) ve biyolojik oksijen ihtiyacı (BOİ) miktarlarının tahmin edilmesi kapsamında derin öğrenme yöntemlerinden LSTM, RNN ve GRU modellerini

karşılaştırmıştır. Çalışılan modeller içerisinde BOİ parametresini GRU modeli, KOİ parametresini ise LSTM modeli başarılı şekilde tahmin edebilmiştir. [5]

Özdemir, yüksek lisans tez çalışmasında, elektrik talebini ARIMA ve XGBoost yöntemleri kullanarak tahmin etmiştir. Çalışma sonucunda ARIMA ve XGBoost modellerinin benzer sonuçlar verdiğini, sonrasında parametler eklenerek veri durumunun düzeltilmesiyle XGBoost modelinin performansı için kayda değer bir artış gösterdiğini tespit etmiştir. [6]

Makridakis ve diğ., araştırma çalışmasında zaman serisi tahmin analizlerinde istatistiksel yaklaşımlara alternatif olarak makine öğrenimi yöntemlerini önermeyi amaçlamışlardır. Çalışmada LSTM, RNN ve sekiz istatistiksel yaklaşım içeren makine öğrenmesi modellerinin performansları karşılaştırılmıştır. Makine öğrenimi tekniklerinin kullanılması sonucu model performansının daha iyi sonuç gösterdiğini tespit etmişlerdir. [7]

1.2 Tezin Amacı

Bu tez çalışmasının iki ana amacı vardır: biri zaman serileri tahmini yapılırken regresyon analizinde karşılaşılan artıkların normal dağılmaması ya da otokorelasyon problemlerine çözüm getirebilecek güçlü teorik yapıya sahip metodlar kullanarak başarılı tahminler üretmek, diğeri ise gelişen teknoloji ile neredeyse tüm makine ve araçlarda elektrik kullanımının artmasıyla elektrik tüketiminin öngörülebilme ihtiyacını karşılamak ve enerji politikalarının planlanması kapsamında katkıda bulunmaktır. Bu amaç doğrultusunda bu çalışma, 2015-2022 yılları arasında Türkiye'de gerçekleşen elektrik tüketim verisi kullanılarak makine öğrenimi, derin öğrenme ve istatistiksel analiz tekniklerinin kapsamlı bir analiz incelemesini sunmaktadır.

1.3 Hipotez

Bu tez çalışmasında, Türkiye'nin elektrik tüketiminin makine öğrenimi ve derin öğrenme yöntemleri ile tahmin edilebileceği hipotezi savunulmaktadır. Bu kapsamda, on katlı çapraz doğrulama yöntemi kullanılmış ve çeşitli hiperparametreler ile deneme yapılarak hipotezin doğruluğu test edilmiştir. Test verisi üzerindeki performanslar karşılaştırılırken en düşük MAE (Ortalama Mutlak

Hata), MSE (Ortalama Kare Hatası) ve RMSE (Kök Ortalama Kare Hatası) hata metriklerine sahip modeller seçilmiştir.

2
ELEKTRİK TÜKETİMİ

2.1 Elektrik Tüketimi Nedir?

Yaşamımızı devam ettirebilmek ve her türlü alanda ilerleyebilmek için kesinlikle elektrik enerjisine ihtiyaç duymaktayız. Durdurulamaz bir hızla gelişen teknoloji ile birlikte elektrik enerjisinin kullanım miktarı ve alanları günden güne artmaktadır. Elektrik enerjisinin belli bir zaman aralığında kullanılmasına elektrik tüketimi denir. Elektrik tüketimi, farklı amaçlarda kullanılan elektrik enerjisinin ne kadar harcandığını belirtmektedir. Bu tüketim türleri evsel, ticari, endüstriyel, ulaşım ve tarımsal gibi çeşitli kategorileri içerebilmektedir. Tüketim kategorilerinin oluşturulması, enerji planlama, politika oluşturma ve sürdürebilirlik hedefi doğrultusunda farklı stratejilerin belirlenmesinde yardımcı olabilir.

Elektrik tüketimi çoğunlukla kilowatt saat (kWh) veya megawatt saat (MWh) cinsinden ölçülmektedir. Bir cihazın çalışma süresi (saat) ile gücü (watt) çarpılarak tüketim miktarı bulunmaktadır. [8]

Sadece ülkemizde değil, tüm dünya olarak elektrik tüketimi hakkında bilgi sahibi olunması, sürdürebilirliğin teşvik edilmesi ve çevresel etkilerin azaltılması konusu tüm insanların sorumluluğundadır. Elektrik enerjisi kullanımlarının araştırma çalışmaları kapsamında takip edilip kontrol edilmesi toplum açısından da büyük önem taşımaktadır.

2.1.1 Dünya Genelinde Elektrik Tüketimi

Dünya çapında elektrik tüketimi değişen bir parameter olup sürekli olarak artmaktadır. Bu artışın arkasında yer alan temel faktörlerden biri, gelişmekte olan ülkelerde artan nüfus ve sanayileşmedir. Bunların yanı sıra, günümüzün bir gerçeği olan dijitalleşme sayesinde gerçekleştirilen teknolojik ilerlemeler, ülkelerin ekonomik büyümesi gibi etkenlere bağlı olarak da kesintisiz bir artış göstermektedir. Gelişmiş ülkelerde, günlük hayatın her alanında elektrik yoğun bir şekilde kullanılırken, gelişmekte olan ülkelerde elektrifikasyon süreci hızla devam etmektedir. Dolayısıyla küresel elektrik tüketiminde genel bir artışa neden olmaktadır. [9]

Çeşitli uluslararası enerji ajansları ve kuruluşlar tarafından periyodik olarak dünya genelindeki elektrik tüketim verileri derlenir ve yayınlanır. Uluslararası karşılaştırmaların yapılması ve enerji politikalarının oluşturulması açısından bu veriler çok büyük önem taşımaktadır. [10]

2.1.2 Türkiye'de Elektrik Tüketimi

Ülkemiz, enerji talebi yüksek bir ülkedir. Ekonomik büyümenin yanı sıra şehirleşme, nüfus büyüklüğü, iklim koşulları, endüstriyel faaliyetler ve enerji politikaları gibi pek çok faktör elektrik tüketimi üzerinde kesintisiz bir artışa sebep olduğunu söyleyebiliriz. Elektrik tüketiminin yoğunlaştığı alanlar genellikle endüstriyel, evsel ve ticari kullanım alanlarıdır. Tüketimi dengelemek ve çevresel etkileri azaltmak için yenilenebilir enerji projelerine yatırım yapılması konusunda çalışmalar yapılabilir. Türkiye İstatistik Kurumu (TÜİK), Enerji ve Tabii Kaynaklar Bakanlığı ve Enerji Piyasası Düzenleme Kurumu (EPDK) gibi resmi kurumlar, ülkemizin elektrik tüketimi, üretimi ve kapasite kullanımı ile ilgili verileri düzenli olarak paylaşmaktadır. Bu veriler, elektrik tüketimindeki değişimleri takip etmek, farklı stratejiler belirleyebilmek ve gelecek senelerde değişen yaşam şartlarına uygun enerji planlaması yapabilmek için araştırma çalışmalarında kullanılmaktadır.

3
KURAMSAL TEMELLER

3.1 Yapay Zekâ

Yapay zekâ (AI), bilgisayar sistemlerinde, insan ve doğadaki diğer organizmaların yeteneklerini taklit ederek istenilen belirli bir görevi gerçekleştirme yeteneği sağlayan bir yapay işletim sistemidir. Bu sistem çoğunlukla karmaşık problemleri çözmek, karar verme, öğrenme, algılama, nesne tanıma ve öneri yapma gibi konularını ele alır.

Günümüz dünyasının bir gerçeği olan dijitalleşme hayatımızda çok önemli bir rol oynamaktadır. Dijital dönüşüm sürecinde yapay zekâ, dijital teknolojileri günden güne artan ihtiyaçlarımız doğrultusunda kullanabilmemiz için büyük bir deneyim sağlamaktadır. Bir çok şirket, yapay zekânın ulaştırdığı tüm bilgilerinden faydalanmak için veri bilimi ekiplerine önemli düzeyde yatırım yapmaktadırlar. [11] Sadece yapılan çalışmalar ile değil, günlük hayatta karşılaştığımız, ilk başta yabancı gelen sonrasında direkt alışabildiğimiz mükemmel kolaylıklar sağlayan yapay zekâ araçları ilgi çekici bir konu olmaya devam edecektir. Yapay zekâ ve bu alandaki teknoloji kesintisiz bir hız ile gelişiyor ve deneyimlediğimiz özellikleri buz dağının sadece görünen yüzü.

3.2 Makine Öğrenmesi

Yapay zekanın alt dalı olan Makine Öğrenmesi (ML), bilgisayar sisteminin belirli bir görevi açık talimatlar olmadan yerine getirmek ya da belirli bir alandaki tecrübesini geliştirmek için çeşitli algoritmaları ve istatistiksel modelleri kullanan geliştirme bilimidir. Makine öğrenmesi, kullanılan veriler ile karmaşık görevlerin çözümünde etkili bir kodlamanın yapılamadığı durumlarda ya da taleplerin değiştiği kurgularda mükemmel bir uygulanabilirlik seçeneği sağlamaktadır. Makine öğrenmesinin, her gün bir yenisinin daha eklendiği oldukça fazla uygulama alanı vardır.

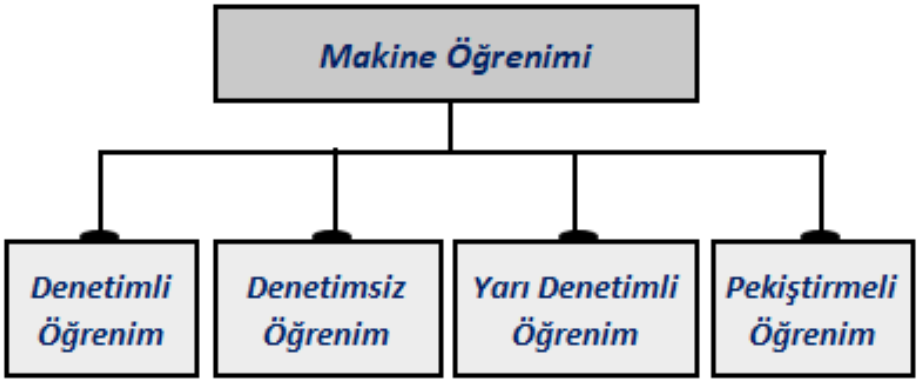
İnsanların belirli bir konuda daha fazla alıştırma yapmasıyla konu ile ilgili kendilerini geliştirmesi, deneyim sahibi olması makine öğrenmesine verilen sık

örneklerden biridir. Daha fazla veri ve deneyim miktarı sayesinde makine öğrenmesinin çözümleri de daha doğru hâle gelecektir.

Yapay zekâ ve makine öğrenimi çoğunlukla bir arada değerlendiriliyor. Bazı durumlarda birbirlerinin yerine kullanılabiliyor. Ancak her zaman aynı anlama gelmezler. Bütün makine öğrenimi çözümlerinin birer yapay zekâ olması, fakat her yapay zekâ çözümlerinin makine öğrenimi olmaması, unutulmaması gereken önemli bir ayrımdır. [11]

3.2.1 Makine Öğrenim Türleri

Makine öğrenmesi; denetimli, denetimsiz, yarı denetimli ve pekiştirmeli olarak dört gruba ayrılmaktadır. Her bir tür, çeşitli görevlerde kendine özgü başarılar göstermektedir.



Şekil 3. 1 Makine öğrenimi türleri

3.2.1.1 Denetimli Öğrenim

Denetimli ya da gözetimli öğrenme (SL) olarak bilinen yöntemde, öğrenim verisi olarak girdiler kullanılır ve girdiler sayesinde öğrenim sağlanır. [12] Algoritmaya önceden belirlenmiş niteliklere göre nasıl kategorize edileceğini öğreten, etiketlenmiş bir biçimde algoritmaya beslenir. Yeterli eğitim uygulanan algoritma, önceden etiketlenmemiş verileri benzer şekilde hazırlayabilir.

3.2.1.2 Denetimsiz Öğrenim

Denetimsiz ya da gözetimsiz öğrenme (UL) olarak bilinen bu yöntem, ilişki tespiti, sıralamak ve kümelemek için geliştirilen makine öğrenim türüdür. [13]

Bu yöntemin amacı, kullanılan verilerin temel yapısı, dağılımı ve veriler hakkında daha fazla bilgi edinilmesinin sağlanmasıdır. Farklı kategorilere sahip veriler arasında kesin bir tanımlamanın yapılamadığı durumlarda oldukça kullanışlıdır.

3.2.1.3 Yarı Denetimli Öğrenim

Yarı denetimli öğrenim, denetimli ve denetimsiz makine öğrenimi tekniklerinin birleştirildiği, bu iki teknik arasında yer alan makine öğrenim türü olarak tanımlanabilir. Bu yöntemde, girdi verisindeki değişkenlerin yapısını keşfetme ve öğrenme kapsamında SL ve UL arasında kalan problemleri çözmeye çalışır.

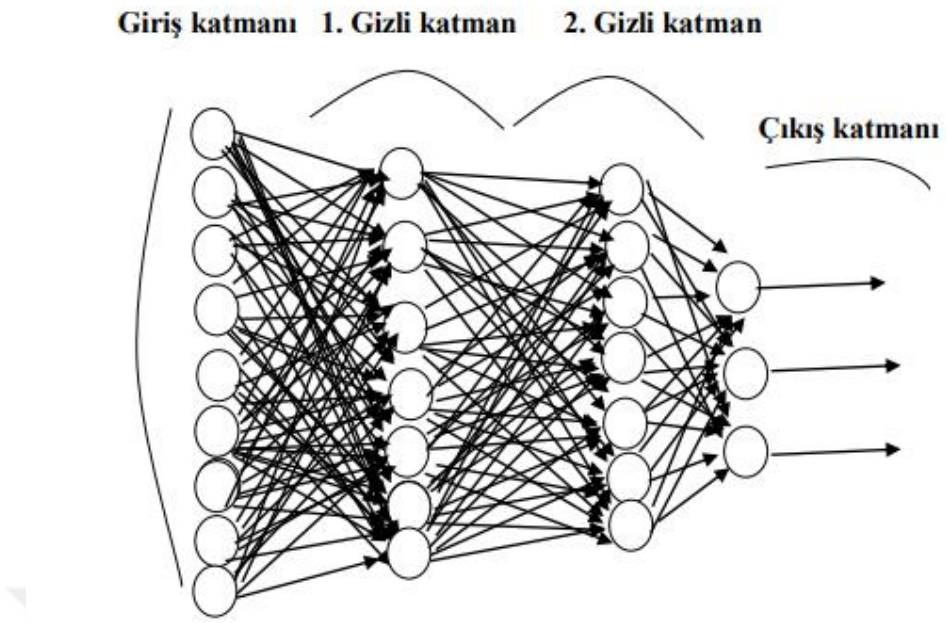
3.2.1.4 Pekiştirmeli Öğrenim

Pekiştirmeli Öğrenim, oldukça karmaşık problemleri çözme kabiliyetine sahip makine öğrenim türüdür. Bu yöntem, deneme ve yanılma üzerine kuruludur. Pekiştirmeli öğrenim türünde oyun benzeri kural kurgusu tasarlanır bu sayese seçimler yapılırken bu seçimin ya güçlendirilmesi ya da caydırılması sağlanır.

3.3 Derin Öğrenme

Derin öğrenme, makine öğrenmesinin bir alt kümesidir. Bu yöntemde, verilerin kullanımına ilişkin makine öğrenimi algoritmaları gibi değil, katman olarak isimlendirilen özel bir durumunu temsil eder. [14] Temel olarak tıpkı insan beyninin çalışma şeklinde modellenen algoritmalardır. Büyük miktardaki verilerin eğitimi, sinir ağlarındaki nöronları yapılandırmayı hedeflemektedir. Bu yöntemde bilgi birden çok veri kaynağından elde edilir ve bir insanın müdahalesine gerek duyulmadan analiz yapabilme yeteneğine sahiptir.

Yöntemin algoritma yapısı yapay sinir ağları kökenlidir. YSA modelinde bulunan katman sayısının arttırılmasıyla derin öğrenme modeli oluşturulmaktadır. Yani, derin öğrenme birden çok katman ve saklı değişkenden oluşmaktadır. Şekil 3.2'de Derin öğrenme algoritmasının mimari yapısı gösterilmiştir. [14]



Şekil 3. 2 Derin öğrenme modeli

Şekil 3.2 incelendiğinde, yapay sinir ağı modeline birden çok yeni gizli katmanlar (ara katman- hidden layer) eklenerek derin öğrenme modeli oluşturulmuş ve bu katmanlar aracılığı ile çıkış katmanınında çözüm sonuçlarının elde edilmesi sağlanmaktadır.

Parametre sayıları gizli katman sayılarına bağlı olarak yükselmektedir. Derin öğrenme modelinde, alt katmandan üst katmana geçildikçe değerler daha fazla ayırt edici özellikler taşır.

3.4 Makine Öğrenmesi ile Derin Öğrenme Arasındaki Temel Farklar

Makine öğrenmesi ve derin öğrenme yöntemlerinin arasındaki farklılıklar konu başlıkları olarak şöyle sıralanabilir; hedeflenen kullanım örnekleri, problem çözme yaklaşımı, eğitim yöntemleri, performans, insan müdahalesi ve altyapı gereksinimleri. [15]

Her iki yöntem oluşturulmuş temel farklılıklar Tablo 3.1’ de özetlenmiştir.

Tablo 3. 1 Makine öğrenmesi ve derin öğrenme farklılıkları

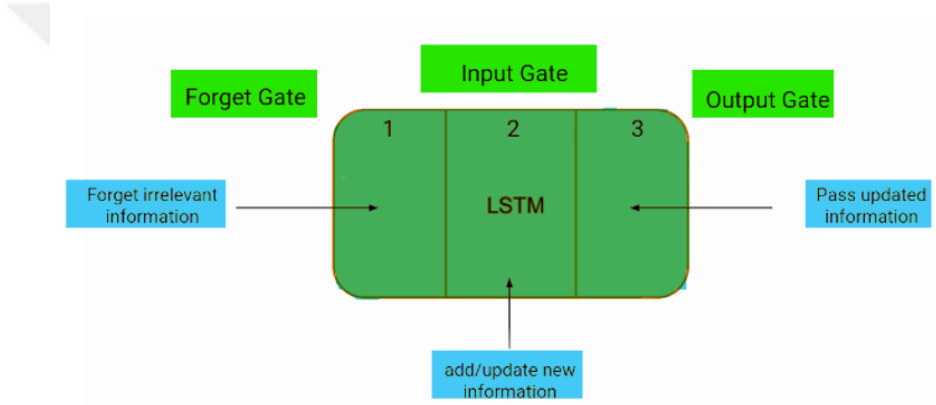
	Makine Öğrenmesi	**Derin Öğrenme**
Nedir?	Yapay zekâ metodolojisidir. Her makine öğrenimi, derin öğrenme değildir.	Gelişmiş bir makine öğrenimi metodolojisidir. Her derin öğrenme, aynı zamanda makine öğrenimidir.
Uygun Olduğu Alanlar	Etiketlenmiş veriler ile çalışan iyi tanımlanmış görevler için ideal bir yöntemdir.	Karmaşık görevler için yapılandırılmamış veriler için anlamlandırılması oldukça ideal bir yöntemdir.
Problem Çözme	İstatistiksel yöntemler ve matematik yöntemleri ile çözülür.	İstatistiksel yöntemler ve matematiği, sinir ağır mimarisi ile birleştirir.
Eğitim	Model eğitimi aşamasında, veri özelliklerini manuel yol ile seçip ayıklamak ve ağırlıklar atamak gerekir.	Model eğitimi aşamasında, hatalar tarafından gelen bilinen geri bildirimler kullanılarak kendi kendine öğrenebilir.
Yapısı	Yapısı daha az karmaşıktır. Daha küçük veri hacmine sahiptir.	Yapısı daha karmaşıktır. Çok daha büyük veri hacmine sahiptir.

3.5 LSTM, XGBoost, RNN, GRU ve ANN Yöntemleri

Çalışmanın bu bölümünde, uygulamada kullanılacak yöntemlere ait tekniklerin özet bir şekilde açıklanmasına ve temel tanımlamalarına yer verilmiştir.

3.5.1 Uzun Kısa Vadeli Bellek (LSTM)

LSTM, tekrarlayan sinir ağı (RNN) olan ve geleneksel RNN'lerin uzun vadedeki bağımlılıkları yakalama kapsamındaki sınırlarını aşmak için tasarlanmıştır. Mimari yapısı gereği bellek hücreleri ve kapıları içermesi sayesinde, uzun sekanslar boyunca ilgiyi bilgiyi seçici bir biçimde tutmasına ve bırakmamasına, ilgisiz bilgiyi ise unutmayı esas alan tekrarlayan sinir ağı modelidir. [16] Teorik yapısından dolayı çok büyük veriler üzerinde çalışabilmektedir. Ancak her ne kadar büyük verilerde başarılı sonuçlar üretse de model eğitim süresi oldukça uzun sürmektedir. Zaman serisi tahmini, doğal dil işleme, ses tanıma vb. kullanım alanlarına örnek olarak verilebilir.



Şekil 3. 3 LSTM mimari yapısı

Şekil 3.3'de LSTM modelinin mimari yapısı verilmiştir. [17] Yöntem, üç ana kapıdan meydana gelmektedir. Unut Kapısı (Forget Gate), yeni girilen bilgilerin ilgili olanları hatırlanmasına, ilgisiz olanların ise unutulmasına karar verdiği kısımdır. Giriş Kapısı (Input Gate), yeni bilginin önem derecesinin değerlendirildiği bölümdür. Ek olarak, hücre durumu bilgisi giriş kapısında güncellenir. Çıkış Kapısı (Output Gate), sonraki gizli katmanda hangi bilginin taşınacağının hesaplandığı bölümdür. [18]

3.5.2 Ekstrem Gradyan Arttırma (XGBoost)

XGBoost, hem sınıflandırma hem de regresyon görevlerinde olağanüstü performansı ile iyi bilinen bir topluluk (ansamble) öğrenme algoritmasıdır. CART (Sınıflandırma ve Regresyon Ağacı)'na dayanır. [6] Bu yöntem, zayıf karar ağaçlarını tekrar tekrar eğitir ve tahminleri birleştirerek güçlü bir model oluşturur.

Bootsraping (Yeniden Örnekleme) yöntemi üzerine kurulu olduğu için büyük verilere tekrarlı örnekleme yaparak başarılı sonuçlar elde edilmesini sağlamaktadır. Model performansı ve öğrenme hızı konusunda pek çok özelleştirme olanakları sunmaktadır. Bu model, en uygun özelliği belirleyerek kaybı en aza indirgemekte oldukça başarılıdır.

$$L_{xgb} = \sum_{i=1}^{N} L(yi, F(xi)) + \sum_{n=1}^{M} \Omega(h_m) \quad (3.1)$$

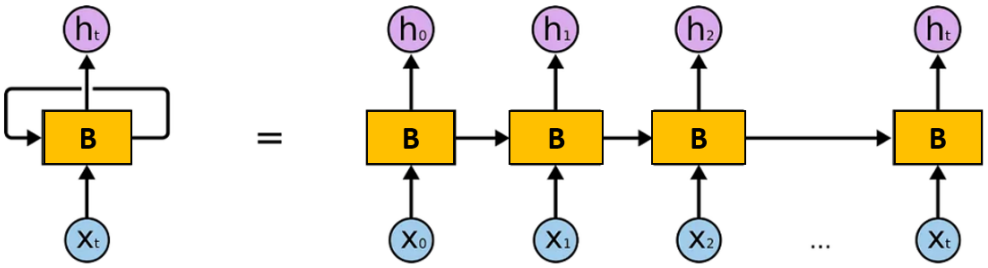
$$\Omega(\mathrm{h}) = \gamma \mathrm{T} + \frac{1}{2}\lambda \parallel w^2 \parallel \quad (3.2)$$

Denklem 3.1 ve 3.2 incelendiğinde, T ağacın yaprak sayısıdır, w ise yaprakların çıktı skorudur. "γ ve λ" katsayıdır. M'inci ağaç eklendiğinde, bir önceki ağaç eğitilmiştir.

3.5.3 Tekrarlayan Sinir Ağları (RNN)

Tekrarlayan sinir ağları, ağırlıkları paylaşan ve bir dizi kademeden geçen ve nöronlardan oluşmuş bir derin öğrenme modelidir. [18] Bu yapı sayesinde ağın, sıralı veri üzerinde etkili bir biçimde çalışılması sağlanır. Bir önceki adımdan gelen bilgi saklanır. Korunan eski bilgi ile yeni gelen bilgi birleştirilerek bir sonraki adıma geçilmesi sağlanır. Farklı bir deyişle RNN, hem mevcut girdiyi işler hem de bir önceki adımdan gelen durumu kaydeder. RNN bu sayede zamansal bağımlılıkları ve sıralı bilgileri öğrenmeyi sağlar.

Bu yöntem, zaman serisi, ses, metin gibi sıralı verilerdeki bağımlılıkları modellemek için tasarlanmıştır. Tekrar tekrar aynı ağırlıklar kullanılarak sıralı veri üzerinde çalışılır.



Şekil 3. 4 RNN mimari yapısı

Şekil 3.4’ te RNN modelinin mimari yapısı verilmiştir. [19] Şekil incelendiğinde çıkan sonucun bir sonraki adımda kullanılacak girdi verisini beslediğini açıkça ortaya koymaktadır. Yani, h_1 için X_0 girdisi ile sonuçlanan h_0 ve X_1 kullanılmaktadır. Sonraki adımlar aynı şekilde ilerlemektedir.

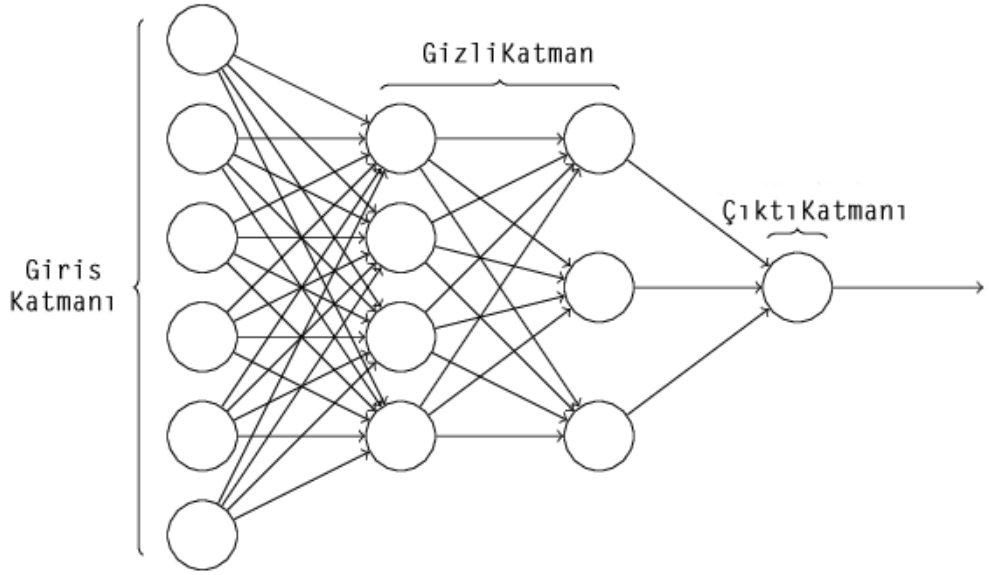
3.5.4 Kapılı Tekrarlayan Birimler (GRU)

GRU yapay sinir ağları yönteminde, daha az kapı (geçit) kullanılarak yine LSTM’ye benzer çözümler üretebilmektedir. [20] Bu modeller, LSTM’ye kıyasla daha az ağırlıkla çalışırlar ve RNN’lerin eksikliklerini gidermeyi hedefleyen RNN mimarisinin farklı bir varyantıdır. GRU’ da sadece seçici bir biçimde gizli durumları güncelleyen güncelleme (update) ve sıfırlamayı sağlayan sıfırlama (reset) geçitleri bulunmaktadır. Bu iki geçit sayesinde ilgili bilgi saklanabilirken ilgisiz bilgi de göz ardı edilebilmektedir.

GRU’ ların yapısı LSTM’ler ile oldukça benzerdir. Aynı LSTM gibi GRU da bilgi akışlarının kontrolü için kapıları kullanır. Bazı iyileştirmeler ve daha basit bir mimariye sahiptir.

3.5.5 Yapay Sinir Ağları (ANN)

Yapay Sinir Ağı, tıpkı bir insan beyninin niteliklerini taklit etmek amacıyla tasarlanmış, girdiyi işleyen, analiz eden ve sonuç çıkaran bir yapay zekânın alt kümesidir. Genellikle tabakalar halinde organize edilen bir dizi yapay nörondan oluşan modeller Yapay Sinir Ağları olarak adlandırılmaktadır. Bu modeller, çoğunlukla regresyon ve sınıflandırma gibi denetimli öğrenme görevlerinde kullanılır.



Şekil 3. 5 Genel bir yapay sinir ağı modeli

Şekil 3.5'te genel bir yapay sinir ağı modelinin mimarisi yapısı verilmiştir. [21] Giriş katmanında "n" adet katman (layer) bulunuyor. Her katmanın kendine özgü hesaplama elemanları vardır. Bunlar tıpkı bir canlının sinir hücrelerini andıran ve farklı sayılarda olabilen araçlardır. Yine Şekil 3.5 yardımıyla bir yapay sinir ağında birden çok girdi katmanın olabileceği ve yapıdaki diğer düğümlere gönderilebilen sadece tek çıktı katmanı bulunacağını görebilmekteyiz.

Çeşitli yapay sinir ağı modellerinde kullanılan terimlerin biyolojik sinir yapısındaki karşılıkları Tablo 3.2'de özetlenmiştir. [5]

Tablo 3. 2 Biyolojik sinir yapısı ile yapay sinir ağı

Yapay Sinir Ağı	**Biyolojik Sinir Sistemi**
Aktivasyon işlemi	Hücre gövdesi
Nod çıktısı	Akson
Nod	Nöron
Toplama İslemi	Dentrit
Ağırlık	Sinaps

3.6 Model Performanslarının Karşılaştırılması

Model performanslarının karşılaştırılması, bir ya da birden çok modelin farklı metrikler kullanılarak değerlendirilmesidir. Bu sayede, en uygun model hangi algoritma ya da parameter ayarlarında oluşturulmuş, bunun tespit edilmesi sağlanabilir.

Bu tez çalışmasında, en düşük MAE, MSE, ve RMSE hata metriklerine sahip modeller tespit edilerek test verisi üzerinde karşılaştırılmıştır.

MAE: Ortalama Mutlak Hata (Mean Absolute Error), veri setindeki gerçek değerler ile tahmin değerleri arasındaki farkların mutlak değerli toplam sonucun n (örnek sayısı)'na bölünmesi ile bulunur. Bu değer "0 ile ∞" arasında değerler almaktadır. MAE değerinin formülle gösterimi Denklem 3.3'te verilmiştir.

$$MAE = \frac{1}{2}\sum_{i=1}^{n}|Y_i - \widehat{Y_i}| \tag{3.3}$$

MSE: Ortalama Kare Hatası (Mean Squared Error), veri setindeki gerçek değerler ile tahmin değerleri arasındaki farkların karelerinin toplam sonucun n (örnek sayısı)'na bölünmesi ile bulunur. MSE değerinin formülle gösterimi Denklem 3.4'te verilmiştir.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Yi - \widehat{Y_i})^2 \tag{3.4}$$

RMSE: Kök Ortalama Kare Hatası (Root Mean Squared Error), ortalama kare hatasının kareköküdür. RMSE daha sık kullanılmaktadır çünkü MSE değeri bazı durumlarda karşılaştıralamayacak kadar çok büyük bulunabiliyor. RMSE sayesinde yorumlamak daha kolaylaşmaktadır. RMSE değerinin formülle gösterimi Denklem 3.5'te verilmiştir.

$$RMSE = \sqrt{\frac{\sum(Y_i - \hat{Y}_i)^2}{N-P}} \tag{3.5}$$

4
UYGULAMA

4.1 Türkiye'de Elektrik Tüketimi Üzerine Bir Uygulama

Tez çalışmasının bu bölümünde, makine öğrenmesi, derin öğrenme ve istatistiksel analiz tekniklerinin kapsamlı bir analiz incelemesini sunmaktadır. Bu teknikler arasında XGBoost, LSTM, RNN, GRU ve ANN bulunmaktadır. Kullanılan yöntemler ve görselleştirmeler Python program yardımıyla gerçekleştirilmiştir. Araştırma materyali olarak Türkiye'deki 31.12.2015 ile 22.03.2022 tarihleri arasındaki elektrik tüketim miktarı kullanılmıştır. Çalışmada kullanılan veri seti EPİAŞ Şeffaflık Platformu web sitesinden elde edilmiştir. [22]

4.1.1 Veri Manipülasyonu ve Görselleştirmesi

Kullanılan veri seti toplam 54571 satırdan oluşmaktadır.

	Tarih	Saat	Tüketim Miktarı (MWh)
0	31.12.2015	00:00	29.590,88
1	31.12.2015	01:00	27.785,43
2	31.12.2015	02:00	26.516,53
3	31.12.2015	03:00	26.091,62
4	31.12.2015	04:00	25.871,54
...	...	...	...
54566	22.03.2022	14:00	45.083,90
54567	22.03.2022	15:00	44.158,84
54568	22.03.2022	16:00	43.608,34
54569	22.03.2022	17:00	42.107,30
54570	22.03.2022	18:00	33.904,92

54571 rows × 3 columns

Şekil 4. 1 Elektrik tüketim miktarı veri seti

Pyrhon programına Excel üzerindeki veri seti aktarılmalıdır. Şekil 4.1'de ham data gösterilmiştir.

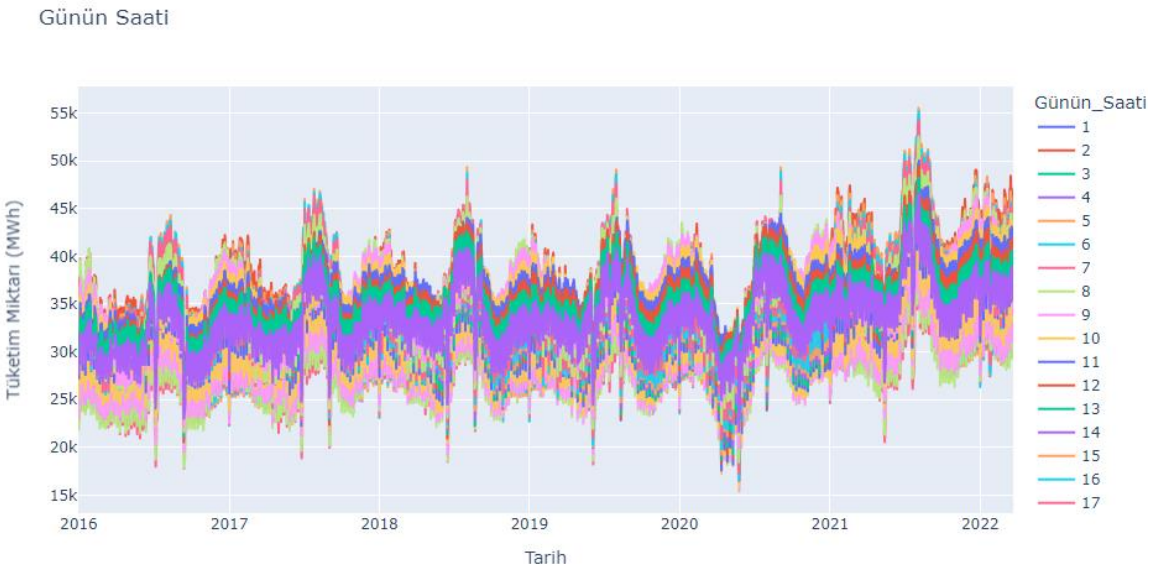
	Tarih	Saat	Tüketim Miktarı (MWh)	Datetime	Günün_Saati	Haftanın_Günü	Ayın_Günü	Yılın_Ayı
0	31.12.2015	00:00	29.590,88	2015-12-31 00:00:00	1	4	31	12
1	31.12.2015	01:00	27.785,43	2015-12-31 01:00:00	2	4	31	12
2	31.12.2015	02:00	26.516,53	2015-12-31 02:00:00	3	4	31	12
3	31.12.2015	03:00	26.091,62	2015-12-31 03:00:00	4	4	31	12
4	31.12.2015	04:00	25.871,54	2015-12-31 04:00:00	5	4	31	12
...	...	...	...	...	...	...	...	...
54566	22.03.2022	14:00	45.083,90	2022-03-22 14:00:00	15	2	22	3
54567	22.03.2022	15:00	44.158,84	2022-03-22 15:00:00	16	2	22	3
54568	22.03.2022	16:00	43.608,34	2022-03-22 16:00:00	17	2	22	3
54569	22.03.2022	17:00	42.107,30	2022-03-22 17:00:00	18	2	22	3
54570	22.03.2022	18:00	33.904,92	2022-03-22 18:00:00	19	2	22	3

54571 rows × 8 columns

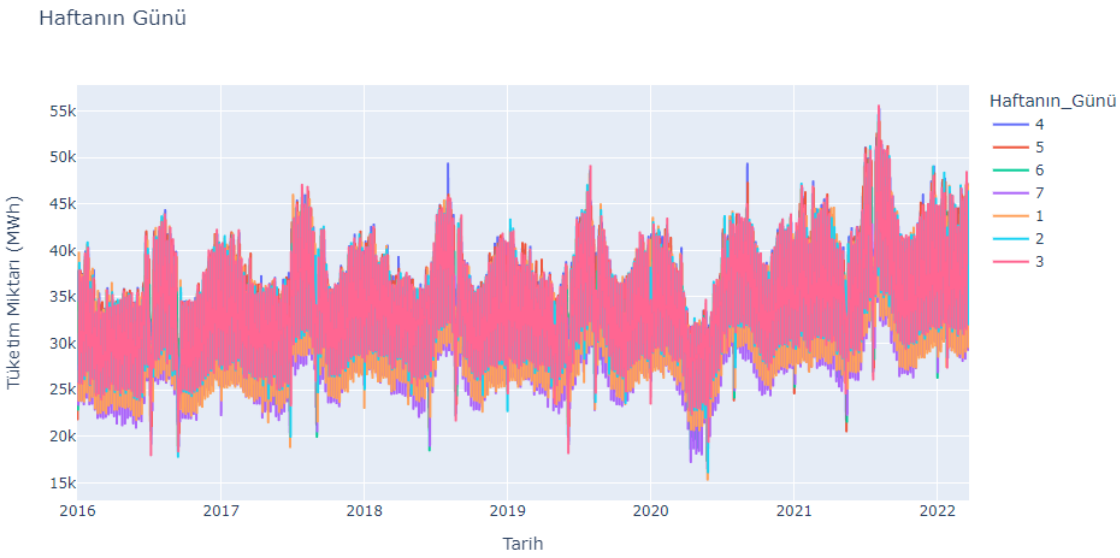
Şekil 4. 2 Yeni sütunların ekli olduğu veri seti

Tarih ve saat bilgisi farklı sütunlarda olduğundan öncelikle iki değişken birleştirilerek Datetime, sonrasında bağımsız değişken olarak kullanılacak günün saati, haftanın günü, ayın günü ve yılın ayı sütunlarının veri setine eklendiği Şekil 4.2'de gösterilmiştir.

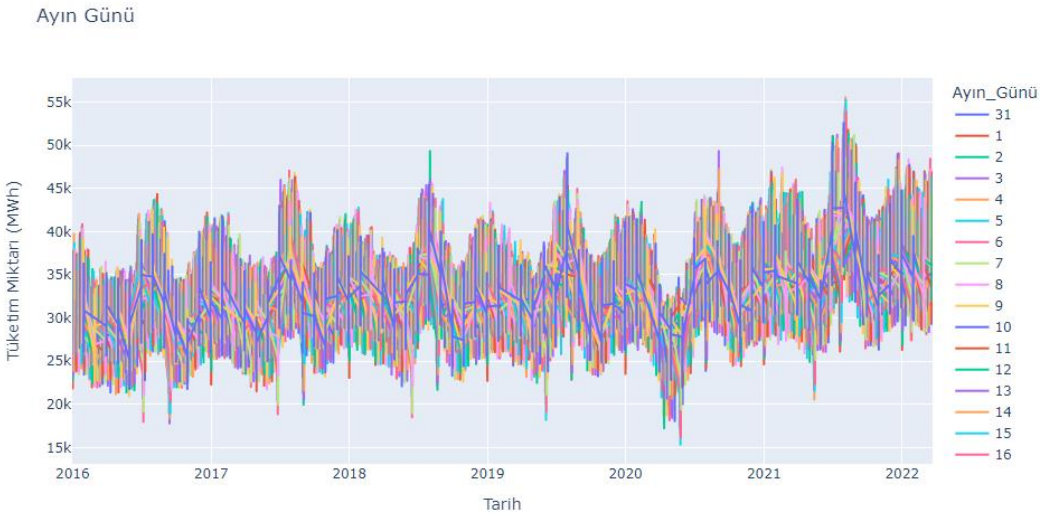
Yeni değişkenlerin eklenmesiyle oluşturulan veri setinde boş değer kontrolü yapılmıştır. Tüketim Miktarı (MWh) değişkeninde 0 değeri olduğu gözlemlenmiştir. 27.03.2016 tarihinde ve saat 02:00'de 0 değeri olarak bir aykırı gözlem mevcuttur. Bu 0 değeri birkaç yol ile düzeltilebilmektedir. Bu analizde bir önceki günün (26.03.2016) saat 02:00 değerini alarak düzeltilmesi sağlanmıştır. Aşağıda, Şekil 4.3'den Şekil 4.6'ye kadar bağımsız değişkenlerin zaman serisi grafikleri gösterilmiştir.



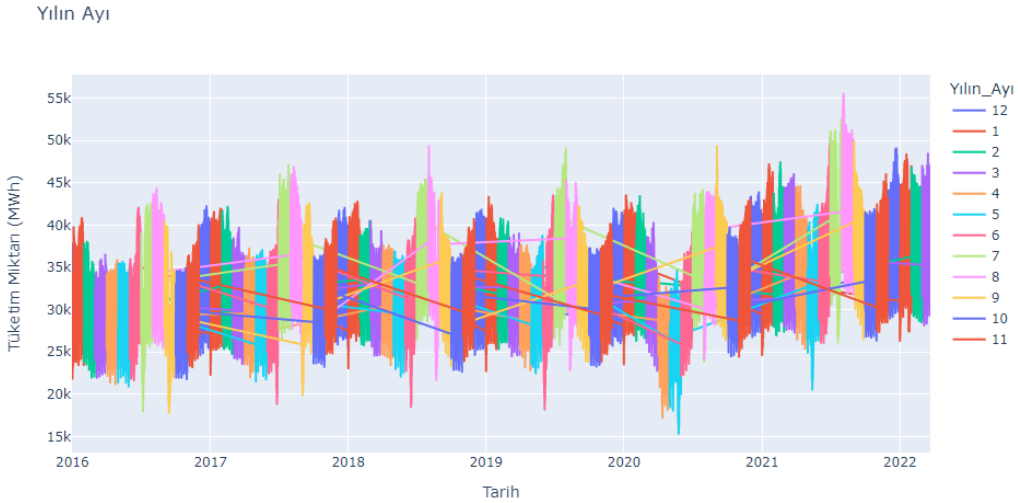
Şekil 4. 3 Günün saati zaman serisi grafiği



Şekil 4. 4 Haftanın günü zaman serisi grafiği

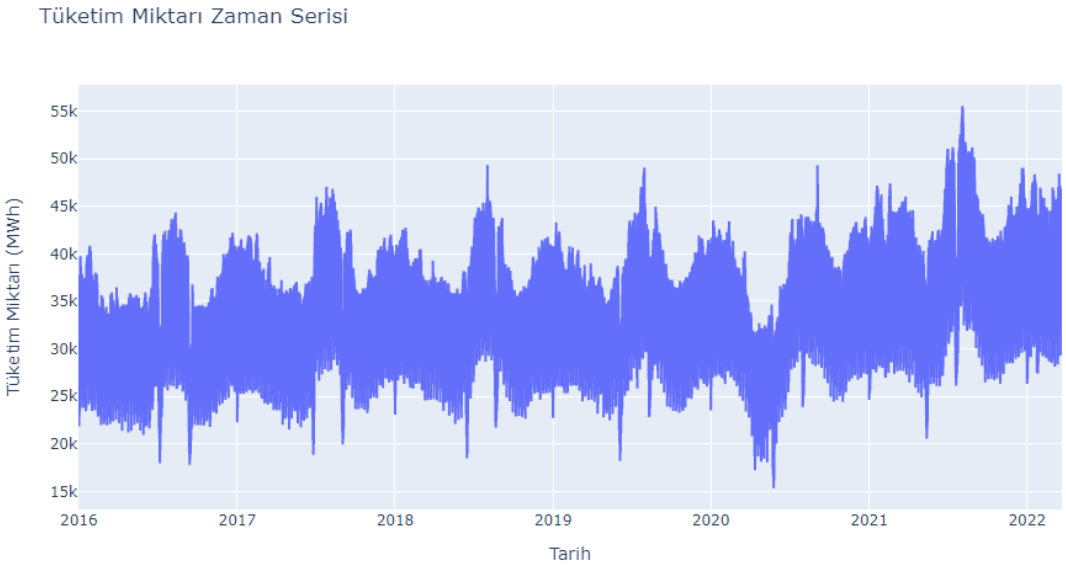


Şekil 4. 5 Ayın günü zaman serisi grafiği



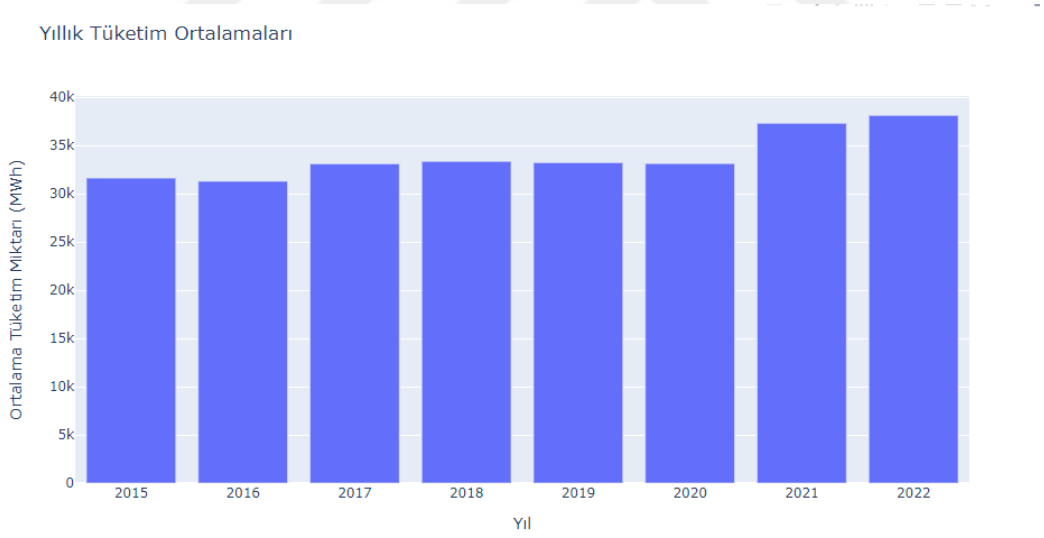
Şekil 4. 6 Yılın ayı zaman serisi grafiği

Tüketim miktarı zaman serisinin grafik ile gösterimi Şekil 4.7’ de verilmiştir.



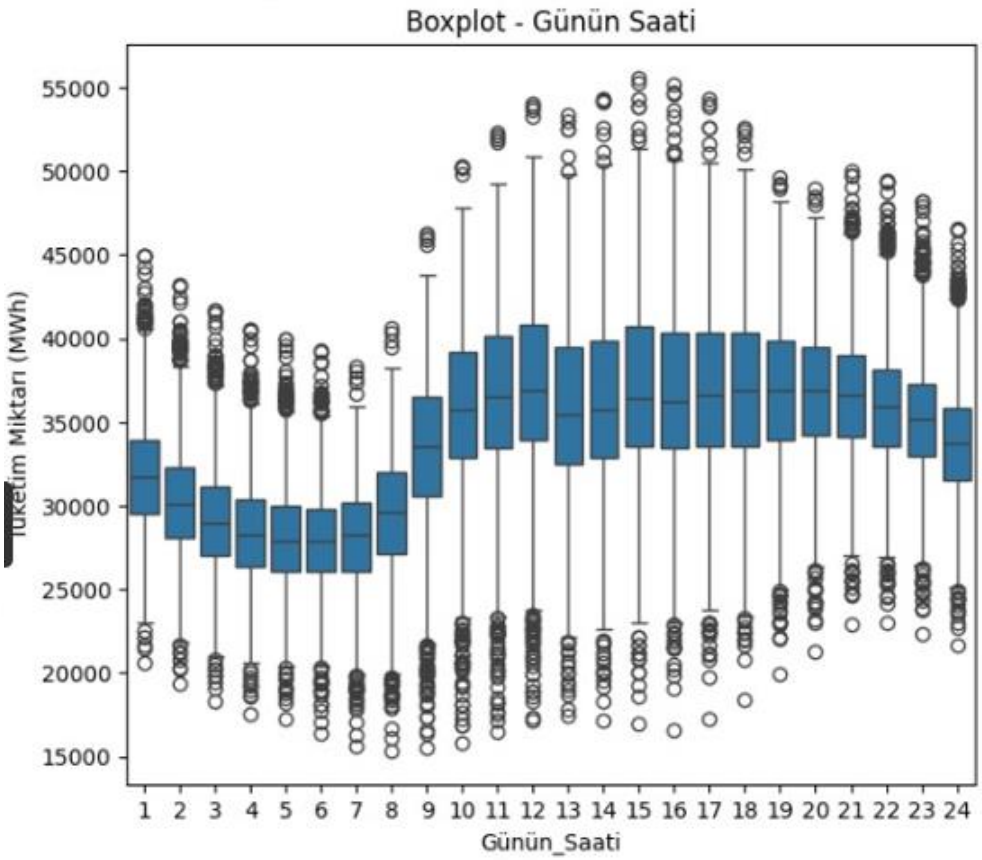
Şekil 4. 7 Tüketim miktarı zaman serisi

Yıllık tüketim ortalamaları incelenmek istediğinde, Şekil 4.8 yardımıyla son iki yılda ortalama tüketim miktarındaki artış görülebilmektedir. 2015 yılında yıllık ortalama miktarı 31.65 MWh iken, 2022 yılında 38.13 MWh'a yükselmiştir.



Şekil 4. 8 Yıllık tüketim ortalamaları

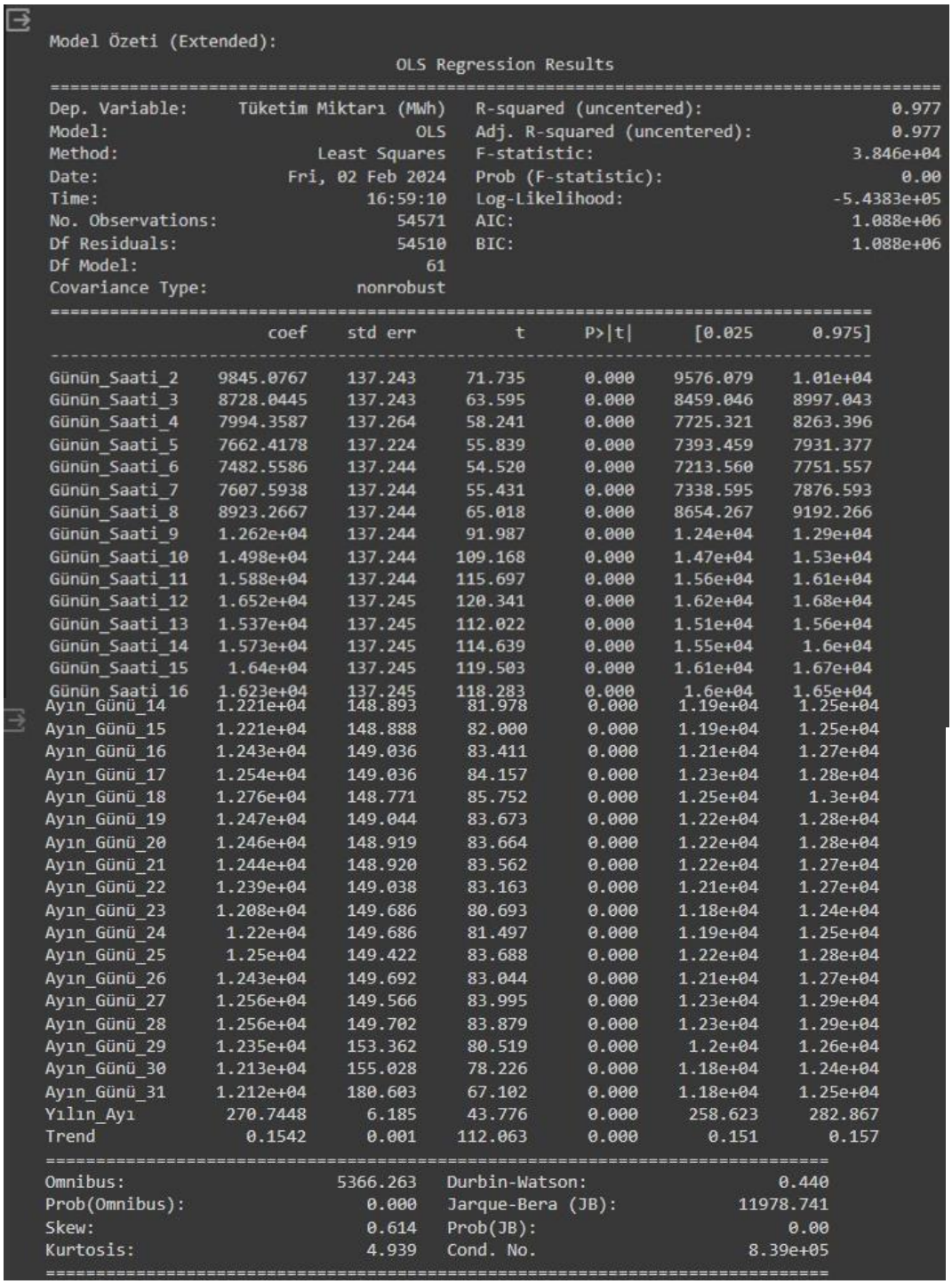
Şekil 4.9 incelendiğinde, günün saati 16 en heterojen, günün saati 24 ise en homojen dağılımı gösterdiği görülmüştür. Ayrıca günün saatine ait tüm veri setinde aykırı değerler mevcuttur.



Şekil 4. 9 Günün saati boxplot grafiği

4.1.2 Regresyon Analizi

Regresyon analizini gerçekleştirebilmek için bazı varsayımların sağlanması gerekmektedir. Ancak varsayımlardan önce ham verideki değişkenlerin tüketim miktarına etkisi aşağıdaki gibi incelenebilir. Değerlendirilen bağımsız değişkenler; Günün Saati, Haftanın Günü, Ayın Günü ve Yılın Ayı olup, bağımlı değişken olarak elektrik tüketimi (MWh) alınmıştır. Bağımsız değişkenler kategorik getirilmiştir. Modelde kullanılan bağımsız değişkenler kendi özelinde n tane değerden oluştuğu için n adet yeni değişken olarak eklenecektir. Ek olarak trend değişkeni de eklenerek yeni bağımsız değişkenler oluşturuldu. Bu değişkenler arasındaki ilişkileri keşfetmek için OLS (Ordinary Least Squares) regresyon analizi yöntemi kullanılmıştır. Modelde yer alan katsayılar ve istatistiksel testler, bağımsız değişkenlerin bağımlı değişken üzerindeki anlamlılığını ve etkilerini değerlendirmek için incelenmiştir.



Model Özeti (Extended):

OLS Regression Results

Dep. Variable:	Tüketim Miktarı (MWh)	R-squared (uncentered):	0.977
Model:	OLS	Adj. R-squared (uncentered):	0.977
Method:	Least Squares	F-statistic:	3.846e+04
Date:	Fri, 02 Feb 2024	Prob (F-statistic):	0.00
Time:	16:59:10	Log-Likelihood:	-5.4383e+05
No. Observations:	54571	AIC:	1.088e+06
Df Residuals:	54510	BIC:	1.088e+06
Df Model:	61		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Günün_Saati_2	9845.0767	137.243	71.735	0.000	9576.079	1.01e+04
Günün_Saati_3	8728.0445	137.243	63.595	0.000	8459.046	8997.043
Günün_Saati_4	7994.3587	137.264	58.241	0.000	7725.321	8263.396
Günün_Saati_5	7662.4178	137.224	55.839	0.000	7393.459	7931.377
Günün_Saati_6	7482.5586	137.244	54.520	0.000	7213.560	7751.557
Günün_Saati_7	7607.5938	137.244	55.431	0.000	7338.595	7876.593
Günün_Saati_8	8923.2667	137.244	65.018	0.000	8654.267	9192.266
Günün_Saati_9	1.262e+04	137.244	91.987	0.000	1.24e+04	1.29e+04
Günün_Saati_10	1.498e+04	137.244	109.168	0.000	1.47e+04	1.53e+04
Günün_Saati_11	1.588e+04	137.244	115.697	0.000	1.56e+04	1.61e+04
Günün_Saati_12	1.652e+04	137.245	120.341	0.000	1.62e+04	1.68e+04
Günün_Saati_13	1.537e+04	137.245	112.022	0.000	1.51e+04	1.56e+04
Günün_Saati_14	1.573e+04	137.245	114.639	0.000	1.55e+04	1.6e+04
Günün_Saati_15	1.64e+04	137.245	119.503	0.000	1.61e+04	1.67e+04
Günün_Saati_16	1.623e+04	137.245	118.283	0.000	1.6e+04	1.65e+04
Ayın_Günü_14	1.221e+04	148.893	81.978	0.000	1.19e+04	1.25e+04
Ayın_Günü_15	1.221e+04	148.888	82.000	0.000	1.19e+04	1.25e+04
Ayın_Günü_16	1.243e+04	149.036	83.411	0.000	1.21e+04	1.27e+04
Ayın_Günü_17	1.254e+04	149.036	84.157	0.000	1.23e+04	1.28e+04
Ayın_Günü_18	1.276e+04	148.771	85.752	0.000	1.25e+04	1.3e+04
Ayın_Günü_19	1.247e+04	149.044	83.673	0.000	1.22e+04	1.28e+04
Ayın_Günü_20	1.246e+04	148.919	83.664	0.000	1.22e+04	1.28e+04
Ayın_Günü_21	1.244e+04	148.920	83.562	0.000	1.22e+04	1.27e+04
Ayın_Günü_22	1.239e+04	149.038	83.163	0.000	1.21e+04	1.27e+04
Ayın_Günü_23	1.208e+04	149.686	80.693	0.000	1.18e+04	1.24e+04
Ayın_Günü_24	1.22e+04	149.686	81.497	0.000	1.19e+04	1.25e+04
Ayın_Günü_25	1.25e+04	149.422	83.688	0.000	1.22e+04	1.28e+04
Ayın_Günü_26	1.243e+04	149.692	83.044	0.000	1.21e+04	1.27e+04
Ayın_Günü_27	1.256e+04	149.566	83.995	0.000	1.23e+04	1.29e+04
Ayın_Günü_28	1.256e+04	149.702	83.879	0.000	1.23e+04	1.29e+04
Ayın_Günü_29	1.235e+04	153.362	80.519	0.000	1.2e+04	1.26e+04
Ayın_Günü_30	1.213e+04	155.028	78.226	0.000	1.18e+04	1.24e+04
Ayın_Günü_31	1.212e+04	180.603	67.102	0.000	1.18e+04	1.25e+04
Yılın_Ayı	270.7448	6.185	43.776	0.000	258.623	282.867
Trend	0.1542	0.001	112.063	0.000	0.151	0.157

Omnibus:	5366.263	Durbin-Watson:	0.440
Prob(Omnibus):	0.000	Jarque-Bera (JB):	11978.741
Skew:	0.614	Prob(JB):	0.00
Kurtosis:	4.939	Cond. No.	8.39e+05

Şekil 4. 10 Regresyon modelinin özet istatistikleri

Şekil 4.10'da verilen regresyon model sonuçlarına göre, tüm bağımsız değişkenlerin elektrik tüketimi üzerinde pozitif ve istatistiksel olarak anlamlı bir etkisi olduğu görülmüştür. Katsayılar, bağımsız değişkenlerin bir birimlik artışının tüketimde ne kadarlık bir değişikliğe neden olacağını göstermektedir.

Örneğin, Günün_Saati_3 değişkeni olan 08:00 saati, 00:00 saatine göre katkı düzeyi 9845.08 daha fazladır ve bu etki istatistiksel olarak anlamlıdır. Ayın_Günü_24 olan yirmi dördüncü gün, birinci güne göre katkı düzeyi 12200 daha fazladır ve bu etki istatistiksel olarak anlamlıdır. Modelin Uyum İyiği incelendiğinde ise; modelin R-Kare değeri 0.977'dır, bu da modelin bağımlı değişkendeki varyansın %97.7'sini açıkladığı anlamına gelmektedir, bu da modelin veriyi oldukça iyi açıkladığını göstermektedir. Modelin F-İstatistiği 3.846e+04 ve p-değeri 0.00'dır, bu da modelin genel olarak istatistiksel olarak anlamlı olduğunu ve en az bir bağımsız değişkenin bağımlı değişkeni açıklamada anlamlı olduğunu göstermektedir. Durbin-Watson istatistiği 0.440'dır ve bu, hata terimlerinde pozitif otokorelasyon olabileceğini göstermektedir. JB testi normal dağılım hipotezini test eder ve bu modeldeki JB değeri 11978.741'dir ve p-değeri 0'dır, bu da hata terimlerinin normal dağılmadığını göstermektedir.

Sonuç olarak, bu çalışma, belirli zaman serisi değişkenlerinin elektrik tüketimi üzerindeki etkilerini başarıyla ortaya koymaktadır. Bununla birlikte, modelin daha iyi optimize edilmesi ve geliştirilmesi için ileri analizlere ihtiyaç duyulabilir. Analiz edilen model, elektrik tüketimini oldukça iyi açıklamaktadır. Ancak, Durbin-Watson istatistiği ve Jarque-Bera testi sonuçları, modelin varsayımlarının tam anlamıyla sağlanmadığını, modelde otokorelasyon ve hata terimlerinin normal dağılmadığına işaret etmektedir. Bu nedenle, modelin daha sağlıklı sonuçlar vermesi için bu sorunların ele alınması ve düzeltilmesi gerekmektedir.



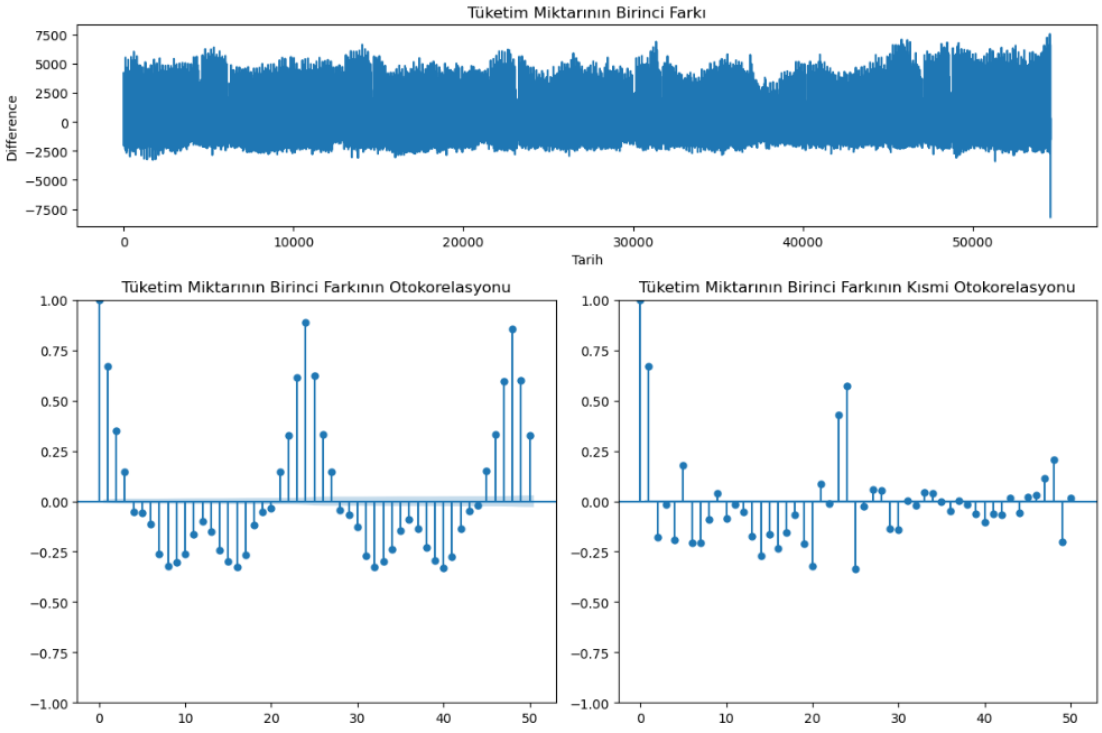
Şekil 4. 11 Augmented dickey-fuller (ADF) testi

Augmented Dickey-Fuller (ADF) testi, zaman serisinin durağan olup olmadığını test etmek için kullanılır. Durağanlık, zaman serisinin ortalama ve varyansının zamanla sabit olduğunu ve otokorelasyon yapısının zamanla değişmediğini gösterir. Şekil 4.11'de Augmented Dickey-Fuller testi incelendiğinde ADF test değeri %95 güven düzeyi olan 0.05 değerinden küçük olduğu için tüketim miktarı durağandır.

	VIF Factor	features
0	1.613489	Günün_Saati_2
1	1.613492	Günün_Saati_3
2	1.613259	Günün_Saati_4
3	1.613736	Günün_Saati_5
4	1.613502	Günün_Saati_6
..	...	...
56	1.467199	Ayın_Günü_29
57	1.455798	Ayın_Günü_30
58	1.297493	Ayın_Günü_31
59	4.204334	Yılın_Ayı
60	3.861803	Trend

Şekil 4. 12 Variance inflation factor (VIF) testi

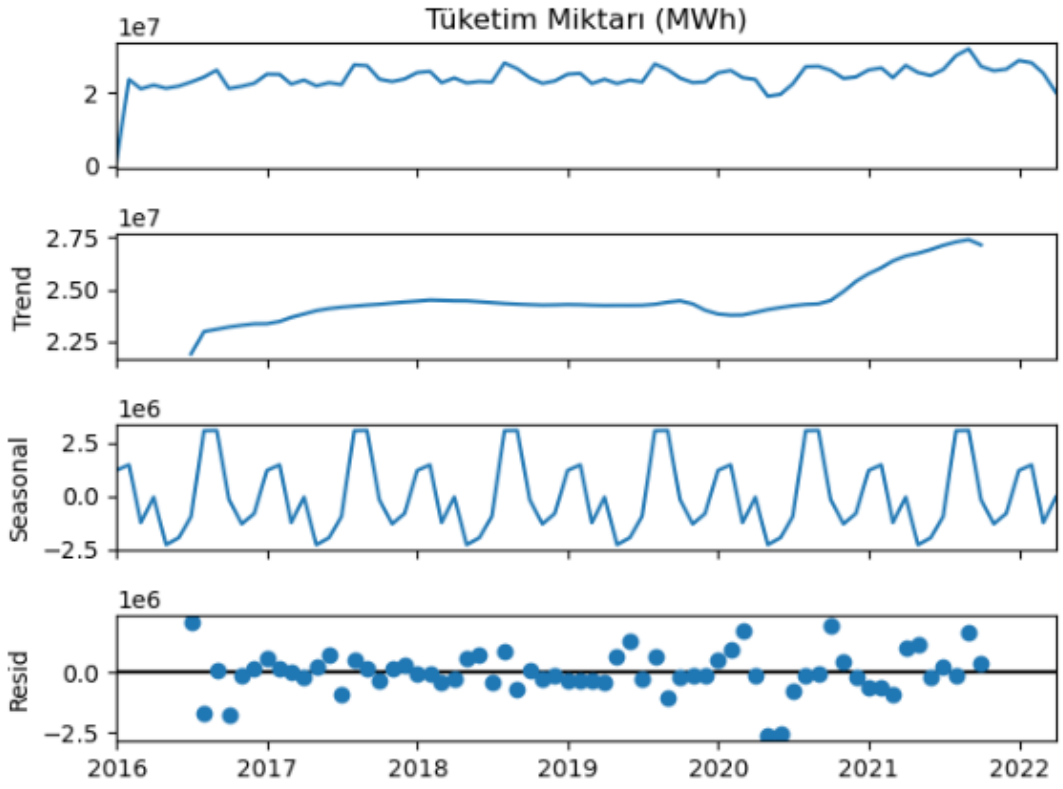
Şekil 4.12'ye bakıldığında tüm bağımsız değişkenlerin VIF değerleri 5'ten çok daha küçüktür. Bu, bağımsız değişkenler arasında çoklu doğrusal bağlantı sorunu olmadığını göstermektedir. Bu durum, regresyon modelinin doğruluğunun ve güvenilirliğinin yüksek olduğunu söylemektedir.



Şekil 4. 13 Tüketim miktarı fark grafikleri

Otokorelasyon ve hata terimlerinin normal dağılmadığını test istatistikleri ile görülmüştü. Şekil 4.13'te verilen grafiksel gösterimde ise pozitif yönlü otokorelasyonun olduğu ve yaklaşık en az 10 farklı gecikmede bu problemin olduğu

görülmektedir. Özellikle ilk grafik incelendiğinde tüketim miktarının 5000 ile sınırlandırıldığını ve durağanlığın ciddi şekilde bozulduğu görülebilmektedir. -2500 ile sınırlandırılan fark değerinin pozitif sonuçlarıda 2500 aralığında olması gerekmektedir. Bu problemi çözebilmek için oldukça fazla çözüm yolu mevcuttur. Zaman serisinin farkı alınabilir, zaman serisindeki trendi kaldırılabilir, LSTM, GRU gibi karmakşık modelleri test edebilecek derin öğrenme modelleri kullanılabilir. Son olarak, tahmin modelinde daha iyi sonuçlar alabilmek için logaritmik dönüşüm veya 0-1 standardize yöntemleri uygulanabilir.



Şekil 4. 14 Veri kümesi trend, mevsimsellik, gürültü grafikleri

Şekil 4.14'de Mevsimsellik (Seasonal) grafiği incelendiğinde, yıllar arasında belirgin bir mevsimsellik gözlemlenmektedir. Trend grafiği, son iki yılda tüketim miktarının artış gösterdiğini açıkça ortaya koymaktadır. Resid grafiği ise, model artıklarının genellikle düzgün bir yapıya sahip olduğunu göstermektedir.

Bu analiz sonuçlarına göre, tahmin modelinin, pozitif otokorelasyon ve artıkların normal dağılımı problemleri çözüldüğünde etkili sonuçlar verebileceğini göstermektedir. Pozitif otokorelasyon problemi ve artıkların normal dağılmaması,

modelin güvenilirliğini ve tahmin başarısını azaltan unsurlardır. Bu iki problemi ele almak için, verilere logaritmik dönüşüm uygulanabileceği gibi, değişkenler 0-1 arasında da standardize edilebilir. Logaritmik dönüşüm, verinin varyansını stabil hale getirmeye yardımcı olabilir ve otokorelasyon sorununu hafifletebilir. Ayrıca, değişkenleri 0-1 arasında standardize etmek, modelin daha dengeli ve sağlam sonuçlar üretmesine yardımcı olabilir. Bu tür dönüşümler, modelin genel performansını ve tahmin doğruluğunu artırabilir, böylece daha güvenilir ve sağlam sonuçlar elde edilebilir.

4.1.3 Tüketim Miktarı (MWh) Tahmini

Elektrik tüketim miktarı tahmini için önceki bölümlerde de bahsedildiği gibi LSTM, XGBoost, RNN, GRU ve ANN teknikleri kullanılmıştır. Bu yöntemler, verinin önceden işlenmesi ve dönüştürülmesi ile bu sorunların üstesinden gelme kapasitesine sahiptir. Örneğin, normalizasyon ya da logaritmik dönüşüm gibi teknikler veriyi daha işlenebilir bir kalıba getirebilir.

Bu çalışmada, modellerin performansını optimize etmek için veri seti [0-1] aralığına standardize edilmiştir. Modeller değerlendirilirken; on katlı çapraz doğrulama yöntemi kullanılmış ve çeşitli hiperparametreler ile denenmiştir. Modellerin performansları karşılaştırılırken en düşük MAE, MSE, ve RMSE hata metrikleri kriterine göre seçilmiştir. Ek olarak, modelin aşırı öğrenme (overfitting) ve yetersiz öğrenme (underfitting) sorunlarını önlemek için hiperparametre ayarlamaları ve veri seti boyutu üzerinde değişiklikler yapılabilir. Bu, modellerin daha dirençli sonuçlar üretmesine yardımcı olabilir.

4.1.4 Hızlı Sonuç Alma (Train, Dev, Test Seti Oluşturma)

Zaman serisi verileri, belirli zaman aralıklarında, kronolojik sıraya uygun ve kesikli biçimde kaydedilmiş bir dizi veri noktasıdır. Bu tür verilerde saniye, dakika, saat, gün vb. ölçümler belirli aralıklarla alınmaktadır. Zaman serisi verileri ile modeli oluştururken modeli eğitmek, doğrulamak ve test etmek için veriyi eğitim (train), doğrulama (dev) ve test (test) setlerine ayrılması gerekmektedir. Zaman serisi verilerinde zaman sırası kritik derecede önemli olduğu için bu tarz verilerde veriyi bölme yöntemi diğerlerine göre biraz daha farklıdır. Veriler çoğunlukla zaman sırasına göre bölünür ve model eğitiminin daha sağlıklı bir şekilde gerçekleştirilmesi sağlanır. Örneğin, eğitim verisi sabit parametrelerle eğitilir,

doğrulama verisi ile çapraz doğrulama uygulanarak farklı hiperparametreler ile en başarılı sonucu veren parametrenin seçilmesi sağlanır ve son olarak test verisi tahmin çözümlerinde bulunur. Bu yöntem büyük verilerde çok daha hızlı model eğitimine olanak sağlar.

Bu çalışmada, veri seti %70 eğitim ve %30 test verisi olacak şekilde bölünmüştür. Eğitim seti, modelin çeşitli parametlerele eğitilmesi için kullanılırken test seti ise modelin genelleme becerisinin değerlendirilme aşamasında kulanılmıştır. Modellerin tahmin işlemleri de bu sıralı veri üzerinden gerçekleştirilmiştir. Kullandığımız veri mevsimsel bir yapıya sahip olsa da zaman serisi verilerinin çoğunluğu belirli bir trend ve mevsimsellik yapısına sahip olmayabilir. Dolayısıyla karmaşık yapılara sahip verilerde parameter eğitimi ve tahmini için bu tür ayırma süreci ile gerçekleştirmek daha doğru çözüm sonuçları elde etmemize yardımcı olacaktır.

4.1.5 Model Eğitimi ve Tahmini

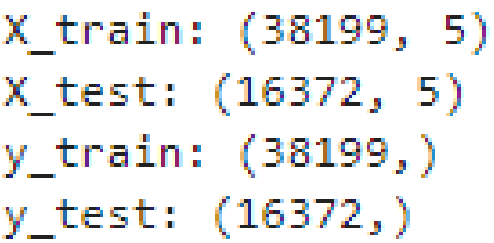
Şekil 4.15'de tüketim miktarı versine uygulanan logaritmik dönüşümü ve tüm verinin ölçeklendirilmiş formatına ait veri seti gösterilmiştir.

	Günün_Saati	Haftanın_Günü	Ayın_Günü	Yılın_Ayı	Yıl	Tüketim Miktarı (MWh)	Diff_Tüketim Miktarı (MWh)	Datetime
0	0.000000	0.500000	1.0	1.000000	0.0	0.510550	NaN	2015-12-31 00:00:00
1	0.043478	0.500000	1.0	1.000000	0.0	0.461662	-0.048889	2015-12-31 01:00:00
2	0.086957	0.500000	1.0	1.000000	0.0	0.425362	-0.036300	2015-12-31 02:00:00
3	0.130435	0.500000	1.0	1.000000	0.0	0.412817	-0.012545	2015-12-31 03:00:00
4	0.173913	0.500000	1.0	1.000000	0.0	0.406239	-0.006578	2015-12-31 04:00:00
...	...	...	...	...	...	...	...	...
54566	0.608696	0.166667	0.7	0.181818	1.0	0.837534	0.005408	2022-03-22 14:00:00
54567	0.652174	0.166667	0.7	0.181818	1.0	0.821434	-0.016100	2022-03-22 15:00:00
54568	0.695652	0.166667	0.7	0.181818	1.0	0.811692	-0.009742	2022-03-22 16:00:00
54569	0.739130	0.166667	0.7	0.181818	1.0	0.784491	-0.027201	2022-03-22 17:00:00
54570	0.782609	0.166667	0.7	0.181818	1.0	0.616237	-0.168253	2022-03-22 18:00:00

54571 rows × 8 columns

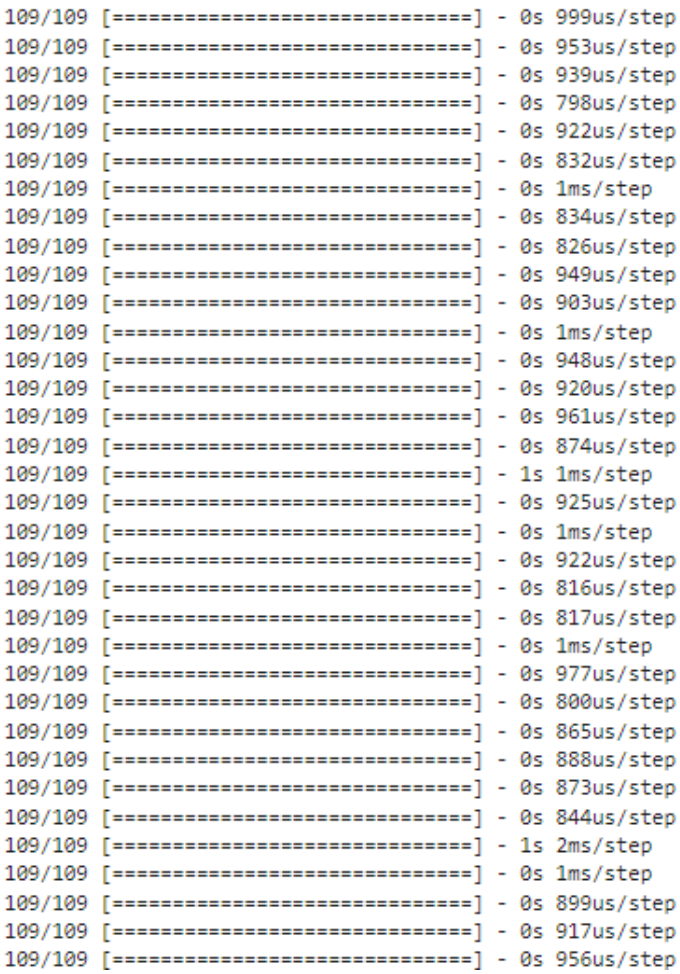
Şekil 4. 15 Logaritması alınan tüketim miktarı ve ölçeklendirilen tüm veri

Veri seti eğitim ve test verileri olarak ayrılmıştır. On katlı çapraz doğrulama adımı uygulanmıştır. Şekil 4.16'da veri seti yapılarının gözlem sayıları eklenmiştir.



Şekil 4. 16 Modelin eğitim aşaması gözlem sayıları

Çalışmada kullanılan yöntemler için tüm modellerde eğitim işlemi yapılmıştır. LSTM modelinin oluşturulması için gerekli analizler Şekil 4.17'de tamamlanmıştır. Geçmiş zaman adımları belirlenirken genellikle 1 veya 3 olarak alınır bu çalışmada geçmiş zaman adımı 3 olarak belirlendi. Ek olarak, bu değer 3 olarak kullanıldığında çok fazla tekrarlama gerektirecektir. Sonuçları elde etmemiz için çok daha uzun süre beklememize sebep olabilir.



Şekil 4. 17 LSTM modeli units değerleri

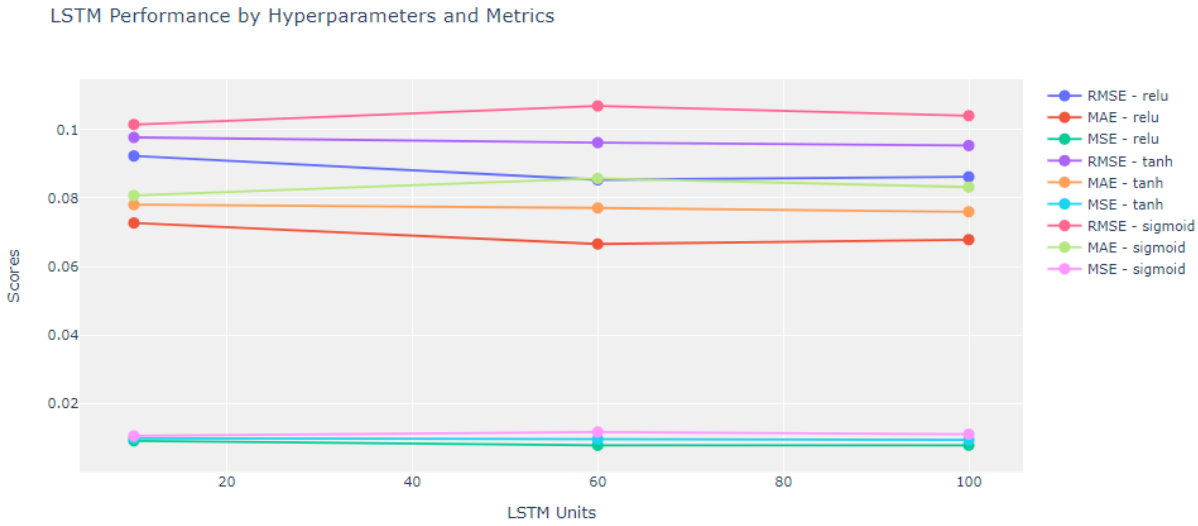
```
109/109 [==============================] - 0s 884us/ste
109/109 [==============================] - 0s 983us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 923us/ste
109/109 [==============================] - 0s 910us/ste
109/109 [==============================] - 0s 893us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 869us/ste
109/109 [==============================] - 0s 881us/ste
109/109 [==============================] - 0s 910us/ste
109/109 [==============================] - 0s 980us/ste
109/109 [==============================] - 0s 983us/ste
109/109 [==============================] - 0s 985us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 905us/ste
109/109 [==============================] - 0s 950us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 909us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 901us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 953us/ste
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 1s 2ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 0s 1ms/step
109/109 [==============================] - 1s 1ms/step
```

Şekil 4.17 LSTM modeli units değerleri (Devamı)

,

4.1.6 Yöntemlerin Model Eğitimindeki (Train Verisi) Hiperparametre Başarıları

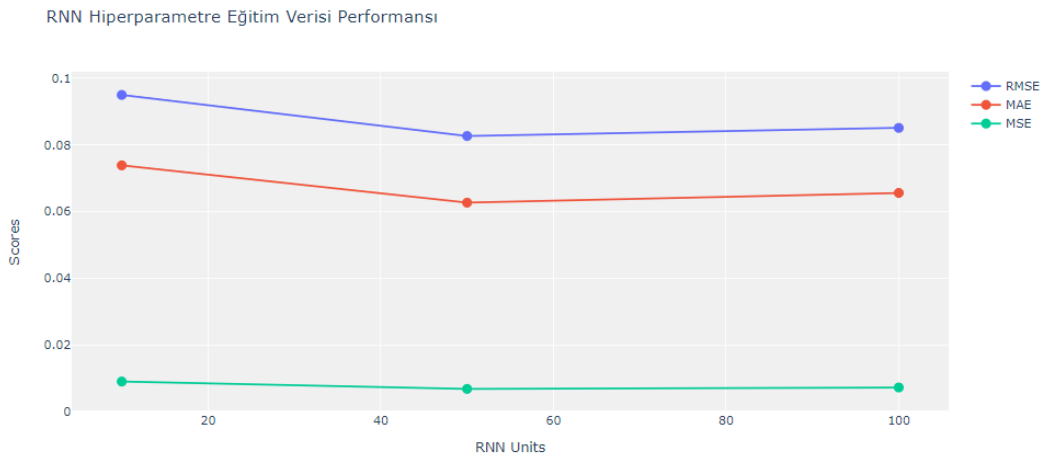
Aşağıda numaralandırılmış şekillerde çalışmada kullanılan tüm modellere ait hiperparametre sonuçları özetlenmiştir. Yöntemlerin en düşük MAE, MSE ve RMSE değerlerine sahip hiperparametre seçilerek test verisi tahminlenecektir.



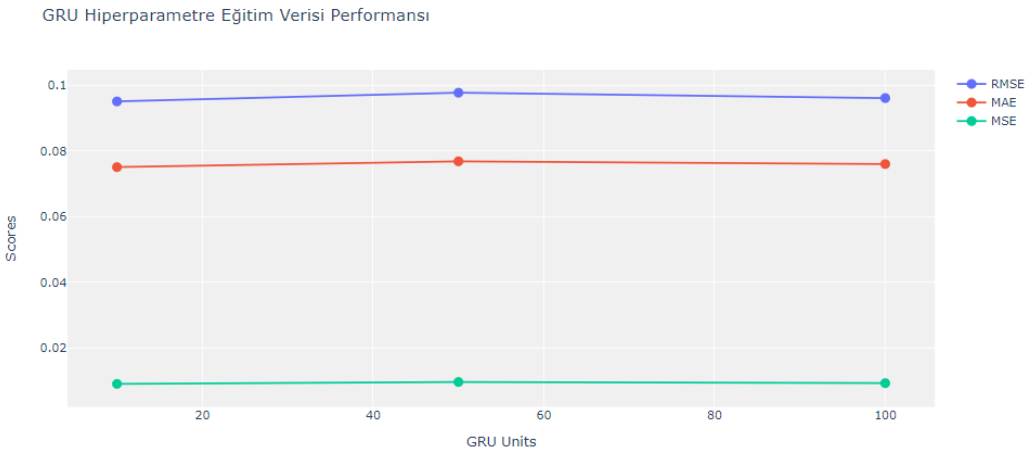
Şekil 4. 18 LSTM modeli hiperparametrelerin performansı



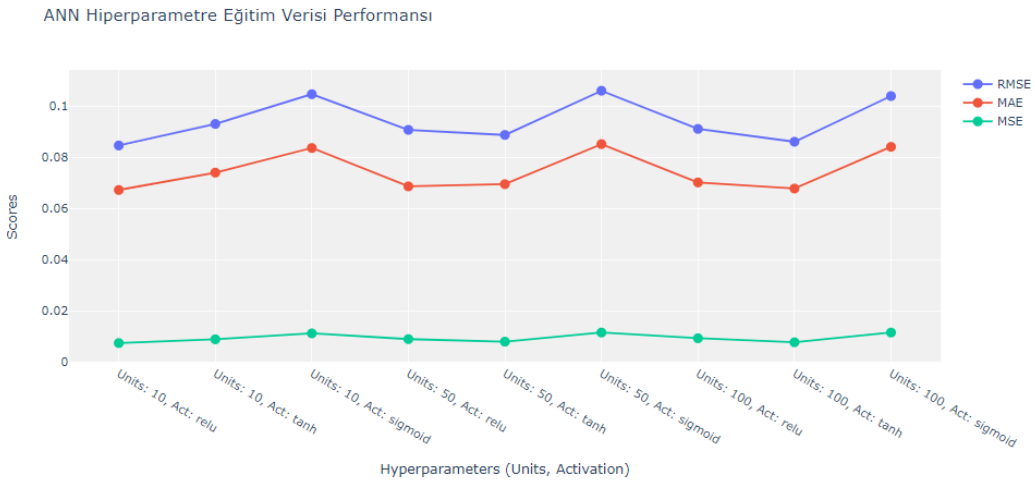
Şekil 4. 19 XGBoost modeli hiperparametrelerin performansı



Şekil 4. 20 RNN modeli hiperparametrelerin performansı

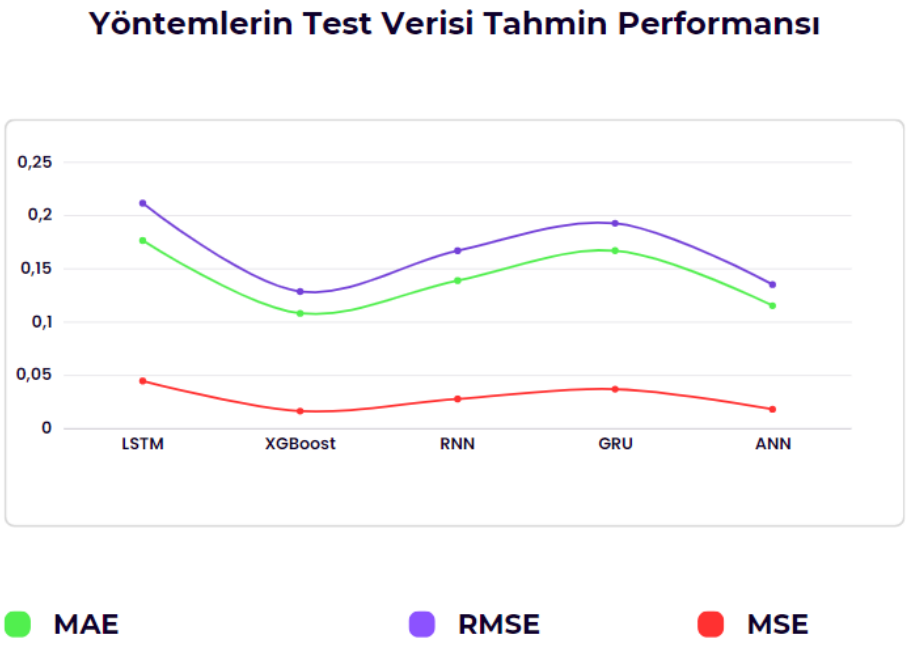


Şekil 4. 21 GRU modeli hiperparametrelerin performansı



Şekil 4. 22 ANN modeli hiperparametrelerin performansı

4.1.7 Yöntemlerin En Doğru Sonucu Veren Hiperparametler Seçilerek Test Verisi Tahmin Başarıları



Şekil 4. 23 Tüm yöntemlerin test verisi tahmin performansları

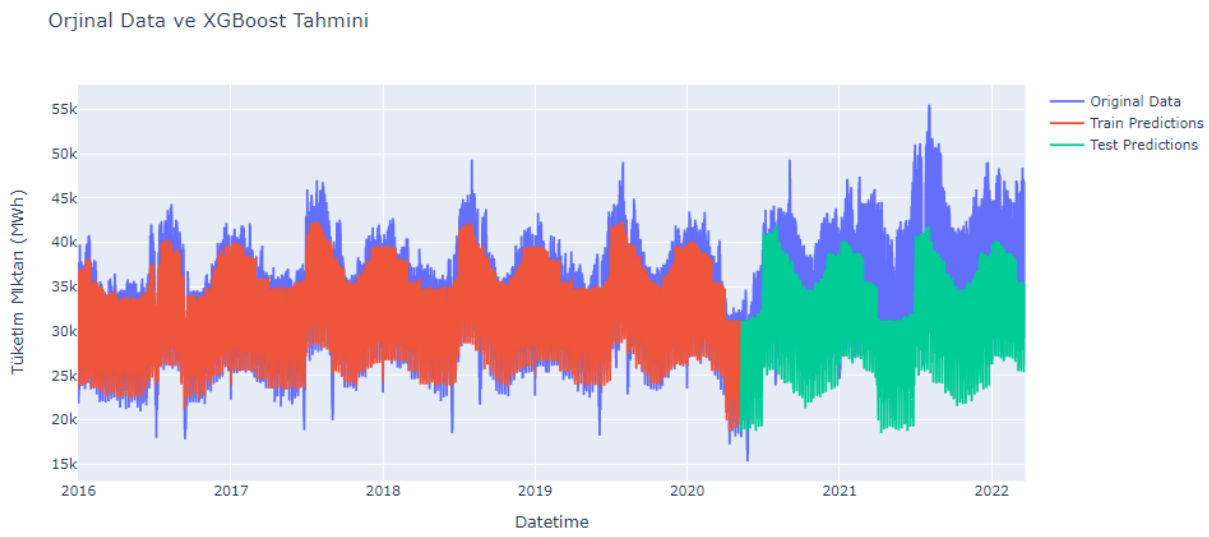
Şekil 4.23 incelendiğinde, çeşitli modellerin ve parametrelerin test verileri üzerindeki performanslarına detaylı bir bakış sunulmaktadır.

Tüm yöntemlere ait hata metrik değerleri Tablo 4.1'de özetlenmiştir.

Her model, optimal (en doğru) parametrelerle seçilmiş ve test verileri üzerinde tahminlerde bulunmuştur. İlerleyen bölüm XGBoost üzerinden devam etmektedir.

Tablo 4. 1 Yöntemlerin hata metrikleri

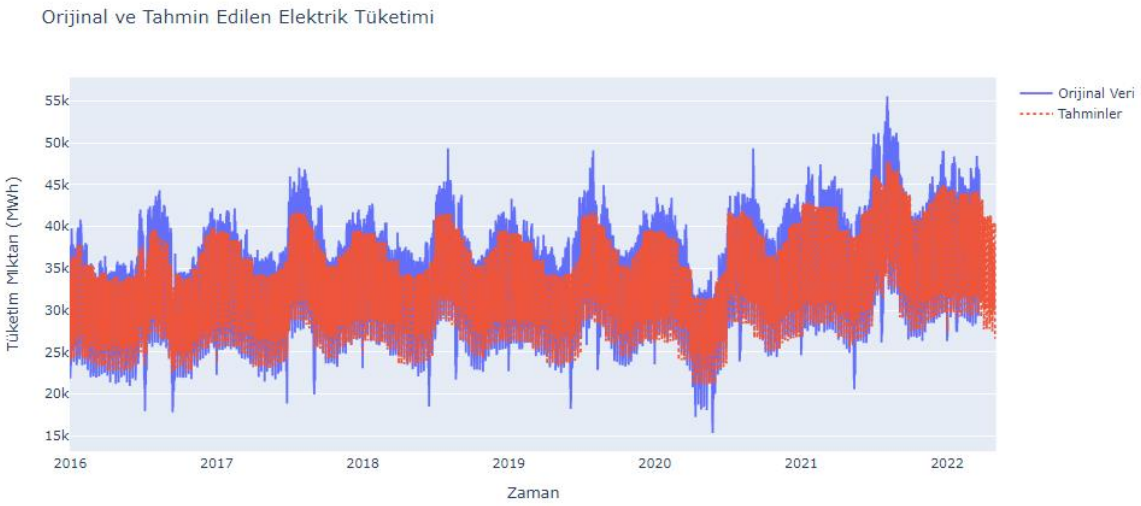
YÖNTEMLER	**MAE**	**MSE**	**RMSE**
LSTM	**0.1766**	**0.0448**	**0.2116**
XGBoost	**0.1082**	**0.0165**	**0.1287**
RNN	**0.1391**	**0.0279**	**0.1671**
GRU	**0.1671**	**0.0372**	**0.1928**
ANN	**0.1154**	**0.0183**	**0.1353**
Lineer Regresyon	**0.7954**	**0.6841**	**0.8271**



Şekil 4. 24 XGBoost model tahminleri ve orjinal veri

Şekil 4.24'te XGBoost modelinin tahminlerini orijinal veriyle birlikte sunmaktadır. Modelin eğitim verisi üzerindeki tahminleri tatmin edici görünse de, test verisi üzerinde aynı başarıyı göstermediği açıktır. Bu, modelin veri setine iyi uyum sağlayamaması, yani underfitting olması sebebiyle olabilir. Underfitting sorununu çözmek için birkaç strateji izlenebilir. Veri setini farklı bir şekilde ayırmayı deneyebiliriz; örneğin, eğitim ve test verilerini %80-%20 yerine başka bir oranda ayırabiliriz. Ölçeklendirme yöntemini değiştirebiliriz. Mevcut durumda 0 ile 1 arasında ölçeklendirme yapılıyor; bunun yerine -1 ile +1 arasında ölçeklendirme yapmayı deneyebiliriz. Modelin hiperparametrelerini optimize edebiliriz. Ancak, bu modelin öğrenme süresini olduça uzatacaktır. Başka bir yol olan train-dev-test ayrımı ile daha başarılı parametre eğitimi yapılabilir. Bu düzenlemelerin ardından modelin performansı tekrar değerlendirilmeli ve gerekli görüldüğünde daha da ayarlamalar yapılmalıdır.

Şekil 4.25'te tüm veri seti ve tahmin edilen elektrik tüketimi zaman serisi verilmiştir.



Şekil 4. 25 Türkiye elektrik tüketim tahmini zaman serisi

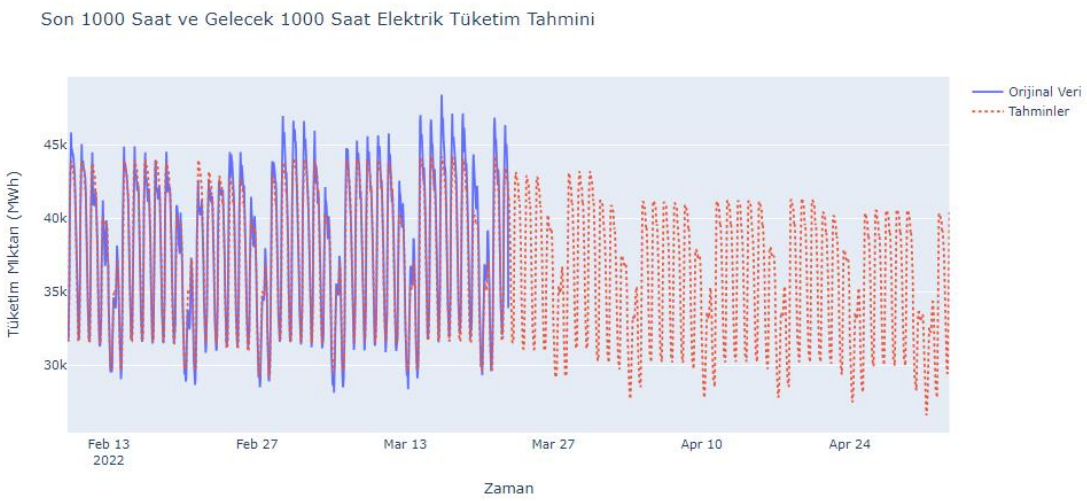
4.1.8 Gelecek Değer Tahmini

Çalışmanın bu bölümünde gelecek 1000 saat için tahmin etme analizleri yapılmıştır. Tüm veri için model eğitme adımları tamamlanmıştır. Şekil 4.26'da tahmin sonuçlarının çıktısı sunulmuştur.

	Tahminler
0	29806.578125
1	28126.904297
2	27094.652344
3	26439.230469
4	26352.882812
...	...
55566	29424.375000
55567	31119.828125
55568	36415.496094
55569	39288.417969
55570	40434.710938

55571 rows × 1 columns

Şekil 4. 26 Gelecek 1000 saat için tahmin sonuçları



Şekil 4. 27 Gelecek 1000 saat için tahmin verisi

Çalışmada gelecek 1000 saat için elektrik tüketim tahmini yapılmıştır. Şekil 4.27 için çok fazla yakınlaştırılması ile 24 Nisan 2022 18:00 tarihinde tahmin edilen tüketim miktarının 34.3828 olduğu Şekil 4.28'de verilmiştir.



Şekil 4. 28 24Nisan 2022 tahmin değeri

EPIAŞ Şeffaflık Platformun'da güncel gerçek zamanlı elektrik tüketim miktarı verileri tekrar çekilmiştir. Şekil 4.28'da bulunan aynı tarih ve saat için güncel veri seti kontrol edildi.

Tablo 4. 2 Gerçek zamanlı tüketim miktarı

Tarih	Saat	Tüketim Miktarı (MWh)
24.04.2022	18:00	34582,42

5
SONUÇ

Bu tez çalışması, Türkiye'nin elektrik tüketiminin tahmin edilmesi ihtiyacını ortaya koymaktadır. Enerji tüketim ihtiyaçlarının gelecek yıllarda çok daha artması, enerji kaynaklarının sürdürebilirliği, güvenliği, enerji maliyetleri ve çevresel etki gibi önemli konularda bilinçli yaklaşım benimsemek hem enerji sektörü hem de toplum açısından önemlidir. Bu kapsamda, elektrik tüketimi takibi ve planlama çalışmaları için elektrik tüketiminin tahmin edilmesi, tüm dünya çapında büyük bir öneme sahiptir.

Elektrik tüketimi, belirli bir zaman aralığında ölçülen genellikle zaman serisi verileri olarak gösterilen bir veri setidir. Makine öğrenmesi, yapay sinir ağları ve derin öğrenme, karmaşık görevlerin çözümünde etkili bir kodlamanın yapılamadığı durumlarda ya da isteklerin değiştiği kurgularda mükemmel bir uygulanabilirlik seçeneği sağlamaktadır. Bu bağlamda, Türkiye'nin elektrik tüketimi tahminini güçlü teorik yapıya sahip metodlar kullanarak başarılı tahminler üretmek için LSTM, XGBoost, RNN, GRU ve ANN teknikleri kullanılmıştır. Model performansı değerlendirme aşamasında en düşük MAE, MSE, ve RMSE hata metriklerine sahip modeller seçilmiştir. Her bir modelin kendine özgü başarıları olmuş, fakat XGBoost, ve Klasik Yapay Sinir Ağı yöntemleri diğerlerine göre daha etkileyici sonuçlar göstermiştir. XGBoost, hızlı öğrenme süreci ile dikkat çekerken; LSTM ise zaman serisi verileri ile çalışmada yaygın olarak kullanılmasına rağmen en doğru sonucu üretememiştir. RNN ve ANN performansı kesinlikle etkileyicidir, fakat enerji tüketimini tahmin etmek söz konusu olduğunda, XGBoost yönteminin seçilmesi daha avantajlı olmuştur. XGBoost modeli, MAE, MSE, ve RMSE için sırasıyla %10.8, %1,6 ve %12,8 olduğu hesaplanmıştır.

Sonuç olarak, çıktı değerlerine bakıldığında, Türkiye'nin elektrik tüketim tahmini için XGBoost modelinin daha üstün performans sergilediği belirlenmiştir. Bu tez çalışması, sadece Türkiye'de değil, diğer ülkelerde de gelecekte yapılacak elektrik tüketimi tahmin çalışmalarına ve enerji politikalarını oluşturma süreçlerine katkıda bulanabilir.

KAYNAKÇA

[1] C. Hamzaçebi, F. Kutay, "Yapay sinir ağları ile Türkiye elektrik enerjisi tüketiminin 2010 yılına kadar tahmini", Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 19(3): 227-233.

[2] S. Yavuz, M. Deveci, "İstatistik normalizasyon tekniklerinin yapay sinir ağın performansına etkisi", Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, (40),167-187.

[3] A.M. Rimeiro, P.R. Carmo, I.R. Silva, D.F. Sadok, T. Lynn, and P.T. Endo, "Short-term firm-level energy-consumption forecasting for energy-intensive manufacturing: a comparison of machine learning and deep learning models" Algorithms. 2020;13(11):274.

[4] G. Oğcu, Ö.F. Demirel, and S.Zaim, " Forecasting Electricity Consumption with neural networks and support vector regression" Procedia – Socail and Behavioral Sciences, vol.58, pp. 1576-1585, 2012.

[5] S. Balli, "Derin tekrarlayan sinir ağı yöntemleri ile erzurum biyolojik atıksu arıtma tesisinin performans tahmini", Yüksek Lisans Tezi, Atatürk Üniversitesi Fen Bilimleri Enstitüsü, Erzurum, 2022.

[6] M.C. Özdemir, "Arima ve xgboost modelleri ile elektrik talep tahmini." Yüksek Lisans Tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü,2022.

[7] AS. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and machine learning forecasting Methods: concerns and ways forward", PLoS ONE 13(3):e0194889.

[8] Elektrik Paketleri Blog, 2018, "Kilowatt (kwh) ve megawatt (mwh)'ın farkı nedir?"

[9] Shura Enerji Dönüşümü, "Elektrifikasyon" ?", https://www.oracle.com/tr/erp/ai-financials/what-is-ai-in-finance/#:~:text=Yapay%20zeka%20nedir%20(AI)%3F,%C3%A7ok%20de%C4%9Ferli%20bir%20i%C5%9F%20varl%C4%B1%C4%9F%C4%B1d%C4%B1r

[10] Wikipedia, 2014, "Dünya enerji tüketimi", https://tr.wikipedia.org/wiki/D%C3%BCnya_enerji_t%C3%BCketimi

[11] Oracle, "Finans alanında yapay zeka nedir?" ?", https://www.oracle.com/tr/erp/ai-financials/what-is-ai-in-finance/#:~:text=Yapay%20zeka%20nedir%20(AI)%3F,%C3%A7ok%20de%C4%9Ferli%20bir%20i%C5%9F%20varl%C4%B1%C4%9F%C4%B1d%C4%B1r.

[12] C. Crisci, B. Ghattas, and G. Perera, " A review of supervised machine learning algorithms and their applications to acological data" Ecological Modelling, vol.240, pp.113-122,2012

[13] M. Usama, J. Qadir, A. Raza, H. Arif, K.A.Yau, Yç Elkhatib,…,A. Al-Fuqaha, "Unsupervised machine learning for networking: Techniques, applications and research challenges" IEEE Access, vol.7, pp. 65579-65615, 2019.

[14] F. Ayayna, "Akciğer kanserinin derin öğrenme yaklaşımları kullanılarak tespit edilmesi", Yüksek Lisans Tezi, Batman Üniversitesi Lisansüstü Eğitim Enstitüsü, Batman, 2023.

[15] AWS Amazon, "Makine öğrenimi ve derin öğrenme arasındaki fark nedir?", https://aws.amazon.com/tr/compare/the-difference-between-machine-learning-and-deep-learning/

[16] C. Olah, Colah' Blog, 2015, "Understanding LSTM Newtorks", https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[17] S. Saxena, Analytics Vidhya, 2024, "What is LSTM? Introduction to long short-term memory", https://www.analyticsvidhya.com/blog/2021/03/introduction-to-long-short-term-memory-lstm/

[18] A. Almalaq, G. Edwards, "A review of deep learning methods applied on load forecasting", 16th IEEE international conference on machine learning and applications (ICMLA), pp. 511-516, 2017.

[19] V. Jitani, Mediumi 2020, "Recurrent neural network: maths in 5 minutes", https://vidishajitani.medium.com/recurrent-neural-network-maths-69214e4d69e1

[20] E. Eğrioğlu, E. Baş, "Zaman Serileri ve Öngörü Yöntemleri (R Uygulamalı), Akademik Yayıncılık, ss.259,2020.

[21] E. Yılmaz, "Yapay sinir ağları ile araç satışlarının tahmin edilmesi", Veri Bilimi Okulu Blog, 2019.

[22] EPIAS Şeffaflık Platformu, "Gerçek Zamanlı Tüketim", https://seffaflik.epias.com.tr/electricity/electricity-consumption/ex-post-consumption/real-time-consumption

TEZDEN ÜRETİLMİŞ YAYINLAR

Konferans Bildirileri

1. G. G. Şimşek, ve D. Hersek, "Makine öğrenmesi yöntemleri ile zaman serisi analizi: Türkiye'de elektrik tüketimi üzerine bir uygulama",12.Uluslararası Akademik Araştırmalar Kongresi ICAR, 19-20 Aralık, 2023.