ELEKTRİK TÜKETİMİNİN DERİN ÖĞRENME YÖNTEMLERİYLE TAHMİN EDİLMESİ

Fatih GÜLER
221402201

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Önder ŞAHİNASLAN

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Şubat, 2024

ELEKTRİK TÜKETİMİNİN DERİN ÖĞRENME YÖNTEMLERİYLE TAHMİN EDİLMESİ

Fatih GÜLER
221402201
Orcid: 0000-0002-2176-8911

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Önder ŞAHİNASLAN

İstanbul
T.C. Maltepe Üniversitesi
Lisansüstü Eğitim Enstitüsü
Şubat, 2024



JÜRİ VE ENSTİTÜ ONAYI

FATİH GÜLER'ın "Elektrik Tüketiminin Derin Öğrenme Yöntemleriyle Tahmin Edilmesi." başlıklı tezi 15.02.2024 tarihinde aşağıdaki jüri tarafından değerlendirilerek "Maltepe Üniversitesi Lisansüstü Eğitim ve Öğretim Yönetmeliği" nin ilgili maddeleri uyarınca Bilgisayar Mühendisliği Anabilim Dalı Yüksek Lisans/Doktora/Sanatta Yeterlik tezi oy birliğiyle/oy çokluğuyla, başarılı/başarısız olarak kabul edilmiştir.

	Unvanı. Adı ve Soyadı	İmza
Üye (Tez Danışmanı)	Dr. Öğr. Üyesi Önder ŞAHİNASLAN Maltepe Üniversitesi	.....
Üye	Dr. Öğr. Üyesi Hüseyin Fehmi Selim BAYRAKLI Milli Savunma Üniversitesi	.....
Üye	Dr. Öğr. Üyesi Erdal GÜVENOĞLU Üniversitesi	.....

Prof. Dr. Selva ERSÖZ KARAKULAKOĞLU
Enstitü Müdürü

ETİK İLKE VE KURALLARA UYUM BEYANI

Bu belge, Yükseköğretim Kurulu tarafından 19.01.2021 tarihli "*Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge*" ile bildirilen 6689 Sayılı Kişisel Verilerin Korunması Kanunu kapsamında gizlenmiştir.

TEŞEKKÜR

Her daim yanımda olan sevgili aileme, kıymetli Amcam Hüseyin GÜLER'e, tez çalışmam sürecinde her türlü desteği veren yöneticilerime, değerli çalışma arkadaşlarıma, çalışmalarım sırasında bilgi ve deneyimleriyle yolumu aydınlatan, samimiyetle her türlü desteğini esirgemeyen kıymetli danışman hocam Dr.Öğr.Üyesi Önder ŞAHİNASLAN'a, teşekkür ederim.

Fatih GÜLER

Şubat, 2024

ÖZET

ELEKTRİK TÜKETİMİNİN DERİN ÖĞRENME YÖNTEMLERİYLE TAHMİN EDİLMESİ

Fatih Güler
Yüksek Lisans Tezi
Bilgisayar Mühendisliği Anabilim Dalı
Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı
Danışman: Dr. Öğr. Üyesi Önder Şahinaslan
Maltepe Üniversitesi Lisansüstü Eğitim Enstitüsü, 2024

Elektrik tüketiminin derin öğrenme yöntemleriyle tahmin edilmesi, enerji sektöründe önemli dönüşüm sağlayan kritik bir teknolojik yaklaşımdır. Günümüzde enerji kaynaklarının sürdürülebilir ve etkili bir şekilde kullanılması, enerji maliyetlerinin düşürülmesi ve çevresel etkilerin minimize edilmesi, toplumlar ve şirketler için temel bir önceliktir. Derin öğrenme, elektrik tüketimi tahmininde önemli bir araçtır. Bu yaklaşım, büyük miktarda veri kullanarak gelecekteki elektrik tüketimini tahmin etme yeteneğine sahiptir. Hava koşulları, mevsimsel değişkenler, önceki tüketim verileri ve diğer faktörleri içeren çeşitli veri kaynaklarından gelen veriyi analiz ederek, derin öğrenme modelleri, elektrik tüketimi tahminlerini hassaslaştırılır. Derin öğrenme yöntemleri aynı zamanda enerji yönetimi sistemlerinin geliştirilmesine de yardımcı olabilir. Anlık elektrik tüketimi verilerini izleyerek enerji tasarrufu fırsatlarını belirleyebilir ve enerji tüketimini optimize edebilirler. Bu, enerji maliyetlerinin azaltılmasına ve enerji kaynaklarının daha verimli kullanılmasına katkı sağlar. Sonuç olarak, elektrik tüketimi tahmini için derin öğrenme, anahtar bir teknoloji olarak öne çıkmaktadır. Enerji sektörünü dönüştürme potansiyeli taşıyan bu yaklaşım, önemli bir rol oynamaktadır. Gelecekte, elektrik tüketiminin daha iyi tahmin edilmesi, enerji verimliliğini artırmak ve sürdürülebilir enerji kullanımını teşvik etmek için hayati bir öneme sahip olacaktır.

Anahtar Sözcükler: Elektrik Tüketimi, Derin Öğrenme, Enerji Yönetimi, Veri Analizi, Tahmin, Enerji Verimliliği, Sürdürülebilirlik.

ABSTRACT

PREDICTION OF ELECTRICITY CONSUMPTION USING DEEP LEARNING METHODS

Fatih Güler
Master Thesis
Department of Computer Engineering
Computer Engineering Master Programme with Thesis
Thesis Advisor: Asst. Prof. Önder Şahinaslan
Maltepe University Graduate School, 2024

Predicting electricity consumption using deep learning methods is a critical technological approach that is driving significant transformation in the energy sector. Today, the sustainable and efficient use of energy resources, the reduction of energy costs, and the minimization of environmental impacts are fundamental priorities for both societies and businesses. Deep learning serves as a valuable tool in electricity consumption prediction. This approach has the capability to forecast future electricity consumption by analyzing a vast amount of data. By scrutinizing data from various sources, including weather conditions, seasonal variations, historical consumption data, and other factors, deep learning models can refine electricity consumption predictions. Furthermore, deep learning methods can also contribute to the enhancement of energy management systems. By monitoring real-time electricity consumption data, they can identify opportunities for energy savings and optimize energy consumption. This, in turn, contributes to cost reduction and more efficient utilization of energy resources. In conclusion, deep learning is emerging as a pivotal technology for electricity consumption prediction. This approach has the potential to transform the energy sector and plays a crucial role in shaping its future. Predicting electricity consumption more accurately will be of vital importance in increasing energy efficiency and promoting sustainable energy usage.

Keywords: Electricity Consumption, Deep Learning, Energy Management, Data Analysis, Prediction, Energy Efficiency, Sustainability

İÇİNDEKİLER

JÜRİ VE ENSTİTÜ ONAYI ii

ETİK İLKE VE KURALLARA UYUM BEYANI iii

TEŞEKKÜR iv

ÖZET v

ABSTRACT vi

İÇİNDEKİLER vii

TABLOLAR LİSTESİ x

ŞEKİLLER LİSTESİ xi

KISALTMALAR xii

SİMGELER LİSTESİ xiii

1. GİRİŞ 1

1.1. Problem 1

1.2. Amaç 2

1.3. Önem 6

1.4. Varsayımlar 7

1.5. Sınırlılıklar 8

1.6. Tanımlar 9

2. GENEL BİLGİLER 11

2.1. Teorik Yaklaşımlar 11
2.2. Literatür Araştırması 12
3. YÖNTEM 23
3.1. Araştırma Modelleri 23
3.2. Eğitim Modelleri 24
3.2.1. KNN 24
3.2.2. Boosting yöntemleri 26
3.2.3. LightGBM 26
3.2.4. XGBoost 29
3.2.5. CatBoost 30
3.2.6. LSTM 32
3.2.7. Grid Search 33
3.2.8. Bayesian optimization 34
3.2.9. Random search 36
3.3. Veriler ve Toplanması 37
3.3.1. Veri çekimi 39
3.3.2. Öznitelik mühendisliği 40
3.3.3. Veri hazırlama süreci 40
3.4. Verilerin Çözümlenmesi ve Yorumlanması 44
3.4.1. LightGBM 44
3.4.2. XGBoost 46

3.4.3. CatBoost 48
3.4.4. KNN 50
3.4.5. LSTM 51
4. BULGULAR VE TARTIŞMA 54
4.1. Bulgular 54
4.2. Tartışma 62
5. SONUÇ VE ÖNERİLER 64
6. KAYNAKÇA 66
ÖZGEÇMİŞ 72

TABLOLAR LİSTESİ

Tablo 1. Faturalanan Tüketim ve Fiili Tüketim 3

Tablo 2. 2022 Yılı Elektrik Piyasası Genel Görünümü 3

Tablo 3. Test Edilen Günler 38

Tablo 4. Test Verisinin Ana Veriye Oranı 38

Tablo 5. Hiper Parametre Optimizasyonu için Seçilen Tarihler 39

Tablo 6. MAPE Değerleri 64

Tablo 7. Ortalama, Medyan ve Standart Sapma Değerleri 65

ŞEKİLLER LİSTESİ

Şekil 1. Kişi Başı Elektrik Enerjisi Tüketimi (kWh/kişi) 6

Şekil 2. KNN Sınıflandırma Örneği 26

Şekil 3. LightGBM Modeli Grafiği 46

Şekil 4. XGBoost Modeli Grafiği 48

Şekil 5. CatBoost Modeli Grafiği 49

Şekil 6. KNN Modeli Grafiği 51

Şekil 7. LSTM Modeli Grafiği 53

Şekil 8. KNN için Alınan Verilerin Tarih ve Saate Göre Grafiği 58

Şekil 9. XGBoost için Alınan Verilerin Tarih ve Saate Göre Grafiği 60

Şekil 10. CatBoost için Alınan Verilerin Tarih ve Saate Göre Grafiği 61

Şekil 11. LightGBM için Alınan Verilerin Tarih ve Saate Göre Grafiği 62

KISALTMALAR

GWh	Gigawatt saat	Gigawatt hour
MW	Megawatt	Megwatt
YEKDEM	Yenilenebilir Enerji Kaynakları Destekleme Mekanizması	
PTF	Piyasa Takas Fiyatı	
MLP	Çok Katmanlı Algılayıcı	Multi-Layer Perceptron
RNN	Tekrarlayan Sinir Ağı	Recurrent Neural Network
MAPE	Ortalama Mutlak Yüzde Hatası	Mean Absolute Percentage Error
ERNN	Elman Tekrarlayan Sinir Ağı	Elman Recurrent Neural Networks
LSTM	Uzun Kısa Süreli Hafıza	Long Short-Term Memory
ESN	Yankı Durumu Ağı	Echo State Network
GRU	Kapılı Tekrarlayan Birimler	Gated Recurrent Units
CNN	Evrişimli Sinir Ağları	Convolutional Neural Networks
TCN	Zamansal Evrişimli Sinir Ağı	Temporal Convolutional Networks

SİMGELER LİSTESİ

$\sum$	:Toplam fonksiyonu
$\boldsymbol{F(x)}$	:Başlangıç modeli
M	:Toplam ağaç sayısı
$\boldsymbol{h_m(x)}$	:m'inci ağacın tahmini
$\boldsymbol{\gamma_m}$	:m'inci ağacın çıktısının öğrenme oranı ile çarpılmış ağırlığı
N	:Veri noktalarının sayısı
L	:Kayıp fonksiyon
$\boldsymbol{y_i}$	:i'nci veri noktasının değeri
$\boldsymbol{F_{m-1}(x_i)}$	:i'nci veri noktası için önceki modelin tahmini
$\boldsymbol{\phi(x_i)}$	:Tahmin fonksiyonu
K	:Toplam ağaç sayısı
$\boldsymbol{f_k}$	:k'inci ağaç
$\boldsymbol{\mathcal{F}}$	:Tüm olası ağaçların uzayı
$\boldsymbol{x_{new}}$	:Yeni parametre seti
$\boldsymbol{y_{new}}$	:Gerçek fonksiyon değeri
$\boldsymbol{D_{new}}$	:Yeni gözlemlerin eski gözlemlerle birleştirilmesi

1. GİRİŞ

Günümüz dünyasında enerji tüketimi ve yönetimi giderek daha karmaşık bir hale gelmekte ve hayati önem arz etmektedir. Teknolojik gelişmeler, endüstriyel büyüme ve artan nüfus nedeniyle elektrik talebi artmaktadır. Bu durum, enerji kaynaklarının sürdürülebilir şekilde kullanılması, maliyetleri düşürmek ve çevresel etkileri azaltmak için önemli olduğunu göstermektedir. Bu bağlamda, enerji yönetimi ve elektrik tüketimi tahmini için derin öğrenme yöntemleri oldukça önemlidir. Derin öğrenme, büyük veri havuzlarını analiz ederek gelecekteki tüketimi tahmin etme ve enerji verimliliği ve sürdürülebilirlik konularında katkıda bulunma yeteneğine sahiptir.

Gelişen teknoloji ve artan toplumsal ihtiyaçlar nedeniyle, elektrik tüketiminin doğru tahmin edilmesi ve yönetilmesi oldukça önemlidir. Bu çalışma, gerilim seviyelerine bağlı olarak çeşitli tüketici gruplarının elektrik tüketimini inceler ve gelecekteki tüketimi tahmin etme tekniklerini değerlendirir. 2018-2023 yıllarına ait elektrik tüketimi verilerini kullanılarak yapılan tahminler, enerji planlaması, kaynakların verimli kullanımı ve çevresel etkilerin azaltılması gibi enerji sektörünün karşılaştığı önemli sorunları çözmeye yardımcı olur.

1.1. Problem

Günümüz dünyasında, enerji tüketimi ve yönetimi giderek artan bir karmaşıklık ve kritiklik düzeyine ulaşmıştır. Hızla büyüyen nüfus, endüstrinin genişlemesi ve teknolojik gelişmeler, elektrik talebinde sürekli bir artışa işaret etmektedir. Bu artışın, enerji kaynaklarının sürdürülebilir ve etkili bir şekilde kullanılması, enerji maliyetlerinin azaltılması ve çevresel etkilerin minimize edilmesi ile nasıl başa çıkılacağı konusu, toplumlar ve işletmeler için giderek daha fazla önem kazanmaktadır.

Enerji, modern toplumların vazgeçilmez bir unsuru haline gelmiştir. Elektrik tüketimi, bu önemli enerji kaynağının kullanımının merkezindedir. Günümüzde, elektrik tüketimi, endüstriyel tesislerden evlere kadar her alanda büyük bir role sahiptir. Bu bağlamda, elektrik tüketiminin etkili bir şekilde yönetilmesi ve tahmin edilmesi, enerji sektöründe

ve daha geniş bir bağlamda toplumsal ve ekonomik sürdürülebilirlik açısından kritik bir rol oynamaktadır.

Bu bağlamda, derin öğrenme yöntemleri, yapay zekâ alanındaki devrim niteliğindeki gelişmeler arasında yer almaktadır. Derin öğrenme, yapay sinir ağlarını daha büyük ve karmaşık veri kümelerini işlemek için kullanma fikrine dayanır. Elektrik tüketimi tahmininde, derin öğrenme yöntemleri, bu büyük veri havuzunu analiz ederek gelecekteki tüketimi tahmin etme yeteneği sunar. Bu, hava koşulları, mevsimsel değişkenler, önceki tüketim verileri ve diğer faktörleri içeren çeşitli veri kaynaklarından gelen bilgilere dayanarak elektrik tüketimi tahminlerini daha kesin hale getirme gücünü taşır.

Derin öğrenme yöntemleri, sadece elektrik tüketimi tahmini ile sınırlı kalmaz; aynı zamanda enerji yönetimi sistemlerinin geliştirilmesine de önemli katkılarda bulunur. Anlık elektrik tüketimi verilerini izleyerek enerji tasarrufu fırsatlarını belirleyebilir ve enerji tüketimini optimize edebilirler. Ayrıca, enerji arzı ve talebini daha iyi dengelemeye yardımcı olurlar, bu da enerji güvenilirliğini artırır.

Derin öğrenme teknolojisinin enerji sektöründeki etkisi, enerji maliyetlerinin azaltılmasının ötesine geçer. Aynı zamanda enerji kaynaklarının daha verimli bir şekilde kullanılmasına ve sürdürülebilir enerji uygulamalarının teşvik edilmesine yardımcı olur.

1.2. Amaç

Gelişen teknoloji ve toplumsal ihtiyaçlar, enerji sektöründe elektrik tüketimi konusunda artan bir karmaşıklık ve önem düzeyine işaret etmektedir. Bu bağlamda, elektrik tüketiminin doğru bir şekilde tahmin edilmesi ve yönetilmesi, hem enerji şirketleri hem de toplumlar için kritik bir öneme sahiptir. Elektrik tüketimi tahmini, enerji planlaması, kaynakların etkili bir şekilde kullanılması, enerji kesintilerinin minimize edilmesi ve çevresel etkilerin azaltılması açısından temel bir faktördür. Bu bağlamda, bu çalışma, 2018-2023 yıllarına ait elektrik tüketimi verilerini ele alarak, farklı tüketici türlerinin dağıtım ve iletim gerilim seviyelerine bağlı tüketim miktarlarını incelemekte ve bu verilere dayalı olarak gelecekteki elektrik tüketimini tahmin etme yöntemlerini değerlendirmeyi amaçlamaktadır. Elektrik tüketiminin bu tür tahminlemeleri, enerji sektörünün daha sürdürülebilir ve verimli bir şekilde yönetilmesine katkı sağlamaktadır.

Bu nedenle, elektrik tüketim tahmini, enerji sektöründe karşılaşılan zorlukların ve fırsatların anlaşılması için kritik bir rol oynamaktadır. Bu çalışmanın amacı, bu önemli konuyu daha ayrıntılı bir şekilde incelemek ve gelecekteki enerji yönetimi stratejilerine ışık tutmaktır.

Tablo 1. Faturalanan Tüketim ve Fiili Tüketim (EPDK, 2023)

Tüketici Türü	Dağıtım Gerilim Seviyesinden Bağlı Tüketicilerin Tüketim Miktarı (GWh)		İletim Gerilim Seviyesinden Bağlı Tüketicilerin Tüketim Miktarı (GWh)		Toplam (GWh)	
	2021	2022	2021	2022	2021	2022
Abone	101.723,92	122.665,91	606,32	1.662,94	102.330,24	124.328,85
Serbest Tüketici	87.572,64	69.949,50	63.154,02	59.343,10	150.726,66	129.292,60
Genel Toplam	189.296,56	192.615,41	63.760,35	61.006,04	253.056,90	253.621,45

Tablo 2. 2022 Yılı Elektrik Piyasası Genel Görünümü (EPDK, 2023)

Konu Başlığı	Birim	2020 Yılı	2021 Yılı	2022 Yılı	2021-2022 Değişim (%)
		Değeri	Değeri	Değeri	
Lisanslı Kurulu Güç	MW	89.067	92.273	95.174	3,1
Lisanslı Üretim*	GWh	294.248	320.081	311.779	-2,6
En Yüksek Ani Puant	MW	49.851	56.303	52.286	-7,1
En Düşük Ani Puant	MW	15.690	20.611	19.451	-5,6
Lisanssız Kurulu Güç	MW	6.823	7.547	8.635	14,4

İhtiyaç Fazlası Satın Alınan Lisanssız Üretim Miktarı	GWh	11.245	11.950	12.258	2,6
Brüt Lisanssız Üretim Miktarı	GWh	11.347	12.217	12.738	4,3
YEKDEM Üretim	MWh	73.482.227	74.156.245	83.852.638	13,1
YEKDEM Ödeme Tutarı	TL	46.335.283.734	61.442.093.854	131.455.875.292	113,9
Fiili Tüketim	GWh	304.999	330.440	327.221	-1
Faturalanan Tüketim	GWh	233.397.377	253.056.905	253.621.451	0,2
Tüketici Sayısı	Adet	46.082.450	47.311.986	48.563.459	2,6
İthalat	GWh	1.888	2.329	6.414	175,4
İhracat	GWh	2.484	4.187	3.710	-11,4
Ortalama YEKDEM fiyatı	TL/MWh	630,56	828,55	1.567,70	89,2
YEKDEM Ek Maliyeti	TL/MWh	108,52	93,98	-317,53	-437,9
Ağırlıklı Ortalama PTF	TL/MWh	285,87	507,85	2.528,12	397,8
Aritmetik Ortalama SMF	TL/MWh	272,29	498,8	2.548,65	411

Tablo 1 ve Tablo 2'deki veriler, enerji sektöründeki önemli göstergeleri ve değişkenleri kapsamaktadır.

Tablo 1’de: "Abone" türündeki tüketicilerin 2021 yılında GWh cinsinden tüketimi 2022 yılına göre önemli bir artış göstermiştir. Aynı şekilde, "Serbest Tüketici" türündeki tüketicilerin tüketimi 2022 yılında düşmüştür. Bu, farklı tüketici türleri arasında tüketim değişikliklerinin olduğunu göstermektedir.

Genel toplam tüketim (Toplam (GWh)), 2021 ve 2022 yılları arasında hemen hemen aynı kalmıştır, bu da toplam tüketimde büyük bir değişiklik olmadığını göstermektedir.

Tablo 2’de: "Lisanslı Kurulu Güç" (MW), 2022 yılına doğru artış göstermiştir. Bu, lisanslı enerji üretim kapasitesinde bir artış olduğunu gösterir.

"Lisanslı Üretim" (GWh) ise 2022 yılında 2021'e göre bir düşüş göstermiştir. Bu, lisanslı üretimdeki miktarın azaldığını yansıtmaktadır.

"YEKDEM Üretim" (MWh) ve "YEKDEM Ödeme Tutarı" (TL), 2022 yılına doğru önemli ölçüde artmıştır. Bu, Yenilenebilir Enerji Kaynakları Destekleme Mekanizması (YEKDEM) çerçevesinde yenilenebilir enerji üretimine olan desteğin arttığını göstermektedir.

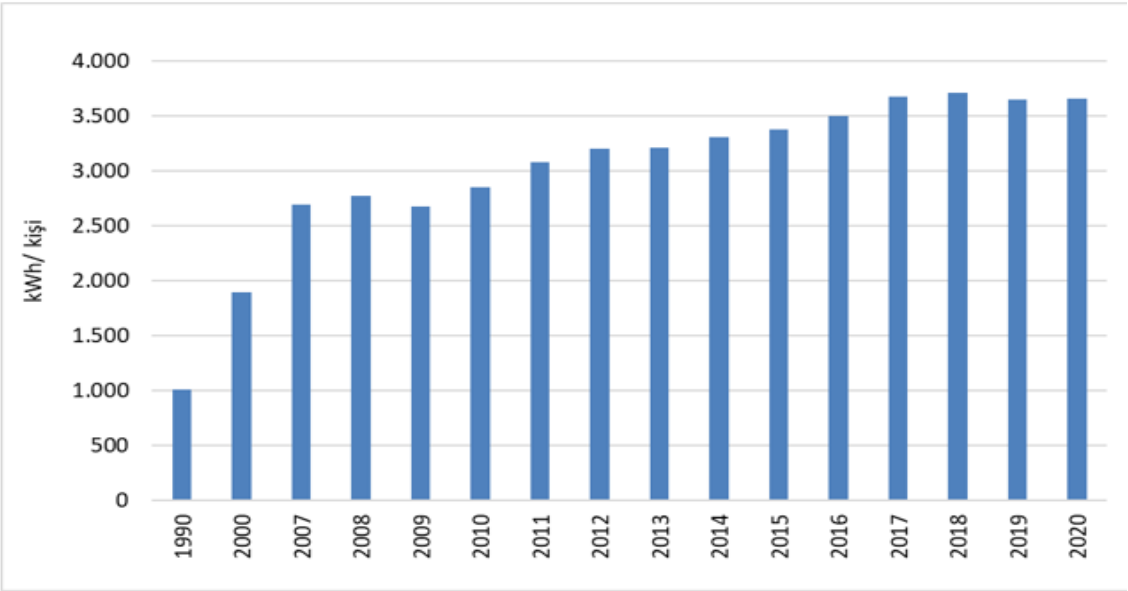
"Fiili Tüketim" (GWh), 2021'den 2022'ye düşmüştür. Bununla birlikte, "Faturalanan Tüketim" (GWh) hemen hemen aynı kalmıştır. Bu durum, faturalanan tüketim ile fiili tüketim arasındaki farkı göstermektedir.

"İthalat" (GWh) 2022 yılında büyük bir artış gösterirken, "İhracat" (GWh) düşmüştür. Bu, enerji ticareti açısından dışa bağımlılık ve dış ticaret dengesindeki değişiklikleri yansıtmaktadır.

"Ortalama YEKDEM Fiyatı" (TL/MWh) 2022 yılında önemli bir artış göstermiş ve YEKDEM üretimine sağlanan ödemenin arttığını göstermektedir.

"Ağırlıklı Ortalama PTF" (TL/MWh) ve "Aritmetik Ortalama SMF" (TL/MWh), 2022 yılında büyük artışlar göstermiştir, bu da enerji fiyatlarının yükseldiğini göstermektedir.

Bu yorumlar, enerji sektöründeki değişkenlikleri ve bu değişkenlerin birbirleriyle ilişkilerini göstermektedir. Bu tür veriler, enerji sektörünün daha iyi anlaşılması, politika geliştirmesi ve enerji planlaması için kritik öneme sahiptir.



Şekil 1. Kişi Başı Elektrik Enerjisi Tüketimi (kWh/kişi) (csb.gov.tr, 2023)

1.3. Önem

Türkiye'de kişi başı elektrik enerjisi tüketimi, 1990'dan 2022'ye kadar olan dönemde önemli değişiklikler göstermiştir. Bu dönemdeki tüketim trendlerini incelediğimizde şu anahtar gözlemler yapılabilir:

Sürekli Artış: 1990 yılından itibaren Türkiye'de kişi başı elektrik enerjisi tüketimi sürekli bir artış göstermiştir. Bu artış, nüfus artışı, ekonomik büyüme ve sanayileşme gibi faktörlere bağlıdır. Türkiye'nin hızla gelişen bir ülke olması, enerji talebinin artmasına neden olmuştur.

Ekonomik Büyüme ile İlişkili: Türkiye'deki kişi başı elektrik tüketimi ile ekonomik büyüme arasında güçlü bir ilişki vardır. Ekonomik büyüme, endüstriyel üretim ve ticaretin artmasına yol açar, bu da daha fazla enerji ihtiyacı anlamına gelir. Bu nedenle, Türkiye'nin ekonomik büyüme süreci, enerji tüketimindeki artışı tetiklemiştir.

Kentleşme ve Sanayileşme: Türkiye'de kentleşme oranının artması ve sanayileşmenin hızlanması, enerji talebini artırmıştır. Kentlerde yaşayan nüfusun artışı, elektrik enerjisi talebinde büyümeye neden olmuştur. Ayrıca, enerji yoğun sanayi sektörleri de enerji tüketimini artıran bir faktördür.

Enerji Verimliliği ve Yenilenebilir Enerji: Son yıllarda Türkiye, enerji verimliliğini arttırmaya ve yenilenebilir enerji kaynaklarına yatırım yapmaya başlamıştır. Bu çabalar, enerji tüketimi artış hızını yavaşlatabilir. Ancak, hâlâ kişi başı enerji tüketimi yüksek bir seviyededir.

İklim Değişikliği ve Sürdürülebilirlik: Küresel iklim değişikliği endişeleri ve sürdürülebilirlik hedefleri, enerji tüketimi konusunda daha dikkatli ve verimli bir yaklaşım benimsemeyi teşvik etmektedir. Türkiye'nin bu bağlamdaki çabaları, gelecekteki enerji tüketimi üzerinde etkili olabilir.

Sonuç olarak, Türkiye'de kişi başı elektrik enerjisi tüketimi, ekonomik büyüme ve kentleşme gibi faktörlerle yakından ilişkilidir. Türkiye, enerji verimliliği, yenilenebilir enerji kullanımı ve sürdürülebilir enerji politikalarına daha fazla odaklanarak enerji tüketimini dengelemeye çalışabilir. Bu çabalar, enerji kaynaklarının daha sürdürülebilir bir şekilde kullanılmasını sağlayabilir ve çevresel etkileri azaltabilir.

1.4. Varsayımlar

Bu çalışmada çalışmanın geçerli kılınması amacıyla bazı varsayımlara yer verilmiştir. Yapılan varsayımlar, bu çalışmanın gerçekleştirilmesi için önem arz etmektedir. Bu varsayımlar gerçekleşmediği takdirde yazılan bu çalışmada değişiklikler öngörülebilir. Bu varsayımlar aşağıdaki gibi yer almaktadır:

- Ekonominin istikrarı: Türkiye ve dünya ekonomisinin bu çalışma yazılırken ki durumunun stabil olması varsayımı yapılmaktadır. Ekonominin mevcut düzen içinde devam etmesi ve büyük çaplı ekonomik buhranların yaşanmaması, bu çalışmada anlatılan konuyu etkilemeyeceği varsayımı yapılmaktadır.

- İklim koşullarının devamlılığı: Elektrik tüketimi mevcut koşullarda iklimin değişikliği ile değişim gösterebilmektedir. Bu bağlamda iklimin büyük ölçülerde değişmeyeceği varsayımı yapılmalıdır. Bu çalışmada ele alınan derin düşüncenin, iklim verilerinin değişmeyeceği kanısında devam etmesi varsayılmalıdır.

- Verilerin doğruluğu: Kullanılan verilerin güvenilirliği esas alınmalıdır. Eğer kullanılan veriler güvenilir değilse, yapılan programların tahmin yüzdesinin düşük olma olasılığı ortaya çıkabilir.

- Enerji politikalarındaki sabitlik: Türkiye ve dünyadaki politikaların sabit kalması varsayımı yapılmaktadır. Ancak, bu politikalarda herhangi bir büyük çaplı değişiklik olursa, yapılan programların sağlığını bozma ihtimali bulunmaktadır.

Yapılan bu varsayımlar yazılan bu çalışmanın yorumlanma çerçevesini oluşturmaktadır. Bu varsayımlarda oluşan herhangi büyük çaplı değişim, bu çalışmada alınan verilerin doğruluğundaki yüzdelerin düşme ihtimalini oluşturur.

1.5. Sınırlılıklar

Bu çalışmanın uygulanması sırasında bazı sınırlamalar vardır. Çalışmanın genel kapsamını tanımlamak için bu kısıtlamalar, elde edilen bulguların ve önerilerin yorumlanmasını kısıtlamaktadır.

Elektrik tüketimi tahmini konusunda yapılan çalışmalar, veri kapsamı ve kalitesinin önemine dikkat çeker. Kullanılan veri setlerindeki eksiklikler veya kalite sorunları, tahmin modellerinin gerçek dünya koşullarına uyumunu zorlaştırabilir. Ayrıca, elektrik tüketimini etkileyen dinamik faktörlerin, örneğin iklim değişiklikleri ve ekonomik dalgalanmaların gelecekte nasıl değişeceğini tahmin etmek oldukça karmaşıktır. Bu da modellerin uzun vadeli performansını etkileyebilir.

Enerji politikalarındaki önemli değişiklikler de enerji tüketimini önemli ölçüde etkileyebilir. Bununla birlikte, bu çalışmalar genellikle enerji politikalarındaki önemli değişikliklerin kapsamlı bir analizini yapmaz; bunun yerine, bu değişikliklerin etkilerini daha geniş bir çerçevede incelemeyi tercih eder. Çalışmaların genel olarak kullanılabilirliği de bir sorun olabilir. Örneğin, Türkiye'deki elektrik tüketimi üzerine yapılan bir çalışmanın sonuçları diğer yerlerde doğrudan kullanılabilir değildir.

Son olarak, uluslararası ekonomik durum ve siyasi istikrarsızlık gibi dışsal faktörler elektrik tüketimini etkileyebilir. Ancak bu tür dışsal faktörlerin kontrolü ve tahmini, bu tür çalışmaların genellikle kapsamı dışında kalır. Bu nedenle, bu faktörlerin elektrik tüketimi üzerindeki etkileri genellikle tahmin modellerinde yeterince dikkate alınmayabilir.

1.6. Tanımlar

Python: Python, genel amaçlı bir programlama dili olup, temiz ve okunabilir bir söz dizimine sahiptir. Açık kaynaklı bir topluluk tarafından geliştirilen bu dil, çeşitli platformlarda kullanılabilir. Veri analizi, yapay zekâ, web geliştirme, bilimsel hesaplamalar ve diğer birçok alanda yaygın olarak kullanılan Python, hem basit hem de karmaşık projeler için uygundur. Python, programcılara esneklik sağlar ve kütüphane ve modüllerin geniş bir ekosistemi sayesinde çeşitli görevleri kolayca yerine getirebilir. Python, öğrenmesi basit ve hızlı bir şekilde verimli yazılım geliştirmenizi kolaylaştırır.

EPİAŞ: Organize Toptan Elektrik Piyasalarında Şeffaflığın Teminine İlişkin Usul Ve Esaslar, Enerji Piyasası Düzenleme Kurulu tarafından 13/05/2016 tarihli ve 6282-4 kararı aracılığıyla Şeffaflık Platformu tarafından yayınlanmalıdır. Şeffaflık Platformu'nun temelini oluşturan ve 28 Mayıs 2016 tarihli 29725 sayılı Resmî Gazetede yayınlanarak yürürlüğe giren bu karar, 06.01.2022 tarihli 10711 Sayılı Kurul Kararı ile güncellenmiştir.

Excel: Excel, Microsoft tarafından geliştirilen ve popüler bir elektronik tablo programıdır. Excel, kullanıcıların verileri düzenlemelerine, hesaplamalarına ve grafikler oluşturmalarına olanak tanır. Elektronik tablolar, verileri hücreler olarak adlandırılan birimlerin birleşim noktalarında saklar ve bu verilere matematiksel formüller ve işlevler uygulayarak otomatik hesaplamalar yapabilir. Ayrıca verileri görsel olarak tablolar, grafikler ve grafikler olarak gösterme yeteneğine sahiptir, bu da veri analizi ve raporlama işlerini kolaylaştırır. Finansal analizden envanter yönetimine kadar birçok alanda kullanılan Excel, verilerin düzenlenmesini, analizini ve sunulmasını çok kolaylaştırır.

Anaconda: Anaconda, bilimsel hesaplamalar ve veri bilimi için kullanılan bir Python dağıtımıdır. Python programlama dilini ve çeşitli veri bilimi kütüphanelerini içeren bir paket yöneticisi ve sanal ortam yöneticisi içerir. Bu, kullanıcıların veri analizi, makine

öğrenimi, bilimsel hesaplamalar ve yapay zekâ gibi karmaşık görevleri kolayca yerine getirmelerine yardımcı olur. Anaconda, veri bilimi topluluğu tarafından yaygın olarak tercih edilen ve çeşitli işletim sistemlerinde kullanılabilen bir platformdur. Jupyter Notebook gibi popüler araçlarla birlikte gelir ve veri analizi ve interaktif kodlama için idealdir.

2. GENEL BİLGİLER

Elektrik tüketimi tahminleri, bir ülkenin ekonomik büyümesi, endüstriyel faaliyetleri ve halkın yaşam kalitesi ile sıkı bir ilişki içindedir. Bu nedenle, bu bölümde, elektrik tüketiminin önemi vurgulanmakta ve bu alandaki tahminlerin enerji politikaları, planlama ve sürdürülebilir kalkınma için temel bir rol oynadığı üzerinde durulmaktadır. Ayrıca, bu bölümde elektrik tüketimi tahmini için kullanılan çeşitli metodolojiler detaylandırılır. Bu yöntemler arasında yapay zekâ ve makine öğrenmesi teknikleri, istatistiksel modeller ve hibrit modeller bulunmakta, her bir yöntemin avantajları, sınırlamaları ve uygulama alanları incelenmektedir. Şahinaslan ve diğerlerine (2023) göre, *"Makine öğrenimi, derin öğrenme algoritmaları kullanarak insan zekâsını taklit eden bir teknolojidir"* (Şahinaslan, Günerkan, ve Şahinaslan, 2023, s. 1).

Elektrik tüketimi tahmininin tarihsel gelişimi ve bu alanda yapılan önemli araştırmalar da bu bölümde özetlenir. Bu özetler, farklı zaman dilimlerinde gerçekleştirilen önemli çalışmaların bulgularını ve enerji tahmin alanında nasıl bir ilerleme kaydedildiğini anlatır. Ayrıca, elektrik tüketimi üzerindeki coğrafi ve ekonomik faktörlerin etkisi de tartışılır; farklı ülkelerde ve farklı ekonomik koşullarda elektrik tüketiminin nasıl değişiklik gösterdiği ve bu değişimlerin tahmin modellerine nasıl entegre edildiği üzerinde durulur.

Son olarak, bu bölümde elektrik tüketimi tahmini alanında gelecekte hangi araştırma yönlerine odaklanılması gerektiği, mevcut yöntemlerin nasıl geliştirilebileceği ve yeni teknolojilerin bu alandaki potansiyeli üzerine görüşler sunulur. Bu sayede, enerji sektöründe karar vericiler, araştırmacılar ve politika yapıcılar için önemli bilgiler aktarılır. Bu alandaki çeşitliliği ve kompleksliği vurgular.

2.1. Teorik Yaklaşımlar

Elektrik tüketimi tahmininde kullanılan teorik yaklaşımlar, zaman serisi verilerinin karmaşık desenlerini ve uzun vadeli bağlantılarını öğrenmek için derin öğrenme gibi yapay sinir ağlarına dayalı gelişmiş tekniklerden yararlanır. Bu modeller, geçmiş verilerin analizi yoluyla gelecekteki eğilimleri tahmin etme yeteneğine sahip olup, elektrik

tüketimi ile ekonomik göstergeler gibi dışsal faktörler arasındaki dinamik ilişkileri modellemek için de kullanılır.

Ayrıca, regresyon analizleri gibi istatistiksel yöntemler, elektrik tüketimini etkileyebilecek dışsal faktörlerin (örneğin, hava durumu veya nüfus artışı) etkisini incelemek için tercih edilir. Bu analizler, belirli faktörlerin tüketim üzerindeki doğrudan etkisini anlamak için son derece faydalıdır.

Elektrik tüketimi tahmininde ayrıca, veri zarflama analizi ve karar verme teknikleri gibi yöntemler de kullanılır. Bu teknikler, verimlilik ve etkinlik analizi yapmak için uygundur ve elektrik tüketimi verilerini analiz ederek etkinlik skorları oluşturur. Aynı zamanda, genetik algoritmalar, bulanık mantık ve yapay sinir ağları gibi yapay zekâ yaklaşımları da elektrik tüketim tahmininde kullanılır. Bu teknikler, veri setlerinin içerdiği karmaşık ilişkileri modellemek ve tahminler yapmak için kullanışlıdır.

Son olarak, elektrik tüketimi tahmininde zaman serisi analizi önemli bir role sahiptir ve geçmiş veri trendlerine dayanarak gelecekteki tüketim seviyelerini tahmin etmek için kullanılır. Bu geniş çaplı yaklaşımlar, elektrik tüketimi tahmininde kullanılan çeşitli teknik ve modellerin temelini oluşturur. Her birinin kendi avantajları ve sınırlamaları bulunmaktadır. Bu modeller, enerji sektöründeki karar vericilere, politika yapıcılara ve araştırmacılara değerli bilgiler sunar.

2.2. Literatür Araştırması

Kaytez (2012) yaptığı bu çalışmada, En Küçük Kareler Destek Vektör Makineleri tekniğinin Türkiye'deki elektrik tüketim tahmininde diğer yöntemlere göre daha doğru olduğunu göstermektedir. Çalışma, günlük sıcaklık değişimleri ve 3000 saatlik geçmiş yük verisinden oluşuyordu. Test verileri, performans, çoklu korelasyon katsayısı ve ortalama yüzde mutlak hata ile değerlendirildi. En Küçük Kareler Destek Vektör Makineleri yönteminin ilk başta çok iyi sonuçlar vermediği vurgulanmıştır. Ancak giriş parametrelerinin değiştirilmesi ile olumlu sonuçlar elde edilebilir. Bu çalışma, Türkiye'de yük tahmininde ve enerji tüketim tahmininde kullanılan çeşitli yöntemlerle karşılaştırıldığında, enerji sektöründe kullanılabilecek çeşitli modellerin geliştirilmesine yardımcı olabilir (Kaytez, 2012).

Uğurel (2013) yaptığı çalışmada, Türkiye'nin elektrik tüketimini ve elde edilen enerji kaynaklarının yıllara göre regresyon modelini ve gelecekte oluşabilecek tüketimin modelini incelemektedir. Çalışma, Türkiye'deki elektrik tüketiminin modellenmesi ve optimizasyonu üzerinedir. İyi bir strateji ile yürütülen süreç optimizasyon çalışmaları, işletmelere maliyet tasarrufu, zaman tasarrufu, artan verimlilik, azalan hatalar ve müşteri memnuniyeti gibi birçok avantaj sağlamaktadır (Uğurel, 2013).

Masaebi (2016) yaptığı çalışmada, İran'ın net elektrik enerjisi tüketimini modellemek için yapay sinir ağlarını kullandı. Modelin eğitim setleri için mutlak yüzde hata oranı %5.76, test setleri için %3.35'tir. Bu bulgular, modelin etkili bir şekilde duyarlı sonuçlar ürettiğini, doğruluğunu ve güvenilir olduğunu göstermektedir. Çalışmada ayrıca çok katmanlı algılayıcı (MLP) yapısı da kullanılmış ve gerçek değerlerle modelin performansı karşılaştırılmıştır (Masaebi, 2016).

Aydın (2016) yazdığı bu çalışmada, İstanbul'da 2016 ile 2019 yılları arasında Türkiye'deki elektrik tüketimini etkileyen meteorolojik faktörleri incelemektedir. Türkiye Elektrik İletim A.Ş. ve meteorolojik istasyonlardan alınan veriler, regresyon analizi yöntemi kullanılarak incelenmiştir. Araştırmaya göre sıcaklık, nem, rüzgâr hızı ve güneşlenme süresi elektrik tüketimini etkiler. Bu sonuçlar, elektrik tüketimi tahminlerinin daha doğru hale gelmesine yardımcı olabilir ve meteorolojik değişkenlerin elektrik tüketimi tahminlerinde dikkate alınması gerektiğini vurgular (Aydın, 2016).

Karabulut (2017) yazdığı bu çalışmada, elektrik tüketimini tahmin etmek için çeşitli teknikler kullanmıştır. Destek vektör makineleri, çoklu doğrusal regresyon, karar ağaçları ve yapay sinir ağları bunlara örnektir. Çalışmada, bu yöntemlerin performansı karşılaştırıldı ve destek vektör makineleri ile en iyi sonuçlar elde edildi (Karabulut, 2017).

Berriel ve diğerleri (2017) yaptığı çalışmada, aylık enerji tüketimini tahmin etmek için derin öğrenme yöntemi kullanmıştır. Yaklaşım, yaklaşık bir milyon müşteriden 9 milyondan fazla örnekle doğrulanmıştır. Çalışma, tamamen bağlı (FC), konvolüsyonel sinir ağları (CNN) ve Uzun Kısa Dönem Bellek (LSTM) ağlarını inceledi. Sonuçlar, LSTM tabanlı modelin hem mutlak hem de göreceli hatalarda en iyi performans gösterdiğini göstermiştir (Berriel, Lopes, Rodrigues, Varejao, ve Oliveira-Santos, 2017).

Li ve diğerleri (2017) yaptığı çalışmada, bina enerji tüketiminin doğru bir şekilde tahmin edilmesini sağlamak amacıyla yığınlanmış otokodlayıcılar (Stacked Autoencoders) ve aşırı öğrenme makinesinin (Extreme Learning Machine) birleştirildiği yeni bir yöntemi sunmaktadır. Önerilen aşırı derin öğrenme yaklaşımı, geleneksel yöntemlere kıyasla daha yüksek tahmin doğruluğu sağlamaktadır. Yığınlanmış otokodlayıcılar, verilerden özelliklerin çıkarılmasında etkili olurken, aşırı öğrenme yaklaşımı hızlı öğrenme ve tahmin yeteneği sunduğu belirtilmiştir (Li, Ding, Zhao, Yi, ve Zhang, 2017).

Amasyali ve El-Gohary (2017) yaptığı çalışmada, derin öğrenmenin bina enerji tüketimi tahmini için nasıl kullanılabileceği ve veri odaklı modellerin enerji verimliliği karar süreçlerindeki önemi vurgulanmaktadır. Araştırmacılar, Deep Neural Network (DNN), Support Vector Machine (SVM), Random Forest (RF) ve doğrusal regresyon gibi dört veri odaklı algoritmayı uygulamış ve test etmişlerdir. Tahminlerin doğruluğu ve hesaplama verimliliği açısından incelenmiştir. Ayrıca, dış hava koşullarına dayalı soğutma enerjisi tüketiminin tahmininde derin öğrenmenin potansiyel faydaları üzerinde durulmuştur (Amasyali ve El-Gohary, 2017).

Koçak (2018) yaptığı çalışmada, Türkiye'deki tüm illerin elektrik kullanımı, veri zarflama analizi kullanılarak değerlendirmiştir. Yapılan analizler, illerin etkinlik skorlarını hesaplayarak, en verimli olan illeri belirlemiştir. Ek olarak, çok ölçütlü karar verme teknikleri kullanılarak illerin elektrik tüketimi, tüketici sayıları ve etkinlikleri için gelecekteki tahminler yapılmıştır. Çalışmada ayrıca, olası etkinsizlikler için proaktif bir yaklaşım benimsenerek eylem planları yapılabileceği belirtildi (Koçak, 2018).

Sezgin (2018) yaptığı bu çalışmada, Türkiye'de elektrik tüketimini hem artıran hem de azaltan unsurları incelemektedir. Çalışmada kişi başına gelir, sanayileşme düzeyi, finansal gelişme, eğitim düzeyi, fiyat, ticaretin açıklığı, karbon emisyonu ve kentleşme kullanılmıştır. Ekonometrik analizler, gelir ve karbon emisyonlarının elektrik tüketimini artırdığı, ancak kentleşme ve fiyatlar üzerinde negatif bir etkisi olduğunu göstermiştir. Çalışma, gelir ve karbon emisyonlarının elektrik tüketimini artırdığını, ancak kentleşme ve fiyatlara zarar verdiğini göstermektedir (Sezgin, 2018).

Tokgöz (2018) yaptığı bu çalışmada, tekrarlayan sinir ağı (RNN) tabanlı bir model geliştirdi ve Türkiye'deki 81 ilin saatlik elektrik tüketim verilerini kullanarak, kısa

dönemli elektrik tüketimini tahmin etti. Model, her bir ilin elektrik tüketimine olan katkısının nüfusuyla doğru orantılı olduğu varsayılarak ağırlık atamasını yaptı. Yapılan araştırmalar, RNN tabanlı modelin diğer yöntemlerden daha iyi sonuçlar verdiğini göstermiştir. Bu çalışma, RNN tabanlı modellerin elektrik tüketimini tahmin etmek için yararlı olabileceğini göstermektedir (Tokgöz, 2018).

Aydınlı (2018) yaptığı çalışmada yapay sinir ağları, soğutma sistemlerinin elektrik tüketimini tahmin etmek için kullanılabilir. Bu tahminler, soğuk hava depolarının enerji verimliliğini artırarak operasyonel maliyetleri düşürebilir. Ayrıca, çağdaş soğuk hava depolarının elektrik tüketim koşullarının önemi vurgulanarak, bu sektörün ekonomik bir girişim olduğu gösterilmiştir. Bununla birlikte, ısıtma sınıflandırma tahmin modelinde sadece %92,126'lık bir başarı oranı kaydedildiğinden, soğutma sistemleri için yapılan tahmin modelinin başarı oranı hakkında bir sonuç elde edilemedi (Aydınlı, 2018).

Özlü (2018) yaptığı çalışmada yapılan deneysel çalışmalar, yapay arı kolonisi algoritması kullanılarak yapılan tahminleme yönteminin mevcut yönteme göre daha iyi sonuçlar verdiğini göstermiştir. Başarı oranları, ilk deneysel çalışmada mesken, ticarethane ve sanayi aboneleri için sırasıyla %90,4, %81,3 ve %96,7 olarak belirlenmiştir. İkinci deneysel çalışmada, mesken, ticarethane ve sanayi aboneleri için başarı oranları sırasıyla %93,4, %86,3 ve %97,2 idi (Özlü, 2018).

Öksüz (2018) yapay sinir ağları kullanan çalışmada, elektrik tüketimini etkileyen faktörleri araştırdı. Çalışma, hava sıcaklığı, nem oranı, rüzgâr hızı, güneş ışınımı, günün saati, haftanın günü ve tatil günü olup olmadığı gibi değişkenlerin elektrik tüketimini tahmin etme üzerindeki etkilerini inceledi. Gerçek değerlerle yapılan tahminler karşılaştırıldığında, tahmin edilen ve gerçekleşen değerlerin tipik olarak birbirlerinden farklı olmadığını görebilirsiniz. MAPE değeri %10,28 olan tahmin, iyi veya doğru tahmin sınıfına girmektedir (Öksüz, 2019).

Yoo ve Alvarez (2018) yaptığı çalışmada, Seoul'deki konut elektrik tüketimini tahmin etmek için sinir ağları kullanarak bir vaka çalışmasını ele almaktadır. Seoul'un özel karakteristiklerinin konut elektrik tüketimini nasıl etkilediği ve son yıllarda meydana gelen tüketim trend değişikliğini açıklamaya çalışmaktadır. Çalışma, gerçek verilere dayalı bir sinir ağı tahmin modeli sunmakta ve farklı veri setleri kullanıldığında modelin

performansını analiz etmektedir. Ayrıca, tahmin üzerinde etkili olan farklı değişkenleri incelemekte ve hangilerinin en yüksek etkiye sahip olduğunu belirlemektedir. Sonuç olarak, çalışma Seoul'deki konut elektrik tüketimini tahmin etmek için sinir ağlarının kullanılabilirliğini ve etkinliğini göstermektedir (Yoo ve Hernández-Álvarez, 2018).

Kim ve Cho (2019) tarafından yazılan bu makalede, derin öğrenme yoluyla elektrik tüketimini tahmin etmek için otomatik kodlayıcıya dayalı bir model önerilir. Geleneksel modellere göre önerilen model, çeşitli durumlarda enerji talebini tahmin etmede daha iyi olduğu ve t-SNE algoritması kullanılarak durumu görselleştirip sonuçları daha açıklayıcı hale getirdiği belirtilmiştir. (Kim ve Cho, 2019).

İnan (2019) yazdığı bu çalışmada, Türkiye'deki ekonomik büyüme ile uluslararası ticaret ve elektrik tüketimi arasındaki ilişkiyi araştırmıştır. Analizler ekonometrik olarak yapılmıştır. Çalışma, 1960'tan 2017'ye kadar olan kişi başına elektrik tüketimi, ithalat ve ihracatın milli gelirdeki payları ve kişi başına gayri safi yurtiçi hasıla verilerini kullanmıştır. Eş bütünleşme testi, değişkenler arasında uzun süreli bir ilişki olduğunu göstermiştir. Elektrik tüketiminin modeldeki bağımlı değişken olduğu uzun dönem katsayıları pozitif ve anlamlı çıkmıştır. Sonuç olarak, Türkiye'deki ekonomik büyüme ile uluslararası ticaret arasında pozitif bir ilişki olduğu ve bu ilişkiyi elektrik tüketiminin de etkilediği ortaya çıkmıştır (İnan, 2019).

Bayar (2019) yazdığı bu çalışmada, elektrik tüketimini tahmin etmek için bir dizi yapay zekâ yöntemi kullanmıştır. Diferansiyel polinom sinir ağı, bulanık mantık, karınca kolonisi optimizasyon modelleri ve hibrit modeller bu yöntemlerden bazılarıdır. Bu yöntemlerin etkinliği karşılaştırıldı ve sonuçlar, klasik tahmin yöntemlerine göre yapay zekâ araçlarının daha doğru tahminler verdiğini gösterdi (Bayar, 2019).

Berksun (2019) yaptığı çalışmada karma Frekanslı Vektör Otoregresyon modeli kullanılarak, Türkiye'deki aylık elektrik tüketimi ile çeyreklik GSYH ve GSYH'nın alt bileşenleri arasındaki ilişki incelenmiştir. Sonuçlar, ekonomik büyüme ile elektrik tüketimi arasındaki ilişkinin farklı olduğunu göstermektedir. Ek olarak, elektrik tüketiminin artmasının tarım sektöründe GSYH üretimini düşürmesine rağmen, diğer sektörlerde GSYH üretimini artırmasına neden olduğu belirtilmektedir. Bu çalışmanın

bulguları, Türkiye'nin sürdürülebilir ekonomik büyümesini desteklemek için bir dizi politika önerisi sunmaktadır (Berksun, 2019).

Ateş (2019) yaptığı çalışmada genetik algoritmalar, bulanık mantık ve yapay sinir ağları dahil olmak üzere çeşitli yapay zeka yaklaşımları kullanmıştır. Yapılan karşılaştırmalı analizler, yapay sinir ağları yönteminin en başarılı olduğunu göstermiştir. Kullanılan veri seti, 2016 yılına ait Türkiye Elektrik İletim A.Ş. (TEİAŞ) tarafından sağlanan saatlik elektrik tüketim verilerini içerir. (Ateş, 2019)

Halaç (2019) yaptığı çalışmada Türkiye'nin kısa dönem elektrik tüketimini tahmin etmek için yapay sinir ağlarını kullanıyor. Çalışma, Facebook'un Prophet kütüphanesi, Twitter'ın anormallik tespiti kütüphanesi ve Long/Short Term Memory (LSTM) yapay sinir ağı uygulamalarıyla geliştirilen tahminleme modellerini içerir. Bu modeller, özellikler üzerinde yapılan dönüşümler yoluyla doğruluklarını geliştirmiştir. Yapılan araştırmalar, yapay sinir ağlarının elektrik tüketimi tahmininde faydalı olabileceğini göstermektedir. Bu teknoloji, elektrik üreticilerinin talebi daha iyi tahmin etmelerine, üretimi daha iyi planlamalarına ve elektrik maliyetlerini daha iyi yönetmelerine olanak tanır. Bu teknolojinin diğer sektörlerde de kullanılması mümkündür. (Halaç, 2019)

Wei ve diğerleri (2019) yaptığı çalışmada, enerji tüketimi tahminleri için geleneksel modeller ile yapay zeka tabanlı modeller arasındaki temel farklar incelenmektedir. Geleneksel modeller genellikle istatistiksel yöntemlere dayanırken, yapay zeka tabanlı modeller daha karmaşık veri analitiği tekniklerini kullanır. Tahmin süresi ve doğruluğu açısından, yapay zeka modelleri genellikle daha uzun vadeli tahminler yapabilir ve daha yüksek doğruluk sağlayabileceği belirtilmiştir (Wei, Li, Pengc, Zeng, ve Lu, 2019).

Mosavi ve Bahmani (2019) yaptığı çalışmada, enerji tüketimi tahmininde makine öğrenimi modellerinin kullanımının önemini vurgulamaktadır. Geleneksel zaman serisi tahmin araçlarına kıyasla, makine öğrenimi modelleri doğruluk, sağlamlık ve genelleme yeteneği açısından büyük bir gelişme sağladığına dikkat çekmiş. Araştırma, farklı makine öğrenimi tekniklerinin, enerji türlerinin, tahmin tiplerinin ve uygulama alanlarının sınıflandırılmasını içeren yeni bir arama ve sınıflandırma yöntemi sunduğunu belirtmektedir (Mosavi ve Bahman, 2019).

Choi ve Lee (2020) yaptığı çalışmada, bir binanın elektrik tüketimini tahmin etmek için derin öğrenme yöntemlerini karşılaştıran bir çalışma gerçekleştirilmiştir. Sonuçlar, LSTM/GRU modellerinin genel olarak iyi performans gösterdiğini göstermektedir. Bununla birlikte, varyasyonlu sequence-to-sequence modellerinin performansı düşüktür. Sonuç olarak, gelecekteki çalışmalar, sırasıyla modellerinin performansını artırmak ve mevsimsel elektrik tüketimi özelliklerini kullanarak daha doğru tahminler yapmak amacıyla tasarlanmıştır. Kore Elektrik Güç Şirketi de bu çalışmaya katkıda bulundu (Choi ve Lee, 2020).

Özkan, Güler ve Aladağ'ın (2020) yazdığı bu çalışmasında, Fourier Analizi, Winters Yöntemi ve En Küçük Kareler Yöntemi ile yapılan tahminlemelerin başarıları incelenmiştir. Bu modeller, değerli ve kıt kaynak olan elektrik enerjisinin planlı tüketim ve stratejik kararların temelini oluşturmak için kullanılmıştır. İki ayrı tahmin algoritmasının, veri seti özellikleri üzerindeki uygunluğu değerlendirilmiş ve benzer özelliklere sahip veri setleri için önerilebilirlikleri araştırılmıştır. Fourier Analizi, Winters Yöntemi ve En Küçük Kareler Yöntemi, elektrik tüketim verileri için uygun tahmin yöntemleri olduğunu göstermiştir. Bu çalışma literatürde yeni bir alt inceleme konusu oluşturmaktadır; veri seti dönemlere ayrılmıştır ve değişkenlik katsayısının değişimi ile model başarı ölçütleri arasındaki ilişki incelenmiştir.

Alanbar ve diğerleri (2020) yaptığı çalışmada, derin öğrenme tekniklerinin kullanıldığı enerji tüketimi tahmini üzerine odaklanmaktadır. Enerji tüketimi tahmini için kullanılan derin öğrenme teknikleri Deep Neural Network (DNN), Deep Belief Network (DBN) ve Recurrent Neural Network (RNN) 'dir. Enerji tüketimi tahmini için en iyi sonucu Long Short-Term Memory (LSTM) verdiği belirtilmiştir (Alanbar, Alfarraj, ve Alghieth, 2020).

Altunkaya (2021) yaptığı çalışmada, Türkiye ve diğer ülkelerdeki tüketim tahmini araştırmaları incelenmiş ve kısa dönem tüketim tahmini çalışmalarının özetleri sunulmuştur. Ayrıca çalışmada, takvim etkisinin ve benzer günlerde tüketimin model üzerindeki etkisini inceledi. Özel günlerde, geçmiş 24 saatlik tüketim verisini iki boyutlu bir görüntü verisine dönüştürerek model başarımını %0.72 oranında artırdı (Altunkaya, 2021).

Nebati, Taş ve Ertaş (2021) yazdığı bu makalede Türkiye'deki elektrik tüketimi verileri kullanılarak gelecekteki enerji talebi tahmin edilmiştir. Regresyon analizi ve zaman serisi teknikleri kullanılarak yapılan çalışmada öncelikle enerji alanındaki araştırmalar incelenmiştir. Ardından Türkiye'deki durum, araştırma metodolojisi, ve uygulama bölümleri incelenmiştir. Regresyon analizi için MAPE değeri %1.30 iken, zaman serisi için %11.63 olarak bulunmuştur. Bu bulgular, regresyon analizinin tahminlerde daha iyi olabileceğini göstermektedir. Elde edilen sonuçlar, TÜİK tarafından belirlenen mevcut değerlerle uyumludur.

Yıldırım (2021) yaptığı bu araştırmada, Yapay Sinir Ağı (YSA) algoritmalarının öğrenim gücünü artırarak başarılı tahminler yapabildiğini, özellikle 3 numaralı YSA modelinin diğer modellere kıyasla en küçük RMSE değerine sahip olduğunu ortaya koymuştur. Bu bulgular, literatürde daha önce söylenenleri doğrulamaktadır ve çalışmanın aylık (dönemsel) tahminler yapmak için farklı bir yaklaşım kullandığını göstermektedir. (Yıldırım, 2021)

Shapi ve diğerleri (2021) yaptığı çalışmada, makine öğrenimi tekniklerini kullanarak elektrik tüketimini doğru bir şekilde tahmin etmeyi amaçlamaktadır. Çeşitli makine öğrenimi modelleri, özellikle K En Yakın Komşu (KNN) modeli, elektrik tüketimini doğru bir şekilde tahmin etmede diğer modellere göre daha başarılı bulunmuştur. Bu çalışma, enerji yönetimi ve maliyet optimizasyonu için bilgiler sunmaktadır (Shapi, Ramli, ve Awalin, 2021).

Zhang ve diğerleri (2021) yaptığı çalışmada, Transformer modelini K-means kümeleme ile birleştirerek güç tüketimini tahmin etmeyi ve anormallikleri tespit etmeyi amaçlamışlardır. Sonuçlar, bu yöntemin genellikle kullanılan LSTM, zaman serisi modelini geride bıraktığını göstermektedir. Araştırma, Fransız evlerindeki gerçek zamanlı elektrik tüketimi verilerini kullanarak yapılmıştır (Zhang, Zhang, Ding, ve Zhang, 2021).

Yıldız (2022) yaptığı çalışmada yüksek mertebeden sinir ağı tasarımı için sembolik regresyon temelli öznitelik çıkarımı kullanır. Enerji tüketim verileri, literatür özeti, yapay sinir ağları, regresyon teknikleri ve parametre optimizasyon teknikleri, elektrik tüketimini tahmin etmek için kullanılan teknikleri içerir. Sembolik regresyon temelli öznitelik

çıkarımı yöntemi, sinir ağı modellerini daha anlamlı hale getirerek modelleme yeteneğini geliştirmiştir. Çalışmanın amacı, zaman serisi tahminlemede öznitelik çıkarımına yeni bir yaklaşım getirerek tahminleme performansını geliştirmek ve hata oranını azaltarak gerçek hayatta verilerin kullanılabilirliğini artırmak. (Yıldız, 2022)

Dincer (2022) tarafından yapılan çalışmalarda, yapay zekâ kullanarak aylık tüketim tahmin modeli oluşturuldu. Kısa dönemli yük tahminleri için çok katmanlı algılayıcı yapısı kullanıldı ve elektrik tüketimini tahmin etmek için gerekli parametreler belirlendi. Son olarak bu parametrelerin performansları ölçüldü. Bu çalışmanın bulguları, yapay zekâ, zaman serileri analizi, regresyon analizi ve yapay sinir ağı gibi çeşitli tekniklerin elektrik tüketim tahminlerinde performanslarının ölçülebileceğini göstermektedir. Bu yöntemler, elektrik tüketim tahminlerinin daha doğru hale gelmesine ve enerji sektöründe daha verimli bir planlama yapılmasına yardımcı olabilir. (Dincer, 2022)

Hajjar (2022) tarafından, çalışmada kullanılan SMOReg modeliyle, Türkiye'nin toplam elektrik tüketimini tahmin etmede %91,42, %2,98, ortalama mutlak yüzde hatası (mean absolute percentage error MAPE) ve ortalama karekök hatası (root mean square error RMSE) için başarı oranları sağladı. Bu çalışma, Türkiye'nin toplam elektrik tüketimini tahmin etmek için makine öğrenimi algoritmalarının kullanılabileceğini gösteriyor. Diğer modellere kıyasla SMOReg modeli daha doğru sonuçlar vermiştir. (Hajjar, 2022)

Marmara (2022) yazdığı bu çalışmada, zaman serileri analizi, yapay sinir ağları ve uzman sistemler gibi çeşitli teknikleri kullanarak, enerji sektöründe elektrik tüketimi tahmini ve maliyet iyileştirmesi konularını ele almıştır. Çalışmada, modelin elektrik tüketimini doğru tahmin ettiğini ve maliyetleri azalttığını göstermiştir. Yapılan analizler, modelin diğer yaklaşımlara göre daha düşük hata oranlarına sahip olduğunu ve farklı saatlerdeki tüketim verilerinin kullanılmasının maliyetleri azalttığını göstermiştir. Bu çalışma, enerji sektöründe gelecekte yapılacak çalışmalara ışık tutacaktır. (Marmara, 2020)

Özcan (2022) tarafından yazılan bu çalışmada, Türkiye'nin elektrik tüketimi verileri kullanılarak, farklı derin öğrenme modelleri test edilmiştir. Elde edilen sonuçlar, doğru hiper parametrelerin kullanılmasıyla kısa vadeli yük tahminlerinde düşük hata tahminleri elde edilebileceğini göstermiştir. Derin öğrenme modellerinin elektrik tüketim tahmini konusunda etkili olduğu sonucuna varılmıştır. Özellikle, bulanık mantık ve ANFIS

modelleri, Türkiye'nin kısa vadeli yük tahmin problemi için alternatif bir yol olarak önerilmiştir. Ayrıca, çalışmada ele alınan konuların yanı sıra, elektrik enerjisi tüketim tahmininin önemi ve Türkiye'nin elektrik tüketimi gibi konular da ele alınmıştır. (Özcan, 2022)

Liu ve Li (2022) yaptığı çalışmada, Birleşik Krallık'ta yıllık elektrik ve enerji tüketimi tahminlerini gerçekleştirmek için geri yayılımlı sinir ağı (BP-NN), çoklu doğrusal regresyon (MLR) ve en küçük kare destek vektör makinesi (LS-SVM) algoritmalarını kullanmaktadır. LS-SVM modelinin eğitim setinde en iyi tahmin performansını sergilediği ve diğer modellere göre daha başarılı olduğu belirtilmektedir (Liu ve Li, 2022).

Movva (2022), okullar için elektrik enerjisi tüketiminin tahmin edilmesine odaklanan bir tez çalışmasının detayları sunulmaktadır. Elektrik talebinin artmasıyla birlikte, doğru tahminlerin önemi vurgulanmaktadır. Farklı zaman serisi tahmin modelleri incelenmekte ve bu modellerin amacı üzerinde durulmaktadır. Tez çalışmasının amacı, elektrik enerjisi tüketimini kısa vadeli olarak tahmin etmek için modeller geliştirmektir (Movva, 2022).

Meng ve diğerleri (2022) yaptığı çalışmada, Çin'deki enerji tüketimini destek vektör regresyon modeli ve Markov Zinciri analizi kullanarak tahmin etmektedir. Sonuçlar, Çin'in enerji tüketiminin artmaya devam edeceğini, ancak gelecek yıllarda daha rasyonel bir yapıya sahip olacağını öngörmektedir. Özellikle, kömür tüketiminin azalacağı ve temiz enerji tüketiminin artacağının tahmin edildiğini belirtilmektedir (Meng, Sun, ve Wang, 2022).

Shaikh ve diğerleri (2022) yaptığı çalışmada, enerji tüketimi tahmininde yapay sinir ağı tabanlı yöntemlerin geliştirilmesi ve yorumlanabilir zaman serileri için nöral taban genişleme analizinin kullanımı hakkında bilgi vermektedir. Çalışmada, enerji tüketimi tahmininde N-BEATS yönteminin nasıl geliştirildiği ve büyük veri setleriyle nasıl kullanıldığı detaylı bir şekilde incelenmektedir (Shaikh, Nazir, Khan, ve Shah, 2022).

Saoud ve diğerleri (2022) yaptığı çalışmada, ev enerji tüketiminin tahmin edilmesi amacıyla Stationary Wavelet Transform (SWT) ve transformer(dönüştürücü) bir araya getirilmesi üzerine bir yöntem önerilmiştir. SWT, sinyali farklı frekans bileşenlerine

ayrıştırarak analiz etmeyi sağlayan bir sinyal işleme tekniğidir. Dönüştürücüler ise derin öğrenme modellerinde kullanılan yapılardır ve özellikle dikkat mekanizmalarını içerirler. Bu çalışmada, ev enerji tüketimi verilerinin daha doğru analiz edilmesi ve tahmin edilmesi için SWT'nin sinyal ayrıştırma yetenekleri ile dönüştürücülerin dikkat mekanizmalarının birleştirilmesi önerilmiştir (Saoud, Marzouqi, ve Hussein, 2022).

Mhlanga (2023) yaptığı çalışmada, yapay zeka ve makine öğrenimi teknolojilerinin gelişmekte olan pazarlardaki enerji üretimi ve tüketimini nasıl iyileştirebileceği incelenmektedir. Şehir planlamacılarının binalardaki enerji tüketimini tahmin etmek için makine öğrenimi modellerini eğitmek için fiziksel özellikler gibi kentsel form karakteristiklerinin enerji yoğunlukları üzerinde etkili olduğu vurgulanmaktadır (Mhlanga, 2023).

Reddy ve diğerleri (2023) yaptığı çalışmada, makine öğrenimi tekniklerini kullanarak elektrik tüketimini doğru bir şekilde tahmin etmeyi amaçlamaktadır. Çeşitli makine öğrenimi modelleri, özellikle K En Yakın Komşu (KNN) modeli, elektrik tüketimini doğru bir şekilde tahmin etmede diğer modellere göre daha başarılı bulunmuştur. Bu çalışma, enerji yönetimi ve maliyet optimizasyonu için bilgiler sunmaktadır (Reddy, ve diğerleri, 2023).

3. YÖNTEM

Bu çalışma, elektrik tüketim verilerini analiz etmek ve tahmin etmek için makine öğrenimi kullanmanın çeşitli yollarını incelemektedir. Zaman serisi özellikleri taşıyan elektrik tüketimi verilerinin analizi, enerji yönetimi ve planlaması için çok önemlidir. Çalışmamızda hem geleneksel hem de çağdaş makine öğrenimi yöntemlerinin kullanılması ile elektrik tüketim verilerinin analizi ve tahmini yapılmıştır.

3.1. Araştırma Modelleri

Çeşitli faktörler, elektrik tüketim verilerini zaman içinde değiştirir. Bu çalışmanın amacı, zaman serisi verilerinin analizi için yaygın olarak kullanılan Python programlama dilindeki kütüphaneleri kullanarak elektrik tüketimini tahmin etmektir.

Çalışmamızda topladığımız elektrik tüketim bilgileri aşağıdaki adımları içeriyordu:

1. Veri Ön İşleme: İlk adımda, veri seti pandas kullanılarak yüklenmiş ve tarih sırasına göre düzenlenmiştir.

2. Özellik Mühendisliği: Tarihle ilgili özellikler (yıl, ay, gün, saat vb.), resmi tatiller, dini bayramlar ve COVID-19 gibi özel durumlar işlenmiştir.

3. Zaman Serisi Özellikleri: Geçmiş tüketim verilerine dayalı gecikme (lag) özellikleri oluşturulmuştur.

4. Model Eğitimi ve Testi: LightGBM, XGBoost, CatBoost, KNN ve LSTM modelleri eğitilmiş ve test edilmiştir.

5. Hiper Parametre Optimizasyonu: Optuna kütüphanesi ile bu modeller için en uygun parametreler belirlenmiştir.

6. Sonuçların Değerlendirilmesi: Her modelin tahmin başarısı MAPE metriği kullanılarak ölçülmüş ve sonuçlar karşılaştırılmıştır.

3.2. Eğitim Modelleri

Yapılan bu çalışmada Verilerin toplanması ve anlamlandırılması için kullanılan programlar aynı zamanda çalışmanın Evren ve Örneklemini oluşturmaktadır. Bu programlar aşağıdaki gibidir:

3.2.1. KNN

k-En Yakın Komşular (k-NN) algoritması, makine öğrenimi ve istatistik için önemli bir araçtır. Bu çok yönlü yaklaşım, hem sınıflandırma hem de regresyon görevlerinde kullanılabilir. Bu da onu denetimli öğrenmede değerli bir araç haline getirir.

k-NN'nin temel amacı, bir veri kümesinin en yakın k veri noktası tarafından verilen bilgilere dayanarak tahminler yapmaktır. Algoritmanın davranışını belirlemede K seçimi çok önemlidir. Sınıflandırmada k-NN, k en yakın komşusu arasındaki çoğunluk oyuna dayanarak bir nesneyi belirli bir sınıfa atar. Bu, nesnenin en yakın komşuları arasında en yaygın olan sınıfa girdiğini gösterir. Özellikle, k değeri 1'e eşitse, nesne en yakın tek komşusunun sınıfına doğrudan atanır. Bu, işlemi basitleştirse de aşırı uyuma neden olabilir.

k-NN algoritması, parametrik olmayan ve tembel öğrenme doğasına sahiptir. Altta yatan fonksiyona yerel olarak yaklaşırken, tüm hesaplamaları fonksiyon değerlendirme noktasına kadar erteler. Özellikler, farklı fiziksel birimleri temsil ettiğinde veya önemli ölçek farklılıkları sergilediğinde, sınıflandırma veya regresyon için mesafeye olan bu güven, eğitim verilerini normalleştirmenin veya standartlaştırmanın önemini vurgulamaktadır. Bu tür bir ön işleme kullanıldığında, k-NN algoritmasının doğruluğu önemli ölçüde artırılabilir.

Komşu veri noktalarının katkılarına ağırlık atama seçeneği, k-NN için bir diğer önemli yöntemdir. Bu nedenle, daha yakın komşular tahmin üzerinde daha büyük bir etkiye sahiptir. Bir ağırlıklandırma şeması tipik olarak, her komşuya 1'in o komşuya olan uzaklığa bölünmesiyle (1/d) bir ağırlık atar. Bu yöntem, her bir komşunun nihai tahmin üzerindeki etkisini hassas hale getirir.

k-NN algoritmasının bir özelliği, verilerin yerel yapısına olan duyarlılığıdır. Bu hassasiyet nedeniyle k-NN, veri dağılımının iyi tanımlanmamış olabileceği veya karar sınırlarının karmaşık ve doğrusal olmadığı durumlarda özellikle iyi çalışır.

Her bir eğitim örneğinin bir sınıf etiketi ile ilişkili bir vektör olarak gösterildiği çok boyutlu bir özellik uzayında k-En Yakın Komşular (k-NN) algoritması çalışır. Algoritmanın eğitim aşaması, özellik vektörlerinin sınıf etiketleriyle birlikte depolanmasını içerir. Bu adım tamamen veri depolamadır.

K parametresi, kullanıcı tarafından tanımlanan bir sabittir ve sınıflandırma aşamasında çok önemlidir. k-NN algoritması, sorgu veya test noktası olarak adlandırılan yeni, etiketsiz bir vektör sunulduğunda, sorgu noktasına en yakın k eğitim örneği arasında en sık görülen etiketi atayarak sınıflandırır. Bir tahmin yapmak için k en yakın komşuların çoğunluk sınıfına bakar.

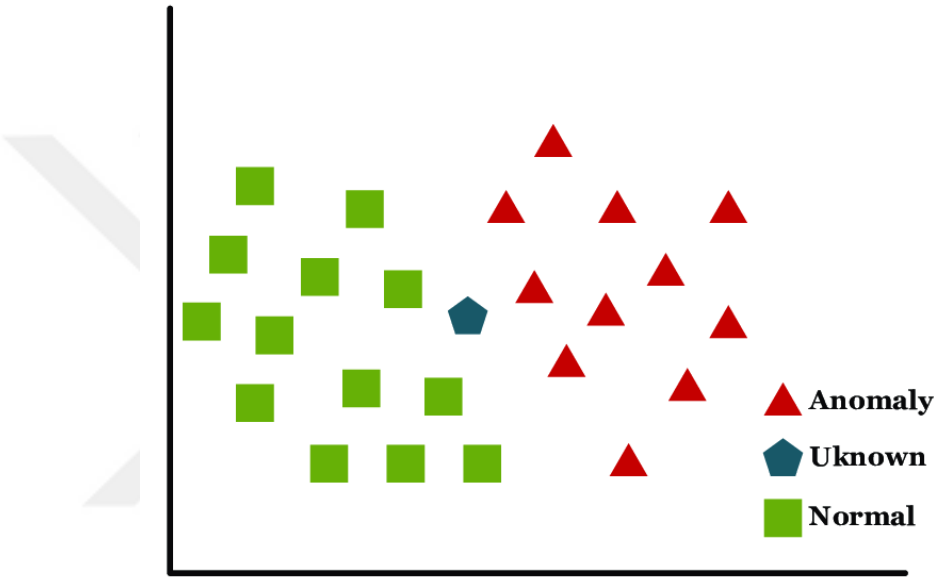
k-NN'de mesafe metriğinin seçimi, özellikle veri noktaları arasındaki yakınlığı bulmak için çok önemlidir. Sürekli değişkenler için Öklid mesafe metriği yaygın olarak kullanılır. Bununla birlikte, metin sınıflandırmasında olduğu gibi ayrık değişkenlerle uğraşırken, örtüşme metriği (veya Hamming mesafesi) gibi başka bir metrik daha uygun olabilir. Pearson ve Spearman gibi korelasyon katsayıları, gen ifadesi mikroarray verileri gibi özel alanlarda mesafe metrikleri olarak kullanılır.

Sınıf dağılımı dengesiz olduğunda, k-NN sınıflandırmasının temel "çoğunluk oylaması" yaklaşımının bir sınırlaması vardır. Daha yüksek frekanslı sınıflar, sayıları nedeniyle k'nın en yakın komşuları arasında temsil edilme olasılıkları daha yüksek olduğundan, bu tür durumlarda tahminlerde daha fazla yer alırlar. Bu sorunu çözmek için ağırlıklandırma şemaları kullanılabilir. Bu şemalar, test noktasının k en yakın komşusuna olan uzaklığını dikkate alır. Bu komşuların her birinin sınıfı (veya regresyon problemlerindeki değeri) test noktasına olan uzaklıkla ters orantılı bir ağırlıkla çarpılır. Bu yöntem, her bir komşunun nihai tahminin sorgu noktasına olan yakınlığına göre değiştirilmesini sağlar.

Veri temsili, çarpıklığı önlemek için soyutlanabilir. Örneğin, kendi kendini organize eden bir haritada (SOM), orijinal eğitim verilerindeki yoğunluklarına bakılmaksızın, her bir düğüm benzer veri noktalarından oluşan bir kümeyi gösterir. k-NN, SOM'a

uygulandığında, sınıflandırma performansını artırabilir ve çarpık sınıf dağılımlarının etkisini azaltabilir.

Özetle, k-En Yakın Komşular, k gibi kullanıcı tarafından belirlenen parametrelere ve seçilen mesafe metriğine bağlı olan esnek bir algoritmadır. Sınıf dağılımına duyarlıdır ve mesafe metriklerini, ağırlıklandırma şemalarını dikkatli bir şekilde incelemek, çeşitli veri türlerinde ve alanlarda önemli ölçüde performansını artırabilir.



Şekil 2. KNN Sınıflandırma Örneği (Researchgate, 2023)

3.2.2. Boosting yöntemleri

Boosting yöntemleri, zayıf sınıflandırıcıları (genellikle doğru sınıflandırma yapma olasılığı yalnızca rastgele tahminlerden biraz daha iyi olan modeller) alıp güçlü sınıflandırıcılar haline getirmek için kullanılan bir dizi algoritmadan oluşur. Bu yaklaşımlar, bir dizi zayıf modeli adım adım eğiterek ve her bir sonraki modeli öncekilerin hatalarını düzeltecek şekilde ayarlayarak çalışır. CatBoost, LightGBM ve XGBoost, en yaygın yükseltme yöntemleridir:

3.2.3. LightGBM

Lightgbm kütüphanesi: LightGBM, Microsoft tarafından geliştirilen, açık kaynaklı bir makine öğrenimi kütüphanesidir ve Gradient Boosting Decision Tree (GBDT)

algoritmalarına dayanır. Ana amacı, büyük veri setleri ve yüksek boyutlu veriler üzerinde hızlı ve verimli bir şekilde çalışmaktır. LightGBM'nin öne çıkan özellikleri ve bilimsel temelleri şu şekilde özetlenebilir.

GBDT Algoritması: LightGBM, zayıf tahminciler olan karar ağaçlarını iteratif olarak eğiterek güçlü bir model oluşturur. Her iterasyonda, model önceki ağaçların hatalarını düzeltmeye odaklanır.

Leaf-wise Büyüme: Geleneksel GBDT'lerin aksine, LightGBM yaprak odaklı büyüme yöntemi kullanır. Bu, özellikle büyük veri setlerinde, daha hızlı ve doğru öğrenme sağlamak için ağaçların derin bölgelerine odaklanır.

Düşük Bellek Kullanımı ve Hız: Veri depolama ve işleme için optimize edilmiş algoritmalar sayesinde, LightGBM büyük veri setlerini az bellek kullanarak etkin bir şekilde işleyebilir. Bu, özellikle büyük ölçekli uygulamalarda avantaj sağlar.

Histogram Tabanlı Özellik Bölünmeleri: LightGBM, özellik değerlerini sabit sayıda kova içine yerleştirerek bölünme noktalarını hızlı bir şekilde hesaplar. Bu yöntem, modelin eğitim süresini önemli ölçüde azaltır.

Kategorik Özellik Desteği: LightGBM, kategorik özellikleri doğrudan işleyebilir, bu da veri ön işleme sürecini kolaylaştırır.

Özellik Önemi ve Hiper parametre Ayarları: Model tarafından kullanılan özelliklerin önem derecesini belirleyebilir ve çeşitli hiper parametre ayarları ile modeli farklı senaryolara göre uyarlayabilir.

LightGBM, sınıflandırma, regresyon ve diğer makine öğrenimi görevlerinde etkili bir şekilde kullanılabilir. Büyük veri setleri ve rekabetçi veri bilimi yarışmalarında sıkça tercih edilen bir araçtır. Hem akademik araştırmalarda, hem de endüstriyel uygulamalarda yaygın olarak kullanılmaktadır.

GBDT'de, her yeni ağaç, önceki ağaçların hatalarını düzeltmeye çalışır. Model $F(x)$ denklem (1.1) de olduğu gibi ifade edilebilir:

$$F(x) = F_0(x) + \sum_{m=1}^{M} \gamma_m h_m(x) \quad (1.1)$$

- $F(x)$: Başlangıç modelidir (genellikle veri setinin çıktılarının ortalaması).
- M: Toplam ağaç sayısıdır.
- $h_m(x)$: m'inci ağacın tahminidir.
- γ_m : m'inci ağacın çıktısının öğrenme oranı ile çarpılmış ağırlığıdır.

Her bir ağaç, hata fonksiyonunu (kayıp fonksiyonunu) azaltacak şekilde denklem (1.2) de olduğu gibi eğitilir:

$$h_m = argmin_h \sum_{i=1}^{N} L\big(y_i, F_{m-1}(x_i) + h(x_i)\big) \quad (1.2)$$

- N: veri noktalarının sayısıdır.
- L : kayıp fonksiyonudur (örneğin, ortalama karesel hata).
- y_i : i'nci veri noktasının gerçek değeri.
- $F_{m-1}(x_i)$: i'nci veri noktası için önceki modelin tahmini.

LightGBM'nin Yenilikleri

LightGBM, GBDT algoritmasına birkaç önemli yenilik getirir:

Leaf-wise Büyüme: Geleneksel GBDT level-wise büyürken, LightGBM leaf-wise büyür. Bu, ağaç büyümesindeki hedefi değiştirir.

Histogram Tabanlı Özellik Bölünmeleri: LightGBM, özellik değerlerini sabit sayıda kova içine yerleştirerek hız kazanır. Bu, sürekli özellikler için bölünme noktalarının hızlı bir şekilde hesaplanmasını sağlar.

Özellik Önemi ve Kategorik Özellik Desteği: LightGBM, her özelliğin modele katkısını hesaplayabilir ve kategorik özellikleri doğrudan işleyebilir, bu da ek kodlama gereksinimini azaltır.

Bu formüller ve tanımlar, LightGBM'nin bilimsel temellerini ve nasıl çalıştığını açıklar. Ancak, pratikte, bu algoritmanın uygulanması, veri setine, kayıp fonksiyonuna ve belirli hiper parametrelere bağlı olarak değişir.

3.2.4. XGBoost

XGBoost (eXtreme Gradient Boosting), makine öğreniminde sıkça kullanılan, yüksek performanslı bir gradient boosting kütüphanesidir. Python dili için kullanılabilir olup, aynı zamanda R, Java ve diğer dillerde de mevcuttur. XGBoost, özellikle yarışma platformları ve endüstriyel uygulamalarda popülerdir. Çünkü hem sınıflandırma hem de regresyon görevlerinde etkili sonuçlar üretebilir. XGBoost'un temelinde Gradient Boosting algoritmaları yatar ve bu algoritma, zayıf tahmincileri (genellikle karar ağaçları) birleştirerek güçlü bir model oluşturur. Bu Modelin özellikleri aşağıdaki gibidir.

1. Gradient Boosting Framework: XGBoost, Gradient Boosting Decision Tree (GBDT) algoritmasının bir uygulamasıdır. Model, zayıf tahmincileri (karar ağaçları) bir araya getirerek, her bir yeni tahmincinin önceki tahmincilerin hatalarını düzeltmeye odaklanır.

2. Regularization: XGBoost, model karmaşıklığını kontrol altında tutmak için hem L1 (Lasso regresyonu) hem de L2 (Ridge regresyonu) düzenlileştirmelerini (regularization) içerir. Bu özellik, modelin aşırı öğrenmesini (overfitting) önlemeye yardımcı olur.

3. Scalability and Efficiency: XGBoost, hesaplama kaynaklarını verimli kullanır ve büyük veri setlerinde yüksek performans gösterir. Paralel ve dağıtık hesaplama desteği, büyük ölçekli veri işleme imkanı sunar.

4. Handling Missing Values: XGBoost, eksik değerleri otomatik olarak işleyebilir. Eksik değerlere sahip özellikler için en iyi bölünme yolu seçilirken, eksik değerlerin nasıl işleneceği konusunda otomatik bir yöntem sunar.

5. Tree Pruning: XGBoost, karar ağaçlarını derinlikten bağımsız olarak budama (pruning) yapar. Bu, ağacın gereksiz dallarını kırparak modelin daha genel ve basit kalmasını sağlar.

6. Built-in Cross-validation: XGBoost, modelin performansını değerlendirmek için yerleşik çapraz doğrulama (cross-validation) özelliğine sahiptir. Bu, optimal hiper parametre ayarlarını bulmada yardımcı olur.

7. Customizable Objective and Evaluation Functions: XGBoost, özelleştirilebilir hedef ve değerlendirme fonksiyonlarına izin verir. Bu, farklı türdeki tahmin görevlerine uyum sağlamasını sağlar.

$$\widehat{y_i} = \phi(x_i) = \sum_{k=1}^{K} f_k(x_i) \ , f_k \epsilon \mathcal{F} \qquad (1.3)$$

- Denklem (1.3) de yi i'nci örnek için tahmin edilen değerdir.
- ϕ(xi) modelin tahmin fonksiyonudur.
- K toplam ağaç sayısıdır.
- fk k'ıncı ağaçtır.
- F tüm olası ağaçların uzayıdır.

3.2.5. CatBoost

CatBoost (Categorical Boosting), Yandex tarafından geliştirilen ve açık kaynak kodlu bir makine öğrenimi kütüphanesidir. Gradient boosting üzerine kurulmuş olan CatBoost, özellikle kategorik verilerin işlenmesindeki yetkinliğiyle tanınır. Python, R, Java gibi çeşitli dillerde kullanılabilir ve hem sınıflandırma hem de regresyon problemlerinde etkili sonuçlar sunar. CatBoost, karmaşık kategorik veri setlerini doğrudan işleyebilme

yeteneğiyle diğer gradient boosting kütüphanelerinden ayrılır. Bu kütüphanenin özellikleri:

1. Symmetric Trees: CatBoost, simetrik ağaçlar kullanır. Bu, modelin daha hızlı ve daha verimli bir şekilde eğitilmesini sağlar.

2. Ordered Boosting: CatBoost'un en önemli yeniliklerinden biri, overfitting'e karşı mücadelede etkili olan "ordered boosting" yöntemidir. Bu yöntemde, modelin eğitimi için kullanılan veri, rastgele alt kümeler halinde bölünür ve her alt küme, modelin geri kalan kısmını eğitmek için kullanılır.

3. Kategorik Değişken İşleme: CatBoost, kategorik özellikleri doğrudan işleyebilir. Özellikle, yüksek kardinaliteli kategorik veriler için etkili bir işleme mekanizması sunar.

4. Robustness to Different Distributions: CatBoost, farklı veri dağılımlarına karşı robust bir yapı sunar. Bu, modelin çeşitli veri setleri üzerinde tutarlı ve güvenilir sonuçlar vermesini sağlar.

5. Model Interpretability: CatBoost, model yorumlanabilirliği açısından çeşitli araçlar sunar. Özellik önemini değerlendirmek ve modelin karar yapısını anlamak için kullanışlı araçlara sahiptir.

6. Efficiency and Scalability: CatBoost, büyük veri setleri ve yüksek boyutlu veriler üzerinde etkili bir şekilde çalışabilir. Hem CPU hem de GPU desteği sunarak yüksek verimlilik ve ölçeklenebilirlik sağlar.

CatBoost'un temel prensibi, bir dizi zayıf tahminciyi (genellikle karar ağaçları) bir araya getirerek güçlü bir tahmin modeli oluşturmaktır. Bu model, denklem (1.4) de olduğu gibi ifade edilebilir:

$$\widehat{y_i} = \sum_{k=1}^{K} f_k(x_i) \ , f_k \epsilon \mathcal{F} \tag{1.4}$$

- Burada $\hat{y}_i$ i'nci örnek için tahmin edilen değerdir.
- K toplam ağaç sayısıdır.
- f_k k'ıncı ağaçtır.
- $\mathcal{F}$ tüm olası ağaçların uzayıdır.

3.2.6. LSTM

LSTM (Long Short-Term Memory), yapay sinir ağlarının özel bir türüdür. Özellikle zaman serisi verileri gibi sıralı verilerin işlenmesi için tasarlanmıştır. Yapay sinir ağlarının temelinde yatan geri yayılım algoritmasıyla eğitilirler ve veriler arasındaki karmaşık ilişkileri öğrenmek amacıyla kullanılırlar. Zaman serisi verileri gibi verilerde, önceki girdilerin sonraki çıktılara etkisi önemlidir. LSTM, uzun süreli bağımlılıkları daha etkili bir şekilde öğrenmek için özel olarak tasarlanmıştır. Bu sayede geçmiş verilerin, gelecekteki tahminlere etkisini daha iyi yakalayabilirler.

LSTM'in temel özellikleri hücre durumu ve üç farklı kapıdan oluşur: giriş kapısı, unutma kapısı ve çıkış kapısı. Hücre durumu, veri dizisi boyunca bilgiyi taşıyan yapıdır ve sadece kapılar aracılığıyla küçük değişikliklere uğrar. Giriş kapısı, hangi bilginin hücre durumuna eklenmesi gerektiğine karar verirken, unutma kapısı hücre durumundan hangi bilginin atılması gerektiğini belirler ve çıkış kapısı, hangi bilginin hücre durumundan alınarak çıktı olarak kullanılacağını kontrol eder.

LSTM, dil modelleme, metin üretimi, zaman serisi tahmini, ses tanıma gibi birçok uygulama alanında başarılı bir şekilde kullanılırken, uzun süreli bağımlılıkları etkili bir şekilde öğrenme ve karmaşık sıralı verileri işleme yeteneği ile dikkat çeker. Ancak eğitim süresinin diğer modellere göre daha uzun olabileceği ve hiper parametre ayarlaması ile model yapılandırmasının karmaşık olabileceği dezavantajları da vardır.

Her yöntem, çeşitli durumlarda ve veri tiplerinde faydalıdır, bu nedenle en uygun yöntemin seçimi, kullanılan veri setinin özelliklerine ve sorunun gereksinimlerine bağlıdır.

3.2.7. Grid Search

Makine öğrenimi modellerinin başarısını artırmak için, doğru hiper parametre kombinasyonlarını bulmak önemlidir. Bu noktada, Grid Search, sistematik bir yaklaşım sunarak hiper parametrelerin optimize edilmesini amaçlayan etkili bir yöntemdir. Makine öğrenimi modellerindeki hiper parametreler, modelin performansını doğrudan etkiler. Grid Search, belirlenen hiper parametre değer kombinasyonlarını sistematik bir şekilde değerlendirerek, en iyi performansı sağlayacak hiper parametre setini bulmayı amaçlar. Avantajları:

Sistematik ve eksiksiz bir arama stratejisi, basit ve kullanımı kolaylığı avantajları arasında yer almaktadır. Dezavantajları ise büyük hiper parametre uzayları için hesaplama maliyeti yüksek olması, ofistike optimizasyon stratejilerine göre daha yavaş konverjans olması.

Sonuç olarak Grid Search, hiper parametre ayarı için etkili bir başlangıç noktası sağlar. Ancak, büyük hiper parametre uzaylarında veya daha hızlı konverjans elde etmek istendiğinde diğer optimizasyon stratejileri de düşünülmelidir. Grid Search, makine öğrenimi modellerinin performansını artırmak ve doğru hiper parametre setini bulmak için kullanışlı bir araçtır.

Temel Adımlar:

1. Hiper parametreler ve Değerler Belirlenir

- $\theta = [\theta_1, \theta_2, \dots, \theta_n]$: Hiper parametreleri temsil eden vektör.
- $\Theta = \{v_{i1}, v_{i2}, \dots, v_{imi}\}$: Her bir hiper parametre için belirlenen değer kümesi.

2. Hiper parametre Uzayı Oluşturulur

- $\Theta = \Theta_1 \times \Theta_1 \times \dots \times \Theta_n$: Tüm hiper parametre uzayı. Tüm olası hiper parametre kombinasyonları içerir.

3.Optimize edilen fonksiyon belirlenir.

- $performance(\theta)$: Modelin performansını içeren metrik

4.Grid Search algoritması uygulanır.

– Her bir hiper parametre kombinasyonu için performans değerlendirilir. Denklem (1.5) de gösterildiği gibidir.

$$performance(\theta_{ij})\ for\ \theta_{ij}\ \epsilon\ \Theta \tag{1.5}$$

5. En İyi Parametre Seti Seçilir

– En iyi performansa sahip hiper parametre kombinasyonu seçilir. Denklem (1.6) de gösterildiği gibidir.

$$\theta_* = argmax\theta_{ij}\epsilon\ \Theta\ performance(\theta_{ij}) \tag{1.6}$$

3.2.8. Bayesian optimization

Bayesian Optimization, birçok uygulamada kullanılan bir optimize etme algoritmasıdır. Bu algoritma, optimize edilen bir fonksiyonun değerlendirmesi maliyetli olduğunda veya parametre uzayı büyük olduğunda özellikle etkili olabilir. Aynı zamanda, enerji kullanım tahmininde de başarıyla kullanılabilecek bir algoritmadır. Enerji kullanım tahmininde, bir sistemin belirli koşullar altında ne kadar enerji tüketeceğini tahmin etmek genellikle karmaşık bir problemdir ve bu tahminler doğru bir şekilde yapılmalıdır. Bu tür bir tahminin doğruluğunu artırmak ve enerji verimliliğini optimize etmek için Bayesian Optimization kullanılabilir. Olasılık tabanlı Naive Bayes sınıflandırma yöntemi aynı zamanda makine öğrenmesi yaklaşımıdır (Şahinaslan, Dalyan, ve Şahinaslan, 2022).

Bayesian Optimization, enerji tüketim tahmini yaparken belirli parametrelerin en iyi değerlerini bulunmasına yardımcı olan etkili bir optimizasyon yöntemidir. Enerji tüketimini belirleyen matematiksel bir fonksiyonun en iyi değerini elde etmek için farklı parametre kombinasyonlarını denemek maliyetli olabilir ve zaman alabilir.

Bayesian Optimization, bu durumda özellikle etkili bir çözümdür. Belirli sayıda deneme yaparak, fonksiyonun en iyi değerine yaklaşan veya bu değeri bulan bir model oluşturur. Bu modeli kullanarak, gerçek fonksiyonu test etmek yerine, hangi parametrelerin daha iyi olma olasılığını tahmin edilebilir. Bu sayede maliyetli denemeleri minimize ederek enerji tüketimini en iyi şekilde tahmin eden parametre kombinasyonuna daha hızlı ve etkili bir şekilde ulaşabiliriz.

1.Optimize Edilen Fonksiyon (Objective Function):$f(x)$

- $f(x)$: Enerji tüketimini belirleyen fonksiyon

2.Gözlemler (Observations):$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$

- (x_i, y_i): i'inci gözlem, x_i parametre setinde yapılan değerlendirme sonucu elde edilen y_i fonksiyon değeri.

3.Bayesian Model (Surrogate Model): $P(y|x, D)$

- $P(y|x, D)$:x parametre setindeki değerlendirmenin sonucu olan y değerinin olasılık dağılımını ifade eden bir model. Genellikle Gaussian Process Regression (GRP) kullanılır.

4.Akıllı Seçim (Acqusition Function):$a(x; D)$

- $a(x; D)$: Optimize edilmemiş alanlarda hangi noktaların değerlendirileceğini seçen bir akıllı seçim fonksiyonu.

5.Fonksiyon Değerlendirmesi (Function Evaluation):$y_{new} = f(x_{new})$

- x_{new}: Akıllı seçim fonksiyonu tarafından seçilen yeni parametre seti.
- y_{new}: Gerçek fonksiyon değeri.

6.Bayesian Model Güncellemesi: $D_{new} = D \cup \{(x_{new}, y_{new})\}$

- D_{new}: Yeni gözlemlerin eski gözlemlerle birleştirilmesi.

3.2.9. Random search

Random Search, bir optimize etme algoritması olup, optimize edilen bir fonksiyonun değerlendirmesi maliyetli veya parametre uzayı büyük olduğunda kullanılan basit bir yöntemdir. Bu algoritma, parametre uzayında rastgele seçilen noktalarda fonksiyonu değerlendirir ve en iyi sonuçları elde edene kadar devam eder.

Bu algoritma ile enerji tüketim tahmini modelini optimize etmeye yönelik basit bir strateji sunar. Ancak, bu yöntem genellikle daha sofistike optimizasyon algoritmalarına kıyasla daha yavaş konverjansa sahiptir. Yani bu algoritma, basit bir optimizasyon stratejisine dayanır. Parametre uzayında rastgele seçilen noktalarda fonksiyonu değerlendirir. Bu, algoritmanın herhangi bir özel bilgi veya önceki bilgi kullanmadan rastgele denemelerle ilerlemesini sağlar. Rastgele seçim stratejisi, en iyi sonuca ulaşmak için parametre uzayını gezme konusunda etkili olmayabilir. Özellikle parametre uzayı büyük veya karmaşık olduğunda, rastgele seçimlerle optimal çözüme ulaşma süreci yavaş olabilir. Bu durum, diğer daha sofistike optimizasyon stratejilerinin kullanılmasının önerilmesine neden olabilir. Random Search'e kıyasla, daha sofistike optimizasyon algoritmaları (örneğin, Bayesian Optimization gibi) genellikle daha hızlı konverjansa sahip olabilir. Bu algoritmalar, parametre uzayında daha akıllı bir arama stratejisi kullanabilir ve daha az sayıda deneme ile daha iyi sonuçlar elde etmeye yönelik tasarlanabilir. Random Search'e kıyasla, daha sofistike optimizasyon algoritmaları (örneğin, Bayesian Optimization gibi) genellikle daha hızlı konverjansa sahip olabilir. Bu algoritmalar, parametre uzayında daha akıllı bir arama stratejisi kullanabilir ve daha az sayıda deneme ile daha iyi sonuçlar elde etmeye yönelik tasarlanabilir.

Random Search basitliği ve kolay kullanımı ile avantaj sağlar, ancak büyük veya karmaşık parametre uzaylarıyla başa çıkmak veya hızlı konverjans elde etmek için daha gelişmiş optimizasyon stratejileri düşünülebilir. Bu durum, problem bağlamına ve kullanılabilir kaynaklara bağlı olarak değişebilir.

Sonuç olarak, Random Search'ün avantajları, basit ve kullanımının kolay olmasıdır, ancak daha karmaşık problemlerde Bayesian Optimization gibi yöntemler genellikle daha etkili olabilir.

Random Search ve Enerji Tüketim Tahmini:

1-Optimize Edilen Fonksiyon (Objective Function): f(x)

- f(x): Enerji tüketimini belirleyen fonksiyon.
- x: Parametre seti. Bu set, enerji tüketimini etkileyen parametreleri içerir.

2-Parametre uzayı ve başlangıç değerleri:

- Parametre uzayı, enerji tüketimini etkileyen farklı parametreleri içerir.
- Başlangıçta rastgele parametre değerleri seçilir.

3-Fonksiyon Değerlendirmesi (Function Evaluation):

- Her iterasyonda Random Search, parametre uzayında rastgele bir noktayı seçer.
- Seçilen parametre seti, enerji tüketim fonksiyonuna uygulanarak değerlendirilir.
- Yani, y=f(x), burada y enerji tüketim tahminini temsil eder.

4-En İyi Parametre Setinin Güncellenmesi:

- Her değerlendirmeden sonra elde edilen enerji tüketim tahminleri karşılaştırılır.
- En iyi sonuçları veren parametre seti kaydedilir.

5-Iterasyonlar:

- Belirli bir durma kriteri sağlanana kadar adımlar tekrarlanır.
- Örneğin, belirli bir maksimum iterasyon sayısı veya belirli bir süre boyunca devam edebilir.

3.3. Veriler ve Toplanması

Veri seti, zaman damgası ve elektrik tüketim değerlerini içermektedir. 1 Ocak 2018 tarihinden başlayarak 25 Aralık 2023 tarihine kadar 52.438 satırdan oluşan toplam 6 yıllık saatlik verileri içermektedir.

Eğitim verisi, verinin başlangıcından test edilecek verinin başlangıç tarihine kadar uzanmaktadır. Test verisi ise eğitim verisinin bittiği tarihten sonraki 48 saati

kapsamaktadır. Örneğin: Test verisi 10 Aralık 00:00 – 11 Aralık 23:00 tarihlerini kapsamaktadır. Bunun sebebi Piyasa Takas Fiyatının (PTF) her gün 14:00'te bir sonraki gün için açıklanmasıdır. Test edilen günlerin tam olarak aralığı Tablo 3'de gösterildiği gibidir.

Tablo 3. Test Edilen Günler

Başlangıç Tarihi	**Bitiş Tarihi**
21 Aralık 2023 00:00	22 Aralık 2023 23:00
22 Aralık 2023 00:00	23 Aralık 2023 23:00
23 Aralık 2023 00:00	24 Aralık 2023 23:00
24 Aralık 2023 00:00	25 Aralık 2023 23:00
25 Aralık 2023 00:00	26 Aralık 2023 23:00

Tablo 3'deki gibi seçilen günler için test verisinin ana veriye oranı ise Tablo 4'deki gibidir.

Tablo 4. Test Verisinin Ana Veriye Oranı

Başlangıç Tarihi	**Bitiş Tarihi**	**Oran (%)**
21 Aralık 2023 00:00	22 Aralık 2023 23:00	0,0915
22 Aralık 2023 00:00	23 Aralık 2023 23:00	0,0915
23 Aralık 2023 00:00	24 Aralık 2023 23:00	0,0915
24 Aralık 2023 00:00	25 Aralık 2023 23:00	0,0915
25 Aralık 2023 00:00	26 Aralık 2023 23:00	0,0915

Tablo 4'de görüldüğü gibi test verileri 48'er saatlik zaman dilimlerini kapsamaktadır. Ana verinin toplamda %0,457 'si test verisi olarak ayrılmıştır. Kullanılan modellerin hiper parametre optimizasyonlarını gerçekleştirebilmek için bir validasyon verisinin oluşturulması gerekmektedir. Bu tarihler Tablo 5'deki gibi seçilmiştir.

Tablo 5. Hiper Parametre Optimizasyonu için Seçilen Tarihler

Başlangıç Tarihi	**Bitiş Tarihi**
19 Aralık 2023 00:00	20 Aralık 2023 23:00
20 Aralık 2023 00:00	21 Aralık 2023 23:00

İlk olarak elektrik tüketim miktarını etkileyen tarih değişkenleri veriye eklenmiştir. Ardından yapılan analizler sonucunda etkileri tespit edilen özel günler (resmi tatiller ve dini bayramlar) ve COVID-19 dönemini işaret eden bir kukla değişken (1,0 olarak oluşturmanın adı kukla değişken) oluşturulmuştur. Oluşturulan değişkenlerin yardımı ile elektrik tüketim değerlerinin saat, haftanın günü, ay, ve yıl gibi tarih bazındaki değişmelerinin; COVID-19, resmi ve dini bayramlar gibi özel gün bazındaki değişmelerinin tespit edilmesine yardımcı olmaktadır. Ek olarak, tüketim verilerinin önceki değerlerine dayanarak gecikme veya gecikme özellikleri oluşturulmuştur. Bu, zaman serisi verilerinin doğasını belirlemek ve modelin daha önceki verileri kullanarak gelecekteki tüketimi tahmin etmesine izin vermek için gereklidir.

Kaynak veriler EPİAŞ'ın "Real Time Consumption" web servisinden yararlanılarak elde edilmiştir. Kullanılan bu kaynak veriler MS Office programlarından biri olan Excel'e aktarılmıştır. Excel üzerinden okuma yapılacak ana veri elde edilip ilgili adımlar sırasıyla uygulanmıştır.

3.3.1. Veri çekimi

EPİAŞ'ın web servislerinden herkesin kullanımına açık olan "Real Time Consumption" yani gerçek zamanlı tüketim endpoint'i aracılığıyla veriler elde edilmiştir. 1 Ocak 2018 tarihinden başlayarak 25 Aralık 2023 tarihine kadar olan saatlik elektrik tüketim verileri

çekilmiştir. Veri içerisinde zaman damgası ve elektrik tüketim değerleri bulunmaktadır. Web servisten elde edilen veri içeriği MS Office programlarından biri olan Excel'e aktarılmıştır. Excel, zaman damgası ve elektrik tüketim değeri olarak iki sütundan, 52.438 satırdan oluşmaktadır.

3.3.2. Öznitelik mühendisliği

Saatlik elektrik tüketim veri seti üzerinde çeşitli zaman serisi analizleri ve öznitelik mühendisliği işlemleri gerçekleştirilir. Bu işlemler şu adımlardan oluşmaktadır.

Zaman Değişkenlerinin Oluşturulması: 'date' sütunundan yararlanarak yıl, ay, ayın günü, haftanın günü ve saat bilgileri elde edilir. Böylece temel zaman öznitelikleri sağlanmış olur.

Resmi Bayramların İşaretlenmesi: Türkiye'deki belirli resmi bayramlar, veri setinde özel işaretler (flag) olarak tanımlanır. Bu işaretler, bayram günlerindeki elektrik tüketimini analiz etmemek için kullanılmıştır.

Dini Bayramların İşaretlenmesi: Kurban ve Ramazan bayramlarının tarihleri özel olarak belirlenir ve bu tarihler veri setinde işaretlenir. Bu işlem, dini bayramların elektrik tüketimi üzerindeki etkilerine dahil etmemek için kullanılmıştır.

Covid-19 Pandemisi Etkisinin İşaretlenmesi: Covid-19 pandemisinin belirli bir dönemdeki etkisi, veri setinde özel bir işaretle gösterilir. Bu, pandeminin elektrik tüketimi üzerindeki etkilerine dahil etmemek için kullanılmıştır.

Zaman Serisi Gecikme Özniteliklerinin Eklenmesi: Elektrik tüketimi verileri için önceki günlerden gelen gecikme (Lag) verileri eklenir. Bu, tüketim verilerinin zaman içindeki değişimini daha iyi anlamaya yardımcı olacaktır.

3.3.3. Veri hazırlama süreci

Elektrik tüketimi verilerinin analiz edilmesi için kullanılacak olan özniteliklerin bir model formülü içinde düzenlenmesi sağlanmıştır. Formül:

consumption ~ C(Yil) + C(Ay) + C(Hafta_Gun) + C(Saat) + Yilbasi + Nisan_23 + Mayis_1 + Mayis_19 + Temmuz_15 + Agustos_30 + Ekim_29 + Ramazan_Bayrami + Kurban_Bayrami + Covid_19 + Lag_2 + Lag_3 + Lag_4 + Lag_5 + Lag_6 + Lag_7

İlk olarak, test veri seti için başlangıç ve bitiş tarihleri belirlenmiş olup, test seti için başlangıç tarihi olarak 2023 yılının 10 Aralık gününü seçilmiştir. Bitiş tarihi ise başlangıç tarihinden bir gün sonrası, yani 2023 yılının 11 Aralık günü olarak belirlenmiştir.

Eğitim setimi, ana veri setinden elde edilmiştir. Bu set, date kolonunun barındırdığı tarihleri, test setinin başlangıç tarihinden küçük olan tüm kayıtları içermektedir. Bu işlem, veri setinin test tarihinden önceki tüm verilerini kapsamakta ve modele öğrenme için gerekli veri tabanını sağlamaktadır.

Test seti de ana veri setindeki data kolonu kullanarak oluşturulmaktadır. Bu set, date kolonunda belirlenen test tarihleri arasında olan kayıtları içermektedir. Yani, test başlangıç tarihi ile test bitiş tarihi arasındaki tarihleri kapsayan veriler bu sete dahil edilmiştir. Test seti, oluşturulacak olan modellerin tahmin performansını değerlendirmede kullanılmaktadır.

Test setinde sadece date ve consumption sütunlarını seçerek yeni bir tablo oluşturulmuştur. Oluşturulan yeni tablo kullanılacak modellerin sonuçlarının analizi ve sunumu için kullanılmaktadır. Bu, modellerin test seti üzerindeki performansını gözlemlemek ve değerlendirmek için bir adımdır.

Oluşturulan formül baz alınarak bağımlı ve bağımsız değişken oluşturulmuştur. Formülde bağımlı ve bağımsız değişken ayrıştırılmasına örnek verecek olursak; **‘y ~ x1 + x2’** formülü, **y**'nin bağımlı değişken ve **x1** ile **x2**'nin bağımsız değişkenler olduğunu belirtir.

Eğitim seti ve test seti üzerinde formül baz alınarak bağımlı ve bağımsız değişken ayrıştırılması yapılmıştır. Test veri setinden bağımlı ve bağımsız değişken elde edilmesi işlemi sırasında eksik değerler (NaN) sıfır ile doldurulmuştur. Bu işlem, modelleme sırasında eksik veri sorununu önlemek için yapılmaktadır.

Test veri seti ile test veri setinden elde edilen bağımsız değişkenlerin sütunları eşleştirilmiştir. Bu işlem, test veri seti ile test veri setinden elde edilen bağımsız değişkenlerin aynı sütun yapısına sahip olmasını sağlar, böylece modelleme sırasında uyumsuzluk olmaz.

Sütunların eşleştirilmesi durumunda test veri setinden elde edilen bağımsız değişkenlerde bazı sütunların eksik olmasına karşın sütunlar sıfırla doldurulmuştur. Bu yöntemle hem test verisi, hem de bağımsız değişkenlerden oluşan veri setinin aynı sütun yapısına sahip olması sağlanmıştır.

Eğitim veri setinden elde edilen bağımsız değişken veri setinde kategorisel, tarihsel(yıl, ay, gün, saat), sayısal (Lag...), resmi ve dini bayramları içeren sütunları birleştirerek bağımsız değişkene atama yapılır. Ardından, eksik değerleri olan satırlar kaldırılmaktadır. Eğitim veri setindeki 'consumption' kolonu bağımlı değişkene atama yapılır.

Eğitim veri setinden elde edilen bağımlı ve bağımsız değişkenleri iki alt kümesine bölmek için kullanılırken, bağımsız değişkenler ve bağımlı değişkenlerin uygun şekilde eşleştirilip eşleşmediğini kontrol etmek amacıyla elde edilen bağımlı değişken indeksi içinde bulunan her bir indeks, aynı zamanda elde edilen bağımsız değişken indeksi içinde de bulunuyorsa, bu, her iki veri setinin doğru şekilde eşleştirilir. Bağımlı değişkenlerdeki eksik değerleri olan satırılar kaldırılır.

Yukarıdaki işlemin tam tersi, test veri seti ve buna bağlı bağımlı ve bağımsız değişkenler için uygulanmıştır. Eğitim ve test veri setlerinde kategorisel ve tarihsel sütunlar metin formatına dönüştürülmüştür.

Eğitim setinde kategorisel verilerin bulunduğu sütunlar yeni bir listede toplanmıştır.

Eğitim veri setinde her bir kategorik sütunun konumu bulunup bu konumlar bir liste olarak saklandı. Yani, her bir kategorik sütunun veri çerçevesinde hangi sıra numarasında bulunduğunu belirlendi.

Catboost modelini eğitirken catboost modelinden çıktığı sağlayacağımız eğitim veri seti ve test veri seti hazırlandı. Catboost eğitim veri seti, eğitim seti üzerinden daha önce alınan bağımlı değişken veri seti, bağımsız değişken veri seti ve kategorisel veri seti

setlerinin birleşimi ile oluşturuldu. Catboost test veri seti ise test veri seti üzerinden daha önce alınan bağımlı değişken veri seti, bağımsız değişken veri seti ve kategorisel veri setlerinin birleşimi ile oluşturuldu.

LSTM (Long Short-Term Memory) modeli için zaman serisi verilerini işlemek ve hazırlamak amacıyla süreç uygulanmıştır. Bu süreç, bir veri setinin önceden belirlenmiş başlangıç ve bitiş tarihleri arasında bölünmesi ve ardından zaman serisi analizi için uygun hale getirilmesini içerir. İşlemin ana adımları şu şekildedir:

- **Veri Setinin Bölünmesi**: Veri seti, belirtilen tarih aralığına göre eğitim ve test setlerine ayrılır. Eğitim seti, başlangıç tarihinden önceki verileri; test seti ise başlangıç ve bitiş tarihleri arasındaki verileri içerir.
- **Veri Hazırlığı**: Veri seti, kategorik, sayısal, ikili ve tarih sütunlarına göre ayrılır. Bu sütunlar, LSTM modeline beslenecek verilerin hazırlanmasında kullanılır.
- **Eksik Veri Kontrolü ve Temizleme**: Eksik verilerin kontrolü yapılır ve sayısal sütunlardaki eksik veriler temizlenir.
- **Kategorik Değişkenlerin Düzenlenmesi**: Kategorik değişkenler, dummy (one-hot encoding) değişkenlere dönüştürülür.
- **Ölçeklendirme**: Tüm veriler, LSTM modelinin daha iyi sonuçlar vermesini sağlamak için 0 ile 1 arasında ölçeklenir.
- **Zaman Serisi Dönüşümü**: veri setini, modelin öğrenmesi için gerekli olan şekilde, yani önceki zaman adımlarını ve tahmin edilecek değeri içerecek şekilde dönüştürür.
- **Eğitim ve Test Setlerinin Ayrılması**: Hazırlanan veri seti, eğitim ve test setlerine ayrılır.
- **Verilerin LSTM Formatına Dönüştürülmesi**: Eğitim ve test setleri, LSTM modelinin girdi formatına uygun olarak yeniden şekillendirilir.
- **Sonuçların Döndürülmesi**: İşlenmiş veri setleri ve yararlı olabilecek ek bilgiler (örneğin, ölçekleyici ve orijinal test veri seti) son kullanıcıya döndürülür.

Bu süreç, veri setinin zaman serisi analizi için uygun hale getirilmesini ve LSTM modelinin eğitimi için gerekli veri yapılarının hazırlanmasını sağlamaktadır.

3.4. Verilerin Çözümlenmesi ve Yorumlanması

Bu bölümünde sunulan çalışma, elektrik tüketim verilerinin kapsamlı bir şekilde analizi ve tahmini için bir yaklaşım sunmaktadır. Bu tür veriler üzerinde derin öğrenme ve makine öğrenimi modelleri incelenmiştir. Enerji tüketimi tahminlerinin iyileştirilmesi için önemli sonuçlar elde edilmiştir. Bu çalışma, elektrik tüketim tahminlerinde derin öğrenme ve makine öğrenimi modellerinin kullanılabilirliğini vurgulamaktadır. Ayrıca, bu alandaki gelecek araştırmalar için de potansiyel olarak kullanılabilecek alanlar da vurgulanmıştır.

Alınan bu veriler Python yardımı ile analiz edilmiştir. Analizi gerçekleştirilirken test edilen günler aşağıdaki gibidir:

– 2023-12-18 ile 2023-12-19
– 2023-12-19 ile 2023-12-20

Analiz edilen sonuçlar şu şekildedir:

3.4.1. LightGBM

İlk olarak LightGBM modeli kullanılmıştır. Bu modelde test amaçlı kullanılan sabit değerler şu şekildedir:

- max_depth: 4
- learning_rate: 0.1
- n_estimators: 100
- max_bin: 200
- num_leaves: 30
- n_jobs: 8
- verbose:-1

Bu parametrelerin açıklamaları aşağıda verilmiştir.

1. max_depth: Karar ağacının maksimum derinliğini belirler. Bu, ağacın kaç düğüme kadar dallanabileceğini kontrol eder.

2. learning_rate: Her ağacın katkısını kontrol eden bir öğrenme oranıdır. Küçük bir öğrenme oranı, modelin daha güvenilir ancak daha yavaş öğrenmesini sağlar.

3. n_estimators: Toplam ağaç sayısını belirler. Daha fazla ağaç, modelin genellikle daha iyi performans göstermesine olanak tanır, ancak aynı zamanda daha fazla hesaplama gücü gerektirir.

4. max_bin: Sayısal özelliklerin bölünebileceği maksimum aralık sayısını belirler. Bu, sayısal özelliklerin daha az hassas olmasını sağlar.

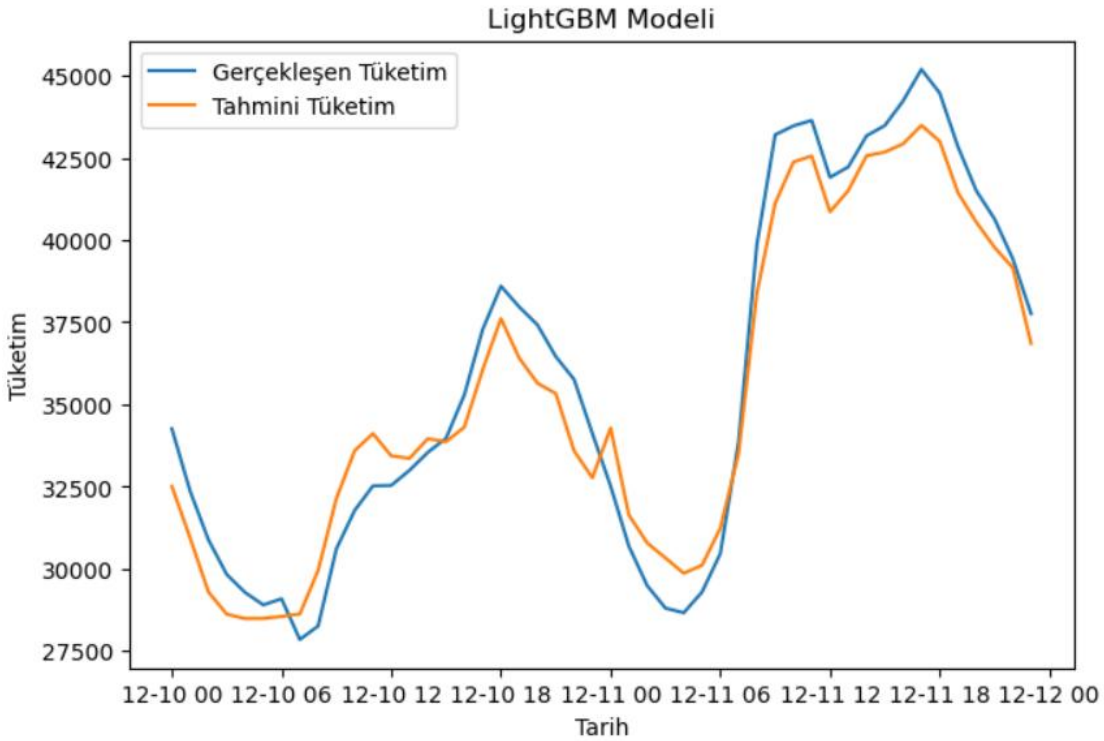
5. num_leaves: Bir ağaçtaki maksimum yaprak sayısını belirler. Daha büyük numara, modelin daha karmaşık ilişkileri öğrenmesine olanak tanır, ancak aynı zamanda aşırı uydurmayı artırabilir.

6. n_jobs: Eğitim sırasında kullanılacak işçi sayısını belirler. Bu, eğitimi hızlandırmak için kullanılan paralel işleme yeteneğini temsil eder.

7. verbose: Eğitim sürecinde çıktıları kontrol eden bir parametredir. -1 değeri, çıktıların kapatılması anlamına gelir.

Bu parametreler, modelin nasıl öğrenme yapacağını ve genel performansını etkileyen önemli faktörleri barındırır. Test amaçlı sabit girilen parametreler sonucunda elde edilen tahminler y_pred olarak kaydedilmiştir ve bir DataFrame'e (return_df) eklenmiştir. Bu tahminlerin gerçek değerlerle karşılaştırılması, modelin performansının değerlendirilmesi için kullanılmıştır.

Python programlama dili yardımıyla Matplotlib kütüphanesi kullanarak bir figür (grafik) oluşturulmuş ve bu figürün boyutları belirlenmiştir. Oluşturulan bu figür aşağıdaki gibidir:



Şekil 3. LightGBM Modeli Grafiği

3.4.2. XGBoost

XGBoost (eXtreme Gradient Boosting) modelinde kullanılan kod parçası, bazı hiper parametreler ile eğitilen bir regresyon modelini göstermektedir. Belirtilen test amaçlı kullanılan sabit değerlerin açıklaması şöyledir:

– n_estimators = 100

Bu, modelin kullanacağı ağaç sayısını belirtir. XGBoost, bir dizi ağaç modeli oluşturarak çalışır ve n_estimators, oluşturulacak toplam ağaç sayısını ifade eder. Burada, model değeri 100 ağaç kullanılmıştır.

– max_depth = 4

Her ağacın maksimum derinliğini belirtir. Derinlik, ağacın kökünden en uzak yaprağa kadar olan en uzun yoldur. Daha büyük bir max_depth değeri, modelin daha karmaşık hale gelmesine ve potansiyel olarak aşırı uyuma (overfitting) yol açmasına neden olabilir. Bu durumda, her ağacın maksimum derinliği 4 olarak belirlenmiştir.

– learning_rate = 0.05

Her ağacın katkısını azaltmak için kullanılan bir ölçeklendirme faktörüdür. Bu, genellikle "adım boyutu" olarak adlandırılır ve her iterasyonda modelin ne kadar "öğreneceğini" (yani, ağırlıklarını ne kadar güncelleyeceğini) belirler. Daha küçük bir öğrenme oranı, genellikle daha kararlı, ancak daha yavaş bir eğitim sürecine yol açar. Burada, öğrenme oranı 0.05 olarak ayarlanmıştır.

– random_state = 42

Modelin eğitimi sırasında kullanılan rastgele sayı üretecinin tohum değeridir. Bu, modelin her seferinde aynı şekilde eğitilmesini sağlar, böylece deneyler tekrarlanabilir. 42 genellikle örnek bir tohum değeri olarak kullanılır.

– model_xgb.fit(x_train.values, y_train.values)

Bu, modelin eğitim veri seti (x_train ve y_train) kullanılarak eğitildiği yerdir. x_train.values ve y_train.values, verilerin NumPy dizisi olarak kullanılması gerektiğini belirtir, çünkü XGBoost genellikle bu formatta veriyle daha iyi çalışır.

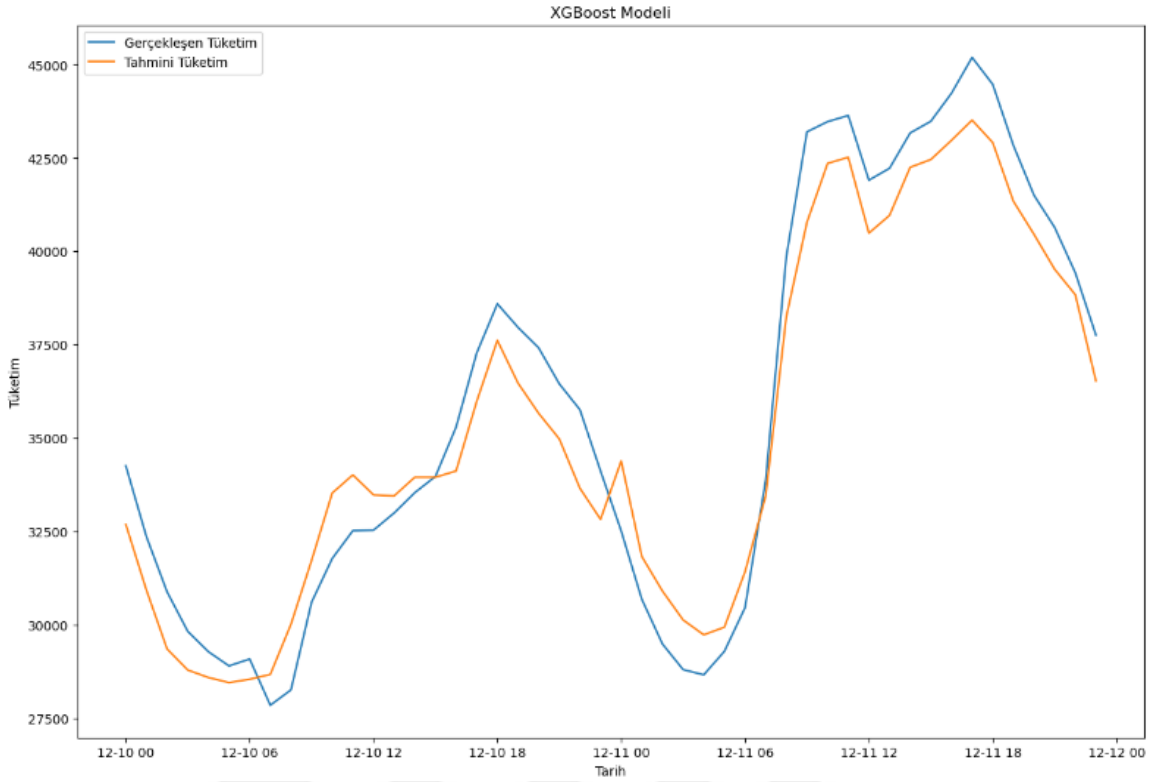
– y_pred = model_xgb.predict(x_test.values)

Eğitilmiş model, test veri seti (x_test) üzerinde tahminler yapmak için kullanılır. Burada, x_test.values kullanılarak modelden tahminler alınır ve bu tahminler y_pred değişkenine atanır.

– return_df["XGBoost"] = y_pred

Model tarafından test amaçlı sabit girilen parametreler sonucunda elde edilen tahminler, return_df adlı bir DataFrame'in "XGBoost" sütununa atanır. Bu, tahmin edilen değerlerin kaydedilmesi ve daha sonraki analiz veya karşılaştırma için kullanılması anlamına gelir.

Oluşturulan bu figür aşağıdaki gibidir:



Şekil 4. XGBoost Modeli Grafiği

3.4.3. CatBoost

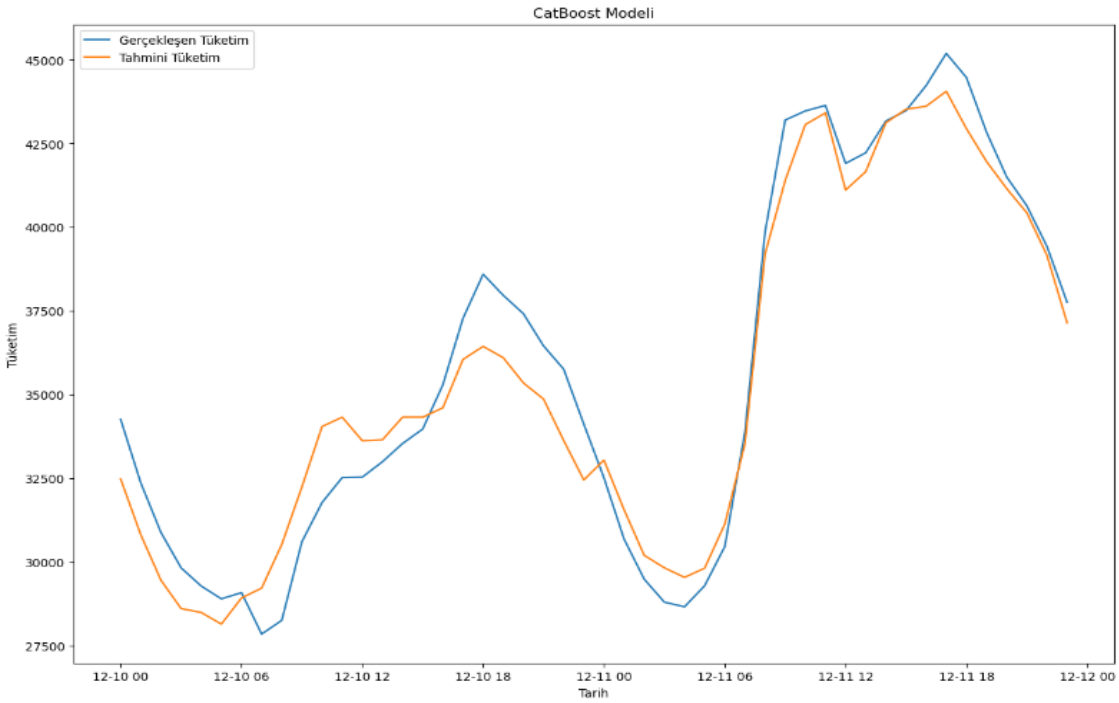
CatBoost (Category Boosting) modelinde kullanılan test amaçlı sabit girilen parametreler bir regresyon modeli oluşturur ve bazı belirli hiper parametreler ile eğitilir. Belirtilen test amaçlı kullanılan sabit değerlerin açıklaması şu şekildedir:

- CatBoostRegressor(max_depth = 4, learning_rate = 1e-1, iterations = 100): CatBoostRegressor, kategorik verilerle iyi başa çıkan ve gradyan artırma algoritması kullanan bir regresyon modelidir.
- max_depth = 4: Her karar ağacının maksimum derinliğini belirtir. Daha büyük bir max_depth değeri, modelin daha karmaşık hale gelmesine yol açabilir.
- learning_rate = 1e-1: Öğrenme oranı, bu durumda 0.1 olarak belirlenmiştir. Bu, her iterasyonda model ağırlıklarının ne kadar güncelleneceğini belirler.
- iterations = 100: Modelin yapacağı toplam iterasyon (veya ağaç) sayısını belirtir.

– model_cat.fit(Train_pool, verbose = 0): Bu, modelin Train_pool veri seti kullanılarak eğitildiği yerdir. Train_pool, genellikle özellikler ve hedef değişken içeren bir veri setidir.
– verbose = 0 parametresi, eğitim sırasında herhangi bir çıktıyı (örneğin, iterasyon durumu veya hata metrikleri) bastırmaması gerektiğini belirtir.
– y_pred = model_cat.predict(Test_pool): Eğitilmiş model, Test_pool veri seti üzerinde tahminler yapmak için kullanılır. Test_pool, genellikle test amaçlı özellikler içeren bir veri setidir.
– return_df["CatBoost"] = y_pred: Model tarafından yapılan tahminler, return_df adlı bir DataFrame'in "CatBoost" sütununa atanır. Bu, tahmin edilen değerlerin kaydedilmesi ve daha sonraki analiz veya karşılaştırma için kullanılması anlamına gelir.

CatBoost, özellikle kategorik verilerin işlenmesi ve modelin karmaşıklığını yönetirken iyi performans gösteren bir makine öğrenimi algoritmasıdır. Bu kod parçası, modelin eğitilmesi ve test veri seti üzerinde tahminler yapılması için temel adımları göstermiştir.

Oluşturulan bu figür aşağıdaki gibidir:



Şekil 5. CatBoost Modeli Grafiği

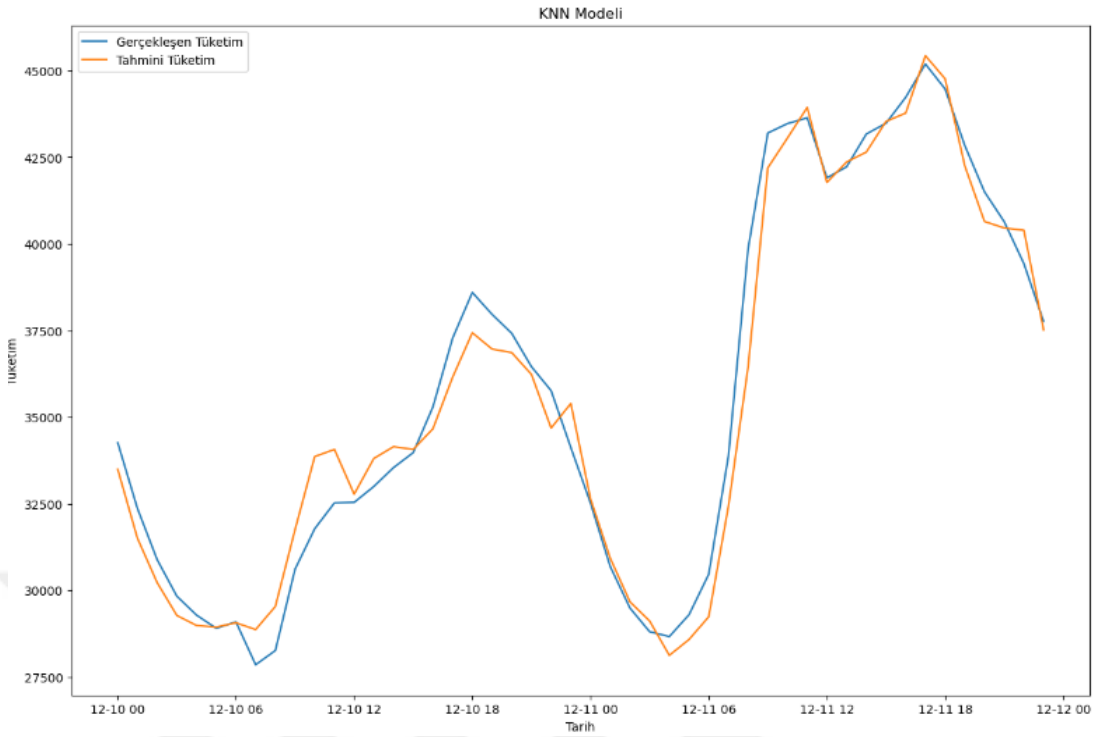
3.4.4. KNN

KNN modelindeki kullanılan kod parçacığı, scikit-learn kütüphanesini kullanarak K-En Yakın Komşu (KNN) algoritması ile bir regresyon modeli oluşturur ve eğitir. Test amaçlı sabit girilen parametrelerin açıklaması şu şekildedir:

- KNeighborsRegressor(n_neighbors = 5): KNeighborsRegressor, K-En Yakın Komşu algoritmasını kullanan bir regresyon modelidir.
- n_neighbors = 5: Bu, tahmin yapılırken dikkate alınacak en yakın komşu sayısını belirtir. Burada, her tahmin için en yakın 5 komşu kullanılacak.
- model.fit(x_train, y_train):Bu, modelin eğitim veri seti (x_train ve y_train) kullanılarak eğitildiği yerdir. x_train, modelin öğrenmesi gereken özelliklerin ve y_train, tahmin etmeye çalıştığı hedef değerlerin olduğu veri setidir.
- y_pred = model.predict(x_test):Eğitilmiş model, test veri seti (x_test) üzerinde tahminler yapmak için kullanılır. Bu, modelin x_test veri setindeki örnekler için hedef değerleri tahmin etmesi anlamına gelir.
- return_df["KNN"] = y_pred:Model tarafından yapılan tahminler, return_df adlı bir DataFrame'in "KNN" sütununa atanır. Bu, modelin tahmin ettiği değerlerin kaydedilmesi ve daha sonraki analiz veya karşılaştırma için kullanılması anlamına gelir.

K-En Yakın Komşu algoritması, veri noktalarının birbirine olan uzaklıklarını temel alarak çalışır. Verilen bir test noktası için, eğitim setindeki en yakın 'k' komşuya bakılır ve bu komşuların hedef değerlerinin ortalaması (regresyon için) veya modu (sınıflandırma için) tahmin olarak kullanılmıştır. Bu yöntem, özellikle veri seti küçük ve basit olduğunda iyi performans gösterebilir. Ancak, çok büyük veri setleri ve yüksek boyutlu özellikler için hesaplama açısından maliyetli olabilir.

Oluşturulan bu figür aşağıdaki gibidir:



Şekil 6. KNN Modeli Grafiği

3.4.5. LSTM

LSTM Modelindeki kod parçası, TensorFlow ve Keras kütüphanelerini kullanarak LSTM (Long Short-Term Memory) tabanlı bir sinir ağı modeli oluşturup, eğitmiştir. Test amaçlı sabit girilen parametrelerin açıklaması şu şekildedir:

- Tohum Değerleri (Seed Values): PYTHONHASHSEED = '0': Python'un hash-based operasyonları için sabit bir başlangıç noktası sağlar. Bu, hash işlemlerinin tekrarlanabilir olmasını sağlar.
- python_random.seed (42): Python'un yerleşik rastgele sayı üreteci için bir tohum değeri ayarlar. Bu, Python'un random kütüphanesi tarafından üretilen rastgele sayıların tekrarlanabilir olmasını sağlar.
- tf.compat.v1.set_random_seed (42): TensorFlow'un rastgele sayı üreteci için bir tohum değeri ayarlar. Bu, TensorFlow işlemlerinde (örneğin, ağırlık başlatmalarında) kullanılan rastgele sayıların tutarlı olmasını sağlar.

– np.random.seed (42): Numpy kütüphanesinin rastgele sayı üreteci için bir tohum değeri ayarlar. Bu, Numpy kullanılarak yapılan rastgele sayı üretiminin tekrarlanabilir olmasını sağlar.

Adam Optimizer Parametreleri:

– learning_rate = 1e-1: Modelin her bir adımda ne kadar "öğreneceğini" belirleyen öğrenme oranıdır. 0.1 değeri, orta derecede hızlı bir öğrenme oranını ifade eder.
– beta_1 = 0.9 ve beta_2 = 0.999: Adam optimizer'ın momentum ve ölçeklendirilmiş gradyan vektörlerini düzeltmek için kullanılan bozulma oranlarıdır.
– amsgrad = False: Adam optimizer'ın AMSGrad varyasyonunu kullanıp kullanmama ayarı. AMSGrad, bazı durumlarda optimizer'ın kararlılığını artırabilir.
– Öğrenme Oranı Azaltıcı (Learning Rate Reducer) Parametreleri:
– factor = 0.8: Eğitim sırasında belirli bir koşul altında öğrenme oranının azaltılacağı faktör.
– patience = 10: İyileşme görülmeden önce beklenmesi gereken ardışık dönem sayısı.
– min_delta = 1e-7: İyileşme olarak kabul edilecek minimum değişiklik miktarı.
– cooldown = 10: Öğrenme oranı azaltıldıktan sonra bir sonraki azaltma öncesi beklenmesi gereken dönem sayısı.

Başlatıcılar (Initializers):

– initializer_1 = initializers.RandomNormal (seed=42) ve initializer_2 = initializers.RandomNormal (seed=15): Ağırlıkları başlatmak için kullanılan rastgele normal dağılım başlatıcıları. Farklı tohum değerleri, başlatıcıların farklı ağırlık değerleri üretmesini sağlar.

LSTM ve Dense Katman Parametreleri:

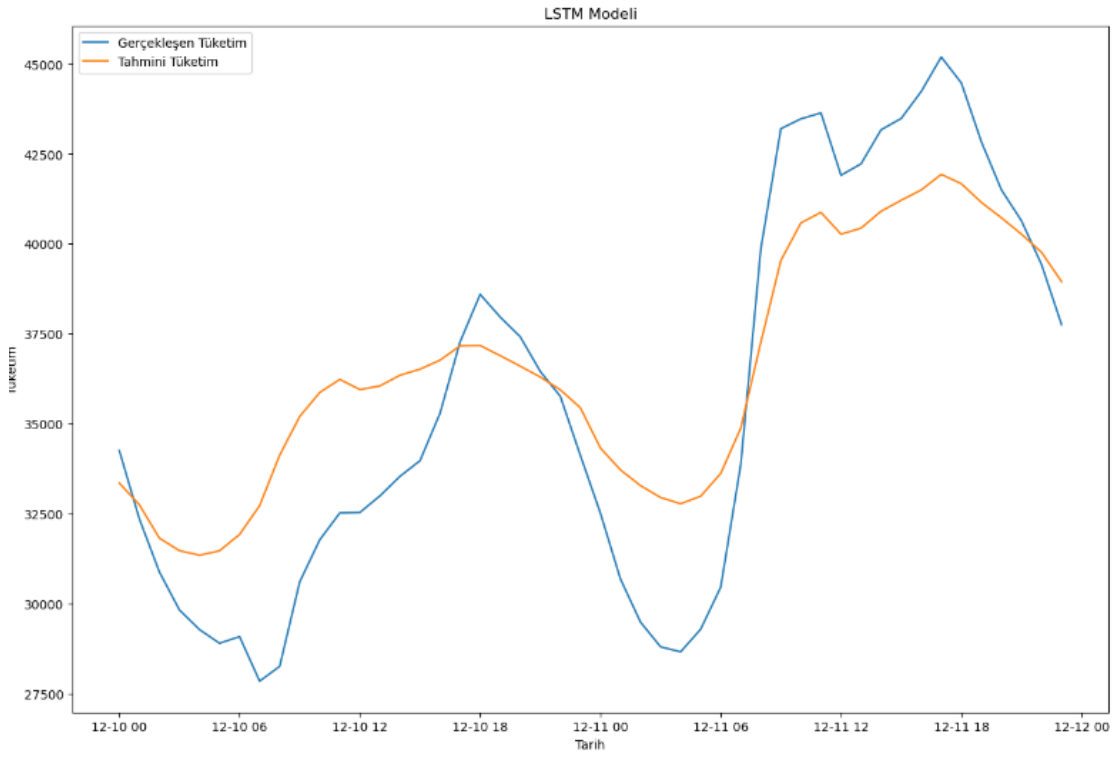
Her LSTM katmanında 25 hücre sayısı bulunmaktadır. Bu, her bir LSTM katmanının karmaşıklığını ve kapasitesini belirler.

– return_sequences = True veya False: Bir LSTM katmanının çıktısının, sonraki katmana diziler (True) veya tek bir çıktı (False) olarak iletilip iletilmeyeceğini belirler.

Eğitim Parametreleri:

– epochs = 10: Modelin veri üzerinden kaç kez geçeceğini belirleyen toplam eğitim dönemi sayısı.
– batch_size = 365*24: Her bir eğitim iterasyonunda modelin eğitim yapacağı örnek sayısı. Bu, genellikle veri setinin boyutuna ve bellek kapasitesine bağlı olarak ayarlanır.
– Bu sabitler, modelin nasıl eğitileceğini, ağırlıklarının nasıl başlatılacağını ve optimize edici davranışını belirleyerek modelin performansını ve tekrarlanabilirliğini etkileyecek önemli faktörlerdir.

Oluşturulan bu figür Şekil 7'de belirtildiği gibidir:



Şekil 7. LSTM Modeli Grafiği

4. BULGULAR VE TARTIŞMA

4.1. Bulgular

XGBoost, Catboost, LightGBM, KNN, LSTM gibi modeller için hiper parametre optimizasyonu gerçekleştirildi. Hiper parametre optimizasyonu gerçekleştirilirken test edilen günler aşağıdaki gibidir:

- 2023-12-18 ile 2023-12-19
- 2023-12-19 ile 2023-12-20

Farklı günler için elde edilen sonuçların ortalamaları minimize edilmeye çalışılmıştır.

Hiper parametre optimizasyonu "Bayesian Optimization" algoritması ile gerçekleştirilmiştir.

"Bayesian Optimization" daha iyi olan noktaya doğru ilerleme mantığında çalıştığı için tercih edilmiştir.

Random Search, yöntemin temel prensibi, arama alanındaki çözümleri rastgele keşfetmektir. Bu yaklaşımın temel avantajı, basit ve uygulanması kolay olmasıdır. Ancak, bu basitlik bazı önemli sınırlamaları da beraberinde getirir.

İlk olarak, Random Search'in en büyük dezavantajlarından biri, arama alanını etkin bir şekilde keşfetmemesidir. Arama alanı büyüdükçe, optimal veya yakın-optimal çözümleri bulma olasılığı azalır. Özellikle yüksek boyutlu veri setleri ve karmaşık modellerle çalışırken, bu yöntemle en iyi sonuçlara ulaşmak oldukça zordur.

Buna ek olarak, Random Search zaman ve kaynak açısından verimsiz olabilir. Optimal bir çözüme ulaşmak için çok sayıda deneme gerekebilir ve bu süreç, özellikle hesaplama kaynakları sınırlı olduğunda, pratik olmayabilir. Ayrıca, her seferinde farklı sonuçlar üretebilmesi, deneylerin tekrarlanabilirliği ve güvenilirliği açısından problemlere yol açabilir.

Bir diğer önemli nokta, Random Search'in önceden elde edilen bilgileri veya bulguları dikkate almamasıdır. Bu, özellikle daha önce yapılan denemelerden elde edilen verilerin var olduğu durumlarda, verimliliği azaltır ve potansiyel olarak daha iyi çözümlere ulaşmayı zorlaştırır.

Son olarak, büyük ölçekli problemlerde Random Search'in etkinliği önemli ölçüde azalabilir. Büyük veri setleri ve karmaşık modellerle çalışırken, rastgele arama yöntemi, arama alanının yalnızca küçük bir bölümünü keşfedebilir ve bu da optimal çözümlerin gözden kaçmasına neden olabilir.

Özetle, Random Search, belirli durumlarda kullanışlı olabilir, ancak sınırlamaları göz önünde bulundurulmalıdır. Özellikle karmaşık ve yüksek boyutlu problemler söz konusu olduğunda, daha gelişmiş ve yönlendirilmiş arama tekniklerinin kullanılması genellikle daha etkili sonuçlar verir.

Random search bu rassallıktan dolayı optimum algoritmayı bulamama ihtimalinden dolayı tercih edilmemiştir.

Grid Search, özellikle makine öğreniminde hiper parametre optimizasyonu için yaygın olarak kullanılan bir yöntemdir. Bu metod, belirlenen hiper parametre değerlerinin tüm olası kombinasyonlarını sistematik bir şekilde deneyerek en iyi sonucu bulmayı amaçlar. Ancak, Grid Search'in bazı dezavantajları da bulunmaktadır.

Birincisi, Grid Search, özellikle büyük parametre uzayları söz konusu olduğunda, oldukça zaman alıcı ve kaynak yoğun bir yöntem olabilir. Her bir parametre kombinasyonu için modeli eğitmek gerektiğinden, parametre sayısı arttıkça gereken hesaplama süresi ve kaynakları da üssel olarak artar. Bu, özellikle büyük veri kümeleri ve karmaşık modellerle çalışırken ciddi bir sınırlama oluşturur.

İkincisi, Grid Search'in parametre uzayını "yoğun" bir şekilde tarayarak çalışması, en iyi çözüme ulaşma garantisi vermez. Örneğin, belirlenen parametre değerleri arasında en iyi sonucu verecek bir kombinasyon olmayabilir. Ayrıca, parametre aralıklarının ve adımlarının doğru bir şekilde seçilmesi gerektiğinden, bu seçimlerin yanlış yapılması durumunda en iyi sonuçlara ulaşılamayabilir.

Üçüncüsü, Grid Search sabit adımlarla ilerlediği için, bazı durumlarda parametre uzayının belirli bölgelerinde gereksiz yere yoğun bir arama yapılabilir. Bu, bazı potansiyel olarak daha verimli bölgelerin göz ardı edilmesine yol açabilir.

Son olarak, Grid Search, parametreler arasındaki etkileşimleri dikkate almaz. Bu, bazı parametrelerin birlikte nasıl etkileşime girdiğinin önemli olduğu durumlarda, yanıltıcı sonuçlara yol açabilir.

Bu dezavantajlara rağmen, Grid Search, modelin performansını farklı parametre kombinasyonları altında değerlendirmek ve bir başlangıç noktası sağlamak için yararlı olabilir. Ancak, daha etkin ve verimli yöntemler araştırıldığında, özellikle büyük veri setleri ve karmaşık modellerle çalışırken, Random Search veya Bayesyan optimizasyonu gibi alternatif yöntemler tercih edilebilir. Random Search' ün tercih edilmeme sebebi açıklandığından dolayı olası en iyi yöntem Bayesyan yöntemidir.

Aşağıda bu çalışmada kullanılan modellerle bu modellere ait parametre ve deneme uzayları verilmiştir.

XGBoost için kullanılan parametreler ve deneme uzayları aşağıdaki gibidir:

- max_depth: [3, 8]
- learning_rate: [0.001, 0.3]
- n_estimators: [100, 2000]
- min_child_weight: [1, 10]
- gamma: [1e-8, 1.0]
- subsample: [0.01, 1.0]
- colsample_bytree: [0.01, 1.0]
- reg_alpha: [1e-8, 1.0]
- reg_lambda: [1e-8, 1.0]

Catboost için kullanılan parametreler ve deneme uzayları aşağıdaki gibidir:

- depth: [3, 5]
- learning_rate: [0.001, 0.3]

- iterations: [100, 2000]
- min_child_samples: [1, 10]
- subsample: [0.01, 1.0]
- colsample_bylevel: [0.01, 1.0]
- l2_leaf_reg: [1e-8, 1.0]

LightGBM için kullanılan parametreler ve deneme uzayları aşağıdaki gibidir:

- max_depth: [3, 5]
- learning_rate: [0.001, 0.3]
- n_estimators: [100, 2000]
- min_child_weight: [1, 10]
- subsample: [0.01, 1.0]
- colsample_bytree: [0.01, 1.0]
- reg_alpha: [1e-8, 1.0]
- reg_lambda: [1e-8, 1.0]

KNN için kullanılan parametreler ve deneme uzayları aşağıdaki gibidir:

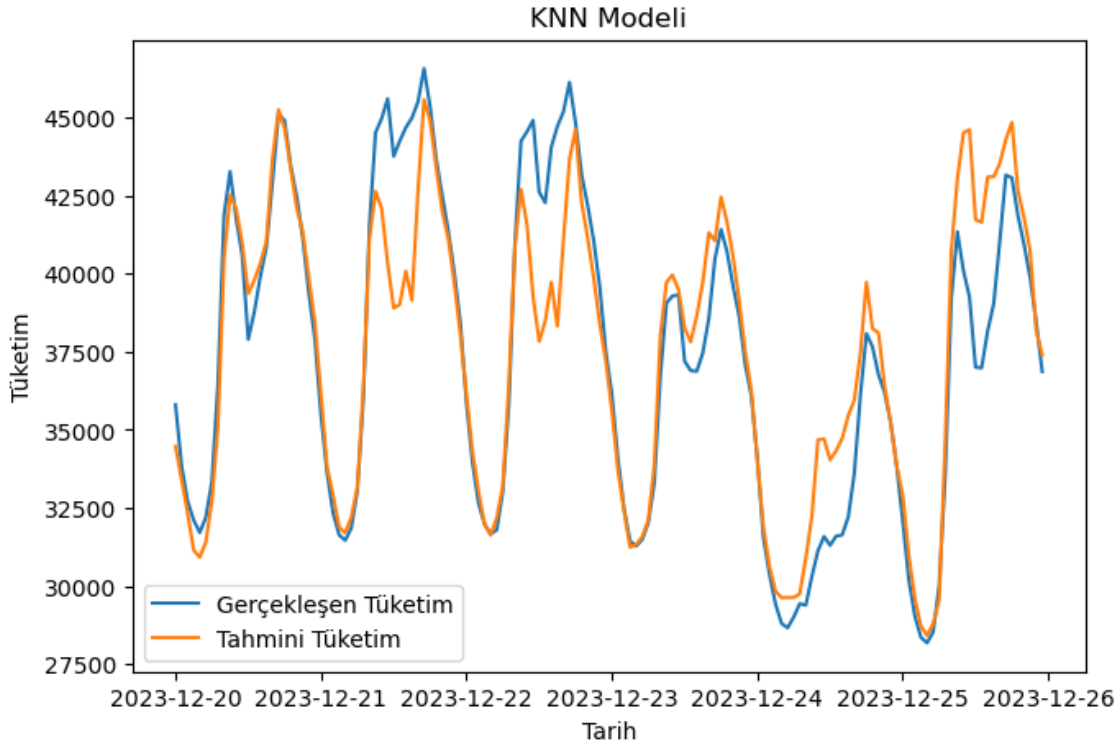
k=3 ile k=50 arasında denemeler yapıldı. Döngü ile en uygun parametre seçildi.

Hiper parametre optimizasyonu sonucunda modeller için en iyi parametre değerleri elde edilmiştir. En iyi parametre değerleri kullanılarak, tahmin edilecek tarih aralıkları aşağıda belirtilmiştir.

- 2023-12-20 ile 2023-12-21
- 2023-12-21 ile 2023-12-22
- 2023-12-22 ile 2023-12-23
- 2023-12-23 ile 2023-12-24
- 2023-12-24 ile 2023-12-25
- 2023-12-25 ile 2023-12-26

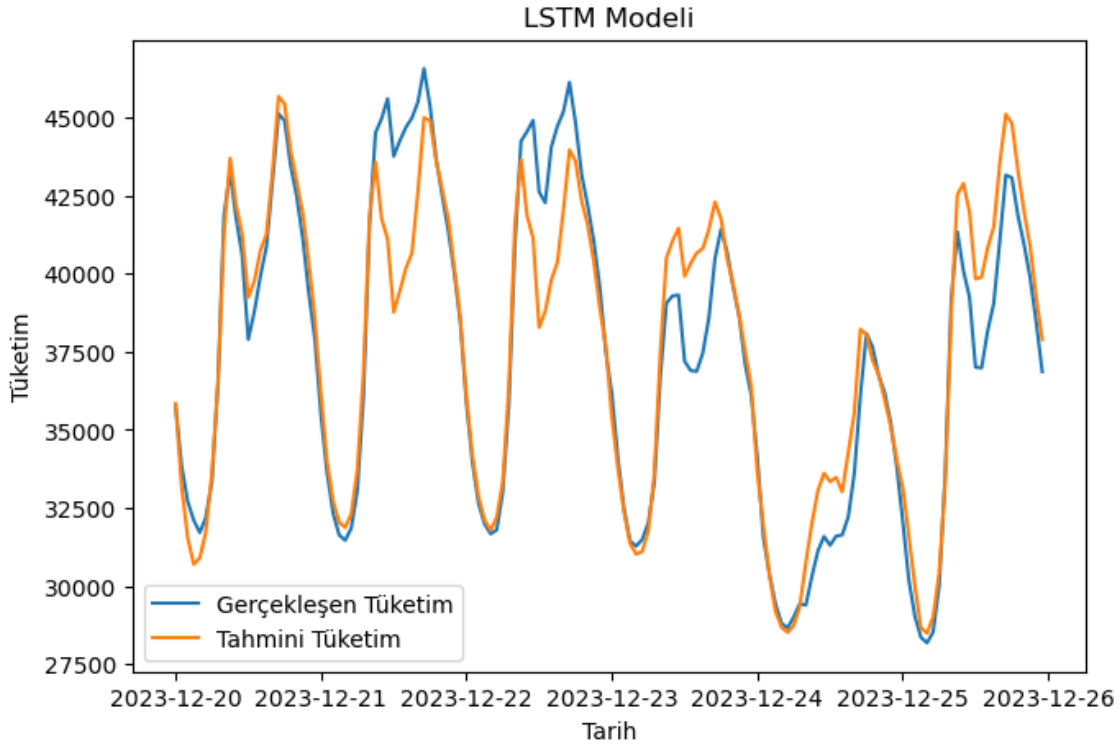
Çalışmada uygulanacak modeller için en iyi parametre sonuçları ve grafikleri aşağıda yer almaktadır.

KNN için en iyi k=13 olarak bulundu. KNN modeli için tahmini tüketim ve gerçekleşen tüketim değerlerine aşağıdaki grafikte yer verilmiştir.



Şekil 8. KNN için Alınan Verilerin Tarih ve Saate Göre Grafiği

LSTM için en iyi parametre değerleri deneme yanılma yöntemi ile gerçekleştirilmiştir. Modelin kullanılabilirliği aştığı ve uzun süren modeller oluşmasını engellemek amacı ile hiper parametre optimizasyonu kullanılmamıştır. Bu tip durumlarla karşılaşıldığında nöron sayısı ve hidden layer sayısı azaltılmıştır. LSTM modeli için tahmini tüketim ve gerçekleşen tüketim değerlerine aşağıdaki grafikte yer verilmiştir.

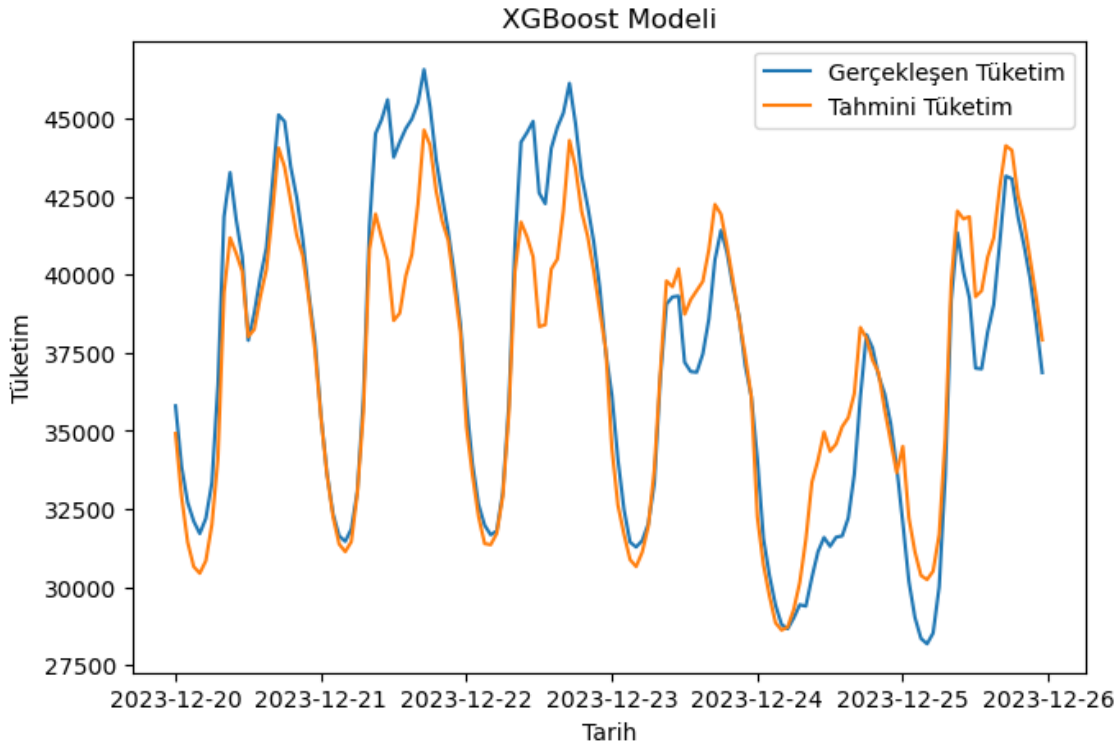


Şekil 9. LSTM için Alınan Verilerin Tarih ve Saate Göre Grafiği

XGBoost için hiper parametre optimizasyonu sonucunda elde edilen en iyi parametre değerleri aşağıdaki gibi bulundu:

- max_depth: 3
- learning_rate: 0.2330742738865523
- n_estimators: 367
- min_child_weight: 10
- gamma: 2.1279185559377696e-05
- subsample: 0.11100864666151612
- colsample_bytree: 0.15828586464248434
- reg_alpha: 2.737665804990315e-06
- reg_lambda:0.007509771340082829

XGBoost modeli için tahmini tüketim ve gerçekleşen tüketim değerlerine aşağıdaki grafikte yer verilmiştir.

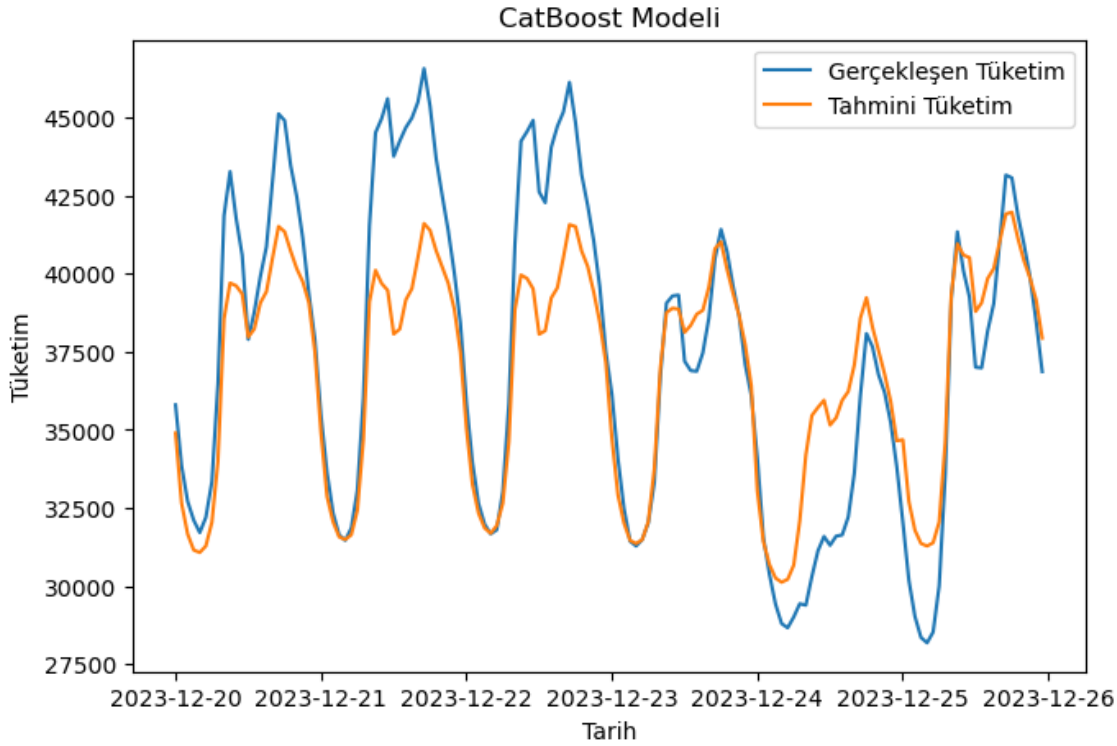


Şekil 9. XGBoost için Alınan Verilerin Tarih ve Saate Göre Grafiği

Catboost için hiper parametre optimizasyonu sonucunda elde edilen en iyi parametre değerleri aşağıdaki gibi bulundu:

- depth: 3
- learning_rate: 0.0014248019009760308
- iterations: 1869
- min_child_samples: 8
- subsample: 0.4677298234995666
- colsample_bylevel: 0.5724827288207895
- l2_leaf_reg: 0.058951696315269475

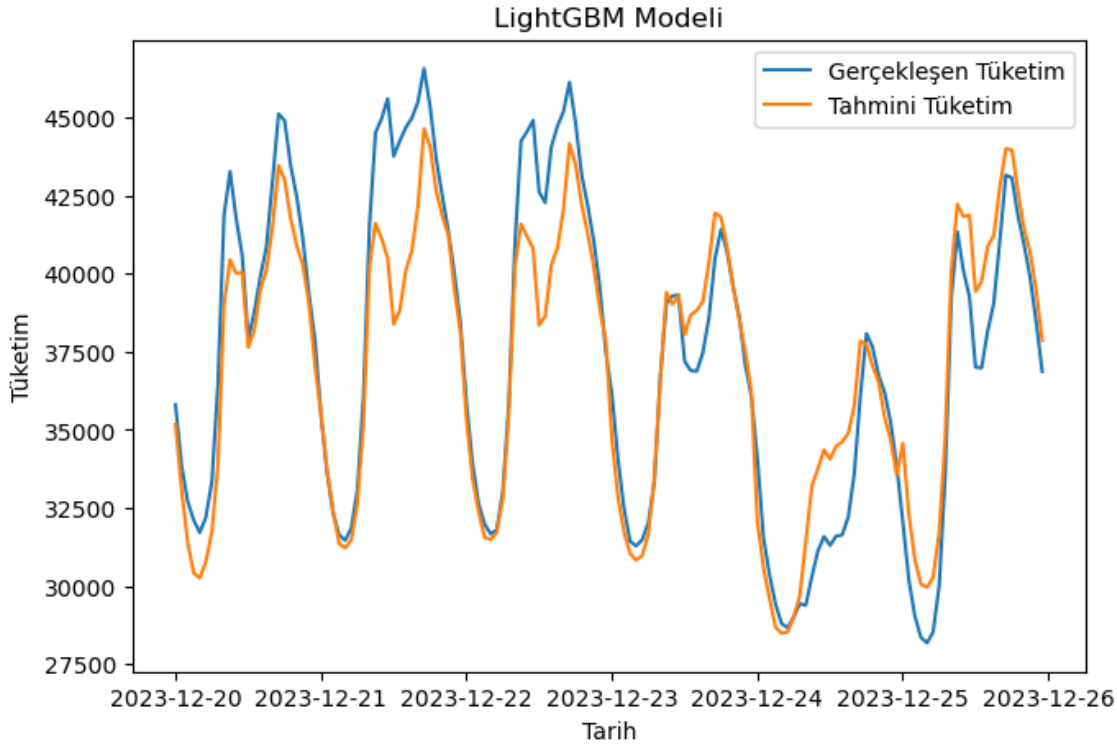
Catboost modeli için tahmini tüketim ve gerçekleşen tüketim değerlerine aşağıdaki grafikte yer verilmiştir.



Şekil 10. CatBoost için Alınan Verilerin Tarih ve Saate Göre Grafiği

LightGBM için hiper parametre optimizasyonu sonucunda elde edilen en iyi parametre değerleri aşağıdaki gibi bulundu:

- max_depth: 4
- learning_rate: 0.054704568231202136
- n_estimators: 301
- min_child_weight: 7
- subsample: 0.3107617564621609
- colsample_bytree: 0.4681018762702122
- reg_alpha: 0.08876966199593905
- reg_lambda: 0.00019395989045328666



Şekil 11. LightGBM için Alınan Verilerin Tarih ve Saate Göre Grafiği

4.2. Tartışma

Epiaş'dan çekilen gerçek zamanlı elektrik tüketim verilerinin ilgili yöntemlerle 20-12-2023 ile 25-12-2023 tarihleri arasındaki elektrik tüketim tahmini değerleri Python kullanılarak öznitelik mühendisliği, hiper parametrik optimizasyon ve algoritma modelleri ile elde edilmiştir. Bu değerlerden MAPE sonuçları çıkarılmıştır. Bunula beraber ortalama, medyan, standart sapma değerleri hesaplanmıştır. Hesaplanan bu değerler sonucunda en iyi model belirlenmiştir.

Literatürde yapılan çalışmalar incelendiğinde, Fatih Özcan'ın 2022 yılında yapmış olduğu "Derin Yapay Sinir Ağları İle Elektrik Tüketim Tahmin" adlı çalışmada çok katmanlı algılayıcı (MLP), elman tekrarlayan sinir ağı (ERNN), uzun kısa süreli bellek (LSTM), yankı durumu ağı (ESN), kapılı tekrarlayan birimler (GRU), kodlayıcı-kod çözücü istiflenmiş uzun kısa süreli bellek (Encoder-Decoder Stacked LSTM), evrişimli sinir ağı (CNN) ve zamansal evrişimli sinir ağı (TCN) modellerini kullandığı görülmektedir. İki çalışmada da ortak kullanılan LSTM modeli karşılaştırıldığında Fatih

Özcan'ın kullanmış olduğu en iyi tahmin performansı parametreleri 4 katmanlı, 128 birimli yapıda elde edilmiştir. Bu çalışmada ise deneme yanılma yoluyla 2 katmanlı, 25 birimli yapıda elde edilmiştir. LSTM için MAPE değerleri kıyaslandığında Fatih Özcan'ın çalışmasında MAPE değerleri grafik üzerinden gösterildiği için doğrudan kıyaslama yapılamamaktadır. Ancak girdi parametreleri denenip sonuçlar kıyaslanabilir.

Başka bir çalışma incelendiğinde, Duygu Altunkaya'nın 2021 yılında yapmış olduğu "Kısa Dönem Elektrik Tüketim Tahmini İçin 2D-CNN Ve Derin Öğrenme Algoritmaları Kullanılarak Hibrit Bir Model Geliştirilmesi" adlı çalışmayla bu çalışmada ortak kullanılan LSTM modeli karşılaştırıldığında, Duygu Altunkaya'nın kullanmış olduğu en başarılı sonucu veren parametreler sırasıyla, öğrenme oranı 0.001, batch boyutu 32, epoch sayısı 30, tanh aktivasyon fonksiyonu ve Adam optimizasyon algoritması kullanılmıştır. Bu çalışmada ise en başarılı sonucu veren parametreler sırasıyla, öğrenme oranı 0.1, batch boyutu 365*24, epoch sayısı 415, tanh aktivasyon fonksiyonu ve Adam optimizasyon algoritması kullanılmıştır. Bu iki çalışma farklı tarihler için tahminde bulunduğu için doğrudan bir kıyaslama yapılamamaktadır. Duygu Altunkaya çalışmasında RNN ve GRU modellerin basit yapısı ile hem daha güvenilir hem de daha doğru tahminler oluşturabilmek için ideal olduğunu belirtmiştir.

5. SONUÇ VE ÖNERİLER

Elektrik tüketimini tahmin etmek, işletmelerin enerji kullanımını artırması, maliyetleri düşürmesi, kaynakları daha iyi planlaması, sürdürülebilirliği teşvik etmesi ve rekabet avantajı sağlamasıyla öne çıkar. Aynı zamanda ülkelerin enerji güvenliğini artırırken, bilimsel araştırmalara veri sağlayarak enerji tüketimi ve iklim değişikliği konularında katkıda bulunur.

Bu çalışmada, Türkiye'de gerçek zamanlı elektrik tüketim değerlerini paylaşan EPİAŞ'dan 01-01-2018 ile 25-01-2023 tarihleri arasındaki saatlik veriler kullanılmıştır. Bu veriler üzerinde öznitelik mühendisliği uygulanmış, PTF her gün saat 14:00'da bir gün sonrası için açıklandığından tahmin edilecek saat aralığı 48 saat olarak belirlenmiş, bayesyen optimizasyonu uygulanmıştır. Kullanılan modeller LightGBM, XGBoost, CatBoost, KNN ve LSTM seçilmiştir. 18-12-2023 ile 19-12-2023 özelinde saatlik veriler test verisi olarak hiper parametrik optimizasyona dâhil edilmiştir. Her bir model için en iyi parametre değerleri elde edilmiştir. En iyi parametre değerleri ile modellerden tahmin sonuçlar elde edilmiştir.

Tablo 6. MAPE Değerleri

Date	LightGBM	XGBoost	CatBoost	KNN	LSTM
2023-12-20	3,82%	3,43%	5,28%	2,76%	2,65%
2023-12-21	4,76%	4,34%	6,00%	3,99%	3,57%
2023-12-22	3,12%	3,25%	3,69%	3,09%	3,28%
2023-12-23	3,69%	4,26%	4,45%	3,17%	2,46%
2023-12-24	4,41%	4,67%	5,68%	4,61%	3,81%

Yukarıdaki tabloda tahmin sonucunda elde edilen ağırlıklı olarak en iyi MAPE değerini LSTM vermiştir.

Tablo 7. Ortalama, Medyan ve Standart Sapma Değerleri

Betimsel İstatistik	**LightGBM**	**XGBoost**	**CatBoost**	**KNN**	**LSTM**
ORTALAMA	3,96%	3,99%	5,02%	3,52%	3,15%
MEDYAN	3,82%	4,26%	5,28%	3,17%	3,28%
STANDART SAPMA	0,64%	0,62%	0,94%	0,76%	0,58%

Yukarıdaki tabloda tahmin sonucunda elde edilen ortalama en iyi model LSTM olduğu görülmüştür. Medyan değeri olarak en iyi KNN fakat LSTM'in standart sapma değeri düşük olduğu için daha güvenilir modeldir. Bu çalışma sonucunda LSTM modelin kullanılması önerilir.

6. KAYNAKÇA

Alanbar, M., Alfarraj, A., ve Alghieth, M. (2020). Energy Consumption Prediction Using Deep Learning Technique. *International Journal of Interactive Mobile Technologies (iJIM)*, 166-177.

Altunkaya, D. (2021). *Kısa dönem elektrik tüketim tahmini için 2D-CNN ve derin öğrenme algoritmaları kullanılarak hibrit bir model geliştirilmesi.* (Yayın no. 692286) [Yüksek Lisans, Konya Gıda ve Tarım Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Amasyali, K., ve El-Gohary, N. (2017). Deep Learning for Building Energy Consumption Prediction. *6th CSCE/CRC International Construction Specialty Conference.* Vancouver: University of Illinois at Urbana-Champaign.

Anaconda. (t.y.). Anaconda Hakkında. 25 Ekim 2023 tarihinde https://www.anaconda.com adresinden alındı.

Ateş, V. (2019). *Türkiye'nin Kısa Dönemli Saatlik Bazda Elektrik Tüketiminin Yapay Zeka Teknikleri Kullanılarak Tahmin Edilmesi.* (Yayın no. 548629) [Doktora Tezi, Kırıkkale Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Aydın, B. D. (2016). *Meteorolojiik Değişkenlerin Elektrik Tüketimi Üzerindeki Etkisi.* (Yayın no. 450916) [Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Aydınlı, M. (2018). *Soğutma Sistemlerinde Elektrik Tüketim Tahmini: Yapay Sinir ağlari Kullanarak Model Kurma.* (Yayın no: 508896). [Yüksek Lisans Tezi, Necmettin Erbakan Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Bayar, E. (2019). *Diferansiyel Polinom Sinir ağı Tekniği İle Elektrik Tüketim Tahmini.* (Yayın no. 535603) [Yüksek Lisans Tezi, Sakarya Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Berksun, D. (2019). *Electricity Consumption and Economic Growth in Turkey: an Mf-var Approach.* (Yayın no. 599409) [Yüksek Lisans,İhsan Doğramacı Bilkent Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Berriel, R. F., Lopes, A. T., Rodrigues, A., Varejao, F. M., ve Oliveira-Santos, T. (2017). Monthly energy consumption forecast: A deep learning approach. *In 2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 4283-4290). IEEE.

Cho, J.-Y. K.-B. (2019). *Electric Energy Consumption Prediction by Deep Learning with State Explainable Autoencoder.* Seul: MDPI.

Choi, H.-J., ve Lee, Y.-J. (2020). Forecasting Building Electricity Power Consumption Using Deep Learning Approach. *Conference: 2020 IEEE International Conference on Big Data and Smart Computing (BigComp).*

Dincer, K. (2022). *Kisa Vadeli (Saatlik) Elektrik Tüketim Tahminlerinde Kullanilmak Üzere Tahmin Parametrelerinin Tespiti.* (Yayın no. 713012) [Yüksek Lisans, Karadeniz Teknik Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Emine Elif Nebati, M. T. (2021). *Türkiye'de Elektrik Tüketiminde Talep Tahmini: Zaman Serisi Ve.* Avrupa Bilim ve Teknoloji Dergisi.

EPİAŞ. (t.y.). EPİAŞ Hakkında. 15 Kasım 2023 tarihinde https://seffaflik.epias.com.tr/about/about adresinden alındı.

Özkan, E., Güler, E., Aladağ, Z, (2020). Elektrik Enerjisi Tüketim Verileri İçin Uygun Tahmin Yöntemi Seçimi. *Kocaeli: Journal of Industrial Engineering.*

Hajjar, K. (2022). *Forecasting of Turkey's Total Electricity Consumption in Sectoral Bases Using Machine Learning Algorithms.* (Yayın no. 734905) [Yüksek Lisans,

İstanbul Kültür Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Halaç, M. A. (2019). *Short Term Electricity Consumption Prediction With Neural Networks.* (Yayın no. 586521) [Yüksek Lisans, Bahçeşehir Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

İnan, D. (2019). *Ekonomik Büyüme, Diş Ticaret Ve Elektrik Tüketimi Arasindaki İlişki: Türkiye Örneği.* (Yayın no. 624037) [Yüksek Lisans Tezi, Dicle Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Karabulut, M. A. (2017). *Elektrik Tüketim Tahmini İçin Bir Karar Destek Sistemi.* (Yayın no. 476750) [Yüksek Lisans Tezi, Gazi Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Kaytez, F. (2012). *En Küçük Kareler Destek Vektör Makineleri ile Türkiye'nin Uzun Dönem Elektrik Tüketim Tahmini Ve Modellemesi.* (Yayın no. 321563). [Doktora Tezi, Gazi Üniversitesi]: https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Kim, J.-Y., ve Cho, S.-B. (2019). *Electric Energy Consumption Prediction by Deep Learning with State Explainable Autoencoder.* Seoul : Yonsei University.

Koçak, İ. (2018). *Türkiye'deki İllerin Elektrik Tüketimlerinin Veri Zarflama Analizi İle Değerlendirilmesi.* (Yayın no. 510819) [Yüksek Lisans Tezi, Gazi Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Li, C., Ding, Z., Zhao, D., Yi, J., ve Zhang, G. (2017). Building Energy Consumption Prediction An Extreme Deep Learning Approach. *Energies*.

Liu, Y., ve Li, J. (2022). Annual Electricity and Energy Consumption Forecasting for the UK Based on Back Propagation Neural Network, Multiple Linear Regression, and Least Square Support Vector Machine. *Processes*, 1-14.

Marmara, S. (2020). *Enerji Sektöründe Elektrik Tüketim Tahmini Ve Maliyet İyileştirmesi Uygulamasi.* (Yayın no. 629721) [Yüksek Lisans, Kocaeli Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Masaebi, P. (2016). *Yapay Sinir ağlari İle İran Elektrik Tüketim Tahmini.* (Yayın no. 435045) [Yüksek Lisans Tezi, Karadeniz Teknik Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Meng, Z., Sun, H., ve Wang, X. (2022). Forecasting Energy Consumption Based on SVR and Markov Model: A Case Study of China. *Frontiers in Environmental Science*.

Mhlanga, D. (2023). Artificial Intelligence and Machine Learning for Energy Consumption and Production in Emerging Markets: A Review. *Energies*.

Mosavi, A., ve Bahman, A. (2019). *Energy consumption prediction using machine learning; a review.* 18 Kasım 2023 tarihinde https://www.academia.edu/ adresinden alındı.

Movva, S. V. (2022). *Electrıcal Energy Consumptıon Predıctıon For Schools.* Luleå University of Technology. 10 Kasım 2023 tarihinde https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1720589&dswid=4177 adresinden alındı.

Öksüz, S. (2019). *Yapay Sinir ağlari İle Denizli İli Mesken Aboneleri Elektrik Tüketim Tahmini.* (Yayın no. 623513) [Yüksek Lisans Tezi, Pamukkale Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Özcan, F. (2022). *Derin Yapay Sinir ağlari İle Elektrik Tüketim Tahmini.* (Yayın no. 766665) [Yüksek Lisans, İzmir Demokrasi Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Özlü, G. (2018). *Yapay Ari Koloni̇si Algoritmasi İle Kayseri Ve Civari Elektrik Tüketim Verilerinin Analizi.* (Yayın no. 529061) [Yüksek Lisans Tezi, Erciyes Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Python. (t.y.). Python Hakkında. 15 Ekim 2023 tarihinde https://www.python.org/about adresinden alındı.

Reddy, G. V., Aitha, L. J., Poojitha, C., Shreya, A. N., Reddy, D. K., ve Meghana, G. S. (2023). Electrıcıty Consumptıon Predıctıon. *E3S Web of Conferences*.

Saoud, L. S., Marzouqi, H. A., ve Hussein, R. (2022). Household Energy Consumption Prediction Using the Stationary Wavelet Transform and Transformers. *IEEE Transactions on Industrial Informatics*.

Sezgin, B. (2018). *Determinants of Electricity Consumption in Turkey.* (Yayın no. 494251) [Yüksek Lisans Tezi, Anadolu Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Shaikh, A. K., Nazir, A., Khan, I., ve Shah, A. S. (2022). *Short term energy consumption forecasting using neural basis expansion analysis for interpretable time series.* Scientific Reports.

Shapi, M. K., Ramli, N. A., ve Awalin, L. J. (2021). Energy consumption prediction by using machine learning for smart building: Case study in Malaysia. *Developments in the Built Environment*, Volume 5.

Şahinaslan, E. (2023). İş Süreci Optimizasyonu: Yöntem, Teknoloji, Riskler ve Fırsatlar. *Akademik İzdüşüm Dergisi*(8(2)), 507-604. 18 Kasım 2023 tarihinde https://dergipark.org.tr/tr/pub/beuiibfaid adresinden alındı.

Şahinaslan, E., Günerkan, M., ve Şahinaslan, Ö. (2023). Makine Öğrenmesinde Kategorik Veri Kodlama Tekniğinin Kullanımına Alternatif Bir Çözüm Yöntemi. *Journal of Intelligent Systems: Theory and Applications, 6(1),* 1-11. doi:https://doi.org/10.38016/jista.1140

Şahinaslan, Ö., Dalyan, H., ve Şahinaslan, E. (2022). Naive Bayes Sınıflandırıcısı Kullanılarak YouTube Verileri Üzerinden Çok Dilli Duygu Analizi. . *Bilişim Teknolojileri Dergisi*,(15(2)), 221-229. doi: https://doi.org/10.17671/gazibtd.999960

Tokgöz, A. (2018). *Recurrent Neural Network Based Approaches for Electricity Consumption Forecasting.* (Yayın no. 511665) [Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Uğurel, G. (2013). *Türkiye'deki Elektrik Tüketiminin Modellemesi Ve Optimizasyonu.* (Yayın no. 341318) [Yüksek Lisans Tezi, İstanbul Ticaret Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Wei, N., Li, C., Pengc, X., Zeng, F., ve Lu, X. (2019). Conventional models and artificial intelligence-based models for energy consumption forecasting: A review. *Journal of Petroleum Science and Engineering*.

Yıldırım, M. (2021). *Bazi Veri Madenciliği Yöntemleri İle Elektrik Tüketim.* (Yayın no. 683657) [Yüksek Lisans, Harran Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Yıldız, S. (2022). *Sembolik Regresyon Temelli Oznitelik Cıkarımı ile Yuksek Mertebeden Sinir Agı Tasarımı: Elektrik Tuketim Tahmini.* (Yayın no. 766132) [Yüksek Lisans, İzmir Demokrasi Üniversitesi]. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp.

Yoo, S. G., ve Hernández-Álvarez, M. (2018). Predicting residential electricity consumption using neural networks: A case study. *Journal of Physics: Conference Series.* Seoul: IOP Publishing.

Zhang, J., Zhang, H., Ding, S., ve Zhang, X. (2021). Power Consumption Predicting and Anomaly Detection Based on Transformer and K-Means. *Frontiers in Energy Research*.

ÖZGEÇMİŞ

Fatih GÜLER

Eğitim

Y.Ls.	2022 - 2024	Maltepe Üniversitesi, Lisansüstü Eğitim Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Bilgisayar Mühendisliği (Tezli)
Ls.	2012-2017	Düzce Üniversitesi, Bilgisayar Mühendisliği

İş/İstihdam

2021-Halen	Kazancı Holding – Yazılım Geliştirme Takım Lideri
2018-2021	Maltepe Üniversitesi – Kıdemli Yazılım Uzmanı