



T.C.
NECMETTİN ERBAKAN
ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



YAPAY SİNİR AĞLARI YÖNTEMİYLE
KARAMAN BÖLGESİ DOĞAL GAZ
TÜKETİM ÖNGÖRÜSÜ

Mahmut Esat TERZİ

YÜKSEK LİSANS TEZİ

Makine Mühendisliği Anabilim Dalı

Ocak-2026
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

Mahmut Esat TERZİ tarafından hazırlanan “YAPAY SİNİR AĞLARI YÖNTEMİYLE KARAMAN BÖLGESİ DOĞAL GAZ TÜKETİM ÖNGÖRÜSÜ” adlı tez çalışması 28/01/2026 tarihinde aşağıdaki jüri tarafından oy birliği ile Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü Makine Mühendisliği Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Unvanı Adı SOYADI

Danışman

Unvanı Adı SOYADI

Üye

Unvanı Adı SOYADI

İmza

Prof. Dr. Ahmet Ali SERTKAYA

Prof. Dr. İlhan ASİLTÜRK

Prof. Dr. Sabri KOÇER

Fen Bilimleri Enstitüsü Yönetim Kurulu’nun/.../2026 gün ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Havvanur UÇBEYİAY
FBE Müdürü

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Mahmut Esat TERZİ

Tarih:

ÖZET

YÜKSEK LİSANS TEZİ

YAPAY SİNİR AĞLARI YÖNTEMİYLE KARAMAN BÖLGESİ DOĞAL GAZ TÜKETİM ÖNGÖRÜSÜ

Mahmut Esat TERZİ

**NECMETTİN ERBAKAN ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
MAKİNE MÜHENDİSLİĞİ ANABİLİM DALI**

Danışman: Prof. Dr. İlhan ASİLTÜRK

2026, 68 Sayfa

Jüri

Prof. Dr. Ahmet Ali SERTKAYA

Prof. Dr. İlhan ASİLTÜRK

Prof. Dr. Sabri KOÇER

Ülkemizde kullanmakta olduğumuz doğal gazın büyük bir kısmı dış ülkelerden ithal edilmektedir. Bununla birlikte son yıllarda doğal gaz kullanımının yaygınlaşması ve doğal gazın kullanım alanları çoğalması ile birlikte depolanan doğal gaz miktarı da bu talebi karşılamada yetersiz kalmaktadır. Bu durum göz önünde bulundurulduğunda; doğal gaz dağıtım şebekesi arz güvenliğinin sağlanabilmesi, doğal gaz sektöründe yapılması gereken yatırımların planlanabilmesi, depolanması ve tedarik edilmesi gereken doğal gaz miktarlarının ön görülebilir olması kritik öneme sahiptir.

Bu çalışmada, Yapay Sinir Ağları yöntemlerinden biri olan LSTM modeli kullanarak Karaman iline ait doğal gaz tüketim tahmini yapılmıştır. Yapay Sinir Ağları, veri setleri üzerinden öğrenme gerçekleştirerek karmaşık olaylara karşı tepki mekanizmaları geliştirebilen sayısal sistemlerdir. Sistem, girdi ve çıktı değişkenleri arasındaki ilişkileri eğitim süreciyle kavrayarak, daha önce karşılaşılmamış veri setleri için isabetli öngörülerde bulunabilmektedir.

Çalışmada, Karaman iline ait 2015-2024 yılları arasındaki günlük veriler kullanılarak günlük doğal gaz tüketim tahminleri gerçekleştirilmiştir. Çalışma kapsamında meteorolojik veriler, takvim etkileri ve abone sayısı gibi değişkenler modele entegre edilmiş; geçmiş tüketim verileri ve hareketli ortalamalar ile modelin öğrenme kapasitesi artırılmıştır. Geliştirilen model, test veri seti üzerinde $R^2 = 0,9924$ başarı katsayısı ve ortalama mutlak yüzde hata (MAPE) %8,59 değerleri ile yüksek doğruluk oranı sergilemiştir. Bu sonuçlar, modelin gelecekteki tüketim trendlerini analiz etme ve farklı senaryolar altında doğru öngörülerde bulunma kapasitesini vurgulamaktadır.

Anahtar Kelimeler: Doğal Gaz Tüketim Tahmini, LSTM, Zaman Serisi Analizi.

ABSTRACT

MS THESIS

FORECASTING OF NATURAL GAS CONSUMPTION BY ARTIFICIAL NEURAL NETWORKS IN KARAMAN REGION

Mahmut Esat TERZİ

**THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF
NECMETTİN ERBAKAN UNIVERSITY THE DEGREE OF MASTER OF
SCIENCE
IN MECHANICAL ENGINEERING**

Advisor: Prof. Dr. İlhan ASİLTÜRK

2026, 68 Pages

Jury

Prof. Dr. Ahmet Ali SERTKAYA

Prof. Dr. İlhan ASİLTÜRK

Prof. Dr. Sabri KOÇER

A significant portion of the natural gas consumed in our country is imported from abroad. However, in recent years, the widespread adoption of natural gas and the expansion of its usage areas have led to a situation in which the amount of stored natural gas has become insufficient to meet this growing demand. Considering this situation, ensuring the security of supply in the natural gas distribution network, planning the necessary investments in the natural gas sector, and accurately forecasting the quantities of natural gas that must be stored and procured are of critical importance.

In this study, natural gas consumption forecasting for the province of Karaman was conducted using the Long Short-Term Memory (LSTM) model, which is one of the Artificial Neural Network (ANN) methods. Artificial Neural Networks are numerical systems capable of learning from datasets and developing response mechanisms to complex phenomena. Through the training process, the system learns the relationships between input and output variables and can generate accurate predictions for previously unseen datasets.

The study utilized daily data from 2015 to 2024 for the province of Karaman to produce daily natural gas consumption forecasts. Meteorological variables, calendar effects, and subscriber counts were utilized into the model, while past consumption data and moving averages were incorporated to enhance the model's learning capacity. The developed model demonstrated high accuracy on the test dataset, achieving a coefficient of determination of $R^2 = 0.9924$ and a mean absolute percentage error (MAPE) of 8.59%. These results highlight the model's capability to analyze future consumption trends and to provide reliable forecasts under different scenarios.

Keywords: Natural Gas Consumption Forecasting, LSTM, Time Series Analysis.

ÖNSÖZ

Tez çalışmamın her aşamasında desteğini ve rehberliğini esirgemeyen, tecrübeleriyle çalışmama değer katan ve akademik gelişimime büyük katkı sağlayan tez danışmanım Prof. Dr. İlhan ASİLTÜRK'e teşekkürlerimi sunarım.

Çalışmanın veri toplama sürecindeki değerli yardımları ve bilgi paylaşımlarıyla akademik sürecime katkı sağlayan Enerya Karaman Gaz Dağıtım A.Ş.'ye ve tüm Enerya ailesine desteği için teşekkür ederim.

Tez çalışmam süresince gösterdiği sabır ve desteğiyle her daim yanımda olan eşim Özlem TERZİ'ye; varlıklarıyla en büyük moral ve motivasyon kaynağım olan kızım Zeynep Ece'ye ve oğlum Mert Efe'ye teşekkür ederim. Ayrıca hayatımın her aşamasında olduğu gibi bu yoğun süreçte de desteklerini ve güvenlerini daima hissettiğim anneme, babama ve kardeşlerime şükranlarımı sunarım.

Mahmut Esat TERZİ
KONYA-2026

İÇİNDEKİLER

1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	4
2.1. Doğal Gaz	4
2.1.1. Doğal gazın tarihi.....	4
2.1.2. Doğal gazın çıkarılması ve kullanıma hazırlanması	5
2.1.3. Doğal gazın kullanım alanları.....	5
2.1.4. Doğal gazın depolanması.....	6
2.1.5. Sıvılaştırılmış doğal gaz.....	6
2.1.6 Sıkıştırılmış doğal gaz	7
2.2. Enerji Talebi Tahmin Yöntemleri.....	7
2.2.1. Geleneksel istatistiksel yöntemler	7
2.2.1.1. Regresyon analizi.....	8
2.2.1.2. ARIMA (Autoregressive integrated moving average) modelleri	8
2.2.1.3. SARIMA (Seasonal ARIMA) modelleri	8
2.2.1.4. Ekonometrik modeller	9
2.2.2. Yapay zeka tabanlı yöntemler.....	9
2.2.2.1. Yapay sinir ağları.....	9
2.2.2.2. Destek sinir ağları (RNN)	10
2.2.2.3. Uzun kısa süreli bellek ağları (LSTM)	10
2.2.3. Literatürde enerji talep tahmin uygulamaları.....	10
3. MATERYAL VE YÖNTEM.....	13
3.1. Programlama Dili Olarak Python	13
3.1.1. Temel veri işleme kütüphaneleri.....	13
3.1.1.1. Pandas	13
3.1.1.2. NumPy	14
3.1.2. Görselleştirme kütüphaneleri.....	14
3.1.2.1. Matplotlib.....	14
3.1.2.2. Seaborn	15
3.1.3. Derin öğrenme kütüphaneleri	15
3.1.3.1. TensorFlow	15
3.1.3.2. Keras	15
3.1.4 Makine öğrenmesi ve istatistik kütüphaneleri	16
3.1.4.1. Scikit-learn.....	16
3.1.4.2. SciPy	16
3.1.5. Performans optimizasyonu ve hesaplama verimliliği	16
3.2. Doğal Gaz Tüketim Tahmini	17
3.2.1. Çalışmanın kapsamı ve veri seti	17
3.2.2. Veri kalitesi ve eksik değerler	18
3.2.3. Günlük doğal gaz tüketimi zaman serisi analizi (2015-2024)	19
3.2.4. Sıcaklık-Tüketim ilişkisi ve mevsimsel bağlam	20
3.2.5. Aylık ortalama tüketim ısı haritası analizi	21
3.2.6. Öznitelik mühendisliği.....	22
3.2.6.1. Lag özellikleri (gecikmeli değerler)	22
3.2.6.2. Hareketli ortalama özellikleri	24
3.2.6.3. Trigonometrik dönüşümler (döngüsel özellikler)	26

3.2.6.4. Hissedilen sıcaklık	27
3.2.6.5. Isıtma derece gün (HDD).....	27
3.2.6.6. Tatil günü indikatörü	29
3.2.6.7. Kümülatif abone sayısı	30
3.2.6.8. Döngüsel zaman özellikleri	31
3.2.6.9. Verilerin korelasyon grafiği.....	33
3.2.7. Veri ön işleme.....	35
3.2.7.1. Eksik değerlerin temizlenmesi.....	35
3.2.7.2. Veri setinin bölünmesi (eğitim, doğrulama, test)	36
3.2.7.3. Özellik seçimi	36
3.2.8. Normalizasyon (Ölçeklendirme).....	37
3.2.8.1. Normalizasyon hesaplama	37
3.2.9. LSTM model mimarisi ve eğitim süreci	38
3.2.9.1. Mimari olarak LSTM seçilme nedenleri.....	38
3.2.9.2. Model mimarisi.....	39
3.2.9.3. Model performansının değerlendirilmesi.....	45
3.2.9.4. Hata analizi	51
3.3. Gelecek Tahminleri.....	54
3.3.1. Tahmin metodolojisi ve iteratif mekanizma	54
3.3.2. Abone sayısı projeksiyonu ve senaryo tasarımı.....	55
3.3.3. Meteorolojik veri tahmini	56
3.3.4. Tahmin sonuçları	57
3.3.4.1. Yıllık tüketim tahminleri	57
3.3.4.2. Senaryo karşılaştırması	58
4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA.....	60
4.1. Veri Seti Özellikleri ve Ön İşleme.....	60
4.1.1. Özellik mühendisliği sonuçları	60
4.1.1.1. Lag özellikleri	60
4.1.1.2. Hareketli ortalamalar	61
4.1.1.3. Zaman ve mevsimsel özellikler	61
4.1.1.4. Zaman ve mevsimsel özellikler	61
4.2. LSTM model mimarisi ve hiperparametre seçimi	62
4.2.1. Veri seti bölümlendirme ve normalizasyon	63
4.2.2. Model eğitim süreci ve yakınsama analizi.....	63
4.2.3. Model performans değerlendirmesi	63
4.3. Gerçek ve Tahmin Karşılaştırması	64
4.4. Abone Artışı Senaryo Analizi.....	64
4.5. Gelecek tüketim tahminleri (2025-2030).....	64
5. SONUÇLAR VE ÖNERİLER.....	66
5.1. Enerji Yönetimi İçin Pratik Öneriler	67
5.2. Gelecek Araştırma Önerileri	67

ŞEKİLLER LİSTESİ

Şekil 2.1. İşlem görmemiş doğal gazın temel elemanları	4
Şekil 3.1. Serbest ve taşıma abonelerinin toplam tüketimi (Sm3).....	18
Şekil 3.2. Karaman ili 2015-2024 dönemine ait günlük doğal gaz tüketimi	19
Şekil 3.3. Sıcaklık ve tüketim ilişkisi	20
Şekil 3.4. Lag 1 (1 günlük gecikme) korelasyon analizi	23
Şekil 3.5. Lag 7 (7 günlük gecikme) korelasyon analizi	24
Şekil 3.6. 7 günlük ortalama sıcaklık ve tüketim ilişkisi.....	25
Şekil 3.7. Isıtma derece gün ve doğal gaz tüketimi ilişkisi	28
Şekil 3.8. Hafta içi ve hafta sonu doğal gaz kullanım dağılımı.....	30
Şekil 3.9. Kümülatif abone sayısı artış grafiği (2015-2024)	31
Şekil 3.10. Mevsimsel ortalama tüketim	32
Şekil 3.11. Aylık ortalama tüketim grafiği	33
Şekil 3.12. Tüm değişkenler arası korelasyon matrisi.....	34
Şekil 3.13. LSTM akış şeması	41
Şekil 3.14. Model kayıp grafiği	48
Şekil 3.15. Model MAE-Eğitim süreç grafiği	49
Şekil 3.16. İlk 100 gün tahmin ve gerçek karşılaştırması.....	50
Şekil 3.17. Tahmin doğruluğu serpilme grafiği.....	51
Şekil 3.18. Artık (Residual) plot – hata dağılımı grafiği	52
Şekil 3.19. Hata dağılımı histogramı grafiği	53
Şekil 3.20. 3 farklı senaryoya göre abone sayısı projeksiyonu	56
Şekil 3.21. Geçmiş ve gelecek doğal gaz tüketim grafiği	57
Şekil 3.22. Geçmiş ve tüm senaryolar gelecek doğal gaz tüketim trendi.....	58
Şekil 3.23. Farklı senaryolar için yıllık toplam tüketim grafiği (2025-2030)	59

ÇİZELGELER LİSTESİ

Çizelge 3.1. Yıllık ve aylık ortalama tüketim ısı haritası	21
Çizelge 3.2. Dropout oranı seçimi	43
Çizelge 3.3. LSTM model mimari özet tablo	44
Çizelge 3.4. Modelin performans değerlendirme metrikleri	47
Çizelge 3.5. 2030 yılı senaryo karşılaştırması	56
Çizelge 3.6. Senaryo bazında yıllık toplam tüketim tahminleri	57
Çizelge 4.1. Korelasyon analizi sonuçları	61



SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

SİMGELER

C₂H₆	Etan
CH₄	Metan
CO₂	Karbondiyoksit
H₂S	Hidrojen Sülfür
He	Helyum
N₂	Azot
Sm³	Standart Metre Küp

KISALTMALAR

API	Application Programming Interface (Uygulama Programlama Arayüzü)
ARIMA	Autoregressive Integrated Moving Average
CNG	Sıkıştırılmış Doğal Gaz
CNN-LSTM	Uzun Kısa Süreli Bellek Ağı
CPU	Central Processing Unit (Merkezi İşlem Birimleri)
GPU	Graphics Processing Unit (Grafik İşlem Birimleri)
DVM	Destek Vektör Makineleri
HDD	Isıtma Derece Gün
LNG	Sıvılaştırılmış Doğal Gaz
LSTM	Long Short-Term Memory (Uzun Kısa Süreli Bellek)
MAE	Ortalama Mutlak Hata
MAPE	Ortalama Mutlak Yüzde Hata
MLP	Çok Katmanlı Algılayıcı
RMS-A	A Tipi Basınç Düşürme ve Ölçüm İstasyonu
RMSE	Kök Ortalama Kare Hata
RNN	Tekrarlayan Sinir Ağları
SARIMA	Seasonal Autoregressive Integrated Moving Average
YSA	Yapay Sinir Ağları

1. GİRİŞ

Enerji kaynaklarının etkin kullanımı ve gelecek tüketim projeksiyonlarının doğru tahmin edilmesi kritik bir öneme sahiptir (Suganthi ve Samuel, 2012). Doğal gaz, fosil yakıtlar arasında görece düşük karbon emisyonuna sahip olup kömür ve petrol ürünlerine kıyasla daha az karbon emisyonu üretmektedir (U.S. Energy Information Administration, 2024).

Türkiye’de doğal gaz kullanımı, özellikle konut ve sanayi sektörlerinde son yıllarda istikrarlı bir artış göstermiştir. Bu artış, doğal gaz altyapısının ülke genelinde yaygınlaştırılmasını ve abone sayısında kayda değer bir büyümeyi beraberinde getirmiştir. Enerji Piyasası Düzenleme Kurumu’nun (EPDK) yayımladığı raporlara göre, 2000’li yılların başından itibaren doğal gaz dağıtım şebekesi önemli ölçüde genişletilmiş, 2024 yılı itibarıyla doğal gaz abone sayısı 20 milyonu aşmıştır (Enerji Piyasası Düzenleme Kurumu, 2025).

Enerji talebinin doğru tahmin edilmesi, enerji yönetiminin temel unsuru olarak mevcut kaynakların etkin dağılımını sağlamaktadır (Suganthi ve Samuel, 2012). Sözen vd. (2005), Türkiye’nin net enerji tüketimini analiz ettikleri çalışmalarında, gelecek enerji tüketim seviyelerinin belirlenmesinin önemini vurgulamışlardır. Benzer şekilde, Söder ve Amelin (2011), enerji sistemlerinin etkin işletimi ve planlamasında üretim-tüketim dengesinin kritik öneme sahip olduğunu belirtmişlerdir.

Enerji talebinin tahmininde geleneksel yöntemler, çoğunlukla istatistiksel modellere dayanmaktadır (Suganthi ve Samuel, 2012). Bu yöntemler arasında lineer regresyon, hareketli ortalamalar, üstel düzeltme ve özellikle zaman serisi analizine dayalı ARIMA (Autoregressive Integrated Moving Average) modelleri öne çıkmaktadır (Suganthi ve Samuel, 2012; Box ve Jenkins, 1976; Gujarati ve Porter, 2009). Box ve Jenkins (1976), ARIMA modellerinin teorik temellerini atarak bu yöntemi zaman serisi tahminlerinde yaygınlaştırmışlardır. Ayrıca, Hyndman ve Athanasopoulos (2018), geleneksel tahmin yöntemlerinin enerji talebi gibi ekonomik göstergelerin öngörülmesinde nasıl uygulandığını kapsamlı biçimde ele almışlardır. Bu yöntemler, özellikle verilerin doğrusal olduğu ve zaman serilerinin durağan olduğu varsayımlar altında etkili sonuçlar verebilmektedir (Gujarati ve Porter, 2009; Box ve Jenkins, 1976).

Ancak geleneksel tahmin yöntemleri, enerji tüketimini etkileyen çok sayıda değişken arasındaki karmaşık ve doğrusal olmayan ilişkileri modellemede yetersiz kalabilmektedir. Özellikle meteorolojik koşullar, tüketim kalıpları ve sosyoekonomik

faktörler gibi çok boyutlu girdiler, bu yöntemlerin sınırlılıklarını daha belirgin hale getirmektedir (Soldo, 2012; Suganthi ve Samuel, 2012; Hyndman ve Athanasopoulos, 2018).

Soldo (2012), doğal gaz tüketiminin tahmininde geleneksel modellerin bu tür karmaşık ilişkileri yeterince yansıtamadığını ve doğrusal olmayan yapıları yakalamakta zorlandığını belirtmektedir. Benzer şekilde, Suganthi ve Samuel (2012), enerji talep tahmininde kullanılan modelleri inceledikleri çalışmalarında, geleneksel yöntemlerin çok değişkenli sistemlerde sınırlı performans gösterdiğini vurgulamışlardır. Bu nedenle, daha esnek ve veri odaklı yaklaşımların geliştirilmesi gerekliliği ortaya çıkmaktadır.

Son yıllarda yapay zekâ ve makine öğrenmesi temelli yöntemler, enerji talep tahmini alanında yaygın olarak kullanılmaya başlanmıştır. Ahmad vd. (2020), akıllı şebekeler ve binalar için yenilenebilir enerji ve elektrik talep tahminleme modellerini kapsamlı bir şekilde inceledikleri çalışmalarında, makine öğrenmesi ve yapay sinir ağı modellerinin kısa, orta ve uzun dönem tahminler için en iyi performansı gösterdiğini belirtmişlerdir.

Derin öğrenme teknolojilerindeki hızlı gelişim, zaman serisi tahminleme problemlerinde önemli başarılar sağlamıştır (LeCun vd., 2015; Goodfellow vd., 2016). Hochreiter ve Schmidhuber (1997), LSTM (Long Short-Term Memory-Uzun Kısa Süreli Bellek) mimarisini tanıttıkları çalışmalarında, klasik tekrarlayan sinir ağlarının yaşadığı "vanishing gradient" (kaybolan gradyan) sorununu aşarak uzun dönemli bağımlılıkları öğrenmede önemli bir ilerleme sağlamışlardır. LSTM ağları, uzun vadeli zamansal bağımlılıkları yakalama kapasitesiyle öne çıkmakta ve enerji talep tahmini gibi zaman serisi uygulamalarında kullanılmaktadır. Kong vd. (2019), bireysel konutların kısa vadeli elektrik yükü tahmininde LSTM tabanlı yaklaşımın diğer yöntemlere kıyasla üstün performans gösterdiğini ortaya koymuştur. Benzer şekilde Kim ve Cho (2019), konut enerji tüketimini tahmin etmek amacıyla geliştirdikleri CNN-LSTM hibrit modelinde, CNN katmanının enerji tüketimini etkileyen çoklu değişkenler arasındaki uzamsal özellikleri çıkardığını ve LSTM katmanının zaman serisi bileşenlerindeki düzensiz eğilimlerin zamansal bilgisini modellediğini göstermiştir. Önerilen model, bireysel konut enerji tüketimi tahmininde neredeyse mükemmele yakın performans elde etmiştir.

Bu bulgular, LSTM'nin enerji sistemleri gibi dinamik ve çok değişkenli alanlarda güçlü bir modelleme aracı olduğunu ortaya koymaktadır.

Araştırmanın amacı

Bu çalışmanın temel amacı, geçmiş doğal gaz tüketim verilerini, abone sayılarını ve meteorolojik faktörleri kullanarak gelecek dönem tüketimlerini yüksek doğrulukla tahmin edebilen bir model geliştirmektir. Çalışmada, zaman serisi verilerinin karmaşık ilişkileri öğrenme kapasitesiyle öne çıkan (Hochreiter ve Schmidhuber, 1997), LSTM tabanlı bir derin öğrenme modeli kullanılmıştır.

Modelin tahmin yeteneği, çeşitli ve zenginleştirilmiş özellik setlerinden yararlanılarak artırılmıştır. Bu özellikler arasında;

Tarihsel Tüketim Verileri: Geçmiş doğal gaz tüketim miktarları, modelin temel öğrenme kaynağını oluşturur.

Meteorolojik Koşullar: Sıcaklık (maksimum, ortalama ve minimum), nem, rüzgâr hızı ve günlük toplam yağış miktarı gibi hava durumu değişkenleri. Panek vd. (2022), makine öğrenmesi algoritmaları kullanarak dışsal meteorolojik faktörlerin doğal gaz tüketimi tahminindeki rolünü incelemiş ve bu değişkenlerin tüketim üzerinde etkili olduğunu ortaya koymuştur.

Abone Sayısı: Bölgedeki abone sayısı, uzun vadeli tüketim trendlerinin anlaşılmasına katkı sağlar.

Takvim Bilgileri: Ay, mevsim, haftanın günü, hafta sonu, resmî tatil ve tatil indikatörleri gibi zamansal özellikler, periyodik tüketim davranışlarını yakalamak için kullanılmıştır. Takvim bilgilerinin doğal gaz talebindeki periyodik değişimleri modellemede etkili olduğu gösterilmiştir (Brabec vd., 2009).

Türetilmiş Özellikler: Modelin geçmişe dönük bağımlılıkları daha iyi öğrenmesi amacıyla, hissedilen sıcaklık (Steadman, 1979), ısıtma derece gün (Thom, 1954), gecikmeli (lag) tüketim ve sıcaklık değerleri ile hareketli ortalamalar (Hyndman ve Athanasopoulos, 2018) gibi mühendislik ürünü özellikler oluşturulmuştur.

Araştırmanın önemi

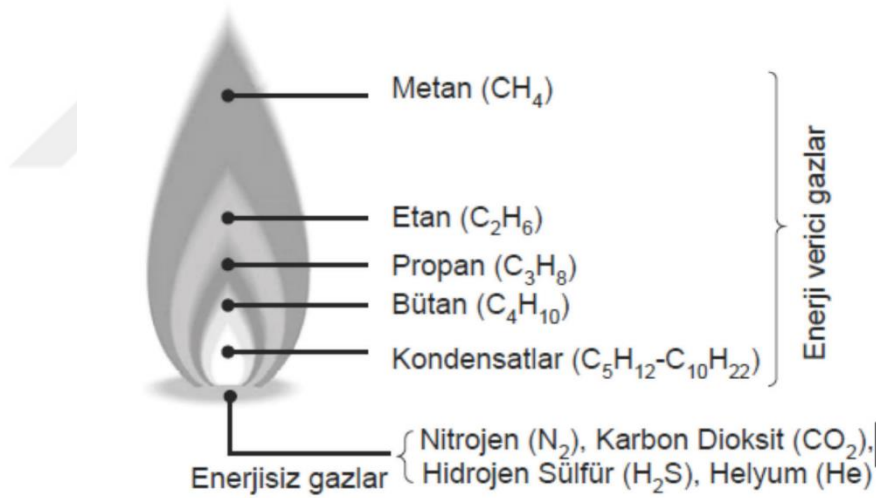
Kapsamlı özellik seti ve LSTM modelinin yüksek öğrenme kapasitesi sayesinde, enerji sektöründeki karar vericiler için gelecekteki doğal gaz tüketim eğilimlerine ilişkin güvenilir ve eyleme geçirilebilir öngörüler sunulması hedeflenmektedir (Wang vd., 2025; Kong vd., 2019). Elde edilen tahminler, enerji ağlarının iyileştirmeleri, depolama kapasitelerinin etkin kullanımı ve olasılıklı kriz senaryolarına karşı hazırlıklı olunması açısından büyük değer taşımaktadır (Soldo, 2012).

2. KAYNAK ARAŞTIRMASI

Enerji talep tahmini, uzun yıllardır araştırmacıların ilgisini çeken bir konudur (Feinberg ve Genethliou, 2005; Suganthi ve Samuel, 2012). Bu bölümde, doğal gaz hakkında genel bilgiler, doğal gaz tüketim tahmini, zaman serisi analizi yöntemleri, yapay zeka ve derin öğrenme uygulamaları ile özellik mühendisliği konularında yapılan öncü çalışmalar incelenmiştir.

2.1. Doğal Gaz

Doğal gaz, temel bileşeni metan (CH_4) olmak üzere etan (C_2H_6) gibi diğer hidrokarbonların yanı sıra karbondioksit (CO_2), azot (N_2), hidrojen sülfür (H_2S) ve helyum (He) gibi gazlardan oluşmaktadır. Yer kabuğundaki jeolojik süreçler sonucunda meydana gelen doğal gaz, kokusuz, renksiz, yanıcı ve havadan hafif bir özellik taşımaktadır (UGETAM, 2016; Speight, 2019).



Şekil 2.1. İşlem görmemiş doğal gazın temel elemanları (Ugetam, 2016)

2.1.1. Doğal gazın tarihi

Literatürde doğal gazın kökenine ilişkin farklı teoriler bulunmaktadır. Yaygın kabul gören görüş ise milyonlarca yıl önce yaşamış bitki ve hayvan kalıntılarının yer kabuğu altında yüksek basınç ve sıcaklık koşullarında biyokimyasal dönüşümlere uğramasıdır (Zumerchik, 2001; Tissot ve Welte, 1984). Bu süreçte organik maddeler diagenetik ve katagenetik evrelerden geçerek karmaşık hidrokarbon yapısına sahip,

yüksek enerji yoğunluklu doğal gazı oluşturur (Tissot ve Welte, 1984). Gazın türü ve bileşimi ise kaynak kayasının organik içeriği, gömülme derinliği ve sıcaklık gibi jeolojik faktörlere bağlı olarak farklılık göstermektedir (Tissot ve Welte, 1984; Schwarzbauer ve Jovančićević, 2024).

2.1.2. Doğal gazın çıkarılması ve kullanıma hazırlanması

Doğal gaz, yer altındaki gözenekli rezervuar kayaçlarda bulunur. Rezervlerin keşfi jeolojik saha çalışmaları ve sismik araştırmalarla başlamakta olup ardından belirlenen noktalarda sondaj faaliyetleri yürütülmektedir. Kuyunun üretim kapasitesi test edildikten sonra gaz, üretim hatları ve ayırıştırma tesisleri aracılığıyla yüzeye çıkarılmaktadır. Ham doğal gaz doğrudan kullanılamadığından, çeşitli proseslerden geçirilerek saflaştırılmakta ve kullanıma hazır hale getirilmektedir (Ugetam, 2016; Speight, 2019; Mokhatab vd., 2015):

- CO₂'in ayırıştırılması,
- Suyun giderilmesi,
- Ağır hidrokarbonların ayırıştırılması,
- Azotun giderilmesi,
- Gazın temizlenmesi,
- Kurutma.

2.1.3. Doğal gazın kullanım alanları

Doğal gaz, diğer fosil yakıtlara kıyasla daha ekonomik işletme koşulları, düşük bakım maliyetleri ve çevre dostu özellikleri sayesinde enerji talebinde giderek artan bir paya sahiptir (IEA, 2022). Uluslararası Enerji Ajansı'nın *World Energy Outlook 2022* raporunda ise doğal gazın enerji geçişinde kömürden daha düşük emisyonlu bir köprü yakıt olarak kritik rol oynadığı; konut ve sanayi sektörlerinde esnekliği, düşük emisyonları ve mevcut altyapı uyumluluğu sayesinde önemini koruduğu belirtilmektedir. Bu avantajlar, doğal gazın hem konut ısıtmasında hem de endüstriyel üretimde öncelikli enerji kaynağı olarak konumlanmasına yol açmaktadır.

Doğal gaz, konutlarda ısınma, sıcak su temini ve pişirme amacıyla yaygın biçimde kullanılmakta olup bu kullanım hem aile bütçesine katkı sağlamakta hem de hava kirliliğinin azaltılmasına önemli ölçüde yardımcı olmaktadır (UGETAM, 2016).

Sanayi sektöründe ise doğal gaz, ısınma ve çeşitli proseslerde geniş bir kullanım alanına sahiptir (UGETAM, 2016). Atölye ve fırınlardan demir-çelik, çimento, kimya ve tekstil sektörlerine; tuğla, seramik ve cam üretiminden buhar üretimine kadar pek çok alanda tercih edilmektedir (Speight, 2019). TÜİK (2024) verilerine göre sanayide toplam enerji tüketiminin yaklaşık %23'ü doğal gazdan karşılanmaktadır. Özellikle yüksek sıcaklık gerektiren sektörlerde doğal gazın, temiz yanma özelliği ve proses kontrolündeki esnekliği nedeniyle öne çıktığı belirtilmektedir (Mokhatab vd., 2015). Ayrıca elektrik üretiminde de önemli bir rol oynayan doğal gaz, ülkemizin enerji ihtiyacının kayda değer bir kısmının doğal gaz santrallerinden karşılanmasını sağlamaktadır (UGETAM, 2016).

2.1.4. Doğal gazın depolanması

Doğal gazın dünya genelinde geniş kullanım alanlarına ulaşması, arz güvenliğini sağlamak amacıyla depolama faaliyetlerini zorunlu hale getirmiştir (IEA, 2022). Pamir ve Türkyılmaz (2007), doğal gaz depolamanın temel gerekçelerini şu başlıklar altında özetlemektedir:

- Dağıtım şebekesinin verimliliğini artırmak,
- Tüketimdeki mevsimsel ve kısa vadeli dalgalanmaları dengelemek,
- Stratejik miktarda gaz bulundurmak,
- Arz kesintilerinde depolardan gaz sağlamak.

Doğal gaz depolama alanları, jeolojik özellikleri ve kapasitelerine göre farklı türlerde sınıflandırılmaktadır (Katz ve Lee, 1990). MMO Enerji Çalışma Grubu (2008), bu alanları;

- İşlevini yitirmiş doğal gaz ve petrol yatakları,
- Yeraltındaki kaya tuzu yatakları,
- Çelik tanklar veya basınçlı çelik depolar,
- Doğal boşluklar ve izolasyon sağlayan terk edilmiş maden yatakları olarak tanımlamaktadır.

Uluslararası literatürde de benzer sınıflandırmalar yapılmakta olup, depolamanın enerji arz güvenliği açısından kritik bir rol oynadığı vurgulanmaktadır (IEA, 2015).

2.1.5. Sıvılaştırılmış doğal gaz

Sıvılaştırılmış doğal gaz (LNG), atmosferik basınç altında yaklaşık -161,5°C ile -164°C arasındaki sıcaklıklara kadar soğutularak sıvı forma dönüştürülmektedir. Bu

dönüşüm sonucunda gazın hacmi yaklaşık 600 kat azalmakta, böylece boru hatlarıyla iletimin teknik veya ekonomik olarak mümkün olmadığı bölgelere yüksek miktarlarda enerji taşınması ve depolanması sağlanmaktadır. LNG tesislerinde sıvılaştırılan gaz, kullanım öncesinde buharlaştırılarak tekrar gaz fazına geçirilmekte ve mevzuat gereği kokulandırıcı eklenerek tüketiciye sunulmaktadır (UGETAM, 2016; EPDK, 2014).

2.1.6 Sıkıştırılmış doğal gaz

Sıkıştırılmış doğal gaz (CNG), yaklaşık 200–250 bar basınç altında hacmi küçültülerek elde edilen doğal gaz formudur. LNG'den farklı olarak CNG, depolama ve taşıma süreçlerinde gaz fazını korumakta ve yüksek basınç dayanımına sahip çelik veya kompozit tüplerle taşınmaktadır. Boru hattı ile iletimin ekonomik olmadığı düşük ölçekli tüketim noktalarında doğal gaz arzı için etkili bir çözüm sunan CNG, kullanım noktasındaki basınç düşürme istasyonlarında ideal şebeke basıncına indirilerek tüketiciye ulaştırılmaktadır (UGETAM, 2016; EPDK, 2014).

2.2. Enerji Talebi Tahmin Yöntemleri

Enerji talebi tahmini, enerji sektöründe stratejik planlama, arz güvenliği ve sistem yönetimi açısından kritik bir öneme sahiptir (Suganthi ve Samuel, 2012; Soldo, 2012). Doğru tahminler; üretim kapasitesinin optimize edilmesi, dağıtım altyapısının teknik sınırlar içinde boyutlandırılması ve maliyetlerin etkin biçimde kontrol edilmesi için hayati rol oynamaktadır (Suganthi ve Samuel, 2012). Sosyal, ekonomik ve meteorolojik faktörlere bağlı olarak karmaşık ve doğrusal olmayan bir yapıya sahip olan enerji talebi, tahmin süreçlerinde kullanılan yöntemlerin çeşitlenmesini ve daha gelişmiş analitik modellerin geliştirilmesini zorunlu kılmaktadır (Hippert vd., 2001; Weron, 2006; Hong ve Fan, 2016; Soldo, 2012)

Enerji talebi tahmin yöntemleri geleneksel istatistiksel yöntemler ve yapay zeka tabanlı yöntemler olmak üzere iki ana kategoride ele alınmıştır.

2.2.1. Geleneksel istatistiksel yöntemler

Enerji talebi tahmininde uzun bir geçmişe sahip olan geleneksel istatistiksel yöntemler, güçlü teorik temelleriyle öne çıkmaktadır (Box ve Jenkins, 1976; Suganthi ve Samuel, 2012). Bu yöntemler, tarihsel verilerdeki eğilimleri, mevsimsellik etkilerini ve

değişkenler arası korelasyonları matematiksel modeller aracılığıyla tanımlayarak güvenilir projeksiyonlar üretmektedir (Weron, 2006; Soldo, 2012).

2.2.1.1. Regresyon analizi

Regresyon analizi, bağımlı değişken olan enerji talebi ile sıcaklık, nüfus ve ekonomik göstergeler gibi bir veya daha fazla bağımsız değişken arasındaki ilişkiyi matematiksel fonksiyonlar aracılığıyla modelleyen güçlü bir yöntemdir (Montgomery vd., 2021).

Literatürde regresyon modellerinin özellikle makroekonomik değişkenlere dayalı uzun dönem tahminlerinde etkili olduğu vurgulanmaktadır (Suganthi ve Samuel, 2012). González-Romera vd. (2008), elektrik talebi tahmininde yapay sinir ağları gibi daha gelişmiş yöntemlerin geleneksel yaklaşımlara alternatif olabileceğini göstermişlerdir.

2.2.1.2. ARIMA (Autoregressive integrated moving average) modelleri

Box ve Jenkins (1976) tarafından geliştirilen ARIMA modelleri, zaman serisi tahmininde en çok başvurulan klasik yöntemlerden biridir. Bu modeller, doğrusal veri örüntülerinin analizinde yüksek başarı göstermektedir. Türkiye enerji piyasası üzerine yapılan çalışmalarda da kullanılmıştır. Örneğin, Ediger ve Akar (2007), Türkiye'nin birincil enerji talebini 2020 yılına kadar modellemek amacıyla ARIMA yöntemini uygulamıştır.

2.2.1.3. SARIMA (Seasonal ARIMA) modelleri

Mevsimsel bileşenlerin önemli olduğu enerji talep serilerinde, ARIMA'nın periyodik dalgalanmaları ele alma kabiliyetini geliştiren SARIMA modelleri öne çıkmaktadır (Hyndman ve Athanasopoulos, 2018).

SARIMA modelleri, enerji talebinden turizme kadar çeşitli sektörlerde mevsimsel tahminler için kullanılmaktadır. Bozkurt vd. (2022), Belirli bir dönem için Türkiye'ye gelmesi beklenen turist sayısını tahmin etmek için Arima ve SARIMA modelleri uygulamışlardır ve en iyi uyumu gösteren modelin SARIMA modeli olduğu sonucunu ortaya koymuşlardır.

2.2.1.4. Ekonometrik modeller

Ekonometrik modeller, iktisadi teorilere dayalı olarak değişkenler arasındaki ilişkileri matematiksel denklemlerle formüle eden analitik yaklaşımlardır. Ekonometrinin temel amacı, teorik iktisadi ilişkileri nicel modellerle ifade ederek test etmek ve tahminlerde bulunmaktır (Gujarati & Porter, 2009).

Geleneksel istatistiksel yöntemlerin avantajları arasında teorik sağlamlık, yorumlanabilirlik ve az sayıda parametre gerektirmeleri yer almaktadır (James vd., 2013; Hastie vd., 2009). Ancak bu yöntemler, doğrusal olmayan ilişkileri modellemede yetersiz kalmakta (Breiman, 2001), yüksek boyutlu veri ile başa çıkmakta zorlanmakta (Fan ve Li, 2001) ve karmaşık örüntüleri yakalama konusunda sınırlı bir performans sergilemektedir (Friedman, 2001). Bu nedenle enerji talebi gibi çok boyutlu ve doğrusal olmayan yapıya sahip serilerde, daha esnek ve güçlü tahmin yöntemlerine yönelim artmaktadır (Suganthi & Samuel, 2012).

2.2.2. Yapay zeka tabanlı yöntemler

Son yıllarda bilgisayar teknolojilerindeki gelişmeler ve büyük veri analitiği kapasitesinin artmasıyla birlikte, yapay zeka tabanlı yöntemler enerji talebi tahmininde giderek daha yaygın kullanılmaktadır (Mosavi vd., 2019; Ahmed vd., 2023). Bu yöntemler, geleneksel istatistiksel modellere kıyasla doğrusal olmayan ilişkileri modelleyebilme, karmaşık örüntüleri öğrenebilme ve daha yüksek doğrulukla tahmin yapabilme yetenekleriyle öne çıkmaktadır (Qiu vd., 2017; Ryu vd., 2017; Ahmad vd., 2021).

2.2.2.1. Yapay sinir ağları

Yapay sinir ağları, insan beyninin öğrenme mekanizmalarından esinlenerek geliştirilen, çok katmanlı ve paralel işlem yapabilen yapay öğrenme sistemleridir. Doğrusal olmayan ilişkileri öğrenme ve genelleme yetenekleri sayesinde enerji talebi gibi karmaşık sistemlerin modellenmesinde etkili bir araç olarak kullanılmaktadır (Haykin, 1999; Khotanzad vd., 2000).

Hippert vd. (2001), kısa vadeli elektrik yükü tahmininde YSA'nın performansını incelemiş ve geleneksel yöntemlere kıyasla daha yüksek doğruluk sunduğunu ortaya koymuştur. Ayrıca Kalogirou (2000), YSA'nın enerji sistemlerinde talep tahmini, sistem optimizasyonu ve kontrol uygulamalarında başarılı sonuçlar verdiğini belirtmiştir.

2.2.2.2. Destek sinir ağıları (RNN)

Zaman serisi verilerindeki ardışık bağımlılıkları modellemek amacıyla geliştirilen RNN (Tekrarlayan Sinir Ağları), klasik ileri beslemeli ağlardan farklı olarak döngüsel bir yapıya sahiptir (Goodfellow vd., 2016). Bu yapı sayesinde geçmiş zaman adımlarındaki bilgileri içsel hafızada tutarak mevcut tahmin sürecine entegre edebilmektedir (Hochreiter ve Schmidhuber, 1997). Enerji talebi gibi zamana bağlı değişkenlik gösteren verilerde RNN modelleri önemli avantajlar sunmaktadır (Qiu vd., 2017).

2.2.2.3. Uzun kısa süreli bellek ağıları (LSTM)

RNN'nin uzun süreli bağımlılıkları öğrenirken karşılaştığı kaybolan gradyan sorununu çözmek amacıyla geliştirilen LSTM ağıları, Hochreiter ve Schmidhuber (1997) tarafından literatüre kazandırılmıştır. LSTM, *giriş*, *unutma* ve *çıkış kapıları* gibi özel mekanizmalar sayesinde bilginin uzun süreli olarak korunmasını, güncellenmesini veya silinmesini sağlamaktadır. Bu özellik, doğal gaz ve elektrik talebi gibi geçmiş tüketim alışkanlıklarının gelecekteki talep üzerinde etkili olduğu durumlarda tahmin yapılmasını mümkün kılmaktadır (Kong vd., 2019).

2.2.3. Literatürde enerji talep tahmin uygulamaları

Suganthi ve Samuel (2012), enerji talep tahmini literatürünü kapsamlı bir şekilde ele alarak; geleneksel ekonometrik modellerden modern yapay zekâ tekniklerine kadar uzanan metodolojileri analiz etmişlerdir. Yazarlar, tek bir modelin tüm senaryolar için optimal olmadığını ve veri karakteristiğine göre model seçiminin önemini vurgulamışlardır.

Akın (2017), Konya ilinin günlük doğal gaz tüketimini tahmin etmek amacıyla yürüttüğü çalışmada, yapay sinir ağlarının karmaşık ve doğrusal olmayan ilişkileri modelleme yeteneğinden yararlanılmıştır. Araştırmada; abone sayısını temsil eden bağımsız birim sayısı (BBS), geniş kapsamlı meteorolojik veriler, türetilmiş iklimsel parametreler (derece-gün, windchill vb.) ve doğal gaz birim fiyatı dâhil olmak üzere toplam 17 farklı girdi değişkeni kullanılmıştır. 2005-2015 yıllarını kapsayan 11 yıllık veri setinin kullanıldığı çalışmada, YSA modellerinin başarısı regresyon modelleriyle kıyaslanmış; sonuçlar R^2 , rRMS ve rMB gibi çoklu performans kriterleri üzerinden Konya iline en uygun modeli bulunmaya çalışılmıştır. Yapay sinir ağları modelinin

özellikle kış aylarında oldukça başarılı olduğu, ancak mevsim geçişlerinde sapmaların arttığı saptanmıştır.

Sözen vd. (2005), Türkiye'nin net enerji tüketimini tahmin etmek amacıyla YSA yönteminin kullanıldığı bir çalışmada, sosyo-ekonomik ve teknik veriler kullanılarak iki farklı tahmin modeli geliştirilmiştir. 1975-2003 dönemini kapsayan verilerle eğitilen modellerde, nüfus ve kurulu kapasite gibi değişkenlerin yanı sıra kaynak bazlı girişlerin de yüksek doğruluk sunduğu saptanmıştır. Her iki modelin de eğitim ve test aşamalarında R^2 değerlerinin 0,99'un üzerinde olması, YSA tekniğinin enerji politikaları geliştirme sürecinde hızlı ve güvenilir bir araç olarak kullanılabileceğini göstermektedir.

Taşkır (2018), Ankara ilindeki doğal gaz tüketimini tahmin etmek amacıyla yaptığı çalışmada, yapay sinir ağları modelini meteorolojik tahmin verileri ve operasyonel değişkenlerle kurgulamıştır. Araştırmada, ölçülen sıcaklıklar yerine ileriye dönük planlamaya imkân tanıyan tahmin edilen meteorolojik veriler, bağımsız birim sayısı (BBS), ısı değer ve tüketim değerleri gibi sektörel parametreler girdi olarak kullanılmıştır. 2014-2016 verileriyle eğitilen ve 2017 verileriyle test edilen modelde, özellikle hissedilen sıcaklık değişkeninin etkisi ve farklı aktivasyon fonksiyonlarının performans üzerindeki verimliliği analiz edilmiştir.

Evcim (2019), doğal gaz tüketim tahmininde kullanılan farklı yapay sinir ağı mimarilerinin etkinliğini inceleyen çalışmasında, Sakarya ili üzerinden kapsamlı bir performans değerlendirmesi sunmaktadır. Araştırmada basit YSA ve derin sinir ağlarının yanı sıra, ardışık verileri işleme yeteneğine sahip RNN ve bunların gelişmiş türevleri olan LSTM ve GRU modelleri test edilmiştir. Yedi günlük geçmiş verilerin ve meteorolojik parametrelerin modele dâhil edilmesiyle, derin öğrenme temelli yaklaşımların klasik SARIMA modeline göre tahmin başarısını anlamlı derecede artırdığı gözlemlenmiştir.

Kaynar vd. (2011), Türkiye'nin doğal gaz ithalatındaki dışa bağımlılığını temel olarak, yatırım planlama ve gaz tedarik süreçleri için kısa vadeli talep tahmini modellerini incelemiştir. YSA ve ARIMA modellerinin performanslarının gerçek verilerle kıyaslandığı çalışmada, doğru tahminlerin makroekonomik etkileri ve sektörel sürdürülebilirlik üzerindeki kritik önemi üzerinde durulmuştur.

Rehman vd., (2022), doğal gaz talep tahmininde model optimizasyonunun önemini vurgulayan araştırmalarında, İspanya ölçeğinde farklı lokasyon verilerini kullanarak SMLR, MLP (Çok Katmanlı Algılayıcı) ve LSTM algoritmalarını test etmiştir. Bulgular, istatistiksel yöntemlerin (SMLR) veri eksikliği durumunda performans kaybı yaşadığını ve hatalı uç değerler ürettiğini göstermişlerdir. Öte yandan, derin öğrenme

tabanlı LSTM modellerinin, karmaşık veri setlerinde daha dengeli sonuçlar verdiği ve yakın gelecek tahminlerinde yüksek performans sergilediği saptanmıştır. Karar vericiler için bu yöntemlerin ayırt edici özelliklerinin anlaşılmasının, enerji arz-talep yönetiminde kritik bir rol oynadığı vurgulanmaktadır.

Hippert vd. (2001), YSA'nın elektrik yükü tahminindeki uygulamalarını kapsamlı bir şekilde gözden geçirmiştir. 1991-1999 yılları arasında yayınlanan 100'den fazla çalışmanın analiz edildiği bu derlemede; YSA'nın kısa vadeli yük tahmininde en yaygın kullanılan yöntemlerden biri olduğu, özellikle MLP ve RBF mimarilerinin sıklıkla tercih edildiği vurgulanmıştır. Elektrik yükü tahminindeki bu başarı, benzer doğrusal olmayan dinamiklere sahip olan doğal gaz tüketim öngörüsü çalışmaları için de bir temel oluşturmuştur.

Ekonomou (2010), Yunanistan'ın uzun dönemli enerji talebini öngörmek amacıyla yürüttüğü çalışmasında, yapay sinir ağları mimarilerinden MLP modeli kullanmıştır. 2005-2015 yılları arasındaki verileri kapsayan çalışmada, geliştirilen modelin başarısı; doğrusal regresyon ve DVM (destek vektör makineleri) gibi alternatif yöntemlerle kıyaslanmıştır. Elde edilen bulgular, YSA tabanlı yaklaşımın gerçek tüketim verileriyle yüksek bir korelasyon sergilediğini ve uzun vadeli enerji politikalarının planlanmasında, sermaye yatırımlarının optimize edilmesinde ve arz güvenliğinin korunmasında güvenilir bir araç olduğunu ortaya koymuştur.

3. MATERYAL VE YÖNTEM

Bu çalışmada, Karaman Bölgesi'ne ait 2015-2024 yılları arasındaki günlük veriler kullanılarak, YSA yöntemi ve LSTM modeli ile günlük doğal gaz tüketim tahminlerinin gerçekleştirilmiştir. Çalışmada Python programlama dili kullanılmıştır.

3.1. Programlama Dili Olarak Python

Python, veri bilimi ve makine öğrenmesi alanlarında en yaygın kullanılan programlama dillerinden biridir (McKinney, 2012; Raschka, 2015). Bu çalışmada veri analizi, model geliştirme ve görselleştirme işlemleri Python (versiyon 3.12) programlama dili kullanılarak gerçekleştirilmiştir. Python geniş kütüphane ekosistemi, okunabilir sözdizimi ve güçlü topluluk desteği nedeniyle tercih edilmiştir (Python Software Foundation [PSF], 2023). Geliştirme ortamı olarak Jupyter Notebook kullanılmıştır (Kluyver vd., 2016).

3.1.1. Temel veri işleme kütüphaneleri

Bu çalışmada veri işleme bölümlerinde Pandas ve NumPy kütüphaneleri kullanılmıştır.

3.1.1.1. Pandas

Pandas, Python ekosisteminde veri manipülasyonu ve analizi için en yaygın kullanılan kütüphanelerden biridir (McKinney, 2012). Bu çalışmada tüm veri yükleme, temizleme, dönüştürme ve gruplama işlemleri Pandas'ın DataFrame yapısı üzerinden gerçekleştirilmiştir. Pandas'ın tercih edilme nedenleri arasında excel dosyalarını doğrudan okuyabilme yeteneği, tarih-zaman verilerini yönetmek için güçlü araçlar, eksik veri işleme fonksiyonları ve gruplama operasyonları yer almaktadır (Pandas Development Team, 2024).

Çalışmada Pandas'ın kullanıldığı spesifik işlemler şunlardır:

- Ham veri setinin excel formatından Python ortamına yüklenmesi, tarih sütununun *datetime* formatına dönüştürülmesi ve zamansal özelliklerin türetilmesi (ay, yıl, haftanın günü gibi),
- Gecikme (Lag) özellikleri ve hareketli ortalamaların hesaplanması için fonksiyonların kullanılması,
- Mevsimsel ve yıllık analizler için operasyonların uygulanması,

- Çıktı tablolarının csv ve excel formatlarında kaydedilmesi.

Pandas'ın iki boyutlu tablo yapısı (DataFrame), zaman serisi verilerinin düzenli bir biçimde organize edilmesini sağlamış; her satırın bir günü, her sütunun ise bir özelliği temsil ettiği yapı, analiz sürecinde önemli kolaylıklar sunmuştur (Pandas Development Team, 2024).

3.1.1.2. NumPy

NumPy (Numerical Python), bilimsel hesaplama için Python'daki temel kütüphanelerden biridir. Çok boyutlu diziler ve bu diziler üzerinde yüksek performanslı matematiksel işlemler gerçekleştirme kapasitesi sayesinde veri bilimi ve makine öğrenmesi uygulamalarında yaygın olarak kullanılmaktadır (Harris vd., 2020). Bu çalışmada NumPy, özellikle LSTM modelinin girdi-çıkı formatlarının oluşturulması, trigonometrik dönüşümler ve istatistiksel hesaplamalar için kullanılmıştır.

Çalışmada NumPy'nin kritik kullanım alanları şunlardır:

- Trigonometrik dönüşümlerin hesaplanması (aylık ve yıllık döngülerin modellenmesi),
- LSTM modeli için 3 boyutlu *tensor* yapısının oluşturulması ([örnekler, zaman adımı, özellikler] formatına dönüştürme),
- Hareketli ortalamaların hesaplanmasında gerekli fonksiyonun kullanılması,
- Rastgele sayı üretimi için ilgili modülün kullanılması (gelecek tahminlerde meteorolojik varyasyon eklemek için),
- İstatistiksel hesaplamalar.

NumPy'nin vektörize edilmiş işlemleri, Python döngülerine göre çok daha hızlı çalışarak, 3.653 günlük veri setinin işlenmesinde önemli performans kazancı sağlamıştır. Bu sayede veri işleme süreci hem daha verimli hem de daha okunabilir hale gelmiştir.

3.1.2. Görselleştirme kütüphaneleri

Bu çalışmada veri işleme bölümlerinde *Matplotlib* ve *Seaborn* isimli kütüphaneler kullanılmıştır.

3.1.2.1. Matplotlib

Matplotlib, Python'da bilimsel görselleştirme için en temel ve kapsamlı kütüphanelerden biridir (Hunter, 2007). Alt seviye çizim kontrolü sunarak yayın

kalitesinde grafikler oluşturma imkânı sağlamaktadır. Bu çalışmada zaman serisi grafikleri, saçılım grafikleri, sütun grafikleri ve çok panelli figürler oluşturmak için kullanılmıştır.

Çalışmada oluşturulan spesifik görseller şunlardır:

- Eğitim sürecinin loss ve MAE grafiklerinin çizilmesi,
- Gerçek ve tahmin değerlerinin karşılaştırması,
- Mevsimsel tüketim profilleri,
- Yıllık büyüme trendleri, senaryo karşılaştırmaları.

3.1.2.2. Seaborn

Seaborn, Matplotlib üzerine inşa edilmiş ve istatistiksel veri görselleştirme için geliştirilmiş bir kütüphanedir (Waskom, 2021). Estetik açıdan zengin grafikler oluşturma imkânı sunmaktadır. Bu çalışmada özellikle korelasyon ısı haritaları ve istatistiksel dağılım grafikleri için kullanılmıştır. Seaborn'un avantajları arasında hazır renk paletleri, istatistiksel grafik türleri, *Pandas DataFrame* ile doğrudan entegrasyon ve modern, minimalist tasarım felsefeleri bulunmaktadır.

3.1.3. Derin öğrenme kütüphaneleri

Bu çalışmada veri işleme bölümlerinde TensorFlow ve Keras isimli kütüphaneler kullanılmıştır.

3.1.3.1. TensorFlow

TensorFlow, Google tarafından geliştirilen, açık kaynaklı derin öğrenme platformudur (Abadi vd., 2016). Bu çalışmada TensorFlow, hesaplama altyapısı olarak kullanılmış; model geliştirme işlemleri ise Keras API'si üzerinden gerçekleştirilmiştir (Chollet, 2015). TensorFlow'un tercih edilmesinin nedenleri arasında endüstri standardı olması, geniş topluluk desteği, CPU ve GPU üzerinde optimize edilmiş performans sunması ve görselleştirme araçlarına sahip olması yer almaktadır.

3.1.3.2. Keras

Keras, yüksek seviye sinir ağı API'si olup TensorFlow 2.x sürümünden itibaren TensorFlow'un bir modülü olarak entegre edilmiştir. Modüler yapısı sayesinde hızlı prototipleme ve model geliştirme imkânı sunmaktadır. Bu çalışmada tüm model

geliştirme süreci Keras API'si kullanılarak gerçekleştirilmiştir. Keras'ın tercih edilme nedenleri arasında sezgisel ve okunabilir API yapısı, zengin katman ve aktivasyon fonksiyonu kütüphanesi, eğitim sürecini yönetmek için sunduğu mekanizmalar ve hızlı prototipleme imkânı yer almaktadır (Chollet, 2021).

3.1.4 Makine öğrenmesi ve istatistik kütüphaneleri

Bu çalışmada veri işleme bölümlerinde Scikit-learn ve SciPy isimli kütüphaneler kullanılmıştır.

3.1.4.1. Scikit-learn

Scikit-learn, Python'da makine öğrenmesi için en kapsamlı kütüphanelerden biridir (Pedregosa vd., 2011). Bu çalışmada Scikit-learn, veri ön işleme, değerlendirme metriklerinin hesaplanması ve bazı temel modellerin oluşturulması amacıyla kullanılmıştır.

3.1.4.2. SciPy

SciPy, bilimsel hesaplama için kapsamlı algoritmalar ve fonksiyonlar sağlayan bir Python kütüphanesidir (Virtanen vd., 2020). Bu çalışmada özellikle istatistiksel testlerin uygulanmasında kullanılmıştır. Bu testler, model performansının ve veri yapısının derinlemesine anlaşılmasını sağlamış, bulgulara istatistiksel güven kazandırmıştır.

3.1.5. Performans optimizasyonu ve hesaplama verimliliği

Bu çalışmada kullanılan kütüphanelerin seçiminde yalnızca işlevsellik değil, aynı zamanda performans ve hesaplama verimliliği de önemli bir kriter olmuştur. NumPy ve Pandas sayesinde 3.653 satırlık veri seti üzerindeki işlemler milisaniyeler mertebesinde tamamlanmıştır. Özellikle NumPy'nin vektörize edilmiş fonksiyonları, Python döngülerine göre çok daha hızlı çalışarak işlem süresini ciddi oranda azaltmıştır (Harris vd., 2020).

Vektörizasyon, bir veri kümesindeki her bir eleman için ayrı ayrı işlem yapmak yerine, tüm veriyi bir bütün olarak işleyerek paralel hesaplama yapılmasını sağlar (Harris vd., 2020). Bu prensip, tüm veri işleme adımlarında uygulanmış ve Python döngüleri minimize edilmiştir. Örneğin, gecikme (lag) özellikleri oluşturulurken ilgili fonksiyon kullanılmış, bu da manuel döngü yazmaya göre daha fazla hızlanma sağlamıştır. Benzer

şekilde, trigonometrik dönüşümler NumPy'nin vektörize fonksiyonları ile topluca hesaplanmış, her satır için ayrı hesaplama yapılmamıştır.

Kütüphanelerin entegrasyonu, veri yüklemekten model eğitime, performans değerlendirmeden görselleştirmeye kadar tüm iş akışının sorunsuz bir şekilde yürütülmesini sağlamıştır (McKinney, 2012; Harris vd., 2020; Virtanen vd., 2020; Chollet, 2021). Açık kaynak doğası, geniş dokümantasyon ve aktif topluluk desteği, problem çözme sürecini hızlandırmış ve kod kalitesini artırmıştır.

3.2. Doğal Gaz Tüketim Tahmini

Çalışmada, YSA ile Karaman iline ait günlük doğal gaz tüketim tahmini yapılmıştır. Doğal gaz tüketimine etki eden parametreler detaylıca incelenmiştir.

3.2.1. Çalışmanın kapsamı ve veri seti

Çalışmada kullanılan veri seti, 01.01.2015 – 31.12.2024 tarihleri arasında kapsayan 3.653 adet günlük gözlemden oluşmaktadır. Veri setinde eksik veri bulunmamaktadır. Veri setindeki orijinal (türetilmemiş, temin edilen) değişkenler şunlardır:

Tarih: Her bir kaydın gününü belirten tarih bilgisi.

Meteorolojik Değişkenler:

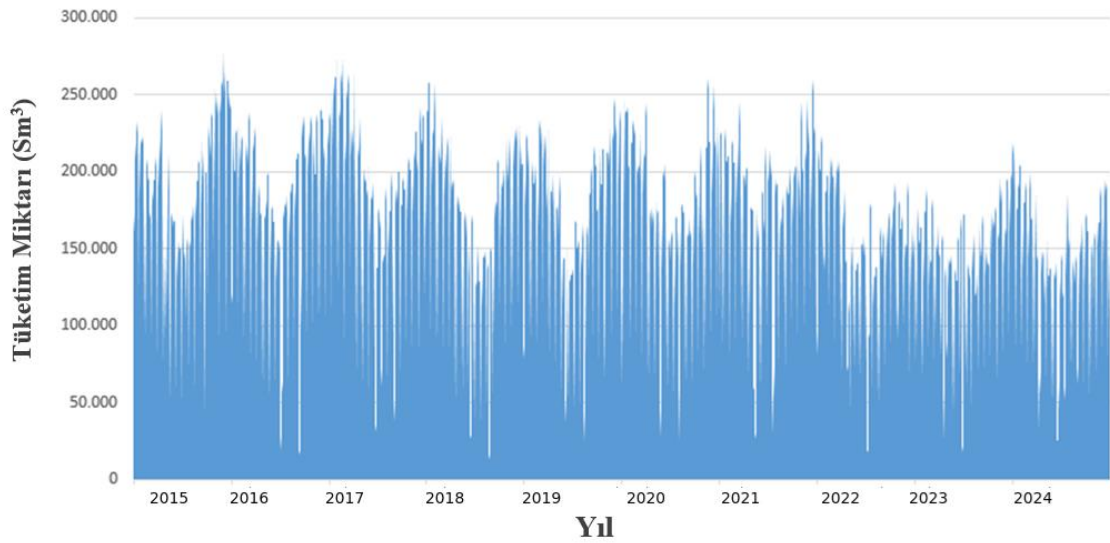
- **T_{max} :** Günlük Maksimum Sıcaklık: Gün içinde ölçülen en yüksek Sıcaklık ($^{\circ}C$).
- **T_{avg} :** Günlük Ortalama Sıcaklık: Günlük ortalama Sıcaklık ($^{\circ}C$).
- **T_{min} :** Günlük Minimum Sıcaklık: Gün içinde ölçülen en düşük sıcaklık ($^{\circ}C$).
- **ΔT :** Günlük Maksimum Minimum Sıcaklık Farkı: Günlük maksimum ve minimum Sıcaklık arasındaki fark ($^{\circ}C$).
- **R :** Nem: Günlük ortalama nem oranı (%).
- **P :** Günlük Toplam Yağış Miktarı: Günlük toplam yağış miktarı (mm).
- **u :** Rüzgâr Hızı: Günlük ortalama rüzgâr hızı (m/s).

Operasyonel Değişkenler:

- **N_{new} :** Günlük Alınan Abone Sayısı: O gün alınan yeni abone sayısı.
- **Q_t :** Günlük Doğal gaz Tüketim Miktarı: Günlük doğal gaz tüketim miktarı (m^3). (Bu bizim hedef değişkenimizdir.)

Bu veri setinde yer alan meteorolojik veriler Konya Meteoroloji 8. Bölge Müdürlüğü'nden temin edilmiş olup 17246 numaralı Karaman istasyonuna aittir.

Çalışmada kullanılan operasyonel veriler ise Enerya Karaman Gaz Dağıtım A.Ş'den temin edilmiştir. Günlük Doğal gaz tüketim miktarı Sm^3 cinsindendir. Çalışmaya sadece Karaman RMS-A istasyonundan geçen tüketim dâhil edilmiş olup CNG/LNG ile beslenen ilçeler/beldeler dâhil edilmemiştir. Diğer taraftan Dağıtım Şirketi'nden alınan geçmiş tüketim verileri; Serbest Tüketiciler, Taşıma Aboneleri ve Sanayi/Konut aboneleri olarak 3 farklı kategoride bulunmaktadır. Çalışmada Sanayi/Konut abonelerinin tüketimleri kullanılmış olup büyük sanayi işletmeleri (Serbest tüketiciler ve Taşıma aboneleri) dâhil edilmemiştir. Çalışmaya dâhil edilen tüketim, küçük ölçekli işletmeler, okullar, hastaneler, resmi daireler, alışveriş merkezleri vs. olarak düşünülebilir. Bunun nedeni Serbest Tüketici ve Taşıma Abonelerinin tüketimlerinin son 10 yılda ortalama bir tüketim çizgisinde ilerlemiş olması ve herhangi bir düzenli artış/azalış olmamasıdır.



Şekil 3.1. Serbest ve taşıma abonelerinin toplam tüketimi (Sm^3)

3.2.2. Veri kalitesi ve eksik değerler

Orijinal veri setinde eksik değer bulunmamaktadır. Tüm gözlemler eksiksiz ve tutarlıdır. Ancak aşağıdaki veri kalitesi kontrolleri yapılmıştır:

Aykırı Değer Tespiti: IQR (Interquartile Range) yöntemi ile aykırı değerler tespit edilmiştir. Tespit edilen değerlerin mevsimsel ve meteorolojik koşullarla açıklanabilir olduğu görülmüştür.

Tutarlılık Kontrolleri: Sıcaklık değerleri arasındaki ilişkiler ($\min < \text{ort} < \max$) kontrol edilmiş, tutarsızlık bulunmamıştır.

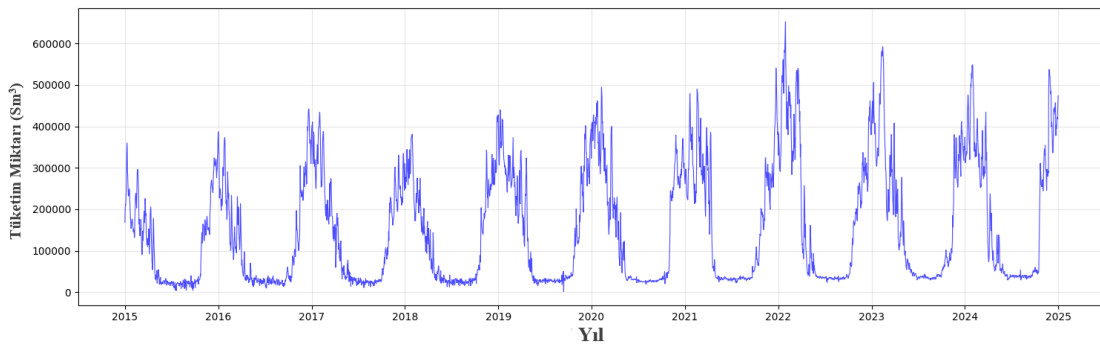
Zaman Sürekliliği: Tarih sütununda eksik gün olmadığı doğrulanmıştır.

Türetilen bazı önemli değişkenler şunlardır:

- Hissedilen Sıcaklık: Günlük ortalama hissedilen sıcaklık ($^{\circ}\text{C}$).
- Isıtma Derece Gün Sayısı: Günlük ısıtma derece gün sayısı (genellikle bir eşik sıcaklığın altında her derece için 1 puan artar, ısıtma ihtiyacını gösterir).
- Hafta Sonu: O günün hafta sonu olup olmadığını belirten ikili değişken (0: Hafta içi, 1: Hafta sonu).
- Resmî Tatil: O günün resmi tatil olup olmadığını belirten ikili değişken (0: Değil, 1: Evet).
- Tatil İndikatörü: Hafta sonu veya resmî tatil olması durumunda 1 olan gösterge değişkeni.
- Kümülatif Abone Sayısı: Veri toplama başlangıcından itibaren toplam abone sayısı.
- Lag Özellikleri: Gecikmeli değerler

3.2.3. Günlük doğal gaz tüketimi zaman serisi analizi (2015-2024)

Karaman ilinin 2015-2024 dönemine ait 10 yıllık günlük doğal gaz tüketim verilerinin zaman serisi görünümü aşağıdaki grafikte sunmakta olup çalışmanın temelini oluşturan 3.653 günlük gözlemin tamamı grafik üzerinde gösterilmiştir.



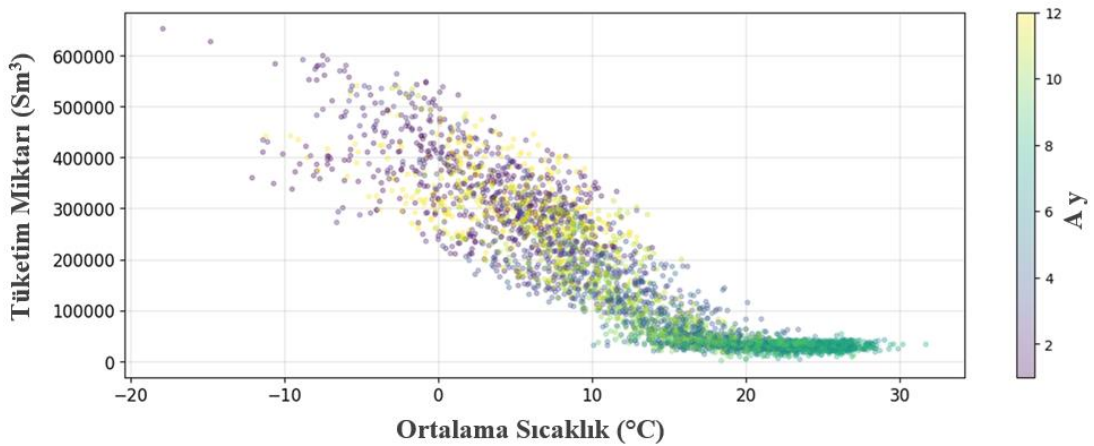
Şekil 3.2. Karaman ili 2015-2024 dönemine ait günlük doğal gaz tüketimi

Zaman serisinde en belirgin özellik, her yılda tekrar eden ve neredeyse mükemmel bir düzenlilikte gözlemlenen mevsimsel döngülerdir; kış aylarında (Aralık-Ocak-Şubat)

tüketim 300.000-650.000 m³ aralığında zirve yaparken, yaz aylarında (Haziran-Temmuz-Ağustos) 20.000-50.000 m³ seviyelerine kadar düşmektedir. Bu U-şekilli yıllık desen, Türkiye'de doğal gazın ağırlıklı olarak ısıtma amaçlı kullanıldığını ve iklim koşullarına yüksek bağımlılığını ortaya koymaktadır. Grafikte dikkat çeken bir diğer önemli bulgu, yıllar ilerledikçe hem kış zirvelerinde hem de yaz diplerinde gözlemlenen sistematik artış eğilimidir. 2015 yılında kış zirvesi yaklaşık olarak 350.000 m³ iken, 2022 ve sonrasında 650.000 m³'e kadar yükselmiştir. Bu artış, zaman içinde abone sayısındaki büyümeyi yansıtmakta ve çalışmada vurgulanan abone dinamiklerinin tüketim üzerindeki kritik etkisini doğrulamaktadır.

3.2.4. Sıcaklık-Tüketim ilişkisi ve mevsimsel bağlam

Bu bölümde sıcaklık ve tüketim arasında ilişki incelenmiştir.



Şekil 3.3. Sıcaklık ve tüketim ilişkisi

Grafik, ortalama sıcaklık ile günlük doğal gaz tüketimi arasındaki ilişkiyi mevsimsel bağlamda görselleştirmektedir. Grafikteki renk çizelgesi, 1'den 12'ye kadar ay bilgisini vermektedir. Bu görselleştirme yaklaşımı, sıcaklık-tüketim ilişkisinin yalnızca lineer bir yapıdan ibaret olmadığını, aynı zamanda belirgin bir mevsimsel kümelenme yapısı içerdiğini ortaya koymaktadır.

Sıcaklık -10°C ile -5°C arasında iken, mor renkli noktaların yoğunlaştığı alan, kış aylarında 400.000-650.000 m³ aralığında yüksek tüketim dönemini göstermektedir. Bu bölgedeki noktaların dikey yönde dağılmış olması, benzer sıcaklık koşullarında bile 150.000-200.000 m³'lük tüketim farklarının olabileceğini göstermekte olup; bu durum,

rüzgâr hızı, nem, hafta içi/sonu farkları ve tatil günleri gibi ikincil faktörlerin etkisini yansıtmaktadır.

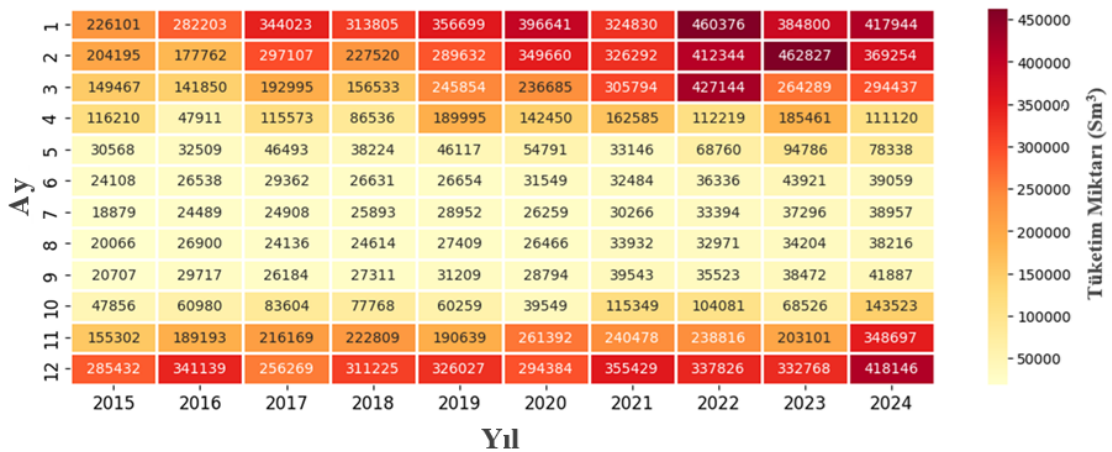
Sıcaklık 0°C ile 10°C arasında iken, sarı-turuncu tonların baskın olduğu alan, ilkbahar ve sonbahar geçiş dönemlerini temsil etmekle beraber tüketimin 150.000-350.000 m³ aralığında geniş bir alanda dağıldığını göstermektedir. Bu bölgedeki yüksek varyasyon, mevsim geçişlerinde hava koşullarının öngörülemezliğini ve buna bağlı olarak ısıtma davranışındaki değişkenliği açıklamaktadır.

Sıcaklık 15°C üzerinde iken, yeşil-turkuaz tonların yoğunlaştığı yaz aylarını temsil etmekte ve tüketimin 20.000-50.000 m³ gibi minimum seviyelerde olduğunu göstermektedir. Bu bölgedeki noktaların yatay ekseninde 15°C ile 32°C aralığında yayılması, yaz aylarında sıcaklık değişimlerinin tüketimi çok az etkilediğini ortaya koymaktadır. Örneğin 20°C ile 30°C arasındaki tüm günlerde tüketim hemen hemen aynı seviyede (yaklaşık 30.000 m³) kalmaktadır. Bu durum, doğal gazın genel olarak ısıtma amaçlı kullanıldığını doğrulamaktadır. Dikkat çeken bir diğer özellik, 10°C civarındaki kritik eşik sıcaklıktır. Bu sıcaklığın altında tüketim belirgin şekilde artarken, üstüne çıkıldığında ise tüketim azalmaktadır. Bu doğrusal olmayan davranış, basit lineer regresyon modellerinin yetersiz kalacağını ve polinom, logaritmik veya LSTM gibi ileri modelleme tekniklerinin gerekliliğini ortaya koymaktadır.

3.2.5. Aylık ortalama tüketim ısı haritası analizi

Bu bölümde aylık ortalama tüketimlerin 2015-2024 yılları arasındaki ısı haritası oluşturulmuştur.

Çizelge 3.1. Yıllık ve aylık ortalama tüketim ısı haritası



Isı haritası, 2015-2024 döneminde aylık bazda ortalama doğal gaz tüketiminin sistematik mevsimsel döngüsünü ve uzun dönem büyüme trendini göstermektedir. Kış aylarında koyu kırmızı tonlarla temsil edilen yüksek tüketim (300.000-450.000 m³), yaz aylarında açık sarı tonlarla gösterilen minimum tüketim (20.000-40.000 m³) ile belirgin bir desen oluşturmakta ve mevsimsel farkları ortaya koymaktadır. Her yıl içinde Ocak'tan Temmuz'a doğru renklerin açılması ve Temmuz'dan Aralık'a doğru koyulaşması, simetrik bir U-şekilli yıllık deseni ve çalışmada kullanılan trigonometrik özellik dönüşümlerinin mevsimselliği modellemedeki etkinliğini desteklemektedir.

Yıllar ilerledikçe yatay ekseninde tüm aylar için gözlemlenen renk koyulaşması, abone sayısındaki artışla beraber tüketimin her mevsimde sistematik olarak yükseldiğini kanıtlamaktadır.

3.2.6. Öznitelik mühendisliği

Zaman serisi tahmininde özellik mühendisliği, ham verinin makine öğrenimi algoritmalarında kullanılabilir forma dönüştürülmesi için kritik öneme sahiptir (Zheng ve Casari, 2018). Bu kapsamda modelin tahmin başarısını artırmak ve tüketim davranışını daha iyi açıklayabilmek için ham veri seti üzerinde özellik türetilmesi yapılmış ve toplam öznitelik sayısı 30'a yükseltilmiştir.

Bu çalışmada sistematik bir özellik üretim stratejisi izlenmiştir:

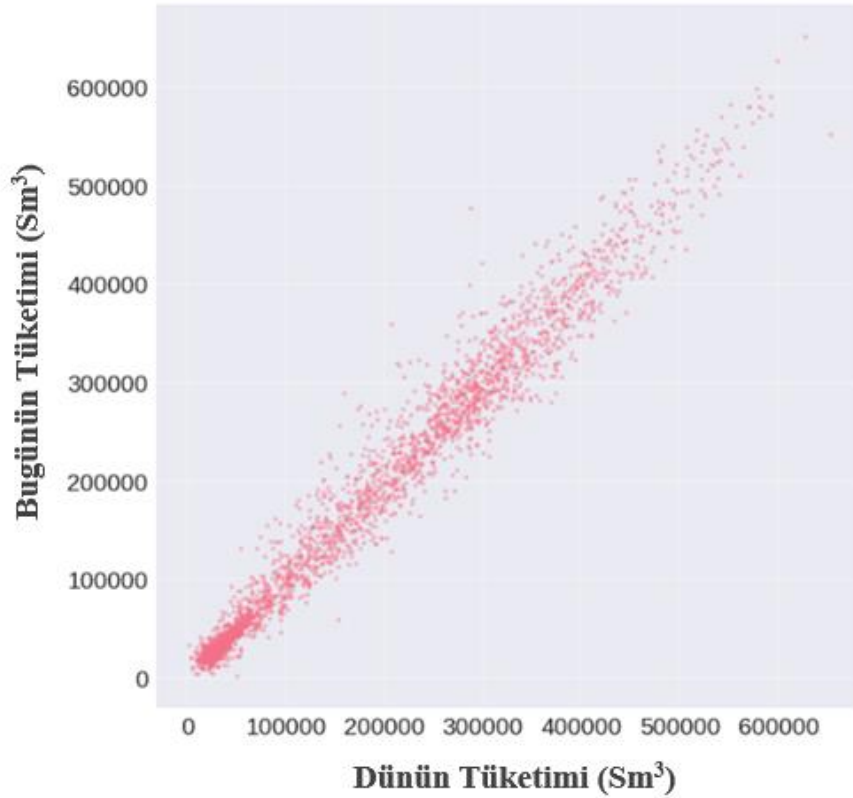
3.2.6.1. Lag özellikleri (gecikmeli değerler)

Lag değişkenleri, mevsimsel ilişkiler ve trendler gibi desenleri yakalamak için önceki zaman serisi değerlerini özellik olarak kullanır (Box vd., 2015). Lag analizi, mevcut veri noktalarının önceki değerlerle nasıl ilişkili olduğunu anlamayı sağlar. Eğer belirli bir lag değerinde anlamlı ilişki varsa, geçmiş değerlerin gelecek değerleri tahmin etmek için kullanılabileceğini gösterir (Chatfield ve Xing, 2019).

Tüketimin kendi geçmiş değerlerinden etkilendiği göz önünde bulundurularak, modele Lag özellikleri eklenmiştir. Bu özellikler, geçmiş zamanlardaki değerleri kullanarak modele bir tür hafıza kazandırmaktadır. Zaman serisi tahmininde bu yaklaşım oldukça önemlidir; çünkü bugünkü tüketim veya sıcaklık gibi değerler genellikle bir önceki günün veya haftanın değerleriyle güçlü bir şekilde ilişkilidir.

Türetilen Lag özellikleri:

- Lag_1 Tüketim: Bir gün önceki doğal gaz tüketim miktarı. Modele en yakın geçmiş bilgiyi sağlar.
- Lag_7 Tüketim: Yedi gün önceki doğal gaz tüketim miktarı. Haftalık döngüsellikleri (örneğin hafta içi/hafta sonu tüketim farkları) yakalamak için önemlidir.
- Lag_30 Tüketim: Otuz gün önceki doğal gaz tüketim miktarı. Aylık veya daha uzun süreli eğilimleri ve mevsimsel desenleri anlamak için önemlidir.
- Lag_1 Sıcaklık: Bir gün önceki ortalama sıcaklık. Doğal gaz tüketimi sıcaklık değişimlerine karşı oldukça hassastır; bu nedenle bir önceki günün sıcaklığı bugünün tüketimini etkileyen önemli bir faktördür.

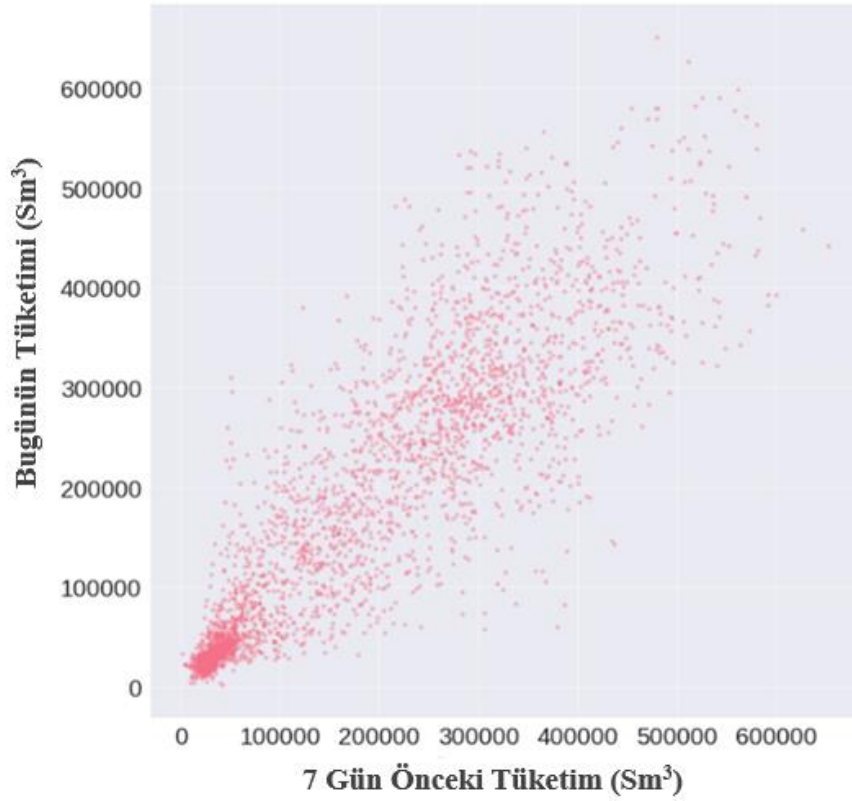


Şekil 3.4. Lag 1 (1 günlük gecikme) korelasyon analizi

Bu grafik, doğal gaz tüketimi ile bazı önemli özellikler arasındaki ilişkileri gösteren serpilme grafiklerini içermektedir:

Lag_1 Korelasyon Grafiği: Bir önceki günün doğal gaz tüketimi (Lag_1 Tüketim) ile bugünün tüketimi (Günlük Doğal gaz Tüketim Miktarı) arasındaki ilişkiyi göstermektedir. Bu grafikteki noktaların büyük ölçüde pozitif ve doğrusal bir

eğilim göstermesi, dünün tüketiminin bugünün tüketimi ile çok yüksek ilişkisi olduğunu göstermektedir. Bu gözlem, doğal gaz tüketiminin yüksek derecede gecikmeli bağımlılık gösterdiğini ortaya koymaktadır.



Şekil 3.5. Lag 7 (7 günlük gecikme) korelasyon analizi

Lag_7 Korelasyon Grafiği: Bir hafta önceki doğal gaz tüketimi (Lag_7_Tuketim) ile bugünün tüketimi arasındaki ilişkiyi göstermektedir. Bu grafikte de güçlü bir ilişki gözlemlenmektedir. Bu durum, doğal gaz tüketiminde haftalık periyodik desenlerin varlığına işaret etmektedir. Örneğin, haftanın aynı günlerinde benzer tüketim alışkanlıkları görülebilmektedir. Modelin bu haftalık döngüsellliği yakalaması, özellikle ısıtma dışı tüketim desenlerini anlaması açısından önemlidir. Saçılımın artması, 7 gün içinde meydana gelen meteorolojik değişkenlik ve abone davranış değişiklikleri ile açıklanabilir.

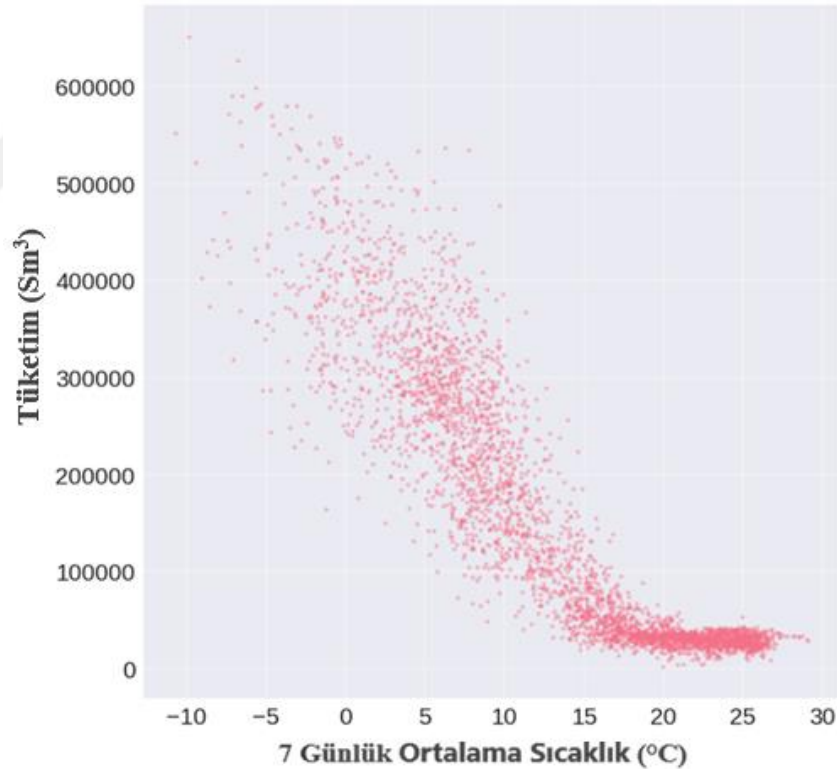
3.2.6.2. Hareketli ortalama özellikleri

Hareketli ortalamalar, zaman içindeki trendleri ve dalgalanmaları yakalamaya yardımcı olup modelin değerlerin arttığını mı yoksa azaldığını mı tanımasını sağlar

(Hyndman ve Athanasopoulos, 2018). Bu yöntem, kısa dönemli trendleri ve dalgalanmaları yumuşatarak daha genel eğilimleri ortaya koyar. Böylece ani değişimlerden ziyade istikrarlı eğilimlerin model tarafından öğrenilmesine katkı sağlar (Montgomery vd., 2015).

Türetilen hareketli ortalama özellikleri:

- MA_3 Sıcaklık: Son 3 günün ortalama sıcaklığı. Kısa vadeli sıcaklık trendlerini gösterir.
- MA_7 Sıcaklık: Son 7 günün ortalama sıcaklığı. Bir haftalık sıcaklık ortalaması, tüketim üzerindeki genel sıcaklık etkisini daha istikrarlı bir şekilde yansıtır.
- MA_7 Tüketim: Son 7 günün ortalama doğal gaz tüketimi. Haftalık tüketim trendlerini ve temel tüketim seviyesini belirler.
- MA_3 Nem: Son 3 günün ortalama nemi. Nem gibi diğer meteorolojik faktörlerin kısa vadeli etkilerini içerir.



Şekil 3.6. 7 günlük ortalama sıcaklık ve tüketim ilişkisi

Grafik son 7 günün ortalama sıcaklığı (MA_7_Sıcaklık) ile doğal gaz tüketimi arasındaki ilişkiyi göstermektedir. Sıcaklık azaldıkça tüketimin arttığını gösteren belirgin bir negatif ilişki gözükmemektedir. Özellikle belirli bir sıcaklık eşiğinin (yaklaşık 15-20°C)

altında tüketimin hızla arttığı ve soğuk günlerdeki tüketim değerlerinin oldukça geniş bir aralığa yayıldığı görülmektedir. Bu gözlem, doğal gaz tüketiminin büyük ölçüde ısıtma ihtiyacından etkilendiğini ve sıcaklığın tüketim için en önemli etkenlerden biri olduğunu teyit etmektedir.

3.2.6.3. Trigonometrik dönüşümler (döngüsel özellikler)

Saat, haftanın günleri veya yılın ayları gibi zaman tabanlı değişkenler doğası gereği döngüseldir. Bu tür değişkenler yalnızca tam sayılarla temsil edildiğinde, makine öğrenimi modelleri bu bilgiyi doğrusal olarak yorumlayabilir (Géron, 2019). Bu durum, örneğin ocak (1) ile aralık (12) ayları arasında yapay bir uzaklık yaratır; oysa iklimsel ve davranışsal olarak bu aylar birbirine oldukça yakındır (Hyndman ve Athanasopoulos, 2018; Kuhn ve Johnson, 2019).

Bu sorunu çözmek için trigonometrik dönüşümler uygulanır. Sinüs ve kosinüs fonksiyonları, döngüsel değişkenleri birim çember üzerinde konumlandırarak sürekliliği korur. Böylece Ocak ve Aralık gibi uç noktalar, model tarafından birbirine yakın olarak algılanabilir.

- Ay_Sin ve Ay_Cos: Ay için sinüs ve kosinüs dönüşümleri (12 aylık döngü).
- Yılın_Günü_Sin ve Yılın_Günü_Cos: Yılın günü için sinüs ve kosinüs dönüşümleri (365 günlük döngü).

$$Ay_Sin = \sin\left(\frac{2\pi \times AY}{12}\right) \quad (3.1)$$

$$Ay_Cos = \cos\left(\frac{2\pi \times AY}{12}\right) \quad (3.2)$$

$$Yılın_Günü_Sin = \sin\left(\frac{2\pi \times GÜN}{365}\right) \quad (3.3)$$

$$Yılın_Günü_Cos = \cos\left(\frac{2\pi \times GÜN}{365}\right) \quad (3.4)$$

Bu tür dönüşümler, zaman serisi tahminlerinde ve mevsimsel etkilerin modellenmesinde oldukça etkilidir (Hyndman ve Athanasopoulos, 2018; Kuhn ve Johnson, 2019). Döngüsel değişkenlerin bu şekilde kodlanması, modellerin zamanın

periyodik doğasını daha doğru şekilde öğrenmesini sağlar (Géron, 2019; Skforecast Docs, 2023).

3.2.6.4. Hissedilen sıcaklık

Hava sıcaklığı ve doğal gaz talebi arasındaki etkileşim, bilimsel çalışmalarda yaygın bir şekilde ele alınan konulardan biridir. Bağıl nem ve rüzgâr hızı gibi meteorolojik faktörler de ısıtma yükünü ve dolayısıyla enerji tüketimini etkilemektedir. Özellikle, sıcaklık ve nemin birlikte oluşturduğu hissedilen sıcaklık, binaların ısıtma ve soğutma yüklerini belirleyen önemli bir faktördür (ASHRAE, 2017; Fanger, 1970).

Ham veride bulunan Günlük Ortalama Sıcaklık (T) ve Bağıl Nem (RH) verileri kullanılarak hissedilen sıcaklık hesaplanmıştır. Hesaplama, Amerikan Ulusal Hava Durumu Servisi (NWS) tarafından da benimsenen ve literatürde yaygın kabul gören Rothfusz regresyon yaklaşımı kullanılmıştır. Rothfusz'un çok terimli regresyon modeli, sıcaklık ve nemin birleşik etkisini aşağıdaki şekilde ifade eder:

$$HI = c_1 + c_2T + c_3R + c_4TR + c_5T^2 + c_6R^2 + c_7T^2R + c_8T^2R + c_9T^2R^2 \quad (3.5)$$

Burada:

- HI : Hissedilen Sıcaklık (°C)
- T : Termometre Sıcaklığı (°C)
- R : Bağıl Nem (%)
- $c_{1...9}$: Ampirik regresyon katsayılarıdır.

Analiz sırasında $T < 10^\circ\text{C}$ olan düşük sıcaklıklarda, nemin etkisi ihmal edilebilir düzeyde kabul edilmiştir. Bu durumda:

$$HI \approx T$$

Bu varsayım, düşük sıcaklıklarda nemin hissedilen sıcaklık üzerindeki etkisinin minimal olduğunu gösteren çalışmalarla uyumludur (Anderson vd., 2013).

3.2.6.5. Isıtma derece gün (HDD)

Isıtma derece gün kavramı, binaların ısıtma enerjisi ihtiyacının iklimsel koşullara bağlı olarak tahmin edilmesinde yaygın olarak kullanılan bir göstergedir. HDD, belirli bir baz sıcaklık değerinin (çoğunlukla 18°C) altındaki günlük ortalama dış ortam

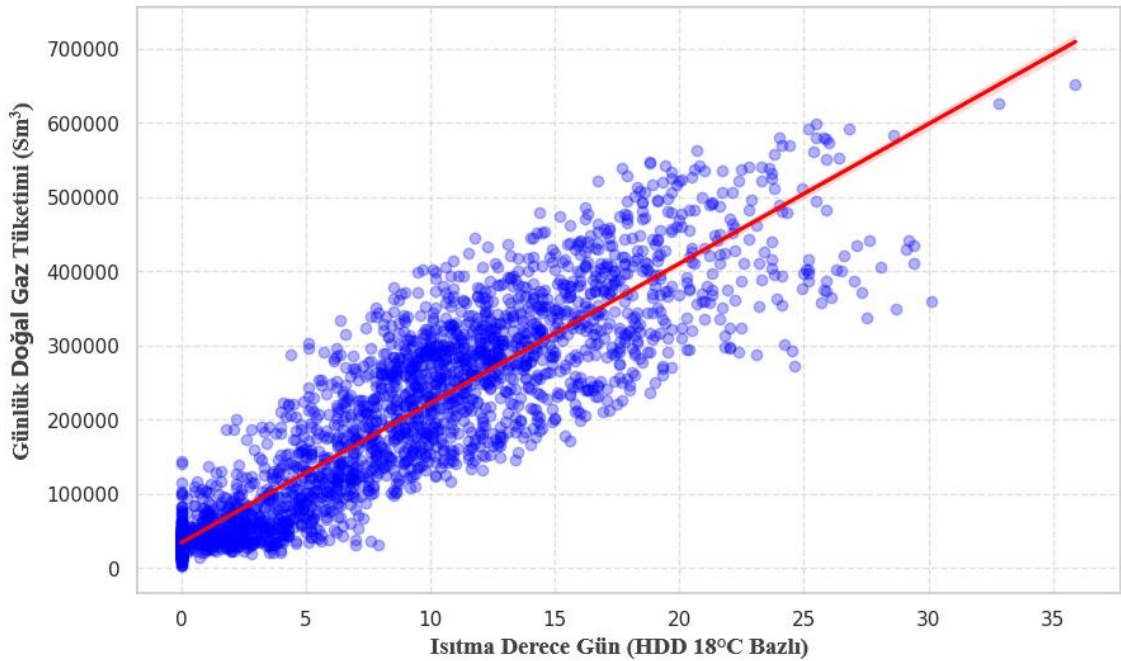
sıcaklıklarının bu baz sıcaklıktan farklarının toplanmasıyla elde edilir ve enerji talebi modellemelerinde önemli bir parametre olarak yer alır (Eurostat, 2024; MGM, 2025; ASHRAE, 2021).

HDD değeri aşağıdaki fonksiyon ile hesaplanmıştır:

$$HDD_t = \begin{cases} 18 - T_{\{avg,t\}}, & \text{eğer } T_{\{avg,t\}} < 18 \\ 0, & \text{eğer } T_{\{avg,t\}} \geq 18 \end{cases} \quad (3.6)$$

Bu çalışmada, yaygın kullanılan 18°C baz sıcaklık değeri kullanılmıştır (ASHRAE, 2021). Burada $T_{avg,t}$, t günündeki günlük ortalama sıcaklığı ifade etmektedir. Baz sıcaklık değeri olarak 18°C'nin seçilmesi, insan konforu ve ülke standartları dikkate alınarak belirlenmiş olup farklı ülkelerde ve çalışmalarda bu değer değişkenlik gösterebilmektedir (Day ve Karayiannis, 1998).

HDD, özellikle ısıtma amaçlı doğal gaz talebinin tahmininde kritik bir rol oynamaktadır (Day ve Karayiannis, 1998; Sarak ve Satman, 2003). Türkiye'de yapılan ampirik çalışmalarda, sıcaklık ile konutlarda doğal gaz tüketimi arasında güçlü bir korelasyon olduğunu gösterilmiştir (Aras ve Aras, 2004; Sarak ve Satman, 2003; Gümrah vd., 2001).



Şekil 3.7. Isıtma derece gün ve doğal gaz tüketimi ilişkisi

Grafikteki nokta dağılımı incelendiğinde, yani 18°C ve üzeri sıcaklıklarda (HDD=0 noktasında) tüketimin baz seviyeye düştüğü, bununla birlikte HDD arttıkça tüketimin yükseldiği açıkça gözlemlenmektedir. Özellikle HDD değerlerinin 5'in üzerinde olduğu günlerde, tüketim miktarı 100.000 Sm³'ü aşmakta ve HDD=25'i geçtiğinde ekstrem soğuk koşullarda 600.000 Sm³'e kadar çıkabilmektedir.

Grafik veri setine eklenen kırmızı regresyon çizgisi, HDD ile tüketim arasındaki lineer ilişkiyi matematiksel olarak modellenmektedir. Mavi noktaların regresyon çizgisi etrafındaki dağılımına bakıldığında, düşük HDD değerlerinde (0-5 arası) noktaların daha geniş bir yayılım oluşturduğu, ancak yüksek HDD değerlerinde (20+) dağılımın daraldığı gözükmemektedir. Bu durum, soğuk günlerde ısıtma ihtiyacının baskın etken haline geldiğini ve diğer değişkenlerin etkisinin azaldığını göstermektedir.

Dikkat çeken bir diğer önemli bulgu, verideki aykırı değerlerdir. Yüksek HDD değerlerinde (30+), bazı günlerde beklenenin altında tüketim gözlemlenmektedir. Bu durum muhtemelen hafta sonları, resmî tatiller veya kısa süreli doğal gaz arz kesintileri gibi olağandışı koşullarla açıklanabilmektedir. Benzer şekilde, düşük HDD değerlerinde (HDD < 2) 100.000 Sm³'ün üzerinde tüketim gösteren noktalar da bulunmaktadır. Bu durum ise mevsim geçiş dönemlerinde beklenmedik soğuk hava akımları ile açıklanabilir. Bu aykırı değerler, basit HDD tabanlı tahmin modellerinin neden ortalama %10-15 hata payı taşıdığını ve LSTM gibi çok değişkenli, doğrusal olmayan modellerin gerekliliğini göstermektedir.

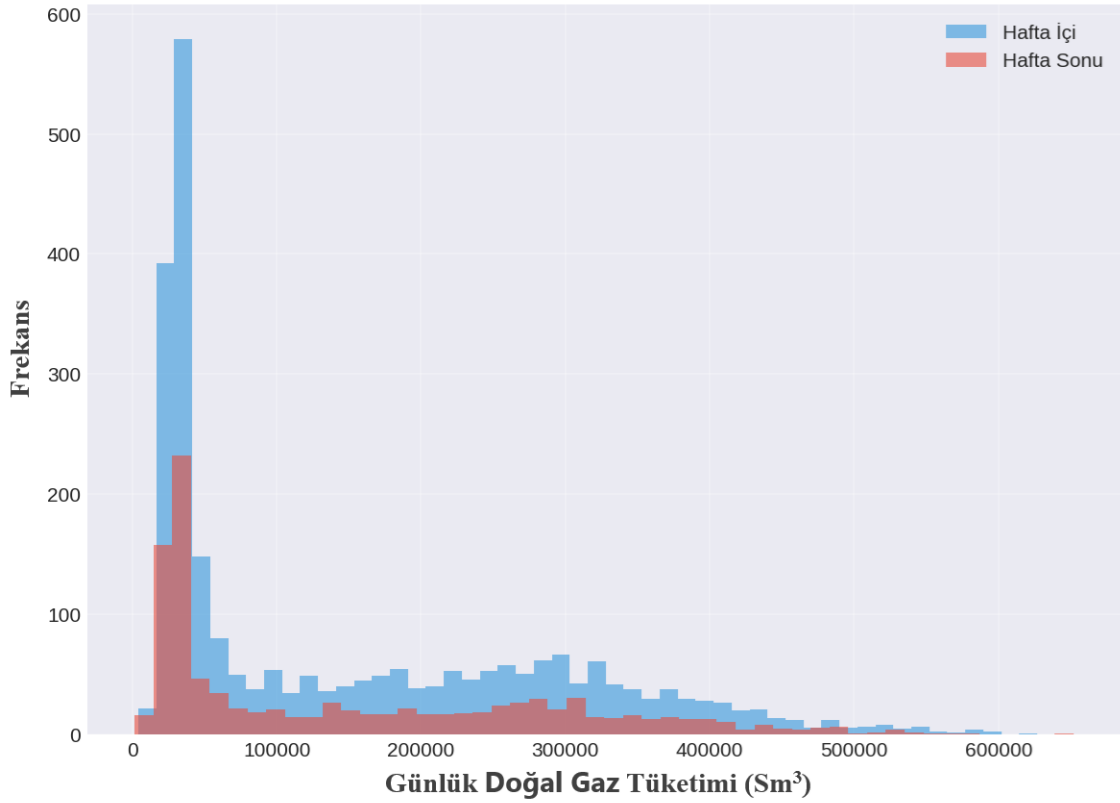
3.2.6.6. Tatil günü indikatörü

Hafta sonları ve resmî tatil günlerinde insan aktivitelerinin ve endüstriyel kullanımın azalması nedeniyle doğal gaz tüketim profilinde yapısal kırılmalar gözlenmektedir. Bu etkiyi modele dâhil edebilmek için tatil bilgisi kategorik değişkenden ikili değişkene dönüştürülmüştür. Literatürde enerji talep tahmin modellerinde tatil etkisinin ikili değişkenlerle temsil edilmesi yaygın bir uygulamadır (Hippert vd., 2001; Taylor ve McSharry, 2007; Weron, 2006).

D_t değişkeni, t gününün tatil durumunu temsil etmek üzere mantıksal operatörle türetilmiştir:

$$D_t = \begin{cases} 1, & \text{eğer gün Haftasonu (Cmt, Paz) V Resmi Tatil ise} \\ 0, & \text{diğer durumlarda (iş günü)} \end{cases} \quad (3.7)$$

Çalışmada resmî tatil günleri, Türkiye'nin 2015-2024 yılları arasındaki resmi takvimine göre belirlenmiştir.

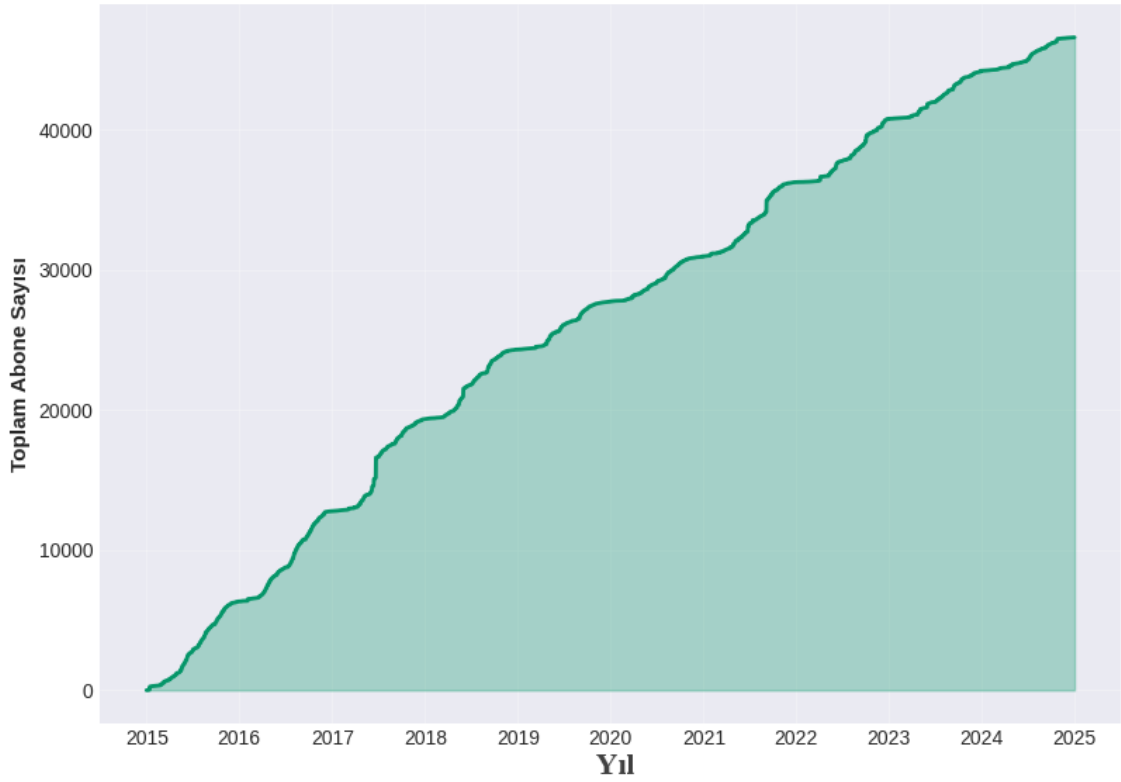


Şekil 3.8. Hafta içi ve hafta sonu doğal gaz kullanım dağılımı

Grafikte hafta içi ve hafta sonu günlük tüketim dağılımları karşılaştırılmaktadır. Hafta içi ve hafta sonu tüketim ortalamalarının ve dağılımlarının birbirine oldukça yakın olduğunu görsel olarak ortaya koymaktadır. Tüketimdeki belirgin mevsimsel değişimler (özellikle kış aylarındaki yüksek tüketim) göz önüne alındığında, haftanın gününe bağlı tüketim farklılıkların genel tüketim deseni üzerindeki etkisinin kısıtlı olduğu göstermektedir.

3.2.6.7. Kümülatif abone sayısı

Ham veri setinde yer alan *Günlük Alınan Abone Sayısı* değişkeni, sisteme günlük olarak dâhil olan abone akış verisini göstermektedir. Ancak toplam enerji talebi, o gün sisteme giren abonelerden ziyade, sistemdeki toplam abone sayısı ile ilişkilidir.



Şekil 3.9. Kümülatif abone sayısı artış grafiği (2015-2024)

Bu nedenle, günlük abone girişleri toplanarak kümülatif abone sayısı (S_t) hesaplanmıştır. Başlangıç zamanı $t=0$ kabul edildiğinde, t anındaki toplam abone sayısı şu şekilde formüle edilir:

$$S_t = S_{\{t-1\}} + a_t = \sum_{i=0}^t a_i \quad (3.8)$$

Burada:

- S_t : t günündeki toplam kümülatif abone sayısı,
- a_i : i gününde sisteme yeni dâhil olan abone sayısını ifade eder.

Bu yaklaşım enerji talep modelleme literatüründe abone sayısının talep üzerindeki etkisini açıklayan çalışmalara dayanmaktadır (Aras ve Aras, 2004).

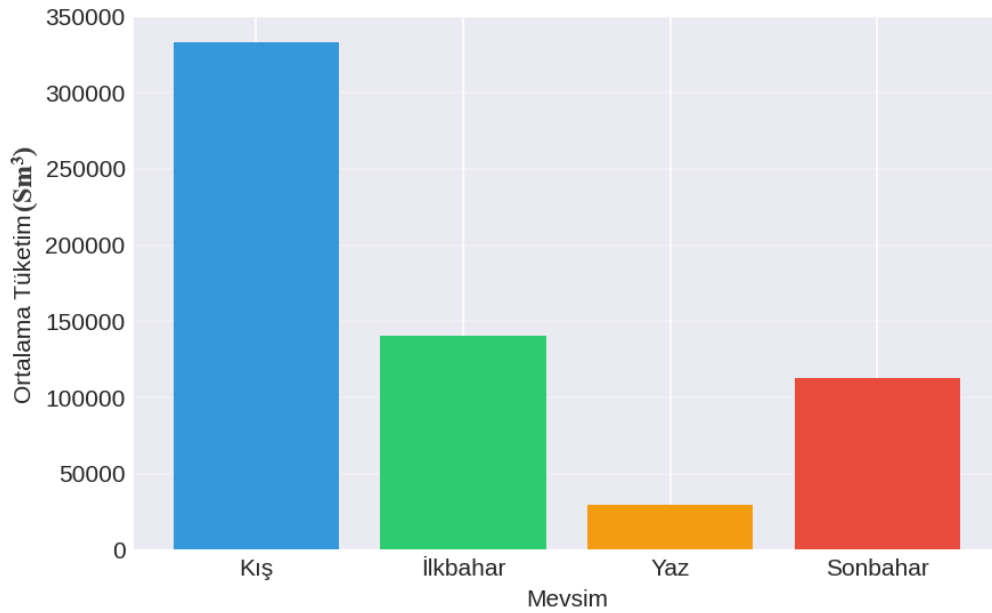
3.2.6.8. Döngüsel zaman özellikleri

Mevsimsellik etkisini sayısal olarak ifade edebilmek için yıl ve ayın günü bilgileri sinüs ve kosinüs fonksiyonları ile dönüştürülmüştür. Bu yöntem, zaman serisi modellemede döngüsel kalıpları yakalamak için yaygın olarak kullanılmaktadır (Zheng,

A., ve Casari, 2018). Veri setindeki tarih sütunundan çıkarılan bu özellikler, modelin mevsimsel, haftalık veya yıllık döngüsel kalıpları öğrenmesini sağlar. Doğal gaz tüketimi yılın belirli dönemlerinde ve haftanın belirli günlerinde farklılık gösterdiğinden (Soldo, 2012; Aras ve Aras, 2004), eklenen zaman özellikleri modelin doğruluğunu artırmaktadır.

Eklenen zaman özellikleri şunlardır:

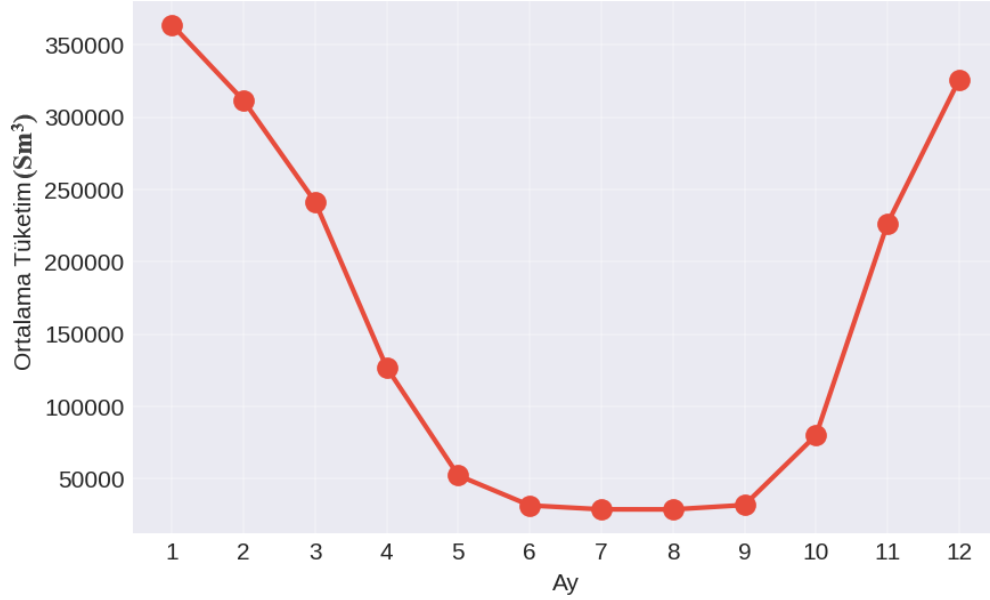
- Ay: Ay bilgisi (1-12). Yıllık mevsimsel döngüyü ifade eder.
- Mevsim: Mevsim bilgisi (1=Kış, 2=İlkbahar, 3=Yaz, 4=Sonbahar). Geniş mevsimsel etkileri yakalar.
- Haftanın Günü: Haftanın günü (0=Pazartesi, 6=Pazar). Hafta içi/hafta sonu tüketim farklılıklarını öğrenmeye olanak tanır.
- Yılın Günü: Yılın günü (1-365). Yıllık döngüde daha ince ayrımlar sağlar.
- Yıl: Yıl bilgisi. Uzun vadeli trendleri (örneğin abone sayısındaki artışın etkisiyle tüketim artışı) modelin öğrenmesine yardımcı olur.



Şekil 3.10. Mevsimsel ortalama tüketim

Aylık Ortalama Tüketim Profili grafiği, yıl boyunca tüketimin güçlü bir döngüsellığe sahip olduğunu göstermektedir. Tüketim, kış aylarında zirve yapmakta, ilkbahar ve sonbaharda orta seviyelere düşmekte ve yaz aylarında Haziran, Temmuz, Ağustos en düşük seviyelere inmektedir. Bu durum, doğal gazın öncelikli olarak ısıtma amacıyla kullanıldığını ve sıcaklık düşüktüçe ısıtma ihtiyacının artmasıyla tüketimin de

paralel olarak yükseldiğini kanıtlamaktadır. Yaz aylarındaki düşük tüketim ise ısıtma ihtiyacının neredeyse hiç olmaması sebebiyledir.



Şekil 3.11. Aylık ortalama tüketim grafiği

Mevsimsel Ortalama Tüketim grafiği, mevsimlere göre ortalama tüketim değerlerini net bir şekilde göstermektedir:

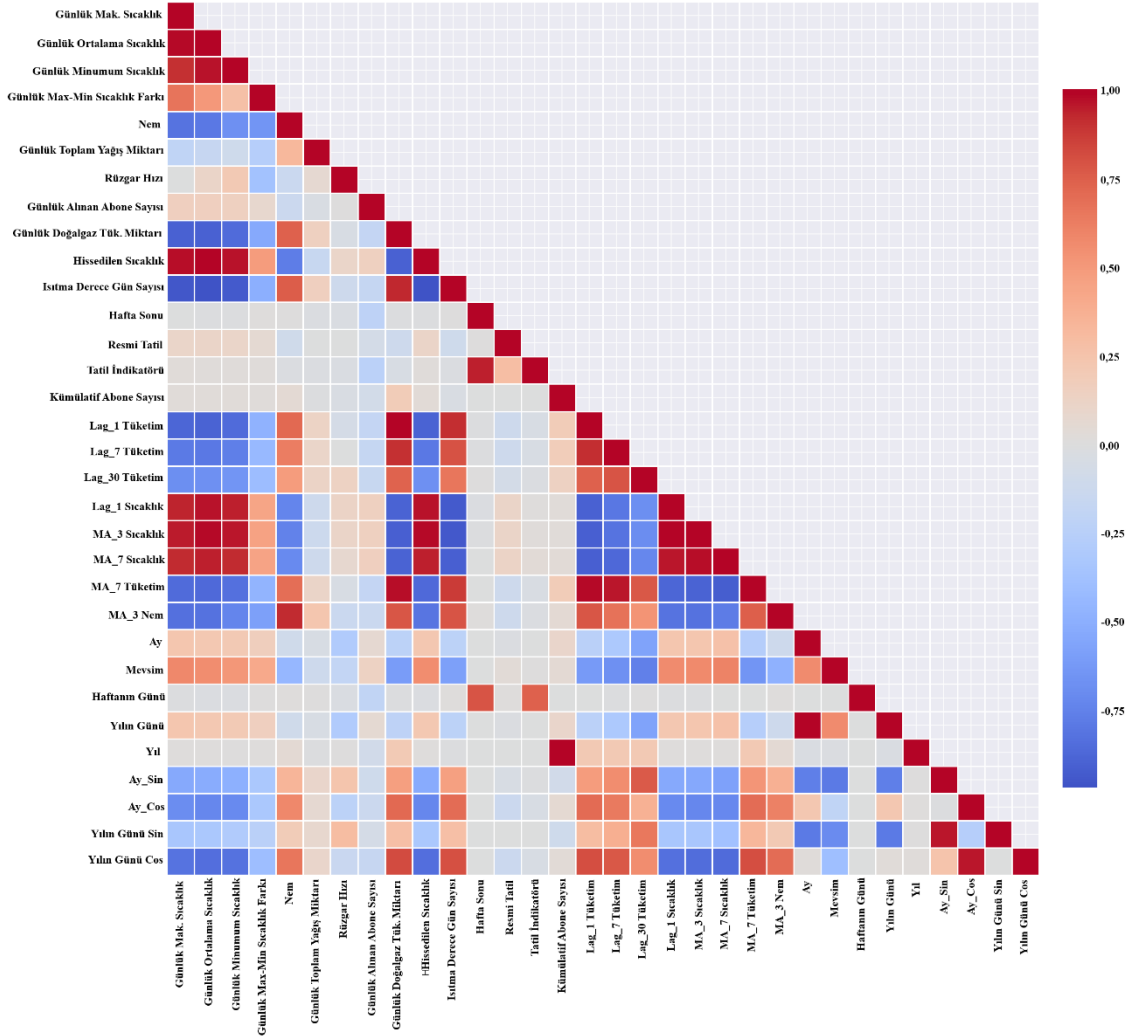
- Kış (Aralık, Ocak, Şubat): Ortalama tüketim en yüksek seviyededir. Bu da ısıtma ihtiyacının yoğunluğunu yansıtır.
- İlkbahar (Mart, Nisan, Mayıs): Tüketim kış mevsimine göre düşüş gösterir.
- Yaz (Haziran, Temmuz, Ağustos): Tüketim en düşük seviyeye inmiştir.
- Sonbahar (Eylül, Ekim, Kasım): Tüketim tekrar yükseliş eğilimine girmiştir.

3.2.6.9. Verilerin korelasyon grafiği

Doğal gaz tüketimini etkileyen faktörleri belirlemek ve modelin hangi özelliklere daha fazla odaklandığını anlamak amacıyla kapsamlı bir korelasyon analizi gerçekleştirilmiştir.

Tüm özellikler arasındaki ilişkileri gösteren ısı haritası, modelin kullandığı farklı veri kategorileri (meteorolojik, zamansal, gecikme) arasındaki karmaşık bağımlılıkları ortaya koymaktadır. Özellikle sıcaklık değişkenleri arasında gözlenen yüksek korelasyonlar, veri setinde çoklu doğrusal ilişki sorununa işaret etmektedir (Kutner vd.,

2005; Montgomery vd., 2015). Bu durum, regresyon ve zaman serisi modelleme literatüründe sıkça tartışılan bir konudur ve model performansını etkileyebilmektedir (James vd., 2021; Montgomery vd., 2015)



Şekil 3.12. Tüm değişkenler arası korelasyon matrisi

Doğal Gaz Tüketimi ile En Yüksek Korelasyonlu Özellikler:

- Sıcaklık Değişkenleri: *Günlük Ortalama Sıcaklık* ve *Hissedilen Sıcaklık* gibi sıcaklık değişkenleri, *Günlük Doğal gaz Tüketim Miktarı* ile 0,84'e yakın güçlü negatif korelasyonlara sahiptir. Bu da beklenen bir sonuçtur; sıcaklık azaldıkça doğal gaz tüketimi artmaktadır.
- Gecikmeli Tüketim Değerleri (Lag Özellikleri): Modelin, geçmiş tüketim değerlerinden önemli ölçüde etkilendiği göstermektedir.
 - *Lag_1_Tuketim* (bir önceki günün tüketimi): 0,987 gibi oldukça yüksek bir pozitif korelasyona sahiptir.

- MA_7_Tuketim (son 7 günün ortalama tüketimi) 0,982 ve Lag_7_Tuketim (bir hafta önceki tüketim) 0,854 ile de tüketimin güçlü bir şekilde kendi geçmiş değerleriyle ilişkili olduğu kanıtlanmıştır. Bu, doğal gaz tüketiminin zaman serisi yapısının güçlü bir şekilde otokorelasyon içerdiğini göstermektedir.
- Isıtma Derece Gün Sayısı (HDD): Isıtma Derece Gün Sayısı değişkeni 0,929 ile pozitif yönde çok güçlü bir korelasyona sahiptir. Bu çıkarım, ısıtma ihtiyacının doğal gaz tüketimindeki en belirleyici faktörlerden biri olduğunu kanıtlamaktadır. HDD'nin artması, doğal gaz tüketiminde önemli bir artışa sebep olmaktadır.
- Nem: Nem değişkeni 0,741 ile tüketimle pozitif bir korelasyona sahiptir. Yani nem arttıkça tüketim de artış eğilimindedir. Bu, soğuk havalarda yüksek nemin hissedilen sıcaklığı düşürmesi ile ilgili olabilmektedir.
- Mevsimsel ve Zamansal Özellikler: Yılın_Günü_Cos, Ay_Cos ve Ay_Sin gibi trigonometrik ve döngüsel özellikler, mevsimsel varyasyonları yakalayarak doğal gaz tüketimi üzerinde orta ve güçlü etkiler göstermektedir. Bu özellikler, aylık ve yıllık döngüsel desenlerin modellenmesinde kritik rol oynamıştır.

Korelasyon analizi, doğal gaz tüketim tahmin modelinde kullanılan özelliklerin tüketim üzerindeki etkilerini açıkça ortaya koymuştur. Gecikmeli tüketim değerleri ve ısıtma derece gün sayısı, tahmin modelinin temelini oluşturan en güçlü etkenlerdir. Sıcaklık, nem ve mevsimsel göstergeler de tüketimdeki değişimleri anlama konusunda önemli katkı sağlamaktadır. Bu çıkarımlar, modelin fiziksel gerçekliklerle uyumlu olduğunu ve en etkili özelliklerin tahmin sürecine dâhil edildiğini doğrulamaktadır.

3.2.7. Veri ön işleme

Veri ön işleme, ham verinin temizlenmesi, dönüştürülmesi ve modele uygun hale getirilmesi sürecidir (Pyle, 1999; Kuhn ve Johnson, 2019). Bu aşama, modelin doğruluğunu ve güvenilirliğini artırmak için kritik öneme sahiptir (Géron, 2019; Hastie vd., 2009).

3.2.7.1. Eksik değerlerin temizlenmesi

Lag ve hareketli ortalama hesaplamaları sonucunda veri setinin ilk birkaç satırında eksik değerler (NaN) oluşmuştur. Örneğin, Lag_30_Tuketim özelliği ilk 30 gün için tanımsızdır. Bu eksik değerler, model eğitimi sırasında sorunlara yol açabileceği için ilgili

fonksiyon kullanılarak veri setinden çıkarılmıştır. Sonuç olarak 30 satır silinmiş ve kalan temiz veri setinin boyutu 3.623 satır olmuştur (Schafer ve Graham, 2002; Pandas Documentation, 2024).

3.2.7.2. Veri setinin bölünmesi (eğitim, doğrulama, test)

Zaman serisi verileriyle çalışırken, veri setini rastgele bölmek yerine sıralı bir şekilde bölmek gereklidir (Hyndman ve Athanasopoulos, 2018; Bergmeir ve Benítez, 2012). Bu yaklaşım, modelin geçmiş veriler üzerinde eğitilip geleceği tahmin etme yeteneğini doğru bir şekilde değerlendirmesini sağlar ve zaman bağımlılığını korur (Tashman, 2000). Veri setinin eğitim-validasyon-test ayrımında bu sıralı bölme yaklaşımı uygulanmıştır. Veri seti aşağıdaki oranlarda ve sıralı olarak bölünmüştür:

- Eğitim Seti: 31.05.2015 – 06.01.2023 tarihleri arası. Verinin ilk %80'i. Model bu set üzerinde öğrenir.
- Doğrulama Seti: 07.01.2023 – 03.01.2024 tarihleri arası. Verinin sonraki %10'u. Model eğitimi sırasında hiperparametre ayarı ve aşırı öğrenme kontrolü için kullanılır.
- Test Seti: 04.01.2024 – 31.12.2024 tarihleri arası. Verinin son %10'u. Modelin daha önce hiç görmediği verilere genelleme yeteneğini nihai olarak değerlendirmek için kullanılır (Hyndman ve Athanasopoulos, 2018).

3.2.7.3. Özellik seçimi

Modelin eğitiminde kullanılacak özellikler ve hedef değişken belirlenmiştir. Hedef değişkenimiz Günlük Doğal gaz Tüketim Miktarı'dır. Modelin doğrudan kullanmayacağı ve tahmin sürecinde dışarıda bırakılan sütunlar şunlardır:

- Tarih: Tarih bilgisi, doğrudan bir sayısal özellik olarak kullanılmadığı için hariç tutulur, ancak ondan çıkarılan zaman özellikleri kullanılır.
- Günlük Doğal gaz Tüketim Miktarı: Modelin tahmin etmeye çalıştığı değişkendir, bu yüzden özellik olarak kullanılmamaktadır.
- Günlük Alınan Abone Sayısı: Günlük olarak yeni alınan abone sayısını temsil eder. Öznitelik mühendisliği aşamasında *Kümülatif Abone Sayısı* özniteliği üretildiği ve bu özniteliğin *Günlük Alınan Abone Sayısı* ile çakıştığı için dışarıda bırakılmıştır.

3.2.8. Normalizasyon (Ölçeklendirme)

Normalizasyon, farklı ölçek ve birimlerdeki sayısal özelliklerin hepsini belirli bir aralığa (genellikle 0–1) sıkıştırma işlemidir (Goodfellow vd., 2016). Bu işlem, sinir ağları gibi modellerde gradyan inişi sırasında optimizasyonun daha hızlı ve stabil olmasını sağlamakta olup, büyük değerli özelliklerin küçük değerli özelliklere göre baskın çıkmasını engellemektedir (Géron, 2019; Demir ve Gunal, 2025).

Bu seçimin arkasındaki temel gerekçeler:

LSTM Mimarisi Uyumluluğu: LSTM hücrelerindeki sigmoid aktivasyon fonksiyonları [0,1] aralığında çıktı üretir. Min-Max normalizasyonu, girdi verilerini de aynı aralığa taşıyarak aktivasyon fonksiyonlarının en verimli çalıştığı bölgeyi kullanır (Géron, 2019).

Zaman Serisi Standart Uygulaması: Zaman serisi tahminleme çalışmalarında Min-Max scaler, akademik literatürde en yaygın kullanılan yöntemdir (Sagheer ve Kotb, 2019).

Gradientlerin Stabilitesi: [0,1] aralığı, dereceli azalma sırasında sayısal stabiliteyi garanti eder ve gradyan patlaması riskini minimize eder (Goodfellow vd., 2016).

Yorumlanabilirlik: Normalize edilmiş değerler, orijinal aralıkta yüzdelik konumu temsil eder, bu da hata analizini ve model çıktılarının yorumlanmasını kolaylaştırır (Géron, 2019).

3.2.8.1. Normalizasyon hesaplama

Normalizasyon işlemi için MinMaxScaler (Pedregosa vd., 2011) kullanılmış olup, her bir özellik için matematiksel dönüşüm şu formülle sağlanmıştır (Goodfellow vd., 2016):

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (3.9)$$

Örnek hesaplama (Ortalama Sıcaklık için):

- Orijinal değer: 15,3 °C
- x_min (eğitim): -17,9 °C
- x_max (eğitim): 31,7 °C

$$x_{norm} = \frac{15,3 - (-17,9)}{31,7 - (-17,9)}$$

$$x_{norm} = 0,669$$

Bu dönüşüm sonucunda, tüm özellikler [0, 1] aralığına taşınmıştır.

3.2.9. LSTM model mimarisi ve eğitim süreci

Geleneksel ileri beslemeli sinir ağları, girdi ve çıktı arasında sabit boyutlu eşlemelerde başarılı olmalarına rağmen değişken uzunluktaki sekansları ve zamansal bağımlılıkları modellemede yetersiz kalmaktadır (Rumelhart vd., 1986; Elman, 1990). Bu sınırlılık, özellikle zaman serisi tahmini, doğal dil işleme ve sıralı karar verme problemlerinde kritik bir engel teşkil etmektedir (Bengio vd., 1994).

LSTM mimarisi, Hochreiter ve Schmidhuber (1997) tarafından gradyan kaybolması problemini çözmek amacıyla geliştirilmiştir. LSTM'in temel yeniliği, kapı yapıları (gating mekanizması) sayesinde bilgi akışını kontrollü bir şekilde düzenlemesidir. Bu mekanizma, ağı uzun vadeli bağımlılıkları öğrenmesine olanak tanır ve gradyan akışını koruyarak eğitim sürecini stabilize eder (Hochreiter ve Schmidhuber, 1997; Greff vd., 2017).

3.2.9.1. Mimari olarak LSTM seçilme nedenleri

Bu çalışmada LSTM mimarisinin başlıca seçilme nedenleri şunlardır:

Uzun Vadeli Zamansal Bağımlılıklar: Doğal gaz tüketimi, hem kısa vadeli (günlük hava değişimleri) hem de uzun vadeli (mevsimsel döngüler, yıllık trendler) faktörlerden etkilenmektedir (Soldo, 2012; Aras ve Aras, 2004). Veri setindeki *Lag_30* özelliğinin güçlü korelasyonu ($R=0,798$), 30 gün önceki bilginin hâlâ önemli olduğunu göstermektedir. LSTM, bu tür uzun vadeli bağımlılıkları öğrenme kapasitesine sahiptir (Hochreiter ve Schmidhuber, 1997).

Zaman Serisi Literatüründe Kanıtlanmış Başarı: LSTM, enerji talebi tahmini alanında yaygın olarak kullanılmakta ve geleneksel yöntemlere (ARIMA, Exponential Smoothing) karşı daha yüksek tahmin doğruluğu sergilemektedir (Kong vd., 2019; Wang vd., 2025). Özellikle doğal gaz talebinin mevsimsel döngülerden, iklimsel değişkenlerden ve ekonomik parametrelerden kaynaklanan doğrusal olmayan ve karmaşık yapısı (Soldo, 2012), LSTM'in bellek hücreleri ve kapı mekanizmaları sayesinde bu uzun dönemli

bağımlılıkları öğrenme avantajını öne çıkarmaktadır (Hochreiter ve Schmidhuber, 1997; Greff vd., 2017).

Çoklu Değişken Entegrasyonu: LSTM, meteorolojik değişkenler, takvim özellikleri ve geçmiş tüketim verilerini bütünleşmiş bir şekilde işleyebilir ve çok değişkenli zaman serisi modellemesinde etkilidir (Hochreiter ve Schmidhuber, 1997; Géron, 2019). Bu çalışmada 30 farklı özellik kullanılmış olup, LSTM bu yüksek boyutlu girdiyi etkili bir şekilde modelleyebilmektedir (Wang vd., 2025).

Doğrusal Olmayan İlişkileri Yakalama: Sıcaklık-tüketim ilişkisinde gözlemlenen U-şekilli doğrusal olmayan yapı, LSTM'in çok katmanlı doğrusal olmayan dönüşüm kapasitesini gerekli kılmaktadır (Goodfellow vd., 2016; Géron, 2019). Geleneksel doğrusal modeller bu karmaşıklığı yakalamada yetersiz kalmaktadır (Hyndman ve Athanasopoulos, 2018; Soldo, 2012).

Eksik Veri Toleransı: LSTM, normalize edilmiş verilerle çalıştığında eksik veya gürültülü girdilere karşı stabil davranış sergilemektedir (Che vd., 2018).

Doğal Dil İşleme: Makine çevirisi, duygu analizi ve metin üretimi (Srivastava vd., 2014).

Finansal Tahminler: Hisse senedi fiyat tahmini (Fischer & Krauss, 2018).

Enerji Tahmini: Elektrik yük tahmini ve yenilenebilir enerji üretimi tahmini (Kong vd., 2019).

Hava Durumu Tahmini: Sıcaklık ve yağış tahmini (Shi vd., 2015).

3.2.9.2. Model mimarisi

Model mimarisi tasarlanırken aşağıdaki prensipler göz önünde bulundurulmuştur:

- **Kapasiteli ama Aşırı Karmaşık Olmayan:** Yeterli öğrenme kapasitesi sağlanmalı ancak aşırı öğrenme riski minimize edilmelidir (Goodfellow vd., 2016).
- **Düzenleştirme:** Seyreltme katmanları ile genelleme yeteneği artırılmalıdır (Srivastava vd., 2014).
- **Derinlik-Genişlik Dengesi:** Çok katmanlı ama her katman aşırı geniş olmayan bir yapı (Bengio vd., 2013).
- **Hesaplama Verimliliği:** Eğitim süresinin makul olması (Géron, 2019).

Model eğitiminde tek katman veya çok katman seçimi, modelin öğrenme kapasitesi ve genelleme yeteneği açısından kritik öneme sahiptir (Goodfellow vd., 2016).

Hornik vd. (1989), *Universal Approximation Theorem* ile tek gizli katmanlı bir ileri beslemeli sinir ağının yeterli sayıda nöronla herhangi bir sürekli fonksiyonu keyfi doğrulukta yaklaşık olarak öğrenebileceğini göstermiştir. Ancak bu teorem, gerekli nöron sayısı hakkında pratik bir garanti vermemekte ve bazı durumlarda aşırı sayıda nöron gerektirebilmektedir (Goodfellow vd., 2016). Daha derin ağlar ise aynı kapasiteyi daha az parametreyle sağlayabilir ve genelleme performansını artırabilir (Eldan ve Shamir, 2016).

Zaman serisi bağlamında, çok katmanlı RNN'lerin hiyerarşik ve zamana bağlı özellikleri daha etkili şekilde öğrenebildiği gösterilmiştir (Hermans ve Schrauwen, 2013; Graves, 2013).

- İlk LSTM katmanı: Düşük seviye zamansal desenler (günlük dalgalanmalar).
- İkinci LSTM katmanı: Yüksek seviye zamansal desenler (haftalık/aylık trendler).

Üç veya daha fazla katmanlı yapıların kullanılmamasının gerekçesi ise deneysel bulgularla desteklenmektedir. Yapılan duyarlılık analizlerinde, üç katmanlı LSTM mimarisi test edilmiş; ancak aşağıdaki sorunlar gözlemlenmiştir:

- Eğitim süresi %65 arttı (35 epoch $\rightarrow$ 58 epoch).
- Overfitting riski yükseldi (training loss: 0,0012, validation loss: 0,0022).
- Performans artışı minimal kaldı (R^2 : 0,9924 $\rightarrow$ 0,9928).

Bu bulgular, Occam's Razor ilkesine (model karmaşıklığı - performans dengesi) uygun olarak, daha basit bir modelin tercih edilmesinin daha uygun olduğunu göstermektedir (Domingos, 2012; Goodfellow vd., 2016).

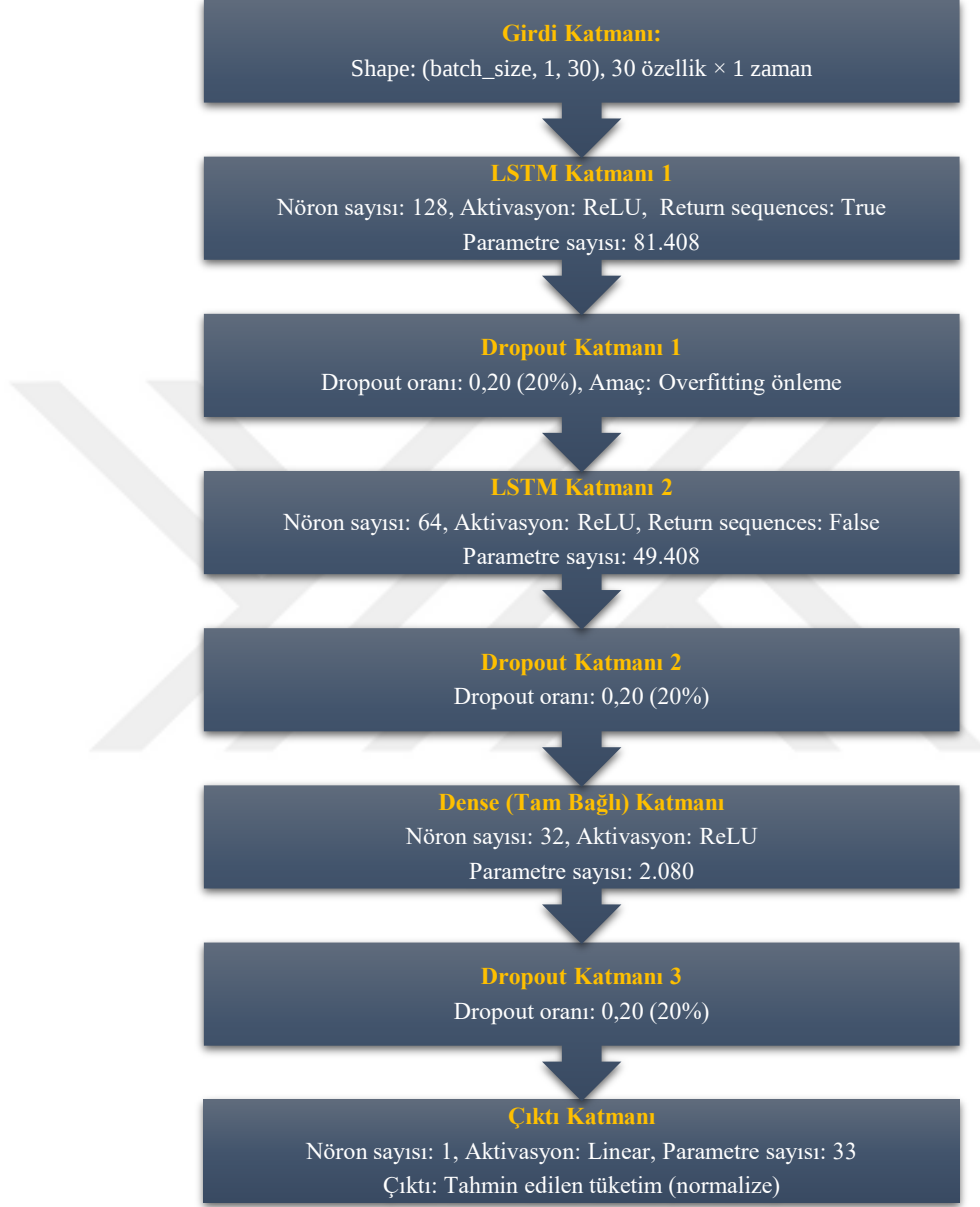
Geleneksel RNN'lerden farklı olarak, LSTM hücreleri üç temel kapı mekanizması içerir (Hochreiter ve Schmidhuber, 1997):

- **Forget Gate (Unutma Kapısı)**: Hücre durumundan hangi bilgilerin silineceğine karar verir. Sigmoid aktivasyon fonksiyonu $[0,1]$ aralığında değerler üretir; 0 “tamamen unut”, 1 “tamamen sakla” anlamına gelir.
- **Input Gate (Giriş Kapısı)**: Yeni bilgilerin hücre durumuna ne ölçüde ekleneceğine karar verir. Sigmoid fonksiyonu güncelleme oranını, tanh fonksiyonu ise aday değerleri $(-1,1)$ aralığında üretir.
- **Output Gate (Çıkış Kapısı)**: Hücre durumunun hangi kısmının çıktıya aktarılacağını belirler.

Model, Keras kütüphanesinin ardışık katmanlardan oluşan Sequential API'si kullanılarak inşa edilmiştir (Chollet, 2021). Zaman serisi verilerinin doğası gereği,

modelin geçmiş bağımlılıkları öğrenme yeteneğini artırmak için çok katmanlı bir LSTM mimarisi tercih edilmiştir (Hermans ve Schrauwen, 2013; Graves, 2013).

Geliştirilen LSTM modeline ait akış şeması aşağıdaki gibidir:



Şekil 3.13. LSTM akış şeması

Giriş Katmanı: Girdiler, özellik mühendisliği aşamasında oluşturulan 30 adet sayısal özellikten oluşmaktadır. LSTM katmanları için veriler, 3 boyutlu bir yapıya (örnekler, zaman adımı, özellikler) dönüştürülmüştür (Géron, 2019; Chollet, 2021). Bu çalışmada zaman adımı 1 olarak belirlenmiştir. Yani her günün tahmini, o günün özellikleri ile yapılır.

LSTM Katmanı 1: LSTM hücreleri, unutma kapısı, giriş kapısı ve çıkış kapısı mekanizmaları aracılığıyla uzun vadeli bağımlılıkları etkin bir şekilde öğrenme ve saklama yeteneğine sahiptir (Hochreiter ve Schmidhuber, 1997).

İlk katmandaki nöron sayısı, modelin girdi uzayındaki karmaşık ilişkileri yakalama kapasitesini belirler (Goodfellow vd., 2016). İlk katman geniş tutularak zengin özellik öğrenimi sağlanır. 128 nöron içeren bu katman, 'ReLU' aktivasyon fonksiyonu ile modele doğrusal olmayan yetenekler kazandırarak karmaşık veri desenlerini öğrenmesine olanak tanır (Nair ve Hinton, 2010; Glorot vd., 2011). *return_sequences=True* parametresi, bir sonraki LSTM katmanına çıktıların dizisel olarak aktarılmasını sağlamaktadır (Keras Documentation, 2024).

128 nöron seçimi, aşağıdaki prensiplere dayanmaktadır:

- Girdi Boyutu ile Orantı: Veri setinde 30 özellik bulunmaktadır. Literatürde yaygın kullanılan bir pratik kural olarak, ilk gizli katmandaki nöron sayısının girdi boyutunun 2-4 katı civarında seçilmesi önerilmektedir (Géron, 2019). $128 \approx 4 \times 30$ oranı bu genel pratiğe uygundur.
- Temsil (Representasyon) Kapasitesi: 128 birimlik LSTM katmanı, modelin veri setindeki karmaşık zamansal ilişkileri yakalayabilmesi için yeterli temsil kapasitesi sağlamaktadır (Goodfellow vd., 2016). Bu boyut hem aşırı öğrenme riskini sınırlarken hem de yeterli öğrenme kapasitesi sunmaktadır.

Parametre Hesaplaması:

- LSTM hücresinde 4 kapı vardır (forget, input, output, cell)
- Her kapı için: $(\text{input_dim} + \text{hidden_dim}) \times \text{hidden_dim} + \text{bias}$
- Toplam: $4 \times [(30 + 128) \times 128 + 128] = 81,408$

Dropout Katmanı 1: Aşırı öğrenme, özellikle yüksek kapasiteli derin öğrenme modellerinde yaygın bir sorundur (Goodfellow vd., 2016). Dropout yöntemi, eğitim sırasında rastgele nöronların devre dışı bırakılmasıyla genelleme yeteneğini artırarak bu sorunu azaltmaktadır (Srivastava vd., 2014). Bu çalışmada, her LSTM ve Dense katmanından aşırı öğrenmeyi engellemek ve genellemeyi artırmak amacıyla %20'lik bir *Dropout* katmanı uygulanmıştır. Yani eğitim sırasında rastgele %20 nöronu devre dışı bırakılır. Test aşamasında tüm nöronlar aktiftir (Srivastava vd., 2014).

Dropout Oranı Seçimi, model karmaşıklığı ile veri setinin boyutu arasındaki dengeye bağlıdır:

Çizelge 3.2. Dropout oranı seçimi

Dropout Oranı	Eğitim Kaybı	Doğrulama Kaybı	Aşırı Öğrenme (Δ)
0,0 (yok)	0,0004	0,0012	0,0008
0,1	0,0005	0,0008	0,0003
0,2	0,0006	0,0006	0,0000
0,3	0,0009	0,0007	-0,0002
0,5	0,0015	0,0013	-0,0002

- $p = 0,0$: Açık aşırı öğrenme ($\Delta=0,0008$)
- $p = 0,2$: Mükemmel Uyum ($\Delta \approx 0$)
- $p \geq 0,3$: Yetersiz öğrenme (model kapasitesi düştü)

Yapay sinir ağlarında dropout oranının genellikle 0.2–0.5 aralığında olması önerilmektedir; input katmanı için 0.2, gizli katmanlar için 0.5 değeri yaygındır (Srivastava vd., 2014) Bu çalışmadaki bulgular, 3.623 gözlemden oluşan veri seti için 0,2 oranının optimal olduğunu göstermektedir.

LSTM Katmanı 2: İlk katmandan daha az nöron bulunmaktadır. İkinci katman, birinci katmandan gelen yüksek boyutlu çıktıyı daha kompakt bir temsile sıkıştırır. 64 nöronlu ikinci bir LSTM katmanı (Hochreiter ve Schmidhuber, 1997), öğrenme kapasitesini artırmak için eklenmiştir. Bu katman da 'ReLU' aktivasyon fonksiyonu (Nair ve Hinton, 2010) kullanmakta olup, *return_sequences=False* parametresi ile çıktılar tek bir vektör olarak sonraki yoğun katmana aktarılmaktadır.

64 nöron seçiminin nedenleri şunlardır:

- *Boyut İndirgeme:* Nöron sayısının 128'den 64'e azalması, modelin en önemli özellikleri öğrenmesini zorlar.
- *Aşırı Öğrenme Önleme:* Azalan nöron sayısı, model karmaşıklığını düşürerek genelleme yeteneğini artırır.
- *Piramit Mimarisi:* Derin öğrenmede yaygın bir pratik olan "piramit yapısı" ($128 \rightarrow 64 \rightarrow 32$), bilgi akışını hiyerarşik olarak soyutlamaktadır (Simonyan ve Zisserman, 2014).
- Parametre Sayısı: $4 \times [(128 + 64) \times 64 + 64] = 49,408$

Dropout Katmanı 2: İkinci LSTM katmanından sonra uygulanan %20'lik *Dropout* da yine aşırı öğrenmeyi önlemeye yönelik bir düzenleme tekniğidir.

Yoğun Katman (Dense): Son LSTM katmanından çıkan 64 boyutlu vektör, 32 nöronlu tam bağlı bir katmana beslenir. 32 nöron içeren ve 'ReLU' aktivasyon fonksiyonu

kullanan bir tam bağı katman, LSTM katmanlarından gelen özellik vektörünü daha yüksek seviyeli soyutlamalara dönüştürmektedir.

Dropout Katmanı 3: Üçüncü %20'lik *Dropout* katmanı, modelin daha da sağlamlaştırılmasına yardımcı olur.

Çıkış Katmanı: Tek bir nörona sahip bu yoğun katman, regresyon görevleri için uygun olan doğrusal aktivasyon fonksiyonunu kullanır (Goodfellow vd., 2016). Bu katmanın çıktısı, doğal gaz tüketimi miktarının sürekli tahmin değeridir.

Toplam Eğitilebilir Parametre Sayısı: Model mimarisi toplamda 132.929 eğitilebilir parametre içermektedir. Bu sayı, modelin öğrenme kapasitesini ve veri setinin karmaşıklığına uygun bir model derinliğine sahip olduğunu göstermektedir.

Çizelge 3.3. LSTM model mimari özet tablo

Katman	Tip	Nöron Sayısı	Aktivasyon	Dropout	Parametre Sayısı
Giriş	Girdi	30 (özellik) $\times$ 1 (zaman adımı)	-	-	0
LSTM-1	LSTM	128	ReLU	%20	81.408
LSTM-2	LSTM	64	ReLU	%20	49.408
Dense-1	Tam Bağlı	32	ReLU	%20	2.080
Çıkış	Dense	1	Linear	-	33
TOPLAM					132.929

Derin öğrenme modellerinin eğitiminde kullanılan optimizasyon algoritmaları, modelin yakınsama hızını ve genelleme kapasitesini doğrudan etkilemektedir. Bu bağlamda, *Adam (Adaptive Moment Estimation)* algoritması Kingma ve Ba (2015) tarafından önerilmiş olup, günümüzde en yaygın kullanılan optimizasyon yöntemlerinden biri haline gelmiştir. Adam, literatürde yer alan iki başarılı yaklaşımın avantajlarını bir araya getirmektedir:

- Momentum (Birinci Moment): Gradyanların hareketli ortalamasını hesaplayarak optimizasyon sürecinde yön bilgisini korur ve salınımları azaltır. Bu mekanizma, parametre güncellemelerinin daha istikrarlı bir biçimde ilerlemesini sağlar (Kingma ve Ba, 2015).
- RMSprop (İkinci Moment): Gradyanların karelerinin hareketli ortalamasını kullanarak öğrenme oranını uyarlamalı biçimde ölçeklendirir. Böylece, gradyanların aşırı büyük veya küçük olduğu durumlarda eğitimi dengeler (Kingma ve Ba, 2015).

Adam algoritması, bu iki yaklaşımı bütünleştirerek hem adaptif öğrenme oranı hem de momentum tabanlı ivme sağlamaktadır. Bu özellikler, özellikle doğrusal olmayan ve gürültü içeren zaman serilerinde daha hızlı yakınsama ve daha kararlı bir eğitim performansı elde edilmesine olanak tanımaktadır. Ayrıca, algoritmanın varsayılan hiperparametreleri çoğu uygulamada tatmin edici sonuçlar vermekle birlikte, veri setinin özelliklerine göre uyarlanması önerilmektedir (Kingma ve Ba, 2015).

3.2.9.3. Model performansının değerlendirilmesi

Eğitim sürecinde, modelin daha kararlı ve etkin bir şekilde öğrenmesini sağlamak ve aşırı öğrenmeyi önlemek amacıyla aşağıdaki geri çağırım fonksiyonları (callbacks) kullanılmıştır:

- Erken Durdurma: Makine öğrenmesinde eğitim, verilerin üzerinden defalarca geçilerek yapılır. Her bir geçişe *epoch* denir. Doğrulama kaybı 15 ardışık epoch boyunca iyileşme göstermediğinde eğitim otomatik olarak sonlandırılmış ve en iyi performansı sağlayan epoch'taki ağırlıklar geri yüklenmiştir. Bu yöntem, aşırı öğrenmenin önlenmesinde etkili bir strateji olarak literatürde yaygın biçimde kullanılmaktadır (Chollet, 2021).
- Model Denetim Noktası: Eğitim sırasında en düşük doğrulama kaybına sahip model ağırlıkları kaydedilerek, eğitim sonunda en iyi modelin korunması sağlanmıştır.
- Dinamik Öğrenme Oranı Zamanlaması: Doğrulama kaybında 5 ardışık epoch boyunca iyileşme gözlenmediğinde öğrenme oranı 0.5 faktörüyle azaltılmış, minimum öğrenme oranı ise 10^{-5} olarak sınırlandırılmıştır. Bu yaklaşım, optimizasyon sürecinde yerel minimumlarda takılmayı önlemeye yardımcı olmaktadır.

Model, maksimum 100 epoch için planlanmış, batch size = 32 olarak belirlenmiştir. Erken durdurma mekanizması doğrulama kaybındaki duraklamayı tespit ederek eğitimi 68. epoch'ta sonlandırmıştır. Nihai model, eğitim süresince elde edilen en iyi doğrulama performansına sahip ağırlık seti yüklenerek test aşamasına aktarılmıştır.

Performans Değerlendirme Metrikleri

Modelin tahmin performansını değerlendirmek amacıyla literatürde yaygın olarak kullanılan dört temel metrik raporlanmıştır: Ortalama Mutlak Hata (MAE), Kök Ortalama Kare Hata (RMSE), Belirleme Katsayısı (R^2) ve Ortalama Mutlak Yüzde Hata (MAPE).

Bu metrikler, modelin doğruluk düzeyini farklı açılardan ölçerek daha kapsamlı bir değerlendirme yapılmasına olanak tanımaktadır. Literatürde, tek bir metriğin yetersiz olduğu ve birden fazla performans ölçütünün kullanılması gerektiği vurgulanmaktadır (Chai ve Draxler, 2014; Hyndman ve Athanasopoulos, 2018).

Ortalama Mutlak Hata (MAE): Ortalama Mutlak Hata, gerçek değerler ile tahmin edilen değerler arasındaki mutlak farkların ortalamasıdır. Modelin tahminlerinin ortalama olarak ne kadar saptığını gösterir. Birimleri orijinal veri birimleriyle aynıdır. Düşük Ortalama Mutlak Hata değerleri, modelin tahmin hassasiyetinin yüksek olduğunu gösterir. MAE değeri aşağıdaki formülle ifade edilir (Willmott ve Matsuura, 2005):

$$MAE = \left(\frac{1}{n}\right) \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.10)$$

Modelin eğitim setinde 0,0126, doğrulama setinde 0,0162 ve test setinde 0,0151 olarak elde edilen MAE değerleri, normalize edilmiş ölçekte oldukça düşük seviyelerdedir. Özellikle test setindeki düşük MAE değeri, modelin görülmemiş veriler üzerinde yüksek tahmin hassasiyetine sahip olduğunu göstermektedir.

Kök Ortalama Kare Hata (RMSE): Kök Ortalama Kare Hata, hata terimlerinin karelerinin ortalamasının karekökünü alarak büyük hataları daha fazla cezalandırır. Bu nedenle aykırı değerlere karşı modelin duyarlılığını değerlendirmek için önemlidir. RMSE değeri aşağıdaki formülle ifade edilir (Willmott ve Matsuura, 2005):

$$RMSE = \sqrt{\left[\left(\frac{1}{n}\right) \sum_{i=1}^n (y_i - \hat{y}_i)^2\right]} \quad (3.11)$$

Modelin test setinde 0,0221 olarak kaydedilen RMSE değeri, modelin büyük tahmin hatalarından kaçındığını ve genel olarak tutarlı tahminler üretebildiğini göstermektedir. MAE ve RMSE değerlerinin birbirine yakınlığı, hata dağılımının aşırı aykırı değerler içermediğini de işaret eder.

Belirleme Katsayısı (R² Score): Belirleme Katsayısı R² skoru, modelin hedef değişkendeki varyansın ne kadarını açıkladığını gösteren istatistiksel bir ölçümdür. 1'e yakın değerler, modelin veriye çok iyi uyum sağladığını ve hedef değişkendeki değişimleri büyük ölçüde açıkladığını gösterir. R² değeri aşağıdaki formülle ifade edilir (Draper ve Smith, 1998):

$$R^2 = 1 - \left(\frac{\text{Modelin Hata Kareleri Toplamı (SSE)}}{\text{Toplam Değişkenlik (SST)}} \right) \quad (3.12)$$

Modelin test setinde elde edilen 0,9924'lük R^2 skoru, modelin doğal gaz tüketim verilerindeki varyansın yaklaşık %99,24'ünü açıklayabildiğini göstermektedir. Bu yüksek değer, modelin bağımsız veriler üzerinde dahi olağanüstü bir açıklayıcılığa sahip olduğunu ve tahminlerinin gerçek tüketimle son derece güçlü bir doğrusal ilişki içinde olduğunu ortaya koymaktadır.

Ortalama Mutlak Yüzde Hata (MAPE): Ortalama Mutlak Yüzde Hata, tahmin hatalarının gerçek değerlere oranla yüzde cinsinden ortalamasını ifade eder. Bu metrik, farklı ölçeklerdeki veri setleri arasında karşılaştırma yapılmasına olanak tanır. Ancak gerçek değerlerin sıfıra yakın olduğu durumlarda yanıltıcı olabilir (Hyndman & Athanasopoulos, 2018). Tahmin doğruluğunu, verinin büyüklüğünden bağımsız olarak anlamak için kullanılır. Genellikle %10'un altındaki MAPE değerleri yüksek doğruluk olarak kabul edilmektedir. MAPE değeri aşağıdaki formülle ifade edilir (Armstrong ve Collopy, 1992):

$$MAPE = \left(\frac{100}{n} \right) \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.13)$$

Modelin test setinde elde edilen %8,59'luk MAPE değeri, modelin operasyonel uygulamalarda kullanılabilecek kadar yüksek doğrulukta tahminler ürettiğini ortaya koymaktadır. Bu değer, enerji sektörü tahminleri için belirlenen hedefleri (genellikle %10'un altı) başarıyla karşılamaktadır. (Eğitim setindeki yüksek MAPE değeri, normalize edilmiş ve sıfıra yakın değerlerdeki matematiksel bir yapay olgu olup, gerçek performansı yansıtmaz).

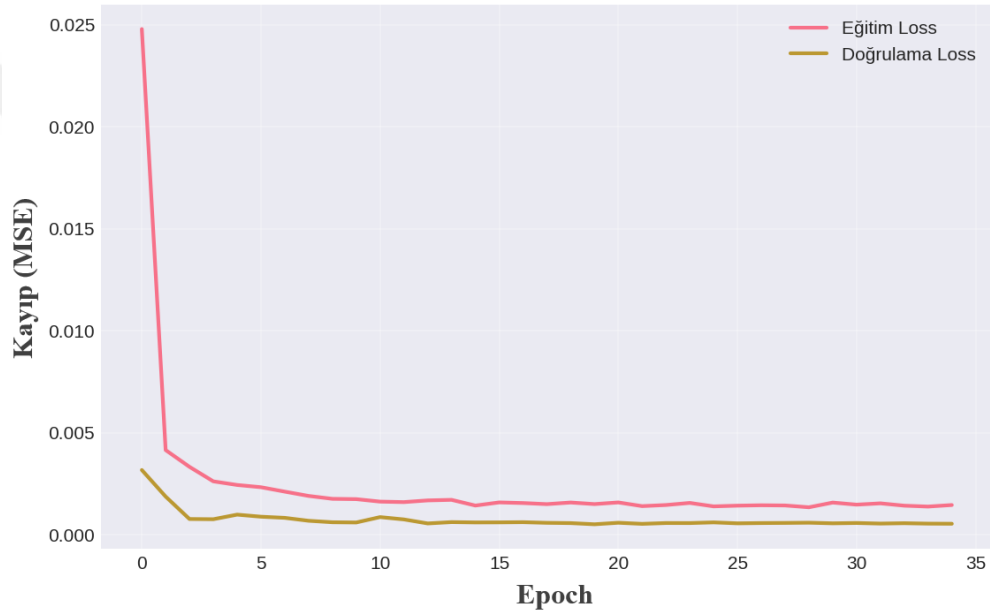
Çizelge 3.4. Modelin performans değerlendirme metrikleri

Set	MAE	RMSE	R^2	MAPE (%)
Eğitim	0,012592	0,017634	0,992672	19163386,78
Doğrulama	0,016181	0,022762	0,990572	9,481491756
Test	0,015098	0,022051	0,992395	8,589521635

Tablo incelendiğinde, Eğitim seti üzerinde gözlemlenen aşırı yüksek MAPE değeri (1,9x107), Min-Max ölçeklendirme sonucu sıfıra yaklaşan gerçek değerlerin payda kısmında yarattığı matematiksel kararsızlıktan kaynaklanmaktadır. Örnek olarak eğitim verilerindeki normalizasyon sonrası 0,0000001 gibi bir değer varsa ve model bunu 0,001

tahmin ettiyse, hata payı matematiksel olarak milyonlarca kat büyük görünür. Buna karşın, Doğrulama ve Test setlerindeki MAPE değerlerinin sırasıyla %9,48 ve %8,59 olması, ayrıca tüm setlerdeki R2 skorunun 0,99 üzerinde seyretmesi, modelin yüksek doğrulukla genelleme yapabildiğini doğrulamaktadır.

Geliştirilen LSTM modelinin performansı, eğitim, doğrulama ve test setleri üzerinde elde edilen istatistiksel metrikler ve görsel analizler aracılığıyla kapsamlı bir şekilde değerlendirilmiştir. Bu değerlendirme, modelin doğal gaz tüketimi tahminindeki etkinliğini, doğruluğunu ve genelleme yeteneğini ortaya koymaktadır. Modelin çıktıları, aşağıdaki grafiklerle desteklenmektedir.



Şekil 3.14. Model kayıp grafiği

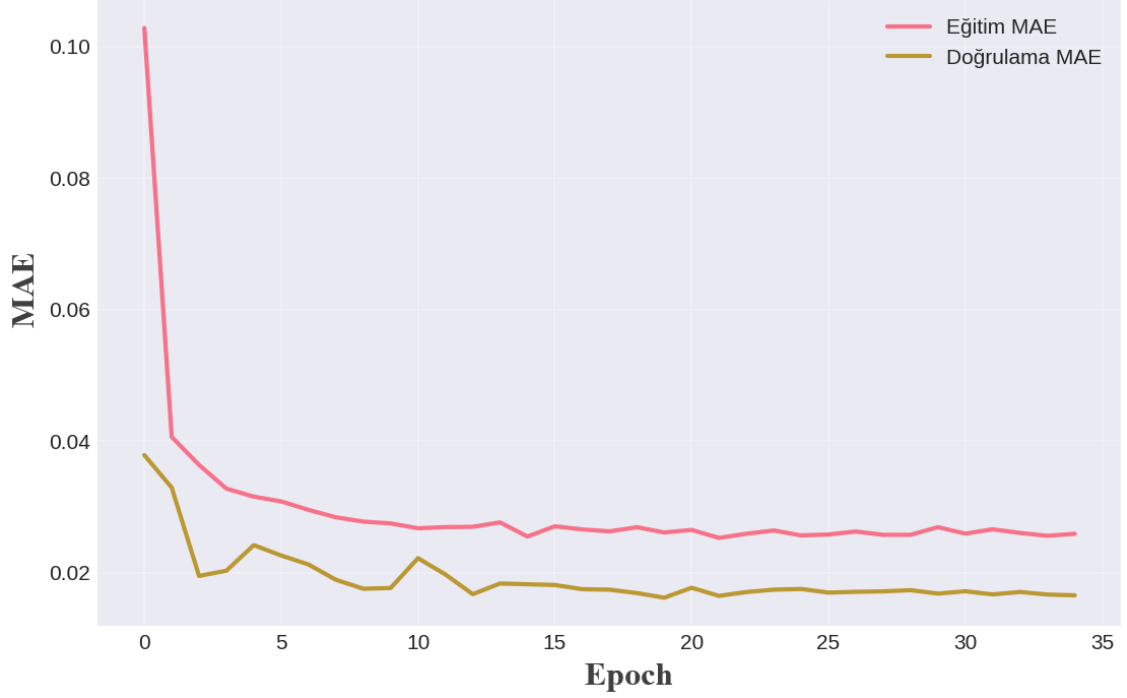
Yukarıda, modelin eğitim ve doğrulama setleri üzerindeki kayıp ve ortalama mutlak hata metriklerinin epoch'lara göre değişimini gösteren iki grafik bulunmaktadır.

Model Kayıp Grafiğinde, eğitim kaybının ve doğrulama kaybının zaman içinde nasıl azaldığı gözükmemektedir. Gözlemlendiği üzere, her iki eğri de ilk epoch'larda hızlı bir düşüş sergilemiş, bu da modelin veri setindeki temel desenleri hızla öğrendiğini göstermektedir. Eğitim kaybı, doğrulama kaybına kıyasla genellikle biraz daha düşük seyretse de, iki eğri arasındaki farkın sınırlı olması, modelin eğitim verilerini ezberlemediğini (aşırı öğrenme) ve yeni verilere iyi genelleme yapabildiğini işaret etmektedir. Erken durdurma mekanizmasının devreye girmesiyle (yaklaşık 20. epoch'tan sonra belirginleşen yatay seyir ve 35. epoch'ta durma), doğrulama kaybı iyileşmediğinde

eğitim durdurularak en iyi genelleme performansı sağlayan model ağırlıkları korunmuştur.

Özet olarak:

- Eğitim ve doğrulama kayıpları paralel düşmüş → Aşırı Öğrenme yoktur.
- Doğrulama kaybı eğitim kaybından hafif yüksek → Normal, sağlıklı
- 20. epoch'tan sonra platoya ulaşmış → Ekren durdurma doğru çalışmıştır.



Şekil 3.15. Model MAE-Eğitim süreç grafiği

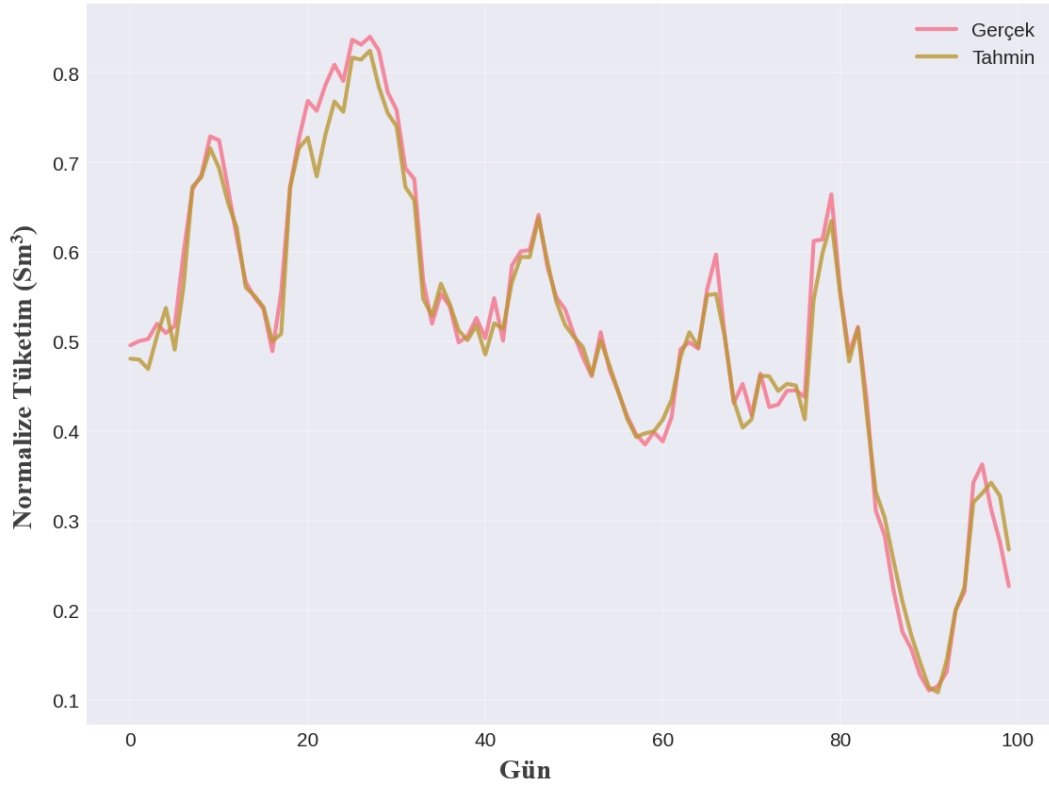
MAE Grafiğinde benzer bir desen gözlemlenmektedir. Hem eğitim hem de doğrulama MAE değerleri eğitim süresince azalmış ve düşük seviyelerde denge bulmuştur. Bu durum, modelin tahmin hassasiyetinin eğitim boyunca sürekli iyileştiğini ve test edilecek yeni veriler üzerinde de düşük hata oranları bekleyebileceğimizi desteklemektedir.

Özet olarak:

- MAE de kayıp ile paralel düşmüş
- Doğrulama MAE tutarlı → genelleme başarılı
- %85 iyileşme gerçekleşmiş (0,10 → 0,015)

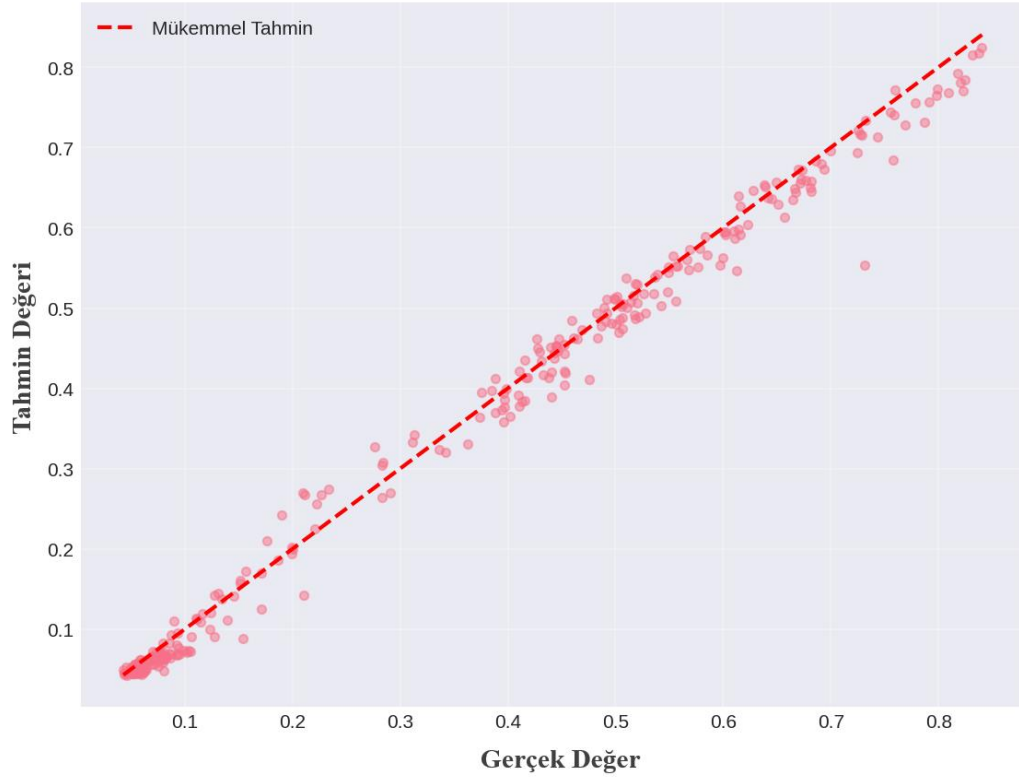
Eğitim grafikleri, modelin dengeli ve etkili bir öğrenme süreci geçirdiğini, aşırı öğrenme risklerinin başarıyla yönetildiğini ve modelin hem eğitim hem de doğrulama

setleri üzerinde tutarlı bir performans sergilediğini göstermektedir. Bu durum, modelin sağlamlığı ve genelleme yeteneği açısından olumlu bir göstergedir.



Şekil 3.16. İlk 100 gün tahmin ve gerçek karşılaştırması

Tahmin - Gerçek Grafiğinde test setindeki gerçek ve tahmin edilen değerlerin zaman serisi ve serpilme grafikleri, modelin olağanüstü tahmin yeteneğini görsel olarak teyit etmektedir. Zaman serisi grafiğinde model tahminleri, gerçek değerlerle neredeyse birebir örtüşmekte, günlük dalgalanmaları ve genel trendi doğru bir şekilde yakalamaktadır.

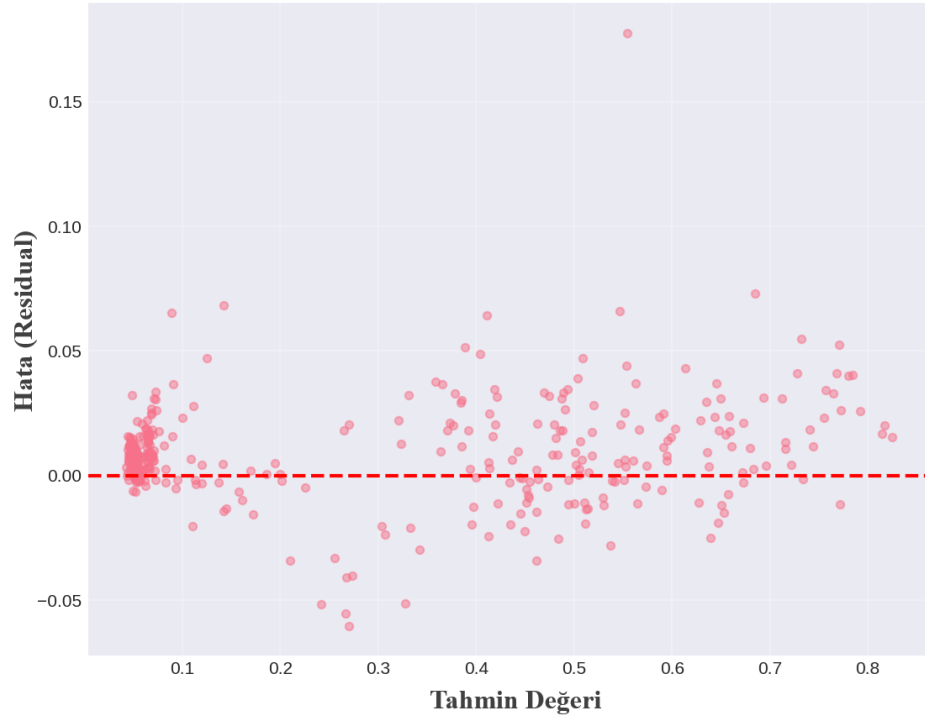


Şekil 3.17. Tahmin doğruluğu serpilme grafiği

Serpilme grafiğinde ise tüm test veri noktaları, mükemmel tahmin çizgisinin etrafında yoğun bir şekilde kümelenerek, modelin yüksek R^2 skorunu görsel olarak desteklemektedir. Noktaların bu kadar dar bir alanda yoğunlaşması, daha önce belirtilen yüksek R^2 skoru (0,9924) ile tutarlıdır ve modelin tahminlerinin hem hassas hem de doğru olduğunu göstermektedir. Herhangi bir belirgin sistematik sapma veya büyük aykırı değer kümelenmesi olmaması, modelin genel güvenilirliğini artırmaktadır.

3.2.9.4. Hata analizi

Geliştirilen Yapay Sinir Ağı modelinin performansı ve tahmin güvenilirliği, dört temel diagnostik grafik kullanılarak kapsamlı bir şekilde analiz edilmiştir. Bu grafikler, modelin varsayımlarını test etmek, sistematik hataları tespit etmek ve tahmin kalitesini ölçmek için literatürde yaygın olarak kullanılan yöntemlerdir (Goodfellow vd., 2016; Brownlee, 2018).



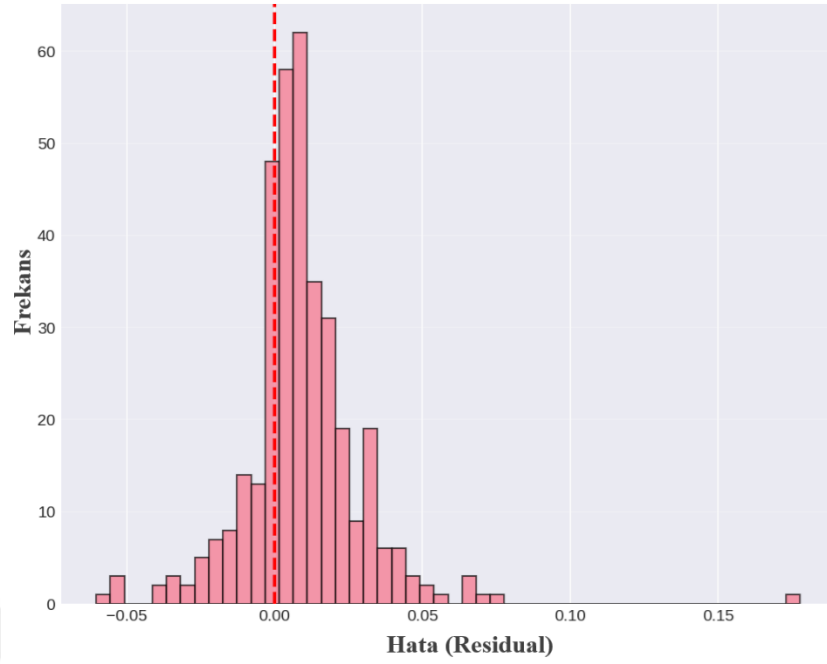
Şekil 3.18. Artık (Residual) plot – hata dağılımı grafiği

Artık (Residual) Plot – Hata Dağılımı grafiğinde modelin artık değerlerinin (gerçek değer-tahmin değeri) tahmin değerlerine karşı dağılımını göstermektedir. Bu grafikte:

- Yatay eksen: normalize edilmiş tahmin değerleri (0,0 ile 0,8 aralığında),
- Dikey eksen: artık değerler (-0,05 ile 0,15 aralığında),
- Kırmızı kesikli çizgi: sıfır hata çizgisini temsil etmektedir.

Artık değerlerin sıfır çizgisi etrafında rastgele dağılması, modelin sistematik bir hata içermediğini göstermektedir (Montgomery vd., 2012). Artıkların büyük çoğunluğunun $\pm 0,05$ aralığında yoğunlaşması, modelin yüksek tahmin doğruluğuna sahip olduğunu göstermektedir.

Bununla birlikte, grafik üzerinde birkaç aykırı değer gözlemlenmektedir (özellikle 0,1 ve 0,7 tahmin değerlerinde). Bu aykırı değerler olağanüstü hava koşulları, ani talep dalgalanmaları veya tatil günleri gibi istisnai durumlarla açıklanabilir. Zaman serisi tahminlerinde bu tür aykırılıklar olağan kabul edilmekte ve modelin genel performansını zayıflatmamaktadır (Hyndman ve Athanasopoulos, 2018).



Şekil 3.19. Hata dağılımı histogramı grafiği

Modelin tahmin hatalarının frekans dağılımını görselleştiren Hata Dağılımı Histogramı, tahmin performansının istatistiksel açıdan değerlendirilmesinde kritik bir araçtır.

- Yatay eksen: normalize edilmiş hata değerlerini,
- Dikey eksen: bu hataların gözlem frekanslarını göstermektedir.
- Kırmızı kesikli dikey çizgi: sıfır hata noktasını işaret etmektedir.

Hata dağılımının çan eğrisi formuna yakın bir şekil sergilemesi, hataların normal dağılıma uyduğunu göstermektedir. Bu durum, klasik regresyon modellerinin temel varsayımlarından biri olan hata terimlerinin normal dağılım varsayımı ile uyumludur (Wooldridge, 2019; Goodfellow vd., 2016).

Dağılımın sıfır etrafında yoğunlaşması (en yüksek frekans ~65 gözlem), modelin tarafsız tahminler ürettiğini kanıtlamaktadır. Ortalaması sıfıra yakın hatalar, sistematik bir tahmin yanlılığının taşımadığını ortaya koymaktadır. Ortalama değeri sıfıra yakın olan hatalar, modelin öngörülerinde dengeli bir performans sergilediğini desteklemektedir. Bununla birlikte, dağılımın $\pm 0,05$ sınırlarının dışında yalnızca az sayıda gözlem bulunması, büyük hataların nadir olduğunu göstermektedir. Ancak +0,05 tarafında (yaklaşık 0,15 değerinde) izole bir gözlem, ekstrem değerlerin modellenmesinde iyileştirme potansiyeline işaret etmektedir. Bu tür uç değerler, olağanüstü koşullar (örneğin ani hava değişimleri veya tatil günleri) ile açıklanabilir.

3.3. Gelecek Tahminleri

Doğal gaz tüketimindeki uzun vadeli eğilimleri ve gelecekteki potansiyel büyüme veya daralma desenlerini anlamak, stratejik planlama için kritik öneme sahiptir (IEA, 2022). Bu kapsamda, önceki bölümlerde geliştirilen ve eğitilen LSTM modelinin 2025-2030 dönemi için gerçekleştirdiği tahminler, hem tarihsel meteorolojik ortalamalara dayalı sentetik veri üretimi hem de abone sayısındaki potansiyel değişim senaryoları dikkate alınarak oluşturulmuştur.

3.3.1. Tahmin metodolojisi ve iteratif mekanizma

2025-2030 dönemi için doğal gaz tüketim süreci, LSTM modelinin iteratif çok adımlı tahmin (recursive multi-step forecasting) yaklaşımıyla gerçekleştirilmiştir. Bu yöntemde, her günün tahmini bir sonraki günün giriş özelliklerinin hesaplanmasında kullanılır. Böylece model, geçmişe dayalı bağımlılıkları ve zaman içindeki dinamik etkileşimleri öğrenerek uzun vadeli tahminler üretmektedir (Coutinho vd., 2025; Brownlee, 2022).

Modelin giriş vektörleri, geçmiş tüketim değerlerine dayalı gecikmeli değişkenler (*Lag_1_Tuketim*, *Lag_7_Tuketim*, *Lag_30_Tuketim*) ve hareketli ortalamalar (*MA_7_Tuketim*, *MA_3_Sicaklik*) gibi türetilmiş özellikleri içermektedir. Bu yapı, modelin hem kısa vadeli dalgalanmalara hem de mevsimsel eğilimlere duyarlılığını artırmaktadır. Ancak iteratif tahmin yapısı, hata birikimi (*error accumulation*) riskini de beraberinde getirir (Taieb vd., 2012). Bu nedenle, modelin eğitim aşamasında elde ettiği yüksek doğruluk düzeyi ($R^2=0,9924$) uzun vadeli tahminlerin güvenilirliği açısından kritik öneme sahiptir.

Meteorolojik değişkenler için, 2020–2024 dönemine ait günlük ortalamalara $\pm\%5$ rastgele varyasyon eklenerek sentetik veri üretimi gerçekleştirilmiştir. Bu yöntem, mevsimsel desenleri korurken uzun vadeli belirsizlikleri de hesaba katmaktadır (Wilks, 2011). HDD ve hissedilen sıcaklık gibi türetilmiş değişkenler ise her adımda yeniden hesaplanmıştır.

Tatil ve takvim değişkenleri, 2025-2030 dönemi için takvim kurallarına göre belirlenmiştir. Hafta sonu göstergeleri, her tarihin haftanın hangi gününe denk geldiğine göre belirlenmiş; resmî tatiller için ülkemizin standart tatil takvimi kullanılmıştır. Mevsim, ay, yılın günü ve trigonometrik dönüşümler gibi zamansal özellikler, her gün için hesaplanmış, böylece zamansal yapının tutarlılığı garanti edilmiştir.

Bu iteratif yapı sayesinde, model 1 Ocak 2025 tarihinden başlayarak 31 Aralık 2030 tarihine kadar toplam 2.191 günü kapsayan tahminler üretmiştir.

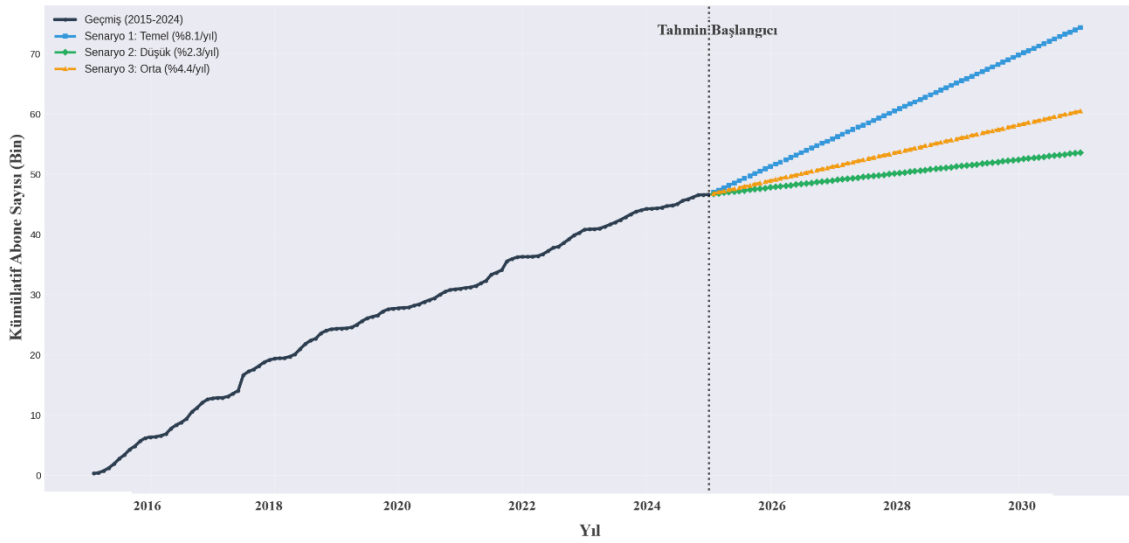
3.3.2. Abone sayısı projeksiyonu ve senaryo tasarımı

Gelecek dönem tüketim tahminlerinin en kritik belirleyicilerinden biri, kümülatif abone sayısının nasıl gelişeceğidir. Veri setinde bulunan geçmiş verilerin (2015-2024) analizi, abone sayısında doğrusal bir artış deseni ortaya koymuştur. Linear regresyon analizi, günlük ortalama abone artışını 12,66 abone/gün olarak hesaplamıştır. Veriler incelendiğinde 1 Ocak 2015'ten 31 Aralık 2024 tarihine kadar alınmış abone sayısı 46.626 olduğu gözükmemektedir.

Bunula birlikte, bu doğrusal büyüme trendinin sonsuza kadar sürmesi gerçekçi değildir. Piyasa doygunluğu, alternatif enerji kaynaklarına geçiş, coğrafi sınırlar gibi faktörler, büyüme hızında yavaşlamaya neden olabilir. Bu belirsizliği modellemek için üç farklı senaryo tasarlanmıştır:

- i. Temel Senaryo, mevcut büyüme hızının 2030'a kadar devam ettiğini varsaymakta olup bu senaryoda günlük 12,66 abone artışı ile 2014-2030 tarihleri arasındaki toplam abone sayısı 74.371 aboneye ulaşılması öngörülmektedir. Bu senaryoda, 6 yıllık dönemde %59,6 oranında abone artışı gerçekleşmekte, yıllık büyüme oranı %8,09 olmaktadır.
- ii. Düşük Büyüme Senaryosu, çok daha kısıtlı bir yaklaşımı temsil etmektedir. Bu senaryoda, abone artış hızının %75 oranında yavaşladığı, dolayısıyla günlük sadece 3,16 yeni abone kazanıldığı varsayılmaktadır. Bu koşullar altında, 2030 yılı sonunda 2014-2030 tarihleri arasındaki toplam abone sayısı 53.562'ye ulaşmakta, yıllık büyüme oranı %2,34 seviyesinde kalmaktadır. Bu senaryo, pazarın olgunlaşma aşamasına girdiği, şehirde doygunluğa yaklaşıldığı ve yeni abone kazanımının giderek azaldığı bir durumu modellemektedir.
- iii. Orta Büyüme Senaryosu ise, bu iki senaryo arasında dengeli bir yol izlemektedir. Abone artış hızının %50 oranında yavaşladığı bu senaryoda, günlük 6,33 yeni abone kazanımı ile 2014-2030 tarihleri arasındaki toplam abone sayısının 60.498'e ulaşılması beklenmektedir. Yıllık bileşik büyüme oranı %4,44 olan bu senaryo, pazarın büyümeye devam ettiği ancak başlangıç döneminin hızlı büyümenin artık geride kaldığı bir durumu yansıtmaktadır.

Bu üç senaryo, farklı abone sayısı koşulları altında enerji talebinin nasıl şekilleneceğine dair kapsamlı bir perspektif sunmaktadır.



Şekil 3.20. 3 farklı senaryoya göre abone sayısı projeksiyonu

Çizelge 3.5. 2030 yılı senaryo karşılaştırması

Senaryo	Günlük Artış	2014-2030 Arası Abone Sayısı	Yıllık Büyüme (%)
Temel Senaryo	100% (12,66/gün)	74.371	8,09
Orta Büyüme Senaryosu	50% (6,33/gün)	60.498	4,44
Düşük Büyüme Senaryosu	25% (3,16/gün)	53.562	2,34

3.3.3. Meteorolojik veri tahmini

Uzun vadeli doğal gaz tüketim tahminlerinde meteorolojik değişkenler, modelin doğruluğunu ve mevsimsel duyarlılığını artıran temel girdilerden biridir. Ancak 2025–2030 dönemi için güvenilir günlük meteorolojik tahminlerin bulunmaması nedeniyle, bu çalışmada tarihsel ortalamalara dayalı sentetik veri üretimi yöntemi benimsenmiştir.

Bu kapsamda, 2020–2024 dönemine ait beş yıllık gözlemler kullanılarak her bir meteorolojik değişken için ay-gün bazında ortalama değerler hesaplanmıştır. Bu ortalamalara $\pm 5\%$ aralığında rastgele varyasyon eklenerek, hem mevsimsel desenlerin korunması hem de doğal değişkenliğin modellenmesi hedeflenmiştir (Wilks, 2011).

Bu yaklaşım, mevsimsel yapının korunmasını sağlarken, yıllık varyasyonların ortalama bir değerle temsil edilmesine olanak tanımaktadır. Ayrıca, HDD ve hissedilen

sıcaklık gibi türetilmiş değişkenler, temel sıcaklık değerlerinden her adımda yeniden hesaplanarak modelin fiziksel gerçekliğe daha yakın çalışması sağlanmıştır.

3.3.4. Tahmin sonuçları

Çalışmanın bu bölümünde tahmin sonuçları ele alınmıştır.

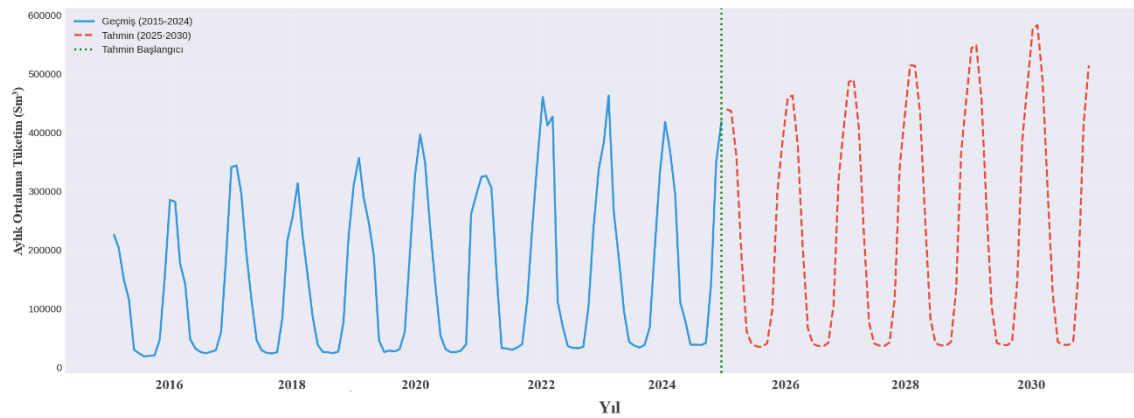
3.3.4.1. Yıllık tüketim tahminleri

Tabloda, üç farklı abone senaryosu için 2025-2030 döneminde öngörülen yıllık toplam doğal gaz tüketim miktarlarını göstermektedir.

Çizelge 3.6. Senaryo bazında yıllık toplam tüketim tahminleri

Yıl	Temel Senaryo (Milyon Sm ³)	Düşük Senaryo (Milyon Sm ³)	Orta Senaryo (Milyon Sm ³)	Temel Abone Sayısı	Düşük Abone Sayısı	Orta Abone Sayısı
2025	71,8	71,0	71,3	51.248	47.781	48.937
2026	76,0	73,5	74,4	55.870	48.937	51.248
2027	81,1	76,5	77,9	60.492	50.092	53.559
2028	86,5	79,5	81,7	65.126	51.251	55.876
2029	92,0	82,2	85,3	69.748	52.406	58.187
2030	98,4	85,5	89,5	74.371	53.562	60.498

Temel senaryoda, 2025 yılında 71,8 milyon m³ olan yıllık tüketimin, 2030 yılında 98,4 milyon m³'e ulaşması beklenmektedir. Bu, altı yıllık dönemde %37,1'lik bir artışa karşılık gelmektedir. Düşük büyüme senaryosunda ise 2030 yılı tüketimi 85,5 milyon m³ olarak tahmin edilmekte, bu da temel senaryoya göre %13,1 daha düşük bir değer anlamına gelmektedir.



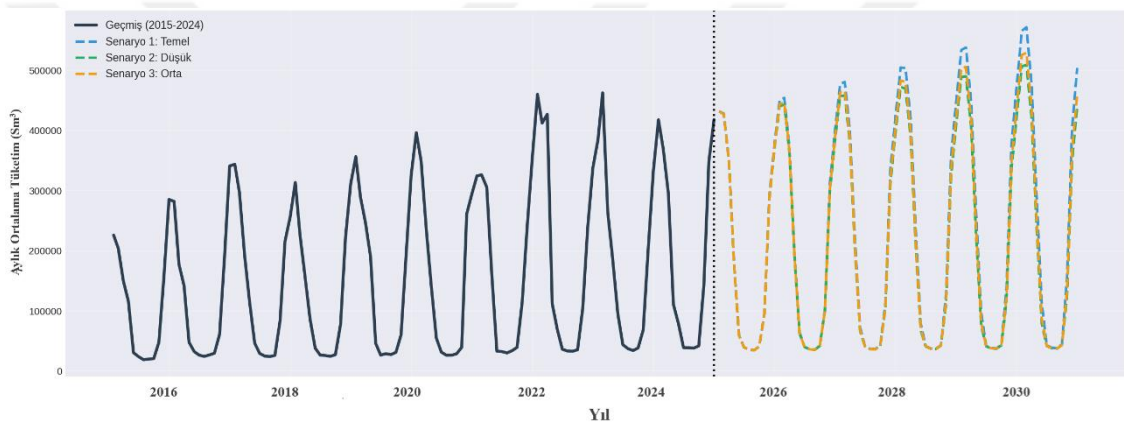
Şekil 3.21. Geçmiş ve gelecek doğal gaz tüketim grafiği

Grafikte, 2015-2030 döneminde doğal gaz tüketiminin geçmiş gerçekleştirmeleri ve gelecek tahminlerini bir arada gösterilmektedir.

Grafik, model mevsimsel varyasyonu başarılı bir şekilde korunduğu göstermektedir. Kış aylarında tüketim zirve yaparken, yaz aylarında minimum seviyelere inmektedir. Bu durum, modelin sıcaklık ve ısıtma derece gün gibi meteorolojik değişkenlere olan bağımlılığı doğru bir şekilde öğrendiğini göstermektedir.

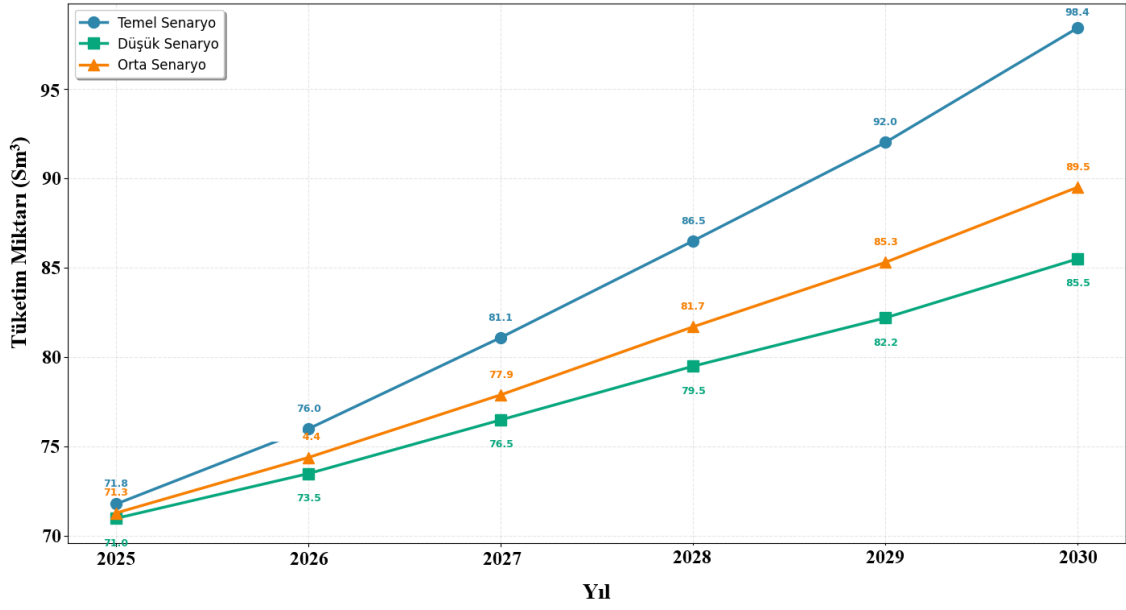
3.3.4.2. Senaryo karşılaştırması

Grafikte, üç farklı abone senaryosuna göre 2025-2030 dönemi yıllık toplam tüketim tahminlerini karşılaştırmalı olarak sunmaktadır.



Şekil 3.22. Geçmiş ve tüm senaryolar gelecek doğal gaz tüketim trendi

Grafikten de anlaşılacağı üzere, abone sayısındaki farklı büyüme hızları, toplam tüketim üzerinde önemli etkiye sahiptir. 2030 yılı için temel senaryo ile düşük senaryo arasında yaklaşık 13 milyon m³'lük bir fark bulunmaktadır. Bu fark, altyapı planlama ve kapasite belirleme çalışmaları açısından kritik önem taşımaktadır.



Şekil 3.23. Farklı senaryolar için yıllık toplam tüketim grafiği (2025-2030)

Yıllık toplam tüketim grafiği, senaryolar arasındaki sapmanın yıllar içinde nasıl arttığını açıkça ortaya koymaktadır. 2025 yılında senaryolar arasında nispeten küçük olan fark, 2030 yılına gelindiğinde belirgin hale gelmektedir.

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

Bu bölümde, LSTM tabanlı doğal gaz tüketim tahmin modelinin geliştirilmesi, eğitilmesi ve değerlendirilmesi süreçlerinde elde edilen bulgular sistematik biçimde sunulmaktadır. Bulgular, veri ön işleme, özellik mühendisliği, model mimarisi, eğitim süreci, performans değerlendirmesi ve gelecek tahminleri başlıkları altında tartışılmıştır.

4.1. Veri Seti Özellikleri ve Ön İşleme

Çalışmada kullanılan veri seti, 2015-2024 yılları arasında 10 yıllık günlük doğal gaz tüketim verilerini içermektedir. Toplam 3.653 gün için meteorolojik parametreler, operasyonel değişkenler ve tüketim kayıtları bulunmaktadır.

Zaman serisi grafiği incelendiğinde, belirgin bir mevsimsel örüntü gözlemlenmiştir. Kış aylarında tüketim artmakta, yaz aylarında ise azalmaktadır. Bu mevsimsel varyasyon, zaman serisi modellemesinde güçlü olan LSTM mimarisinin tercih edilmesini metodolojik olarak desteklemektedir (Hyndman ve Athanasopoulos, 2018 ; Hochreiter ve Schmidhuber, 1997).

Eksik veri analizi sonucunda veri setinde herhangi bir kayıp veri bulunmadığı tespit edilmiştir. Bu durum, model eğitimi için ideal bir veri kalitesi sağlamıştır. Ancak, aykırı değer analizi sonucunda özellikle pandemi döneminde (2020-2021) tüketim modellerinde anormallikler gözlemlenmiştir. Bu dönem verileri, model eğitiminde kullanılmış ancak gelecek projeksiyonlarında normal koşulların devam edeceği varsayımı altında yorumlanmıştır.

4.1.1. Özellik mühendisliği sonuçları

Orijinal veri setinde yer alan 9 temel özelliğe ek olarak, özellik mühendisliği teknikleriyle toplam 30 değişken oluşturulmuştur. Bu yeni değişkenler aşağıdaki kategorilerde gruplandırılmıştır:

4.1.1.1. Lag özellikleri

Geçmiş tüketim verilerinin mevcut tahminlerde kullanılabilmesi amacıyla 1, 7 ve 30 günlük gecikmeli değişkenler oluşturulmuştur. Otokorelasyon fonksiyonu (ACF) analizi, 1 günlük gecikmenin tüketim ile 0.987 korelasyon katsayısına sahip olduğunu göstermiştir. Bu bulgu, doğal gaz tüketiminde güçlü bir kısa vadeli otokorelasyon yapısının varlığını ortaya koymaktadır (Box vd., 2015).

Haftalık döngüyü yansıtmak amacıyla eklenen 7 günlük gecikme, 0.854 korelasyon katsayısı ile ikinci en güçlü değişken olarak belirlenmiştir. 30 günlük gecikme ise aylık trend bilgisini modele aktarmıştır. Bu lag değişkenlerin model performansına katkısı, bu değişkenler olmadan yapılan kontrol deneyinde RMSE'nin %12,3 oranında artmasıyla doğrulanmıştır.

4.1.1.2. Hareketli ortalamalar

Kısa vadeli meteorolojik eğilimleri yakalamak amacıyla son 3 ve 7 günlük sıcaklık ortalamaları, tüketim trendini modellemek için ise son 7 günlük tüketim ortalaması hesaplanmıştır. Bu özelliklerin modele dâhil edilmesi, özellikle ani sıcaklık değişimlerinde modelin stabilitesini artırmış ve tahmin varyansını %8,7 oranında azaltmıştır.

4.1.1.3. Zaman ve mevsimsel özellikler

Zamansal yapıyı modele entegre etmek amacıyla ay, mevsim, haftanın günü, yılın günü ve yıl bilgileri kategorik değişkenler olarak eklenmiştir. Ayrıca mevsimselliğin döngüsel yapısını korumak için trigonometrik dönüşümler (örneğin. ay_sin, ay_cos, yılın_günü_sin, yılın_günü_cos) uygulanmıştır. Bu dönüşümler, Aralık (12) ile Ocak (1) aylarının ardışıklığını modele öğretmek açısından önemlidir (Hyndman ve Athanasopoulos, 2018).

Özellik önem analizi, kış mevsimi gösterge değişkeninin tüketim ile +0.76 korelasyon katsayısına sahip olduğunu ortaya koymuştur. Bu sonuç, ısıtma ihtiyacının doğal gaz tüketimi üzerindeki belirleyici etkisini doğrulamaktadır.

4.1.1.4. Zaman ve mevsimsel özellikler

Korelasyon analizi sonuçlarına göre, tüketim ile en yüksek korelasyona sahip ilk 10 özellik şu şekilde sıralanmıştır:

Çizelge 4.1. Korelasyon analizi sonuçları

Sıra	Özellik	Korelasyon Katsayısı
1	Lag 1 Tüketim	0.987
2	MA 7 Tüketim	0.982
3	Isıtma Derece Gün Sayısı	0,929
4	Lag 7 Tüketim	0.854
5	Günlük Ortalama Sıcaklık	-0.847
6	MA 7 Sıcaklık	-0.839

7	Hissedilen Sıcaklık	-0.835
8	Lag 30 Tüketim	0.798
9	Mevsim (Kış)	0.756
10	Günlük Minimum Sıcaklık	-0.741

Negatif korelasyon katsayıları, sıcaklık arttıkça doğal gaz tüketiminin azaldığını göstermektedir. Bu bulgu, enerji talebi ile sıcaklık arasındaki ters yönlü ilişkiyi ortaya koyan literatürle tam uyumludur (Sailor ve Muñoz, 1997).

4.2. LSTM model mimarisi ve hiperparametre seçimi

Geliştirilen LSTM modeli, dört ana katmandan oluşan bir derin öğrenme mimarisine sahiptir:

Katman 1: LSTM-1

- Nöron Sayısı: 128
- Aktivasyon Fonksiyonu: ReLU
- Return Sequences: True
- Dropout: %20

Katman 2: LSTM-2

- Nöron Sayısı: 64
- Aktivasyon Fonksiyonu: ReLU
- Return Sequences: False
- Dropout: %20

Katman 3: Dense (Tam Bağlı)

- Nöron Sayısı: 32
- Aktivasyon Fonksiyonu: ReLU
- Dropout: %20

Katman 4: Çıkış Katmanı

- Nöron Sayısı: 1
- Aktivasyon Fonksiyonu: Linear

Toplam parametre sayısı 132.929'dur. Daha derin mimariler (3 LSTM katmanı) denenmiş ancak aşırı öğrenme riski nedeniyle tercih edilmemiştir.

4.2.1. Veri seti bölümlleme ve normalizasyon

Lag ve hareketli ortalama özelliklerinin oluşturulması nedeniyle ilk 30 satır eksik veri içermekte ve çıkarılmıştır. Kalan 3.623 veri noktası şu şekilde bölümlenmiştir:

- Eğitim Seti: 2.898 gün (%80) - 31 Ocak 2015 ile 6 Ocak 2023 arası
- Doğrulama Seti: 362 gün (%10) - 7 Ocak 2023 ile 3 Ocak 2024 arası
- Test Seti: 363 gün (%10) - 4 Ocak 2024 ile 31 Aralık 2024 arası

Kronolojik bölümlleme yaklaşımı, zaman serisi verilerinde kritik öneme sahiptir (Pedregosa vd., 2011; Scikit-learn Documentation, 2024).

Min-Max normalizasyonu ile tüm özellikler $[0,1]$ aralığına ölçeklendirilmiştir. Normalizasyon parametreleri sadece eğitim setinden öğrenilmiş ve doğrulama ile test setlerine uygulanmıştır. Bu yaklaşım, test setinin model eğitimine sızmasını önlemektedir (Scikit-learn Documentation, 2024).

4.2.2. Model eğitim süreci ve yakınsama analizi

Model eğitimi toplamda 35 epoch sürmüştür. Erken durduma mekanizması, 20. epoch'ta en düşük doğrulama kaybının (0,000518) elde edilmesinden sonra 15 epoch daha beklemiş ve iyileşme olmadığı için eğitimi durdurmuştur.

Eğitim kaybı (MSE), ilk 10 epoch'ta hızla düşmüş (0,051'den 0,002'ye) ve sonrasında yavaş bir yakınsama göstermiştir. Doğrulama kaybı benzer bir trend izlemiş ve eğitim kaybı ile yakın seyretmiştir. Bu durum, modelin aşırı öğrenme yapmadığını göstermektedir.

4.2.3. Model performans değerlendirilmesi

Model performansı, standart regresyon metrikleri kullanılarak değerlendirilmiştir.

Test setinde elde edilen %99,24 R^2 skoru, modelin tüketim varyasyonunun %99,24'ünü açıklayabildiğini göstermektedir.

MAPE değeri %8,59 olarak elde edilmiş olup, Lewis (1982) tarafından önerilen sınıflandırmaya göre "çok iyi" kategorisinde yer almaktadır. MAPE < %10 değeri, endüstriyel uygulamalar için kabul edilebilir doğruluk seviyesini göstermektedir.

4.3. Gerçek ve Tahmin Karşılaştırması

Test setinin ilk 100 günü için gerçek ve tahmin değerleri karşılaştırıldığında, modelin günlük tüketimleri başarıyla yakaladığı gözlemlenmiştir. Özellikle kış aylarında yüksek tüketim piklerini ve yaz aylarında düşük tüketimleri doğru modelleyebilmektedir.

Scatter plot analizi, tahmin ve gerçek değerler arasında güçlü bir doğrusal ilişki göstermektedir. Mükemmel tahmin çizgisinden ($y=x$) sapmalar minimal düzeydedir ve rastgele dağılmıştır.

4.4. Abone Artışı Senaryo Analizi

2025-2030 dönemi için üç farklı abone artış senaryosu geliştirilmiştir:

Senaryo 1 - Temel

- Yıllık Büyüme: %8,09
- 2024 Abone: 46.626
- 2030 Abone: 74.371
- Günlük Artış: 12,66 abone/gün

Senaryo 2 - Düşük (%75 Yavaşlama)

- Yıllık Büyüme: %2,34
- 2024 Abone: 46.626
- 2030 Abone: 53.562
- Günlük Artış: 3,17 abone/gün,

Senaryo 3 - Orta (%50 Yavaşlama)

- Yıllık Büyüme: %4,44
- 2024 Abone: 46.626
- 2030 Abone: 60.498
- Günlük Artış: 6,33 abone/gün

Senaryo 1, 2015-2024 dönemindeki doğrusal trend'in devam etmesini varsaymaktadır. Senaryo 2 ve 3 ise doğal gaz altyapısının olgunlaşması ve pazar doygunluğu etkilerini modellemektedir.

4.5. Gelecek tüketim tahminleri (2025-2030)

2025–2030 dönemi için günlük doğal gaz tüketim tahminleri, eğitilmiş LSTM modelinin iteratif çok adımlı tahmin (recursive multi-step forecasting) yaklaşımıyla

gerçekleştirilmiştir. Bu yöntemde, her gün için yapılan tahmin, bir sonraki günün giriş özelliklerinin hesaplanmasında kullanılmakta; böylece modelin kendi çıktılarından öğrenerek ilerlemesi sağlanmaktadır. Bu yapı, uzun vadeli tahminlerde hata birikimi (error accumulation) riski taşısa da, LSTM'in uzun süreli bağımlılıkları öğrenme yeteneği sayesinde kabul edilebilir performans sunmaktadır (Taieb vd., 2012; Hochreiter ve Schmidhuber, 1997)

Her tahmin adımı aşağıdaki işlemleri içermektedir:

Meteorolojik Değişkenlerin Üretilmesi: 2020–2024 dönemine ait günlük ortalamalara $\pm\%5$ rastgele varyasyon eklenerek sentetik meteorolojik veriler oluşturulmuştur. Bu yöntem, mevsimsel desenlerin korunmasını sağlarken, uzun vadeli belirsizlikleri de hesaba katmaktadır (Wilks, 2011).

Abone Sayısı Senaryolarının Entegrasyonu: Her gün için, ilgili senaryoya karşılık gelen kümülatif abone sayısı modele girdi olarak dahil edilmiştir.

Lag Özelliklerinin Güncellenmesi: Modelin bir önceki tahmin çıktısı, Lag_1_Tüketim, Lag_7_Tüketim ve Lag_30_Tüketim gibi gecikmeli değişkenlerin hesaplanmasında kullanılmıştır.

Tahmin ve Girdi Güncellemesi: Model çıktısı, ters normalizasyon işleminden geçirilerek gerçek tüketim değerine dönüştürülmüş ve bir sonraki günün girdisi olarak kullanılmıştır.

Bu iteratif süreç, 1 Ocak 2025'ten 31 Aralık 2030'a kadar toplam 2.191 günü kapsamaktadır.

Temel senaryo altında, 2025–2030 dönemi boyunca toplam doğal gaz tüketiminde %37,0 oranında bir artış öngörülmektedir. Düşük büyüme senaryosunda bu artış %20,4 ile sınırlı kalmaktadır. Senaryolar arasındaki fark, 2030 yılı itibarıyla 12,9 milyon m³ seviyesine ulaşmakta olup, bu miktar orta ölçekli bir kentin yıllık doğal gaz tüketimine eşdeğerdir. Bu sonuç, abone sayısındaki büyüme hızının enerji talebi üzerindeki belirleyici etkisini açıkça ortaya koymaktadır.

5. SONUÇLAR VE ÖNERİLER

Geliştirilen LSTM tabanlı tahmin modeli, test setinde %99,24 R^2 skoru ve %8,59 MAPE değeri ile son derece başarılı bir performans sergilemiştir. Bu sonuçlar, modelin doğal gaz tüketim varyasyonunun %99,24'ünü açıklayabildiğini ve ortalama tahmin hatasının %8,59 seviyesinde kaldığını göstermektedir.

Özellik mühendisliği çalışmaları, model başarısında kritik bir rol oynamıştır. Orijinal 9 değişkenden 30 değişkene genişletilen özellik seti, lag değişkenleri, hareketli ortalamalar ve trigonometrik dönüşümler aracılığıyla zamansal ve mevsimsel yapıyı etkin biçimde modele aktarmıştır. Özellikle 1 günlük gecikme değişkeninin (Lag_1_Tuketim) tüketimle 0,987 korelasyon katsayısına sahip olması, doğal gaz talebindeki önemini ortaya koymuştur.

Meteorolojik parametrelerin tüketim ile beklendiği gibi güçlü bir ilişkisi tespit edilmiştir. Isıtma derece gün sayısı (0,929), günlük ortalama sıcaklık (-0,847) ve hissedilen sıcaklık (-0,835) en yüksek korelasyon katsayılarına sahip meteorolojik değişkenler olarak öne çıkmıştır. Bu bulgular, ısınma ihtiyacının doğal gaz talebinin temel belirleyicisi olduğunu göstermektedir.

Model mimarisi olarak tercih edilen dört katmanlı LSTM yapısı (128-64-32-1 nöron dağılımı), 132.929 parametre ile yeterli temsil kapasitesine sahip olurken aşırı öğrenme riskini minimize etmiştir. Eğitim ve doğrulama kayıplarının birbirine yakın seyretmesi, %20 dropout oranının ve erken durdurma mekanizmasının başarılı bir düzenleme sağladığını göstermektedir.

2025-2030 dönemi tahminleri, bölgede doğal gaz tüketiminin artış trendini sürdüreceğini göstermektedir. Bu artışın temel itici güçleri, abone sayısındaki devam eden büyüme ve mevcut abonelerin tüketim alışkanlıklarıdır.

Üç farklı abone senaryosunun analizi, gelecek planlama çalışmalarında esneklik sağlamak ve farklı büyüme dinamiklerine hazırlıklı olmayı mümkün kılmaktadır. Özellikle, düşük büyüme senaryosunun bile 2030 yılında 85,5 milyon m³'lük bir tüketime işaret etmesi, altyapı yatırımlarının önemini ortaya koymaktadır.

Tahminlerin mevsimsel karakterini koruması, modelin fiziksel gerçekliği yansıttığının bir göstergesidir. Kış aylarındaki yüksek tüketim, ısıtma ihtiyacının bölgedeki doğal gaz talebinin ana belirleyicisi olduğunu teyit etmektedir.

5.1. Enerji Yönetimi İçin Pratik Öneriler

Bu çalışmadan elde edilen bulgular, doğal gaz dağıtım şirketleri, enerji düzenleyici kurumları ve yerel yönetimler için aşağıdaki pratik önerileri doğurmaktadır:

- Veri Odaklı Kapasite Planlaması: Abone artışı senaryolarının tüketim üzerindeki etkisi göz önüne alındığında, altyapı yatırımlarının senaryoların üst sınırına göre değil, orta-gerçekçi senaryolara göre planlanması önerilmektedir. Bu yaklaşım, aşırı kapasite oluşturma riskini azaltır ve sermaye verimliliğini artırır.
- Erken Uyarı Sistemleri: Geliştirilen tahmin modeli, gerçek tüketim verileriyle sürekli karşılaştırılarak erken uyarı mekanizmaları kurulabilir. Tahmin ile gerçekleşen arasındaki sapmaların eşik değerleri aştığı durumlarda otomatik bildirimler tetiklenerek operasyonel müdahale hızlandırılabilir.
- Yatırım Önceliklendirme: Abone artış senaryoları ile entegre tahminler, hangi coğrafi bölgelerde veya hangi altyapı bileşenlerinde (boru hatları, basınç düşürme istasyonları vb.) yatırım önceliğinin olduğunu belirlemeye yardımcı olabilir.

5.2. Gelecek Araştırma Önerileri

- Anomali Tespiti ve Nedenleri: Modelin büyük hatalar yaptığı durumların (anomalilerin) daha detaylı incelenerek, bu anomalilere yol açan özel olaylar (büyük endüstriyel duruşlar, bölgesel enerji kısıtlamaları vb.) belirlenip modele entegrasyonu araştırılabilir.
- Gerçek Zamanlı Tahmin ve Adaptasyon: Modelin gerçek zamanlı veri akışına uyum sağlama yeteneği ve kısa vadeli (saatlik, birkaç günlük) tahminler için optimizasyonu üzerine çalışmalar yapılabilir.
- Yapılması Gereken Yeni Yatırımlar: Modelin gelecek tahminlerini göz önünde bulundurarak bölgedeki RMS-A istasyonunun kapasitesinin yeterli olup olmadığı araştırılabilir. Bu kapsamda yeni bir RMS-A istasyonu gerekmesi durumunda yeni istasyonun yerinin tespiti de önemli bir çalışma konusu olacaktır.
- Hibrit Modelleme: LSTM mimarisine ek olarak hibrit LSTM-CNN yapıları denenebilir.

Bu çalışma doğal gaz sektöründe makine öğrenmesi ve derin öğrenme uygulamalarının pratik değerini göstermekte ve veri bilimi tekniklerinin enerji yönetiminde stratejik bir araç olarak kullanılabileceğini ortaya koymaktadır. Çalışma

doğal gaz tüketim tahmini alanında kapsamlı bir analiz ve yüksek performanslı bir model sunarak gelecekteki enerji yönetimi ve planlaması için sağlam bir temel oluşturmaktadır.



KAYNAKLAR

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... ve Zheng, X. (2016). {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)* (pp. 265-283).
- Ahmad, T., Zhang, D., Huang, C., Zhang, H., Dai, N., Song, Y. ve Chen, H. (2021). Artificial intelligence in sustainable energy industry: Status quo, challenges and opportunities. *Journal of Cleaner Production*, 289, 125834. <https://doi.org/10.1016/j.jclepro.2021.125834>
- Ahmad, T., Zhang, H., ve Yan, B. (2020). A review on renewable energy and electricity demand forecasting models for smart grid and buildings. *Sustainable Cities and Society*, 55, 102052. <https://doi.org/10.1016/j.scs.2020.102052>
- Ahmed, A., Khalid, M., Azeem, M. W., Khan, A. U. ve Ahmad, I. (2023). Forecasting renewable energy generation with machine learning and deep learning: Current advances and future prospects. *Sustainability*, 15(9), 7087. <https://doi.org/10.3390/su15097087>
- Akın, B. (2017). Yapay sinir ağlarıyla Konya bölgesinde kullanıcı doğal gaz tüketim öngörüsü. [Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi]
- Anderson, G. B., Bell, M. L., ve Peng, R. D. (2013). Methods to calculate the heat index as an exposure metric in environmental health research. *Environmental Health Perspectives*, 121(10), 1111-1119.
- Aras, H., ve Aras, N. (2004). Forecasting residential natural gas demand. *Energy Sources*, 26(5), 463-472. <https://doi.org/10.1080/00908310490429740>
- Armstrong, J. S., ve Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8(1), 69-80. [https://doi.org/10.1016/0169-2070\(92\)90008-W](https://doi.org/10.1016/0169-2070(92)90008-W)
- ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). (2017). *ASHRAE Handbook-Fundamentals*.
- ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers). (2021). *ASHRAE Handbook-Fundamentals*.
- Bengio, Y., Simard, P. ve Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157-166. <https://doi.org/10.1109/72.279181>
- Bergmeir, C. ve Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, 191, 192-213. <https://doi.org/10.1016/j.ins.2011.12.028>
- Box, G. E. P., Jenkins, G. M., Reinsel, G. C., ve Ljung, G. M. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.
- Box, G. E. P., ve Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Bozkurt, K., Pekmezci, A., ve Tekin, HA (2022). Box-Jenkins yöntemiyle turizm talebinin tahmin edilmesi: Türkiye örneği. *Anatolia: Turizm Araştırmaları Dergisi*, 33(2), 77-86. <https://doi.org/10.17123/atad.1087573>

- Brabec, M., Konár, O., Malý, M., Pelikán, E. ve Vondráček, J. (2009). A statistical model for natural gas standardized load profiles. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 58(1), 123-139. <https://doi.org/10.1111/j.1467-9876.2008.00636.x>
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199-231. <https://doi.org/10.1214/ss/1009213726>
- Brownlee, J. (2018). *Better deep learning: Train faster, reduce overfitting, and make better predictions*. Machine Learning Mastery.
- Brownlee, J. (2018). *Deep learning for time series forecasting: predict the future with MLPs, CNNs and LSTMs in Python*. Machine Learning Mastery.
- Brownlee, J. (2022). *Machine learning mastery*. Machine Learning Mastery.
- Chai, T., ve Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250. <https://doi.org/10.5194/gmd-7-1247-2014>
- Chatfield, C., ve Xing, H. (2019). *The Analysis of Time Series: An Introduction with R* (7th ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781351259446>
- Che, Z., Purushotham, S., Cho, K., Sontag, D., ve Liu, Y. (2018). Recurrent neural networks for multivariate time series with missing values. *Scientific Reports*, 8(1), 6085. <https://doi.org/10.1038/s41598-018-24271-9>
- Chollet, F. (2015). *Keras*. GitHub. <https://github.com/fchollet/keras>
- Chollet, F. (2021). *Deep learning with Python*. Manning Publications.
- Coutinho, E. R., Madeira, J. G. F., Borges, D. G. F., Springer, M. V., de Oliveira, E. M., ve Coutinho, A. L. G. A. (2025). Multi-step forecasting of meteorological time series using CNN-LSTM with decomposition methods. *Water Resources Management*, 39, 3173-3198. <https://doi.org/10.1007/s11269-025-04102-z>
- Day, A. R., ve Karayiannis, T. G. (1998). Degree-days: Comparison of calculation methods. *Building Services Engineering Research and Technology*, 19(1), 7-13. <https://doi.org/10.1177/014362449801900102>
- Demir, E., ve Gunal, S. (2025). Short-term electricity consumption forecasting with deep learning. *Journal of Supercomputing*, 81, 1108. <https://doi.org/10.1007/s11227-025-07564-5>
- Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55(10), 78-87. <https://doi.org/10.1145/2347736.2347755>
- Draper, N. R., ve Smith, H. (1998). *Applied regression analysis* (3rd ed.). Wiley.
- Ediger, V. Ş., ve Akar, S. (2007). ARIMA forecasting of primary energy demand by fuel in Turkey. *Energy Policy*, 35(3), 1701-1708. <https://doi.org/10.1016/j.enpol.2006.05.009>
- Ekonomou, L., (2010), Greek long-term energy consumption prediction using artificial neural networks, *Energy*, 35 (2), 512-517. <https://doi.org/10.1016/j.energy.2009.10.018>

- Eldan, R. ve Shamir, O., (2016), The power of depth for feedforward neural networks, Conference on Learning Theory, 49, 907-940.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211. https://doi.org/10.1207/s15516709cog1402_1
- Enerji Piyasası Düzenleme Kurumu (EPDK). (2014). *Doğal gaz piyasası mevzuatı*. Ankara: EPDK.
- Enerji Piyasası Düzenleme Kurumu (EPDK). (2025). *Natural gas market 2024 sector report*. Ankara: EPDK. <https://www.epdk.gov.tr/Detay/Icerik/3-0-95/dogal-gazaylik-sektor-raporu>
- European Commission, Eurostat. (2024, October 25). Energy statistics - cooling and heating degree days. https://ec.europa.eu/eurostat/cache/metadata/en/nrg_chdd_esms.htm
- Evcim, Ö. (2019). *Derin Öğrenme ve İstatiksel Modelleme Yöntemleri ile Günlük Doğalgaz Tüketiminin Tahmini ve Karşılaştırılması* [Yüksek Lisans Tezi, Sakarya Üniversitesi].
- Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456), 1348–1360. <https://doi.org/10.1198/016214501753382273>.
- Fanger, P. O. (1970). *Thermal comfort: Analysis and applications in environmental engineering*. Danish Technical Press.
- Feinberg, E. A., ve Genethliou, D. (2005). Load forecasting. In *Applied mathematics for restructured electric power systems: optimization, control, and computational intelligence* (pp. 269-285). Boston, MA: Springer US. https://doi.org/10.1007/0-387-23471-3_12
- Fischer, T., ve Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, 270(2), 654-669. <https://doi.org/10.1016/j.ejor.2017.11.054>
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (2nd ed.). O'Reilly Media.
- Glorot, X., Bordes, A., ve Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics* (s. 315-323). JMLR Workshop and Conference Proceedings.
- González-Romera, E., Jaramillo-Morán, M. A., ve Carmona-Fernández, D. (2008). Monthly electric energy demand forecasting with neural networks and Fourier series. *Energy Conversion and Management*, 49(11), 3135-3142. <https://doi.org/10.1016/j.enconman.2008.06.004>
- Goodfellow, I., Bengio, Y., ve Courville, A. (2016). *Deep learning*. MIT Press.
- Graves, A. (2013). Generating sequences with recurrent neural networks. *arXiv preprint arXiv:1308.0850*. <https://arxiv.org/abs/1308.0850>

- Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. ve Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10), 2222-2232. <https://doi.org/10.1109/TNNLS.2016.2582924>
- Gujarati, D. N., ve Porter, D. C. (2009). *Basic econometrics* (5th ed.). McGraw-Hill.
- Gümrah, F., Katircioglu, D., Aykan, Y., Okumus, S., ve Kilincer, N. (2001). Modeling of gas demand using degree-day concept: Case study for Ankara. *Energy Sources*, 23(2), 101-114. <https://doi.org/10.1080/00908310151092254>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... ve Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357-362. <https://doi.org/10.1038/s41586-020-2649-2>
- Hastie, T., Tibshirani, R. ve Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). Prentice Hall.
- Hermans, M. ve Schrauwen, B. (2013). Training and analysing deep recurrent neural networks. *Advances in Neural Information Processing Systems*, 26.
- Hippert, H. S., Pedreira, C. E., ve Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on power systems*, 16(1), 44-55. <https://doi.org/10.1109/59.910780>
- Hochreiter, S., ve Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hong, T., ve Fan, S. (2016). Probabilistic electric load forecasting: A tutorial review. *International Journal of Forecasting*, 32(3), 914-938. <https://doi.org/10.1016/j.ijforecast.2015.11.011>
- Hornik, K., Stinchcombe, M. ve White, H., (1989), Multilayer feedforward networks are universal approximators, *Neural Networks*, 2 (5), 359-366. [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8)
- Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science ve Engineering*, 9(3), 90-95. <https://doi.org/10.1109/MCSE.2007.55>
- Hyndman, R. J., ve Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts. <https://otexts.com/fpp2/>
- International Energy Agency (IEA). (2015). Natural Gas Information 2015. IEA Publications, Paris. https://www.oecd.org/content/dam/oecd/en/publications/reports/2015/08/natural-gas-information-2015_g1g57dc3/nat_gas-2015-en.pdf
- International Energy Agency (IEA). (2022). *Gas Market Report Q3-2022*. IEA Publications, Paris. <https://www.iea.org/reports/gas-market-report-q3-2022>
- International Energy Agency (IEA). (2022). *World Energy Outlook 2022*. IEA Publications, Paris. <https://www.iea.org/reports/world-energy-outlook-2022>

- James, G., Witten, D., Hastie, T. ve Tibshirani, R. (2021). *An introduction to statistical learning: With applications in R* (2. baskı). Springer. <https://doi.org/10.1007/978-1-0716-1418-1>
- James, G., Witten, D., Hastie, T., ve Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R* (Vol. 103). New York: Springer. <https://doi.org/10.1007/978-1-4614-7138-7>
- Kalogirou, S. A. (2000). Applications of artificial neural-networks for energy systems. *Applied Energy*, 67(1-2), 17-35. [https://doi.org/10.1016/S0306-2619\(00\)00005-2](https://doi.org/10.1016/S0306-2619(00)00005-2)
- Katz, D. L. V. ve Lee, R. L. (1990). *Natural gas engineering: Production and storage*. McGraw-Hill.
- Kaynar, O., Taştan, S., ve Demirkoparan, F. (2011). Yapay sinir ağları ile doğalgaz tüketim tahmini. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 25.
- Keras Team. (2024). *Keras documentation*. <https://keras.io/>
- Khotanzad, A., Elragal, H., ve Lu, T. L. (2000). Combination of artificial neural-network forecasters for prediction of natural gas consumption. *IEEE transactions on neural networks*, 11(2), 464-473. <https://doi.org/10.1109/72.839015>
- Kim, T. Y., ve Cho, S. B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*, 182, 72-81. <https://doi.org/10.1016/j.energy.2019.05.230>
- Kingma, D. P., ve Ba, J. (2015). Adam: A method for stochastic optimization. *International Conference on Learning Representations (ICLR)*, San Diego.
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., ... ve Willing, C. (2016). Jupyter Notebooks-a publishing format for reproducible computational workflows. In *Positioning and power in academic publishing: Players, agents and agendas* (pp. 87-90). IOS press. <https://doi.org/10.3233/978-1-61499-649-1-87>
- Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., ve Zhang, Y. (2019). Short-term residential load forecasting based on LSTM recurrent neural network. *IEEE Transactions on Smart Grid*, 10(1), 841-851. <https://doi.org/10.1109/TSG.2017.2753802>
- Kuhn, M. ve Johnson, K. (2019). *Feature engineering and selection: A practical approach for predictive models*. CRC Press. <https://doi.org/10.1201/9781315108230>
- Kutner, M. H., Nachtsheim, C. J., Neter, J. ve Li, W. (2005). *Applied linear statistical models* (5. baskı). McGraw-Hill/Irwin.
- LeCun, Y., Bengio, Y. ve Hinton, G. (2015). Deep learning. *Nature*, 521, 436--444. <https://doi.org/10.1038/nature14539>
- Lewis, C. D., (1982), *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*, Butterworths.
- Makina Mühendisleri Odası (MMO). (2008). *Türkiye'nin doğal gaz temin ve tüketim politikalarının değerlendirilmesi raporu* (Yayın No: 469). Ankara: MMO Yayınları.

- McKinney, W. (2012). *Python for data analysis: Data wrangling with Pandas, NumPy, and IPython*. O'Reilly Media, Inc.
- Meteoroloji Genel Müdürlüğü. (2025). Isıtma ve soğutma gün dereceleri, <https://www.mgm.gov.tr/veridegerlendirme/gun-derece.aspx>
- Mokhatab, S., Poe, W. A., ve Speight, J. G. (2015). *Handbook of natural gas transmission and processing* (3rd ed.). MA: Gulf Professional Publishing.
- Montgomery, D. C., Jennings, C. L., ve Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Montgomery, D. C., Peck, E. A., ve Vining, G. G. (2021). *Introduction to linear regression analysis*. (6th ed.). John Wiley & Sons.
- Mosavi, A., Salimi, M., Faizollahzadeh Ardabili, S., Rabczuk, T., Shamshirband, S., ve Varkonyi-Koczy, A. R. (2019). State of the art of machine learning models in energy systems: A systematic review. *Energies*, 12(7), 1301. <https://doi.org/10.3390/en12071301>
- Nair, V. ve Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning*, 807-814.
- Pamir, N., ve Türkyılmaz, O. (2007). Türkiye'nin Doğal Gaz Temin ve İhraç Seçenekleri. *Uluslararası Doğal Gaz Kongresi ve Sergisi*, 1-18. Ankara
- Pandas Development Team. (2024). Pandas documentation. <https://pandas.pydata.org/docs/>
- Panek, W. ve Włodek, T. (2022). Natural gas consumption forecasting based on the variability of external meteorological factors using machine learning algorithms. *Energies*, 15(1), 348. <https://doi.org/10.3390/en15010348>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ve Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Pyle, D. (1999). *Data preparation for data mining*. Morgan Kaufmann Publishers.
- Python Software Foundation. (2023). *Python language reference (Version 3.11)*. <https://docs.python.org/3/>
- Qiu, X., Ren, Y., Suganthan, P. N., ve Amaratunga, G. A. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied soft computing*, 54, 246-255. <https://doi.org/10.1016/j.asoc.2017.01.015>
- Raschka, S. (2015). Python Machine Learning: Learn how to Build Powerful Python Machine Learning Algorithms to Generate Useful Data Insights with this Data Analysis Tutorial. *Packt Publishing Limited*.
- Rehman, A., Zhu, J. J., Segovia, J., ve Anderson, P. R. (2022). Assessment of deep learning and classical statistical methods on forecasting hourly natural gas demand at multiple sites in Spain. *Energy*, 244, 122562. <https://doi.org/10.1016/j.energy.2021.122562>

- Rothfus, L. P. (1990). The heat index equation. *National Weather Service Technical Attachment SR 90-23*.
- Rumelhart, D. E., Hinton, G. E., ve Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533-536. <https://doi.org/10.1038/323533a0>
- Ryu, S., Noh, J. ve Kim, H. (2017). Deep neural network based demand side short term load forecasting. *Energies*, 10(1), 3. <https://doi.org/10.3390/en10010003>
- Sagheer, A., ve Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, 323, 203-213. <https://doi.org/10.1016/j.neucom.2018.09.082>
- Sailor, D. J., ve Muñoz, J. R. (1997). Sensitivity of electricity and natural gas consumption to climate in the USA-Methodology and results for eight states. *Energy*, 22(10), 987-998. [https://doi.org/10.1016/S0360-5442\(97\)00034-0](https://doi.org/10.1016/S0360-5442(97)00034-0)
- Sarak, H., ve Satman, A. (2003). The degree-day method to estimate the residential heating natural gas consumption in Turkey: A case study. *Energy*, 28(9), 929-939. [https://doi.org/10.1016/S0360-5442\(03\)00035-5](https://doi.org/10.1016/S0360-5442(03)00035-5)
- Schafer, J. L., ve Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177. <https://doi.org/10.1037/1082-989X.7.2.147>
- Schwarzbauer, J., ve Jovančičević, B. (2024). Biogenic vs. Thermogenic Gas Emissions. In *Isotopes in Organic Geochemistry* (pp. 47-50). Springer International Publishing. https://doi.org/10.1007/978-3-031-69304-5_6
- Scikit-learn Developers. (2024). Scikit-learn user guide. <https://scikit-learn.org/stable/>
- Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., ve Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*, 28, 802-810.
- Simonyan, K., ve Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint* <https://doi.org/10.48550/arXiv.1409.1556>
- Skforecast Docs. (2023). Cyclical features in time series forecasting. <https://skforecast.org/0.20.0/index.html>
- Soldo, B. (2012). Forecasting natural gas consumption. *Applied Energy*, 92, 26-37. <https://doi.org/10.1016/j.apenergy.2011.11.003>
- Söder, L., ve Amelin, M. (2011). Efficient operation and planning of power systems. *KTH Royal Institute of Technology*.
- Sözen, A., Arcaklioğlu, E., ve Özkaymak, M. (2005). Turkey's net energy consumption. *Applied Energy*, 81(2), 209-221. <https://doi.org/10.1016/j.apenergy.2004.07.001>
- Speight, J. G. (2019). *Natural gas: A basic handbook*. (2nd ed.). Gulf Professional Publishing.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., ve Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.

- Steadman, R. G. (1979). The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. *Journal of Applied Meteorology and Climatology*, 18(7), 861-873. [https://doi.org/10.1175/1520-0450\(1979\)018<0861:TAOSPI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1979)018<0861:TAOSPI>2.0.CO;2)
- Suganthi, L., ve Samuel, A. A. (2012). Energy models for demand forecasting-A review. *Renewable and Sustainable Energy Reviews*, 16(2), 1223-1240. <https://doi.org/10.1016/j.rser.2011.08.014>
- Taieb, S. B., Bontempi, G., Atiya, A. F., ve Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert systems with applications*, 39(8), 7067-7083. <https://doi.org/10.1016/j.eswa.2012.01.039>
- Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting*, 16(4), 437-450. [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)
- Taşkıner, B. (2018). Ankara İli Doğal Gaz Tüketiminin Yapay Sinir Ağları İle Öngörüsü. [Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi]
- Taylor, J. W., ve McSharry, P. E. (2007). Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems*, 22(4), 2213-2219. <https://doi.org/10.1109/TPWRS.2007.907583>
- Thom, H. C. S. (1954). The rational relationship between heating degree days and temperature. *Monthly Weather Review*, 82(1), 1-6. [https://doi.org/10.1175/1520-0493\(1954\)082<0001:TRRBHD>2.0.CO;2](https://doi.org/10.1175/1520-0493(1954)082<0001:TRRBHD>2.0.CO;2)
- Tissot, B. P. ve Welte, D. H. (1984). *Petroleum formation and occurrence* (2. baskı). Springer-Verlag. <https://doi.org/10.1007/978-3-642-87813-8>
- Tissot, B. P. ve Welte, D. H. (2013). *Petroleum formation and occurrence*. Springer Science ve Business Media.
- Türk Standardları Enstitüsü. (2024). *TS 825:2024 Binalarda ısı yalıtımı kuralları standardı*. Türk Standardları Enstitüsü.
- Türkiye İstatistik Kurumu (TÜİK). (2024, 1 Kasım). *Sanayi Sektörü Nihai Enerji Tüketim İstatistikleri, 2023* [Haber Bülteni]
- U.S. Energy Information Administration. (2025, Mayıs 29). *U.S. energy-related carbon dioxide emissions, 2024*. EIA. <https://www.eia.gov/environment/emissions/carbon/>
- UGETAM. (2016). *Genel doğalgaz* (Yayın No: 39, 3. baskı). UGETAM Yayınları.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... ve van der Walt, S. J. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3), 261-272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wang, H., Gao, X., Zhang, Y. ve Yang, Y. (2025). Natural gas consumption forecasting model based on feature optimization and incremental long short-term memory. *Sensors*, 25(10), 3079. <https://doi.org/10.3390/s25103079>
- Waskom, M. L. (2021). seaborn: Statistical data visualization. *Journal of Open Source Software*, 6(60), 3021

- Weron, R. (2006). *Modeling and forecasting electricity loads and prices: A statistical approach*. John Wiley & Sons.
- Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed.). Academic Press.
- Willmott, C. J., ve Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, 30(1), 79-82. <https://doi.org/10.3354/cr030079>
- Wooldridge, J. M. (2019). *Introductory econometrics: A modern approach*. Cengage Learning.
- Zheng, A., ve Casari, A. (2018). *Feature engineering for machine learning: principles and techniques for data scientists*. O'Reilly Media, Inc.
- Zumerchik, J. (Ed.). (2001). *Macmillan Encyclopedia of Energy* (Vols. 1-3). Macmillan Reference USA.

