



T.C.
NECMETTİN ERBAKAN
ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ



ELEKTRİKLİ ULAŞIM SİSTEMLERİNDE
ENERJİ TÜKETİMİNİN TAHMİNİ

SHIR HUSSAIN AHMADI

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Ekim-2025
KONYA
Her Hakkı Saklıdır

TEZ KABUL VE ONAYI

SHIR HUSSAIN AHMADI tarafından hazırlanan “ELEKTRİKLİ ULAŞIM SİSTEMLERİNDE ENERJİ TÜKETİMİNİN TAHMİNİ” adlı tez çalışması 22/10/2025 tarihinde aşağıdaki jüri tarafından oy birliği ile Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü Elektrik ve Elektronik Mühendisliği Anabilim Dalı YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

İmza

Başkan

Dr. Öğr. Üyesi Ali Osman ÖZKAN

.....

Danışman

Prof. Dr. Muciz ÖZCAN

.....

Üye

Doç. Dr. Mümtaz MUTLUER

.....

Fen Bilimleri Enstitüsü Yönetim Kurulu’nun/.../20.. 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.

SHIR HUSSAIN AHMADI

Tarih:

ÖZET

YÜKSEK LİSANS TEZİ

ELEKTRİKLİ ULAŞIM SİSTEMLERİNDE ENERJİ TÜKETİMİNİN TAHMİNİ

SHIR HUSSAIN AHMADI

Necmettin Erbakan Üniversitesi Fen Bilimleri Enstitüsü
Elektrik-Elektronik Mühendisliği Anabilim Dalı

Danışman: Prof.Dr. Muciz ÖZCAN

2025, 66 Sayfa

Jüri

Prof.Dr. Muciz ÖZCAN

Doç. Dr. Mümtaz MUTLUER

Dr. Öğr. Üyesi Ali Osman ÖZKAN

Bu tez çalışmasında, kentsel ulaşımın karbonsuzlaştırılmasında kilit role sahip olan Elektrikli Otobüslerin (EB) anlık enerji tüketimini yüksek doğrulukla öngörebilen bir makine öğrenmesi modeli geliştirilmiştir. Çalışmada, İsviçre'nin Zürih kentinde hizmet veren elektrikli trolleybüs filosuna ait, kamuya açık ve yüksek çözünürlüklü ZTBus veri setinden yararlanılmıştır. Enerji tüketimini tahmin etmek amacıyla, temel bir model olarak Çoklu Doğrusal Regresyon (MLR) ile XGBoost, LightGBM ve HistGradientBoosting gibi üç farklı gelişmiş gradyan artırma (gradient boosting) algoritması karşılaştırmalı olarak değerlendirilmiştir. Modeller, aracın anlık güç talebini (*electric_powerDemand*) tahmin etmek üzere, araç dinamiği, operasyonel durum ve çevresel koşulları içeren on bir adet girdi özelliği kullanılarak eğitilmiştir.

Yapılan deneysel çalışmalar, zaman serisi çapraz doğrulama (TimeSeriesSplit) yöntemiyle değerlendirildiğinde, gradyan artırma modellerinin temel MLR modelinden (ortalama $R^2 \approx 0.72$) bariz şekilde daha üstün bir performans sergilediğini göstermiştir. Bu modeller arasında LightGBM, ortalama $R^2 \approx 0.977$ skoru ile en yüksek tahmin doğruluğuna ulaşmıştır. Bu bulgu, EB'lerin enerji tüketim dinamiklerinin doğrusal olmayan ve karmaşık bir yapıya sahip olduğunu teyit etmektedir. Nihai model üzerinde yapılan özellik önem analizi, *traction_tractionForce* (çekiş kuvveti) ve *odometry_vehicleSpeed* (araç hızı) değişkenlerinin modelin performansını belirlemede en baskın faktörler olduğunu ortaya koymuştur.

Çalışmanın en önemli çıktısı olarak, en iyi performansı gösteren nihai model, daha önce hiç görmediği bağımsız bir test (holdout) verisi üzerinde test edilmiştir. Bu son değerlendirmede model, anlık güç talebini %97,59'luk bir R^2 skoru ile tahmin etmiş ve bir tam görevin toplam enerji tüketimini ise %0.08 gibi son derece düşük bir hata payıyla öngörmeyi başarmıştır. Bu sonuçlar, yüksek frekanslı verilerin gelişmiş makine öğrenmesi modelleriyle entegrasyonunun, elektrikli otobüslerde enerji tüketimi tahmininin doğruluğunu kayda değer biçimde artırdığını kanıtlamaktadır. Geliştirilen bu model, filo yönetimi, güzergâh optimizasyonu ve şarj altyapısı planlaması gibi alanlarda karar destek sistemleri için güvenilir bir temel oluşturma potansiyeline sahiptir.

Anahtar Kelimeler: Çekış kuvveti, Elektrikli Otobüs, Enerji tüketimi tahmini, Gradyan artırma, LightGBM, Makine öğrenmesi, Özellik mühendisliği, ZTBus veri seti

ABSTRACT

MS THESIS

PREDICTION OF ENERGY CONSUMPTION IN ELECTRIC TRANSPORTATION SYSTEMS

SHIR HUSSAIN AHMADI

THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF
NECMETTİN ERBAKAN UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
IN ELECTRICAL AND ELECTRONICS ENGINEERING

Advisor: Prof.Dr. Muciz ÖZCAN

2025, 66 Pages

Jury

Prof.Dr. Muciz ÖZCAN

Assoc.Prof.Dr. Mümtaz MUTLUER

Asst.Prof.Dr. Ali Osman ÖZKAN

In this thesis study, a machine learning model is developed to predict the instantaneous energy consumption of Electric Buses (EBs) with high accuracy, which play a key role in the decarbonization of urban transportation. This study utilizes the comprehensive, open-access, and high-resolution ZTBus dataset from the electric trolleybus fleet operating in Zurich, Switzerland. To predict energy consumption, a baseline Multiple Linear Regression (MLR) model was comparatively evaluated against three advanced gradient boosting algorithms: XGBoost, LightGBM, and HistGradientBoosting. The models were trained using eleven input features encompassing vehicle dynamics, operational status, and environmental conditions to predict the instantaneous power demand (*electric powerDemand*) as the target variable.

Experimental results, evaluated using time-series cross-validation (TimeSeriesSplit), demonstrate that the gradient boosting models significantly outperform the baseline MLR model (average $R^2 \approx 0.72$). Among these advanced models, LightGBM achieved the highest prediction accuracy with an average R^2 score of approximately 0.977. This finding confirms that the energy consumption dynamics of EBs have a non-linear and complex structure. Feature importance analysis on the final model revealed that *traction_tractionForce* and *odometry_vehicleSpeed* are the most dominant factors in determining the model's performance.

As the most significant output of this study, the best-performing final model was tested on an independent holdout dataset it had never seen before. In this final evaluation, the model successfully predicted the instantaneous power demand with an R^2 score of 97.59% and forecasted the total energy consumption of a complete mission with an exceptionally low error margin of 0.08%. These results prove that the integration of high-frequency data with advanced machine learning models significantly enhances the accuracy of energy consumption prediction for electric buses. The developed model has the potential to serve as a reliable foundation for decision support systems in areas such as fleet management, route optimization, and charging infrastructure planning.

Keywords: Traction force, Electric Bus, Energy consumption prediction, Gradient boosting, LightGBM, Machine learning, Feature engineering, ZTBus dataset.

ÖNSÖZ

Kentsel ulaşımın karbonsuzlaştırılmasında kilit rol oynayan elektrikli otobüslerin enerji tüketiminin doğru tahmin edilmesi, bu tezin ana motivasyon kaynağı olmuştur.

Bu araştırma sürecinde, değerli bilgi ve tecrübeleriyle bana her zaman yol gösteren kıymetli danışmanım Prof. Dr. Muciz ÖZCAN'a en derin şükranlarımı sunarım.

Tez savunma jürimde yer alarak yapıcı eleştirileriyle çalışmama katkı sağlayan Doç. Dr. Mümtaz MUTLUER'e ve Dr. Öğr. Üyesi Ali Osman ÖZKAN'a teşekkür ederim.

Yüksek lisans eğitimim boyunca desteklerini gördüğüm Necmettin Erbakan Üniversitesi Elektrik-Elektronik Mühendisliği Anabilim Dalı'ndaki tüm hocalarıma teşekkür borçluyum.

Son olarak, bu yoğun süreçte sabırları ve manevi destekleriyle her zaman yanımda olan sevgili aileme ve dostlarıma sonsuz teşekkür ederim.

Bu çalışmanın, elektrikli ulaşım sistemleri alanındaki araştırmalara ve pratik uygulamalara katkı sağlamasını dilerim.

Shir Hussain Ahmadi
KONYA-2025

İÇİNDEKİLER

ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
ŞEKİLLER LİSTESİ	ix
ÇİZELGELER LİSTESİ	x
SİMGELER VE KISALTMALAR	xi
1. GİRİŞ	1
1.1. Elektrikli Araçlara Geçiş: Sürücü Faktörleri ve Hızlanan Talep	1
1.2. Elektrikli Araç Kullanımında Menzil Kaygısı	2
1.3. Toplu Taşımada Elektrikli Otobüslerde Kullanılan Enerji Modellemeleri	3
1.4. Tez Çalışmasının Amacı	5
2. KAYNAK ARAŞTIRMASI	7
2.1. EV Enerji Tüketimi Modelleme Yaklaşımları	7
2.1.1. Fizik temelli (analitik) modeller	7
2.1.2. Veri odaklı modeller	9
2.1.3. Hibrit modelleme yaklaşımları	11
2.2. EV’lerde Enerji Tüketimini Belirleyen Faktörler	14
2.2.1. Araç özellikleri	14
2.2.2. Sürüş tarzı ve sürücü davranışları	15
2.2.3. Çevresel koşullar	16
2.2.4. Rota topografyası ve operasyonel koşullar	17
2.3. EB Enerji Tüketimi Modellemesinde Güncel Durum ve Araştırma Eğilimleri ..	20
2.3.1. Otobüslere özgü belirlenen faktörler	20
2.3.2. EV araştırmalarında kullanılan veri setleri	21
2.3.3. EB’lerde menzil belirlemede yapılan çalışmalardan elde edilen bulgular ve doğruluk oranları.....	22
2.4. EV’lerde Menzil Belirlemede kullanılan Algoritmalarındaki Araştırma Boşlukları	25
3. MATERYAL VE YÖNTEM.....	29
3.1. Veri Seti ve Özellikleri	29
3.2. Araştırma Metodolojisi Akış Şeması	33
3.3. Veri Ön İşleme Süreci.....	35
3.3.1. Meta veri filtreleme ve geçerli görev listesinin oluşturulması	35
3.3.2. Ana veri seti için temsili örneklem seçimi.....	36

3.3.3. Verilerin ön işlenmesi ve özellik mühendisliği	38
3.4. Modelleme ve Değerlendirme Yaklaşımı	42
3.4.1. Değerlendirilen makine öğrenmesi algoritmaları	42
3.4.2. Model değerlendirme stratejisi	43
3.4.3. Performans değerlendirme metrikleri	45
3.5. Nihai Modelin Seçimi ve Eğitimi	47
3.5.1. Modellerin karşılaştırmalı performans analizi	47
3.5.2. En iyi modelin seçimi ve gerekçesi	48
3.5.3. Nihai modelin tüm veri seti üzerinde eğitilmesi ve saklanması	48
4. ARAŞTIRMA BULGULARI VE TARTIŞMA	50
4.1. Keşifsel Veri Analizi ve Değişken İlişkileri	50
4.2. Modellerin Karşılaştırmalı Performans Analizi	52
4.2.1. Tahmin doğruluğu ve hata oranlarının değerlendirilmesi	53
4.2.2. Eğitim süresi ve hesaplama verimliliği açısından karşılaştırma	54
4.3. En İyi Modelin Seçimi ve Özellik Önem Analizi	54
4.3.1. En iyi modelin seçimi ve gerekçesi	54
4.3.2. Nihai modelin özellik önem analizi	54
4.4. Nihai Modelin Bağımsız Veri Seti (Holdout) Üzerinde Test Edilmesi	57
4.4.1. Anlık güç talebi tahmin performansı	57
4.4.2. Toplam görev enerjisi tüketimi tahmin sonuçları	59
4.5. Tartışma	60
4.5.1. Bulguların genel değerlendirmesi ve literatürdeki yeri	60
4.5.2. Model performansının ve algoritmik seçimin literatürle karşılaştırılması	60
4.5.3. Metodolojik yaklaşımın ve özellik öneminin derinlemesine analizi	61
5. SONUÇLAR VE ÖNERİLER	63
5.1. Sonuçlar	63
5.2. Araştırmanın Kısıtları	64
5.3. Gelecek Çalışmalar İçin Öneriler	65
6. KAYNAKLAR	68

ŞEKİLLER LİSTESİ

Şekil 3.1. Veri setinde kapsanan dokuz otobüs hattının güzergahı (Widmer vd., 2023)	30
Şekil 3.2. Araştırma metodolojisi akış şeması.....	34
Şekil 3.3. itcs_numberOfPassengers sütununa ileri yönlü doldurma (forward fill) yönteminin uygulanmasının etkisi: ham veriler (kırmızı noktalar) ve temizlenmiş sürekli veriler (mavi çizgi).....	40
Şekil 4.1. Nihai lightgbm modeli için özellik önem dağılımı	56
Şekil 4.2. Holdout veri seti için gerçek ve tahmin edilen değerlerin karşılaştırılması ...	59



ÇİZELGELER LİSTESİ

Çizelge 2.1. EV'lerin enerji tüketimi modelleme yaklaşımlarının karşılaştırması.....	13
Çizelge 2.2. EV enerji tüketimini etkileyen temel faktörler	19
Çizelge 2.3. EV enerji tüketimi modellemesine dair önemli çalışmaların özeti	23
Çizelge 3.1. Modellemede kullanılan değişkenlerin istatistiksel özeti.....	32
Çizelge 3.2. Ana veri seti ve temsili örneklemedeki görevlerin mevsimsel dağılımının karşılaştırılması.....	37
Çizelge 3.3. Ana veri seti ve temsili örneklemedeki görevlerin güzergâh dağılımının karşılaştırılması.....	38
Çizelge 3.4. Özellik mühendisliği ile türetilen yeni değişkenler.....	41
Çizelge 3.5. Modellerin zaman serisi çapraz doğrulama ile karşılaştırmalı performans sonuçları.....	47
Çizelge 4.1. Modellemede kullanılan değişkenler arasındaki pearson korelasyon matrisi	51
Çizelge 4.2. Nihai lightgbm modeli için özellik önem puanları.....	55
Çizelge 4.3. Nihai modelin bağımsız test (holdout) veri seti üzerindeki performans metrikleri.....	58
Çizelge 4.4. Nihai Modelin Holdout Görevi Üzerindeki Toplam Enerji Tüketimi Tahmin Performansı	59

SİMGELER VE KISALTMALAR

Simgeler

CO_2	Karbondioksit
C_d	Aerodinamik sürtünme katsayısı
A	Ön kesit alanı
n	Test setindeki toplam örnek sayısı
y_i	Gerçek değer
$\hat{y}_i$	Model tarafından tahmin edilen değer
$\bar{y}$	Gerçek değerlerin ortalaması
R^2	Belirlilik Katsayısı (Coefficient of Determination)
ΔR^2	R^2 değerindeki değişim/düşüş
W	Güç Birimi (Watt)

Kısaltmalar

ANN	Artificial Neural Network (Yapay Sinir Ağları)
API	Application Programming Interface (Uygulama Programlama Arayüzü)
CAN	Controller Area Network (Kontrolör Alan Ağı)
CSV	Comma-Separated Values (Virgülle Ayrılmış Değerler)
DLNA	Deep Learning Network with Autoencoders (Otomatik Kodlayıcılarla Derin Öğrenme Ağı)
EB	Electric Bus (Elektrikli Otobüs)
EV	Electric Vehicle (Elektrikli Araç)
GNSS	Global Navigation Satellite System (Küresel Navigasyon Uydu Sistemi)
GPR	Gaussian Process Regression (Gauss Süreç Regresyonu)
GPS	Global Positioning System (Küresel Konumlama Sistemi)
HVAC	Heating Ventilation & Air Conditioning (Isıtma, Havalandırma ve İklimlendirme)
LightGBM	Light Gradient Boosting Machine (Hafif Gradyan Artırma Makinesi)
LSTM	Long Short-Term Memory (Uzun Kısa Süreli Bellek)
MAE	Mean Absolute Error (Ortalama Mutlak Hata)
MAPE	Mean Absolute Percentage Error (Ortalama Mutlak Yüzde Hata)
ML	Machine Learning (Makine Öğrenmesi)
MLR	Multiple Linear Regression (Çoklu Doğrusal Regresyon)
PINN	Physics-Informed Neural Networks (Fizik Bilgilendirmeli Sinir Ağları)
RF	Random Forest (Rastgele Ormanlar)
RMSE	Root Mean Squared Error (Hataların Ortalama Karekökü)
SOC	State of Charge (Şarj Durumu)
SVM	Support Vector Machine (Destek Vektör Makineleri)
SVR	Support Vector Regression (Destek Vektör Regresyonu)
VCU	Vehicle Control Unit (Merkezi Kontrol Ünitesi)
XGBoost	eXtreme Gradient Boosting (Aşırı Gradyan Artırma)

1. GİRİŞ

1.1. Elektrikli Araçlara Geçiş: Sürücü Faktörleri ve Hızlanan Talep

Küresel ulaşım sektöründe Elektrikli Araçların (EV) kullanımı her geçen gün hızla Artmaktadır. Bu sektörde sürdürülebilirlik ve enerji güvenliğinin artırılması çalışmaları bu araçların kullanımının artırılmasında çok önemli bir kilit rol almaktadır. Bu çalışmaların bu araçlar üzerine yaptığı olumlu katkılar araç pazarında yer alan EV sayısının hızla artmasından anlaşılabılır. 2021 yılında dünyada 6,7 milyon EV satıldı ve bu rakam, toplam araç satışlarının %8,3'ünü oluşturmaktadır. 2030'a kadar EV'lerin küresel satışların %70'ine ulaşabileceği öngörülmektedir (Chen vd., 2022) . Pazarda yer alan bu EV araçlar içten yanmalı araçlardan belirgin bir şekilde uzaklaşıldığını ortaya koymaktadır. Pazarda yer alan EV'lerin araçların artması kullanıcılar bu araçları kullanırken gereksinim duydukları enerji kullanımını anlama ve yönetme ihtiyacını her geçen gün artırıyor.

İçten yanmalı araçlardan EV geçişin temel nedenleri, iklim değişikliğiyle acil mücadele eylem planları geliştirmek ve enerji güvenliğinde dışa bağımlığın azaltılması sağlamaktır. Birçok ülke Çin ve Avrupa Birliği dâhil karbon emisyonlarını düşürmek için katı eylem planları hazırlayıp kendilerine kısa ve uzun vadede gerçekleştirilecek birtakım hedefler belirlemişlerdir. Belirlenen hedeflerin biride ulaşım sektöründe hali hazırda kullanılan araçların elektrikleştirilmesidir (Jiang vd., 2023). Araştırmalar gösteriyor ki Ulaşım sektörü, küresel CO₂ emisyonlarının önemli bir bölümünden sorumludur (Xu vd., 2023). EV'lerin kullanımlarının yaygınlaştırılması kısa vadede ülkelerin emisyonlarını azaltmak için pratik bir çözüm olmaktadır (Acharyaviriya vd., 2023; Liu vd., 2018). EV'ler çevresel faydaların yanı sıra, ülkelerin fosil yakıt bağımlılığını da azaltarak ülkemizde olduğu gibi enerjide dışa bağımlı olan ülkelerin enerji güvenliğini de artırır (Ayman vd., 2020).

Hızlı EV büyümesi, çevre ve enerji hedeflerine hizmet ederken elektrik şebekesi üzerinde mevcut hatların kurulu gücünün artması vb. birtakım yeni sorunlar oluşmasına sebep olmaktadır. Tek bir EV, bazen bir hanenin birkaç günde tüketeceği elektrik enerjisini çok kısa bir sürede kullanımına neden olur (Foiadelli vd., 2018). Birçok EV'nin aynı anda, özellikle talebin zirve yaptığı saatlerde (Puant zamanı) şarj edilmesi yerel ve bölgesel şebekeleri zorlayarak enerji arz güvenliğini tehlikeye düşürebilir (Zhang vd., 2024). Bunu engellemek için EV şarj alışkanlıklarını doğru anlamak ve uygun saatlerde bu araçların şarj edilmesini sağlamak gerekir. EV'lerin ne zaman, nerede ve ne kadar

enerji tüketeceğini tahmin etmek artık yalnızca sürücüler için önemli değildir. Bu durum aynı zamanda elektrik enerjisi üreten ve dağıtımını yapan firmalarının şebeke yönetimi, dinamik yük dengelemesi, şarj altyapısı planlamasının yanı sıra enerji talebini tahmin eden şirketler için de kritik hâle gelmiştir. Başka bir deyişle, EV enerjisi kullanımını önceden tahmin edebilmek, elektrik şebekelerinde enerji arz güvenliğindeki kararlılığı ve verimliliğini korumada hayati önem taşır.

1.2. Elektrikli Araç Kullanımında Menzil Kaygısı

Elektrikli araçların tüm avantajlarının ve artan kullanımın yaygınlığına rağmen, “menzil kaygısı” olarak adlandırılan önemli bir psikolojik engel, EV’lerin daha geniş kabulünü hâlâ sınırlamaktadır (Modi vd., 2020; Zhang vd., 2020). Menzil kaygısı, sürücülerin EV bataryasının hedefe veya bir şarj istasyonuna varmadan biteceği korkusu olarak tanımlanabilir (Petersen vd., 2022; Wu vd., 2023). Bu kaygı, güncel EV teknolojisi ve altyapısındaki bazı faktörlerden kaynaklanmaktadır. EV’lerin menzili içten yanmalı araçlara göre şarj edilme süresinin fazla olması sebebiyle daha sınırlıdır. Bir başka sorunda bu araçların şarj edilmesi için gereken Kamuya açık şarj istasyonlarının az olması veya zor bulunur olmasıdır. EV araçların şarj sürelerinin içten yanmalı araçların yakıt ikmalinden genellikle daha uzun süre alması da başka bir sorun olarak ortaya çıkmaktadır (Maity ve Sarkar, 2023; Modi ve Bhattacharya, 2022).

EV’lerin doğru ve güvenilir gerçek zamanlı enerji tüketimi tahmin sistemleri, menzil kaygısını azaltmanın temel aracı olarak görülmektedir (Modi vd., 2020; Shen vd., 2022). Sürücülere EV’lerinin kullanımları esnasında batarya sisteminde bulunan mevcut enerjinin seviyesi, planlanan yolculuk için beklenen enerji kullanımı ve kalan menzil hakkında net, güvenilir veriler sunmak oldukça önemlidir. Bu bilgiler, her türlü seyahatte EV kullanımı konusunda sürücülerin menzil konusundaki çekincelerine çözüm üreterek EV’lere olan güveni büyük ölçüde artırabilir (Chen vd., 2024; Petersen vd., 2022). Burada sunulan öngörücü araçlar, kullanıcıların yolculukları esnasında ihtiyaç duydukları araç şarj istasyonlarına zamanında ulaşarak sorunsuz bir şekilde seyahatlerini planlamaları ve ayarlamalarına neden olur. Bu şekilde yapılan öngörüler sürücülerin EV kullanımları esnasında oluşması muhtemel batarya bitim stresini hafifletmesine olanak tanır.

Bu menzil tahmin sisteminin güvenilirliğinin yüksek olması sürücülerin menzil kaygısını azaltarak EV’lere sürücülerin güvenilirliğini sağlamada önemli bir rol alırlar. Uygulamada ise bazen bu sistemlerin yanlış enerji tüketim tahmini yaptıkları görülmektedir (Petersen vd., 2022; Shen vd., 2022), bu durum taşıt kullanıcılarında menzil

kaygısını azaltmak yerine daha da artırarak sürücünün strese girmesini artırabilir. Bu tip sistemlerde yapılan sürekli hatalı tahminler ister istemez sürücülerin bu araçlara olan güvenini hızla sarsar. Bu sistemlerin neden olduğu hatalı menzil tahmini bu araçları kullanmak isteyen kişiler üzerinde güven kaybı döngüsü oluşmasına, sürücülerin bu tip taşıtlara daha temkinli yaklaşmasına veya bazı gerçekleştirmek istedikleri seyahatlerde bu araçların kullanılmamasına yol açar. Bu durum, sonuçta EV üzerindeki menzil kaygısını pekiştirir ve sonuçta EV'lerin yaygınlaşması önünde bir bariyer oluşturarak yaygınlaşma hızını yavaşlatır. Tutarlı ve doğru tahminler sunan modeller geliştirip bunları uygulamaya koymak ise az önce bahsedilen çekinceleri ortadan kaldırıp kullanıcı güvenini artırır. Bu güven artışı hala pil teknolojisi ve hızlı şarj altyapısı tam gelişmemiş olmasına rağmen bu tip araçların seyahatlerde kullanımının benimsenme oranını artırır. Burada bahsedilen nedenlerden, tahmin modellerinin doğruluğu ve kullanıcı güvenin sağlanması durumunun EV'lerin ulaşımda tercih edilme dönüşüm başarısını sağlamak için kritik önem arz etmektedir.

1.3. Toplu Taşımada Elektrikli Otobüslerde Kullanılan Enerji Modellemeleri

Toplu taşıma filolarının, özellikle kentsel otobüslerin elektrikleştirilmesi sürdürülebilir çevre politikalarına destek olmak için şehir içi ulaşımda yapılabilecek temel stratejilerinden birisidir. Toplu ulaşımda Elektrikli otobüslerin (EB'ler) kullanımının sağlanmasıyla içten yanmalı motorların sebep olduğu çevreye partikül madde bırakımı ve çevre ve insan sağlığı üzerinde olumsuz etkilere sebep olan sera gazı emisyonları engellenebilir. Bu şekilde nüfus yoğun olduğu yerleşim yerlerinde hava kalitesinin iyileşmesine ve iklim mücadelesine pozitif katkı sağlanır (Abdelaty ve Mohamed, 2021; Pamula ve Pamula, 2022). Bu kısımda belirtilen EB'lerin avantajlardan bu taşıtların toplu taşımaya kullanımı dünya çapında artmaktadır; özellikle Çin ve Avrupa'da kayda değer artışlar ve iddialı filo dönüşüm hedefleri ortaya koyulduğu görülmektedir (Würtz vd., 2024; Zhang vd., 2024). Bu taşıtlara olan taleplerdeki ivme artışı ve çevre üzerinde görülen olumlu faydalarına rağmen, EB'ler için özel olarak tasarlanmış ayrıntılı, doğru ve pratik enerji tüketiminin belirlenmesi için kullanılan mevcut modellerin geliştirilmesi ve bu modellerin test edilmeleri konularında yapılan çalışmalarda kayda değer bir araştırma boşluğu olduğu ve bu alanla da yapılan çalışmalardan anlaşılmaktadır (Abdelaty vd., 2021; Chen vd., 2021; Jiang vd., 2023). Günümüzde yapılan araştırmaların çoğu binek araçların enerji modeli üzerine odaklanılmış olup, elde edilen modellerin EB'lere doğrudan uygulanması veya basitçe

uyarlanması konularında hala yetersiz kaldığı görülmektedir. EB'leri, özel araçlardan ayıran bazı karakteristik özelliklere sahiptirler. Bunlar arasında sayabileceğimi özellikleri sabit güzergâhlar, yolcu gereksinimlerine bağlı sık duruş ve kalkışlar, araç kütlelerini etkileyen büyük ve değişken yolcu yükleri, daha büyük kabin hacimleri ve sık kapı açılmalar nedeniyle özellikle Isıtma, Havalandırma ve İklimlendirme (HVAC) sistemleri için artan yardımcı güç gereksinimleri ve zorlu sürüş döngüleri bulunur (Abdelaty vd., 2021; Liu ve Liang, 2022; Xu vd., 2023). Mevcut birçok EV enerji modeli bu otobüslere özgü tüm bu faktörleri kapsamamaktadır (Jiang vd., 2023), bu kısımda belirtilen EB'lerin kullanımı esnasında karşılaşılabilecek özellikler toplu taşıma işletmecilerinin önceden erişemeyeceği ayrıntılı verilere dayanmaktadır (Abdelaty ve Mohamed, 2022).

EB enerji modellerindeki araştırma boşluğu, fizik temelli ve makine öğrenmesi (ML) yaklaşımlarının eksikliğinden ziyade, otobüse özgü karmaşık verilerin elde edilmesinden kaynaklanan zorluklarından oluşmaktadır. Örneğin, bu taşıtlarda yolcu yükü sürekli değişir bu durumda EB'nin toplam ağırlığını değiştirerek aracın hareket ettirilmesi esnasında kullanılan enerji miktarının değişimini önemli ölçüde etkiler (Würtz vd., 2024; Zhang vd., 2024). EB'lere gerçek zamanlı yolcu sayım sistemleri kurmak, toplu taşıma firmaları için hem operasyonel hem de finansal açıdan zorlayıcı bir sorun olabilir. Benzer şekilde, sık kapı açılmalarının HVAC enerji talebi üzerindeki kesin etkisini farklı sıcaklık ve nem seviyelerinde ölçmek, detaylı işletme kayıtları ve çevresel veriler gerektirir. Bu verilere her zaman kolayca erişilebilmesi veya kullanılan sistemlere entegre edilmeleri çok kolay olmamaktadır (Widmer vd., 2023). Yapay Sinir Ağları (ANN) ve Uzun Kısa Süreli Bellek (LSTM) ağları gibi gelişmiş yöntemler genel EV enerji tahmini için yaygın şekilde kullanılsa da (Chen vd., 2021; De Cauwer vd., 2017), bunların EB'lere başarılı uyarlanması, yalnızca otobüslere has olan bu karmaşık değişkenlerin güvenilir şekilde ele alınmasına bağlıdır. Dolayısıyla bu alandaki bilimsel boşluk, tamamen yeni modelleme yöntemleri bulmaktan ziyade, mevcut güçlü modelleri otobüse özgü değişkenleri göz önünde tutarak bu verilerin mevcut modelleme sistemlerine uyarlanarak test edilmesi gerekmektedir. Bu tez çalışmasında hem değişken yolcu yükleri ve kapı açılmaları gibi operasyonel belirsizliklerin sisteme dahil edildiği hem de ZTBus veri seti gibi ayrıntılı kayıtların etkin bir biçimde kullanıldığı EB'lerin menzil belirlenmesinde kullanılan modellerin incelemesi amaçlanmıştır. Bu tez çalışmasında, ZTBus veri setinden elde edilen yüksek frekanslı ve gerçek dünya operasyonel verilerine dayanan, veri odaklı bir öngörü modelinin geliştirilmesi ve

titizlikle değerdendirilmesi önerilmektedir. Bu kapsamda, LightGBM gibi gelişmiş gradyan artırma (gradient boosting) algoritmalarının performansı, geleneksel bir doğrusal modele karşı sistematik olarak karşılaştırılacaktır. Çalışmanın temel hedefi, saniyelik ayrıntılı verilerden ve sağlam bir doğrulama metodolojisinden (TimeSeriesSplit) yararlanarak, yalnızca anlık ve toplam enerji tüketimini olağanüstü yüksek bir doğrulukla tahmin etmekle kalmayıp, aynı zamanda bu tüketimi yönlendiren temel faktörlere dair içgörüler sunan bir model oluşturulabileceğini kanıtlamaktır. Nihai amaç, elektrikli otobüs filo yönetimindeki gerçek dünya uygulamaları için güvenilir bir araç olarak hizmet edebilecek, doğrulanmış bir modelleme yaklaşımı sunmaktır.

1.4. Tez Çalışmasının Amacı

Bu tez çalışması, hızla gelişen elektrikli araçların kullanımı esnasında kullanılan menzil belirleme modellerinde kullanılan algoritmaların iyileştirilmesi ve bunların kullanımı esnasında doğru enerji tüketimi tahminine duyulan ihtiyacın iyileştirilmesi üzerine odaklanılmıştır. Bu çalışmada, akademik literatürde yapılan çalışmalar ve ZTBus veri kümesine dayalı bilgiler temel alınmıştır (Widmer vd., 2023). ZTBus veri kümesi, İsviçre, Zürih'te çalışan bir elektrikli şehir otobüsü filosunda zamana dayalı ayrıntılı işletme verilerini kaydettiği için buradan alınacak veriler önemlidir. Bu sistemden alınan Veri seti aracın toplam elektrik güç talebini, tahrik sistemi davranışını, odometri ve GPS verilerini, çevresel sıcaklık gibi koşulları, kapı açılma kayıtlarını ve dinamik yolcu sayımlarını içermektedir. Araçların hizmet verdiği sürede çeşitli mevsimler ve işletme koşulları göz önünde bulundurularak sistematik şekilde toplanan bu veriler, EB'lere yönelik enerji tüketimi modellerinin geliştirilmesi, kalibrasyonu ve titizlikle doğrulanması için oldukça önemli uygulama verilerini içermektedir. Böylesine kapsamlı ve uygulamadan elde edilen gerçek veri setlerinden elde edilen bulgular, EB enerji kullanımını etkileyen faktörler arasındaki ayrıntılı etkileşimlerin belirlenmesi için geliştirilecek menzil belirleme de kullanılacak modellerin oluşturulması için şarttır.

Bu tezin birincil amacı, elektrikli ulaşırma sistemlerinde enerji tüketimi tahminini algoritmalarını geliştirerek iyileştirmek ve daha gerçekçi tahminler sunmayı hedeflemektir. Bunu gerçekleştirmek için, EB'lerin özgün nitelikleri ve kentsel karbonsuzlaştırmadaki kilit rolleri dikkate alınarak öngörücü modeller geliştirilecek, eleştirel olarak değerdendirilecek ve elde edilen sonuçlar iyileştirilecektir. Araştırmada, ZTBus veri kümesinden ve akademik literatürde bildirilen mevcut yöntemler, temel etkenler ve bulgulara dair içgörüler titizlikle işlenerek bunlar arasındaki ilişkiler

birleştirilmiştir. Bu ilişkilerin geliştirilmesiyle araç parametreleri, işletme desenleri ve değişen çevresel koşullar arasındaki karmaşık, doğrusal olmayan bağlantıları mümkün olduğunca doğru bir şekilde yansıtan güçlü girdi özelliklerine sahip olan bu mimari model sayesinde menzil belirlenmesinde kullanılan sistemin daha doğru ve güvenilir sonuçlar sunmasında yardımcı olacaktır. Tezin nihai amacı menzil belirlemede kullanılacak sistemlerin daha güvenilir ve pratik olarak uygulanabilecek şekilde sonuçlar elde etmektir. Yapılacak çalışmadan elde edilen sonuçlarla sürücülerin menzil kaygısı azaltılacak ve sistem EB filo operasyonlarını daha yüksek verimlilik ve düşük maliyetle işletilmesi için optimize edilerek sürdürülebilir, çevre dostu şehir içi ulaşım çözümlerine yönelik geniş çaplı geçişi desteklenecektir.



2. KAYNAK ARAŞTIRMASI

2.1. EV Enerji Tüketimi Modelleme Yaklaşımları

EV’lerde enerji tüketimini tahmin etmek karmaşık bir iştir ve araştırmacılar bunu çeşitli modelleme yöntemleriyle çözmeye çalışmaktadırlar. Bu yöntemlerin her birinin kendine özgü ilkeleri, avantajları ve sınırlamaları vardır. Bu iş için kullanılan yöntemler;

- i) Fizik temelli (analitik) modeller,
- ii) Veri odaklı modeller,
- iii) Hibrit modeller,

Olmak üzere üç ana gruba ayrılırlar. Bu temel yaklaşımları anlamak için özellikle EB uygulamaları alandaki ilerlemeleri ve zorlukları bu bağlamlar içinde değerlendirmek çok önemlidir.

2.1.1. Fizik temelli (analitik) modeller

Fizik temelli modeller, analitik veya beyaz kutu modeller olarak da adlandırılır, bir aracın enerji tüketimini hesaplamak için başlıca Newton mekaniği ve termodinamik olmak üzere temel fizik yasaları kullanır (Huang vd., 2023; Shen vd., 2023). Kinetik olarak hareket eden araç üzerindeki kuvvetleri tanımlayan denklemler kullanılır. Hareket esnasında aracın yuvarlanma direnci (lastik deformasyonu ve yol yüzeyi etkileşimi nedeniyle), aerodinamik sürüklenme (hızla artan hava direnci), yerçekimi kuvvetleri (yokuş iniş ve çıkışlarda) ve atalet kuvvetleri (ivmelenme ve frenleme sırasında) (Wu vd., 2023) göz önüne alınır. Ayrıca sistem tasarımı elektrik motoru, güç elektroniği (inverter/dönüştürücü), batarya sistemi ve aktarma organları (şanzıman, akslar) gibi güç aktarım bileşenlerinin enerji dönüşüm verimliliklerini içerir (Shen vd., 2022).

Araç enerji tüketimi, performans ve menzil tahminlerinin doğruluğu, kullanılan matematiksel modelin girdilerine doğrudan bağlıdır. Özellikle elektrikli ve hibrit araçlarda, sistemin dinamik davranışlarını gerçeğe yakın şekilde temsil edebilmek için modele entegre edilen parametrelerin araca özgü ve hassas biçimde tanımlanmış olması gerekmektedir. Bu bağlamda; aerodinamik sürüklenme katsayısı (C_d), yuvarlanma direnci katsayısı (C_{rr}), aracın kütle dağılımı ve sistem bileşenlerinin verimlilik haritaları kritik rol oynamaktadır. Örneğin, motor verimliliği genellikle tork ve motor hızı eksenlerinde değişkenlik gösterirken, batarya verimliliği akım ve şarj durumu (SOC) gibi değişkenlere bağlı olarak farklılaşmaktadır.

Bu modeller genellikle başlangıç aşamasında araç tasarımının bileşen boyutlandırılması, bu tasarım parametrelerinin (örn. araç kütlesi, aerodinamik sürüklenme katsayısı, batarya kapasitesi, yol eğimi) enerji performansına teorik etkisinin değerlendirilmesi ve elde edilen tasarımın standart sürüş döngüleri altında araç davranışlarının simüle edilmesi için kullanılırlar (Liu vd., 2021; Shen vd., 2023).

Fizik temelli modellerin en önemli avantajı, yüksek derecede yorumlanabilme özelliğine sahip olmalarıdır. Bu modeller, denenmiş ve iyi bilinen fiziksel denklemler üzerine kurulu olduklarından, mühendis ve araştırmacıların belirli bir parametre değişikliğinin enerji tüketimini nasıl etkilediğini doğrudan görmesini hesaplayarak yorumlayabilmesine olanak sağlar. Ayrıca, araçlar için bilinmesi gereken anahtar parametrelerin ve bunların alması gereken değerleri yoğun şekilde tahmin edilebilmesi, uygulamadan elde edilmiş olan gerçek işletme verilerinin sınırlı olması nedeniyle makul araştırmacılara önceden doğru tahminlerde bulunmalarına imkân sağlayabilirler (Liu vd., 2021). Bu parametrelerin elde edilmesi zor, maliyetli olabilir ve araç yaşı, bakım durumu, lastik basıncı veya ortam sıcaklığı gibi faktörlerle bu durum önemli ölçüde değişebilir (Jiang vd., 2023; Zhang vd., 2020).

Salt fizik-temelli modeller, dinamik trafik etkileşimleri, sürücü davranışları (örneğin ivmelenme ve frenlemedeki agresiflik düzeyi) ile yardımcı sistemlerin (örneğin HVAC) değişken kullanımını içeren gerçek sürüş koşullarının tüm rastlantısal ve karmaşık doğasını tam anlamıyla yansıtmakta genellikle yetersiz kalır. Bu tür etkilerin doğru biçimde modellenmesi, ya kapsamlı veri temelli kalibrasyon süreçlerini ya da belirli düzeyde basitleştirici varsayımların yapılmasını gerektirir

Salt fizik-temelli modeller, dinamik trafik etkileşimleri, sürücü davranışları (örneğin ivmelenme ve frenleme agresifliği) ile HVAC gibi yardımcı sistemlerin değişken kullanımı gibi, gerçek sürüş koşullarının rastlantısal ve karmaşık doğasını tam olarak yansıtmakta genellikle yetersiz kalmaktadır. Bu tür unsurların doğru biçimde modellenmesi, kapsamlı kalibrasyon süreçleri veya çeşitli basitleştirici varsayımlar gerektirir; bu da farklı işletme senaryolarında modelin tahmin doğruluğunu zayıflatabilir (Modi vd., 2020; Shen vd., 2023). Menzil belirlemede EB’lerde yolcu biniş-inişlerinden kaynaklanan aracın yükünün dolasıyla kütlesinin değişimi ve yüksek sıcaklıklarda araç içinde artan sıcaklığın düşürülmesi için kullanılan HVAC sistemlerinin devreye alınması da aracın enerji tüketiminde ciddi dalgalanmalara yol açmaktadır; bu kısımda belirtilmeye çalışılan bu vb. dinamik faktörler yalnızca fiziksel parametrelere

dayalı modellerin araç menzili tahmininde kullanılması sistemin doğru veriler sunabilme yeteneğini azaltır ve modelleme sürecini zorlaştırır (Kaya vd., 2024).

2.1.2. Veri odaklı modeller

Fizik temelli yaklaşımların aksine, veri odaklı modeller istatistiksel yöntemler ve ML algoritmalarını kullanarak geçmiş işletme verilerinden doğrudan veriler arasındaki, korelasyonları ve öngörüşel ilişkiler belirlemekte kullanılır (Liu vd., 2021; Zhang vd., 2020). Bu modeller, hareketi ve enerji dönüşümünde kullanılan fiziksel yasalar ve bunlardan elde edilen matematiksel formülasyonlara ihtiyaç duyulmaz. Veri odaklı modellerde bu işlemler taşıta ait sistemsel giriş özellikleri (örn. hız, ivmelenme, ortam sıcaklığı, yol eğimi) anlık değerleri kullanılarak ile taşıtların hareket ettirilmesinde kullanılacak çıktı değişkeni olan enerji tüketimi arasındaki geliştirilecek algoritmalara göre elde edilecek karmaşık eşlemeyi “öğrenirler”.

Veri odaklı enerji tüketimi modellemede kullanılan teknikler ise çok çeşitlidir ve sürekli olarak geliştirilmektedirler. Veri odaklı olarak kullanılan basit yöntemler arasında, girişler ile çıktı arasında doğrusal bir ilişkiyi kurmada kullanılan yöntemler doğrusal regresyon ve çoklu doğrusal regresyon (MLR) yöntemleridir (Abdelaty ve Mohamed, 2021; De Cauwer vd., 2017). Ancak EV enerji tüketimi ile giriş verileri arasında doğrusal olmayan ilişkiye sahip olduklarından, bunların yerine genellikle uygulamada daha gelişmiş ML algoritmaları tercih edilerek kullanılır. Bu kapsamda kullanılan yöntemler arasında; veri sınıflandırma ve regresyon problemlerinde etkili olan Yapay Sinir Ağları (ANN) (Chen vd., 2021; Pamula ve Pamula, 2022), veri kümesi üzerinde optimal ayırma düzlemini belirleyen Destek Vektör Makineleri (SVM) ve regresyon karşılığı olan Destek Vektör Regresyonu (SVR) (Li vd., 2022), karar verme sürecinin doğruluğunu artırmak amacıyla birden fazla karar ağacını birleştiren Rastgele Ormanlar (RF) gibi ansambl öğrenme yöntemleri (Achariyaviriya vd., 2023; Lin vd., 2023), ve model hatalarını aşamalı şekilde azaltan Gradyan Artırma Makineleri (örneğin XGBoost, LightGBM) (Abdelaty vd., 2021; Zhang vd., 2024) yer almaktadır.

Bunlara ek olarak, Derin Öğrenme Ağları (DLN) zamansal veri modellemede öne çıkmaktadır. Özellikle Uzun Kısa Süreli Bellek (LSTM) ağları, hız profilleri gibi zaman serilerini yüksek doğrulukla modelleyebilme potansiyeline sahiptir (Chen vd., 2021; Xu vd., 2023). Gauss Süreç Regresyonu (GPR) ise yalnızca tahminler sunmakla kalmayıp belirsizlik ölçüsü de sağlayan olasılıksal bir yaklaşım sunar (Jiang vd., 2023; Karadag vd., 2023).

Veri odaklı modellerin temel gücü, saf fiziksel denklemlerle doğru şekilde temsil edilmesi zor veya imkansız olan çok sayıda veri arasındaki karmaşık, doğrusal olmayan ilişkileri ve etkileşimleri modelleyebilme imkanı sunmasıdır (Jiang vd., 2023). Yeterince büyük ve temsil edici veri kümeleri üzerinde eğitildiklerinde, bu modeller genellikle gerçek dünya değişkenliğini yakalamada kullanılan fizik temelli modellere göre daha üstün tahmin doğruluğu elde edebilirler (Abdelaty vd., 2021). Bu modeller, sürüş tarzı veya trafik sıkışıklığı gibi faktörlerin etkilerini, bu olgular için belirlenmesi gereken alt modellere ihtiyaç duymadan sağlanan veri kümelerinden dolaylı olarak öğrenebilirler.

Veri odaklı yaklaşımlar birçok avantaj sunsa da, bazı önemli sınırlılıkları da beraberinde getirmektedir. Bu yöntemlere yönelik en yaygın eleştirilerden biri, özellikle derin sinir ağları gibi karmaşık modellerin "kara kutu" niteliğidir. Modelin bir öngöründe bulunma süreci, genellikle insan tarafından doğrudan anlaşılabilir bir biçimde açıklanamaz. Örneğin, derin öğrenme modelleri çoğu zaman elde ettikleri sonuçların altında yatan nedensel mekanizmalar veya fiziksel ilkeler hakkında sınırlı içgörü sağlar (Wang vd., 2021).

Bu şeffaflık eksikliği, özellikle tahminin gerekçesinin de sonuç kadar önemli olduğu uygulamalarda, modelin güvenilirliği ve kullanıcı tarafından benimsenmesi açısından ciddi bir engel teşkil edebilir. Ayrıca, bu modellerin genel başarımı, büyük ölçüde eğitim verisinin kalitesine, miktarına ve temsiliyetine bağlıdır (Zhang vd., 2020; Zhu vd., 2025). Verideki eksiklikler, önyargılar veya dengesizlikler, modelin genelleme yeteneğini olumsuz yönde etkileyebilir ve doğruluğu daha düşük olan sonuçlar elde edilmesine sebep olabilir. Özellikle derin öğrenme mimarileri gibi karmaşık modellerin eğitimi hesaplama açısından yoğun ve önemli donanım kaynaklarına ihtiyaç duyulmasına sebep olabilir.

Bu modellerin bir diğer önemli sınırlılığı ise, dikkatli şekilde geliştirilmediklerinde ve uygun doğrulama yöntemleri (örneğin çapraz doğrulama, tutma test setleri) kullanılmadığında, eğitim verisindeki gürültüleri gerçek desenlerle karıştırarak aşırı uyum (overfitting) riski taşımalarıdır. Bu durum, modelin eğitim verisinde yüksek doğruluk sergilerken, yeni ve gerçek dünya verilerinde başarımının ciddi şekilde düşmesine neden olabilir. Dolayısıyla, veri odaklı modellerin güvenilirliği yalnızca verinin kalitesine değil, aynı zamanda model geliştirme ve doğrulama süreçlerinin titizlikle yürütülmesine de bağlıdır.

2.1.3. Hibrit modelleme yaklaşımları

Fizik temelli ve veri odaklı paradigmaların tamamlayıcı güçlü ve zayıf yönlerini göz önünde tutan araştırmacılar, giderek bunların karışımı olan ve her sistemin güçlü yönlerini kullanan hibrit modelleme yaklaşımlarını benimsemektedirler (Huang vd., 2023; Li vd., 2022). Bu yöntemler, her iki paradigmanın öğelerini bir araya getirerek, tek başına hiçbir yaklaşımın sunamadığı sağlamlık, doğruluk ve yorumlanabilirlik sunan modeller oluşturmayı hedeflemektedirler. Hibrit modellemede temel amaç, güvenilir olduğu durumlar için fiziksel bilgiyi kullanmak, daha karmaşık ve belirsiz durumlarda ise uygun sonuçlar elde etmek üzere veri odaklı tekniklerden yararlanmaktır. Bu modellerin temel gerekçesi, fiziksel yasaların sağladığı yorumlanabilirlik ile makine öğreniminin karmaşık, doğrusal olmayan etkileşimleri ve gerçek veri içindeki nüansları öğrenebilme esnekliği arasında denge kurmaktır (Jiang vd., 2023). Hibrit modelleme özellikle bazı fiziksel ilişkilerin iyi tanımlanabildiği, ancak diğer etmenlerin (örneğin sürücü davranışı, verimlilik eğrileri veya ortam koşullarının etkisi) karmaşık, belirsiz ya da eksik veriyle temsil edildiği durumlarda önemli avantajlar sunar. Tam kapsamlı fiziksel model kalibrasyonu için gerekli tüm veriler mevcut değilse veya bu tür modellerin oluşturulması yüksek maliyetliyse, hibrit modeller doğruluk, genellenebilirlik ve yorumlanabilirlik açısından dengeli bir çözüm sağlayabilir.

Hibrit modellerin çeşitli uygulama biçimleri bulunmaktadır. Örneğin, fizik temelli bir model, ilk enerji tüketimi tahminlerini üretmek ya da anlamlı girdi özellikleri çıkarmak için kullanılabilir; bu çıktılar, artık hataları öğrenerek tahminleri iyileştiren bir veri tabanlı modele (örneğin yapay sinir ağı veya regresyon modeli) beslenebilir (Huang vd., 2023; Li vd., 2022). Bu yapı, özellikle modelin doğruluğunu artırmak ve sistematik sapmaları azaltmak için etkilidir.

Bir diğer hibrit yaklaşım olan Fizik Bilgilendirmeli Sinir Ağları (PINN), fiziksel yasaları genellikle diferansiyel denklemler biçiminde doğrudan ağı kayıp fonksiyonuna entegre ederek, öğrenme sürecinin fiziksel tutarlılığını güvence altına alır. Ayrıca, fizik-temelli modeller, gerçek veri miktarının yetersiz veya toplanmasının maliyetli olduğu durumlarda büyük ölçekli sentetik veri kümeleri üretmek amacıyla da kullanılabilir; bu sentetik veriler, veri odaklı modellerin eğitimi için önemli bir kaynak sunar (Abdelaty vd., 2021).

Bazı hibrit yaklaşımlar ise, güç aktarım sistemleri gibi yapılar için fiziksel ilkeleri temel alırken, batarya veya elektrik motoru gibi bileşenlerin verimlilik karakteristiklerini

farklı çalışma koşulları altında modellemek ya da bu bileşenlere ait parametreleri kalibre etmek için veri tabanlı yöntemlerden yararlanır (Shen vd., 2023).

EV enerji kullanımının belirlenmesinde bu üç modelleme paradigmasının fizik temelli, veri odaklı ve hibrit incelenmesi, evrensel olarak kullanılabilecek bir yaklaşım sunmadığı söylenebilir. Her paradigmanın başarısı, uygulama bağlamına, veri özelliklerine ve kullanılabilirliğine, hesaplama kaynaklarına ve modelleme çabalarının önceliklerine (tahmin doğruluğu mu yoksa yorumlanabilirlik mi) bağlı olarak önemli ölçüde değişir. Bu durum, “No Free Lunch” teoreminde de hiçbir algoritmanın tüm problemler için gereken en iyi bir çözümü sunmadığı belirtilmiştir.

Fizik temelli modeller, fizik yasalarına dayalı olarak yüksek yorumlanabilirlik sunarken ve temel parametreler çok iyi bilindiğinde sınırlı veriyle bile uygun sonuçlar elde edilebilir. Ancak gerçekte uygulama koşullarında çok sayıda modellenmemiş etki veya belirsiz parametre içeren karmaşık verilerin bir arada kullanılması durumu bu algoritmaların doğru sonuçlar elde edememesine sebep olur (Jiang vd., 2023; Modi vd., 2020). Rastlantısal sürücü davranışı veya yardımcı sistemlerin detaylı etkilerini uygun kalibrasyon işlemlerine tabi tutulmazlarsa doğru sonuçlar vermekte zorlanabilirler.

Veri odaklı modeller, verilerden karmaşık, doğrusal olmayan desenleri keşfetmede ve büyük, temsil edici veri setleriyle yüksek doğruluk elde etmede başarılıdır (Wang vd., 2021). Ancak genellikle “kara kutu” gibi davranarak öngörülerinin nedenlerini anlaşılmaz kılar ve performansları veri kalitesi ve veri miktarına sıkı sıkıya bağlıdır (Zhang vd., 2020).

Hibrit modeller, her iki yaklaşımın en iyi yönlerini bir araya getirerek, fizik temelli modellerin şeffaflığı ile veri odaklı yöntemlerin karmaşıklık ve belirsizlikle başa çıkma yeteneğini bir arada sunmayı hedeflemektedirler (Huang vd., 2023). Buna karşın, bu iki farklı paradigmanın etkin şekilde bütünleştirilmesi, model tasarımı ve kalibrasyonunda sistem karmaşıklıklarını da beraberinde getirir.

EV uygulamaları için hangi modelleme paradigmasının kullanılması gerektiğinin belirlenmesi durumu oldukça karmaşık olan bir süreçtir. Örneğin EB’lerin Yolcu yükündeki değişimleri taşıt kütlesini önemli ölçüde etkilemektedir, yine bunlarda kullanılan HVAC sistemlerinin çevresel koşullara bağlı yüksek enerji çekişi ve kent içi güzergâhlardaki sık “Dur-Kalk” döngüleri gibi otobüse özgü sürekli olarak değişebilen faktörler, saf fiziksel modellerin kullanımının ve tüm değişen koşullar için doğru bir biçimde kalibre edilmesini zorlaştırır. Bu nedenle, EB’lerin menzil tahmini edilmesi işleminde yeterli veri varsa veri odaklı algoritmalar veya hibrit yaklaşım algoritmalarının

yapılan bilimsel çalışmalarda sıkça tercih edildiği görülür (Abdelaty vd., 2021; Chen vd., 2021). Dolayısıyla, bu tez kapsamında modelleme paradigması seçimi; araştırma hedefleri, mevcut veri (örn. ZTBus veri kümesi) ve tahmin gücü ile açıklayıcı netlik arasındaki denge göz önünde bulundurularak elde edilen sonuçlar gerekçeleriyle beraber açıklanmaya çalışılmıştır. Çizelge 2.1’de EV’lerin enerji tüketimini belirlemede kullanılan modelleme yaklaşımlarının karşılaştırılması verilmiştir.

Çizelge 2.1. EV’lerin enerji tüketimi modelleme yaklaşımlarının karşılaştırması

Başlık	Fizik Tabanlı Modeller	Veri Odaklı Modeller	Hibrit Modeller
Temel Prensipler/ Metodoloji	- Araç kütlesi, aerodinamik sürtünme vb. fiziksel parametreler - Denklemler - tabanlı enerji hesaplaması - Farklı hız/ivme koşullarında simülasyon	- Geçmiş sürüş verilerinden kalıp çıkarma - İstatistiksel veya ML yöntemleri - Modelin veriyle eğitilmesi	- Fizik tabanlı modele ML “düzeltmesi” ekleme - Hata kalıplarını öğrenme - İki yaklaşımın kombinasyonu
Tipik Girdi Özellikleri	- Kütle, aerodinamik katsayı - Verimlilik haritaları - Hız/ivmelenme ve yol eğimi	- Hız, ivmelenme, yol eğimi - Batarya SOC, Sıcaklık, HVAC kullanımı - Yolcu yükü, sürüş tarzı	- Fiziksel parametreler (kütle, sürtünme vb.) - Operasyonel veriler (Hız, SOC, Sıcaklık vb.)
Güçlü Yönler	- Yüksek yorumlanabilirlik sunar. - Az veriyle bile çalışabilir.	- Karmaşık, doğrusal olmayan ilişkileri yakalar. - Gerçek koşullara hızlı uyum sağlar.	- Açıklanabilirlik ve doğruluk arasında denge sağlar. - Bilinen fiziki kullanıp nüansları öğrenir.
Sınırlamalar	- Pek çok kesin araç verisine ihtiyaç duyar. - Gerçek sürüş tarzı ve ek yükleri yakalamakta zorlanır.	“Kara kutu” olup açıklaması zordur. - Çok miktarda ve temiz veri gerektirir.	- Geliştirmesi ve ayarlaması karmaşık olabilir. - Her iki yöntemin eksilerini taşıyabilir.
Örnek Çalışmalar (Yazar, Yıl)	(Shen vd., 2022) (Fiori vd., 2016)	(De Cauwer vd., 2017) (Jiang vd., 2023) (Chen vd., 2021) (Pamula ve Pamula, 2022) (Zhang vd., 2020) (Zhu vd., 2025)	(Huang vd., 2023) (Li vd., 2022) (Jiang vd., 2023)

2.2. EV'lerde Enerji Tüketimini Belirleyen Faktörler

EV'lerin enerji tüketimi sabit bir değere bağlı olmayıp, birbiriyle etkileşim halindeki çok sayıda faktöre bağlı olarak önemli ölçüde değişkenlik göstermektedir. Bu faktörler; aracın teknik özellikleri, sürüş tarzı ve sürücü davranışları, güzergâhın topografik yapısı ve çevresel koşullar gibi ana kategorilere ayrılabilir. Dolayısıyla, bu etkenlerin ve aralarındaki karmaşık ilişkilerin doğru bir şekilde anlaşılması, güvenilir enerji tüketimi tahmin modelleri geliştirmenin temelini oluşturmaktadır.

2.2.1. Araç özellikleri

EV'lerin enerji tüketimini belirleyen bu faktörler, temel olarak aracın fiziksel tasarımı ve operasyonel durumuyla ilgilidir.

- **Araç Kütlesi:** Aracın boş ağırlığı ile yolcu ve yük gibi değişken unsurları içeren toplam kütlesi, enerji tüketimini doğrudan etkileyen en temel faktörlerden biridir. Kütledeki artış, aracın ivmelenmesi, yokuş tırmanması ve yuvarlanma direncini aşması için gereken enerji miktarını artırmaktadır (Abdelaty ve Mohamed, 2021; Würtz vd., 2024). Bu durum, özellikle yolcu sayısındaki sürekli değişimin toplam kütleyi anlık olarak etkilediği için EB'ler için kritik bir rol oynamaktadır (Y. Liu ve Liang, 2022).
- **Aerodinamik Tasarım:** Aracın aerodinamik sürtünme katsayısı (C_d) ile ön kesit alanının (A) bir fonksiyonu olan aerodinamik direnç, özellikle yüksek hızlarda enerji tüketimini önemli ölçüde etkilemektedir. Bu direnci yenmek için gereken enerji, hızın karesiyle orantılı olarak arttığından, otoyol gibi yüksek hızlı sürüş senaryolarında menzile belirlemede en belirleyici bir faktör haline gelmektedir (Abdelaty ve Mohamed, 2021; Li vd., 2021).
- **Batarya Özellikleri:** Batarya sisteminin genel sağlık durumu, anlık özelliklerini belirlemede ve enerji verimliliği açısından kritik bir rol oynamaktadır. Bataryanın SOC değeri, mevcut enerji seviyesini göstermesinin yanı sıra deşarj ve rejeneratif frenleme anındaki iç direnci ve verimliliği de etkilemektedir (Qin vd., 2021; Zhang vd., 2020). Benzer şekilde, batarya sıcaklığı da performansı doğrudan etkileyen faktörler arasında yer almaktadır. Batarya sıcaklığının optimum aralığın dışındaki durumları bataryanın verimliliği düşürmekte, güç çıkışını sınırlamakta ve bataryanın yıpranma sürecini hızlandırmaktadır (Liu vd., 2018; Xu vd., 2023). Uzun vadede ise kapasite kaybı ve iç direnç artışı olarak tanımlanan batarya yaşlanması olarak adlandırılan duruma neden olmaktadır. Bu durumda aynı rota için gereken enerji miktarını artmasına neden olmaktadır.

- **Güç Aktarma Organı Verimliliği:** Elektrik motoru, invertör ve aktarma organları gibi güç aktarım sistemi bileşenlerinin, bataryadaki kimyasal enerjiyi tekerleklerde mekanik işe dönüştürme verimliliği, toplam enerji tüketiminde belirleyicidir. Bu sistemin verimliliği, tork, hız ve sıcaklık gibi anlık çalışma koşullarına bağlı olarak sürekli değişkenlik göstermektedir (Jiang vd., 2023; Shen vd., 2022). Yapılan çalışmalar, ortam sıcaklığının batarya performansından bağımsız olarak doğrudan güç aktarım organlarının verimliliğini de etkileyebildiğini ortaya koymaktadır (Jiang vd., 2023).

- **Yardımcı Sistemler (HVAC ve Diğerleri):** HVAC gibi yardımcı sistemler, bir EV'nin toplam enerji sarfiyatının önemli bir bölümünü oluşturabilmektedir. Özellikle aşırı sıcak veya soğuk hava koşullarında kabin içi konforu sağlamak için harcanan enerji, toplam tüketimi ciddi oranda artırmaktadır (Abdelaty vd., 2021; Würtz vd., 2024; Zhang vd., 2024). Bu etki, daha geniş kabin hacmine sahip olan ve kapıları sık sık açılan EB'lerde daha da belirgin rol oynamaktadır. Bunların yanı sıra aydınlatma, hidrolik direksiyon, bilgi-eğlence sistemleri ve diğer kontrol üniteleri de toplam enerji tüketimine ek bir yük getirmektedir (Abdelaty ve Mohamed, 2021).

2.2.2. Sürüş tarzı ve sürücü davranışları

Sürücü tarafından benimsenen sürüş tarzı ve sergilenen davranışlar, aracın enerji tüketimi üzerinde son derece belirleyici bir rol oynamaktadır.

- **Hız Profilleri:** Aracın zaman içindeki hız değişimi olarak tanımlanan hız profili, enerji tüketimini doğrudan etkilemektedir. Ortalama hız bir gösterge olmakla birlikte, hızdaki dalgalanmalar, dur-kalk döngülerinin sıklığı ve hız limitlerine uyum gibi faktörler de önemlidir. Sürekli yüksek hızlar, aerodinamik sürtünmeyi katlanarak artırdığı için enerji sarfiyatını yükseltmektedir (Zhang vd., 2020). Buna karşılık, yoğun şehir içi trafikteki düşük hızlar da uzun süreli yardımcı sistem kullanımı ve motorun verimsiz çalışma noktalarında kalması nedeniyle enerji tüketimini artırabilmektedir (Jiang vd., 2023).

- **Hızlanma ve Yavaşlama Desenleri:** Sürücünün hızlanma ve yavaşlama anlarındaki agresiflik seviyesi, enerji tüketimini belirleyen kritik bir faktördür. Sert ve ani hızlanmalar bataryadan anlık olarak yüksek güç çekilmesine yol açarken, ani frenlemeler ise daha yumuşak ve öngörülü yavaşlama manevralarına kıyasla rejeneratif frenlemeden elde edilecek enerji geri kazanımının verimini düşürmektedir (Chen vd., 2024; Maity ve

Sarkar, 2023). Bu nedenle, birçok bilimsel çalışmada sürüş tarzı, enerji modellemesinde kilit bir değişken olarak ele alınmaktadır (Guo vd., 2020; Würtz vd., 2024).

- **Rejeneratif Frenleme:** EV'lerin temel avantajlarından biri olan rejeneratif frenleme, yavaşlama ve frenleme esnasında aracın kinetik ve potansiyel enerjisinin bir kısmını tekrar elektrik enerjisine dönüştürerek bataryayı şarj etme yeteneğidir. Bu enerji geri kazanımının miktarı ve verimliliği, özellikle sık dur-kalk manevralarının yapıldığı kentsel sürüş koşullarında net enerji tüketimini önemli ölçüde düşürebilmektedir (Miri vd., 2021; Zhang vd., 2020). Rejenerasyonun verimliliği; yavaşlama ivmesi, bataryanın SOC değeri, batarya sıcaklığı ve motorun anlık çalışma karakteristikleri gibi bir dizi faktöre bağlı olarak değişkenlik göstermektedir.

2.2.3. Çevresel koşullar

Aracın içinde bulunduğu dış ortam koşulları, enerji tüketimini hem doğrudan hem de dolaylı yollarla önemli ölçüde etkilemektedir.

- **Ortam Sıcaklığı:** Ortam sıcaklığı, enerji tüketimini birden çok mekanizma üzerinden etkileyen en önemli çevresel faktörlerden biridir. Sıcaklık, bataryanın elektrokimyasal performansını, verimliliğini ve rejeneratif frenleme sırasındaki şarj kabul kapasitesini doğrudan etkilemektedir (Zhang vd., 2020). Bununla birlikte en büyük etki, HVAC sisteminin kullanımını tetiklemesidir. Çok soğuk havalarda ısıtma, çok sıcak havalarda ise iklimlendirme sisteminin devreye girmesi, yüksek enerji tüketimine yol açmaktadır (Würtz vd., 2024). Yapılan çalışmalar, enerji tüketimi ile ortam sıcaklığı arasında “U-şeklinde” bir ilişki olduğunu ve en yüksek tüketimin aşırı soğuk ve sıcak koşullarda gerçekleştiğini ortaya koymaktadır (Jiang vd., 2023).

- **Hava Koşulları:** Sıcaklığın yanı sıra yağmur, kar ve buzlanma gibi diğer hava olayları da enerji tüketimini etkilemektedir. Bu tür koşullar, yol yüzeyindeki sürtünmeyi artırarak yuvarlanma direncini yükseltebilmektedir. Ayrıca, sürücüyü daha temkinli bir sürüş tarzı benimsemeye (daha düşük hızlar ve daha yumuşak manevralar) yönlendirerek enerji talep profilini de değiştirmektedir (De Cauwer vd., 2017; Wang vd., 2024). Yağışlı havalar, aynı zamanda silecekler ve buğu çözücüler gibi yardımcı sistemlerin daha fazla kullanılmasına neden olarak ek bir enerji yükü oluşturmaktadır.

- **Rüzgâr Hızı ve Yönü:** Rüzgârın hızı ve yönü, aracın maruz kaldığı toplam aerodinamik yükü değiştirerek enerji tüketimini etkilemektedir. Cepheden esen rüzgâr (karşı rüzgâr) aerodinamik sürüklenmeyi artırarak aynı hızı korumak için daha fazla enerji harcanmasını gerektirirken, arkadan esen rüzgâr bu etkiyi bir miktar azaltmaktadır (Zhang

vd., 2020). Yandan esen rüzgârlar ise hem sürüklenmeyi artırabilmekte hem de aracı rotada tutmak için ek direksiyon manevraları gerektirdiğinden enerji tüketimini etkileyebilmektedir (Lin vd., 2023).

2.2.4. Rota topografyası ve operasyonel koşullar

Güzergâhın fiziksel özellikleri ile aracın bu güzergâhtaki operasyonel koşulları da enerji tüketimi üzerinde önemli bir etkiye sahiptir.

- **Yol Eğimi/Yükselti Değişimleri:** Yolun eğimi ve güzergâh boyunca meydana gelen yükselti değişimleri, enerji tüketimini doğrudan etkilemektedir. Yokuş yukarı sürüş, yer çekimi kuvvetini yenmek için ilave enerji gerektirmekte ve engebeli arazilerde toplam tüketimin önemli bir bileşenini oluşturmaktadır (De Cauwer vd., 2017; Wang vd., 2024). Buna karşılık, yokuş aşağı inişlerde potansiyel enerjinin bir kısmı rejeneratif frenleme ile geri kazanılarak tırmanışta harcanan enerjinin bir bölümü telafi edilebilmektedir. Bu nedenle, yalnızca toplam yükselti değişimi değil, aynı zamanda tırmanış ve inişlerin profili (uzun ve hafif eğimlere karşı kısa ve dik rampalar) de sonuç üzerinde belirleyici olmaktadır.
- **Yol Türü ve Koşulları:** Yol yüzeyinin cinsi (asfalt, beton vb.) ve fiziksel durumu (düzgün, bozuk, çukurlu) yuvarlanma direncini doğrudan etkilemektedir. Yolun türü (örneğin kentsel anayol, otoyol veya kırsal yol) ise beklenen hız profillerini, trafik yoğunluğunu ve kavşak sıklığını tanımlayarak enerji tüketimini dolaylı olarak şekillendirmektedir (Chen vd., 2022; Yufang vd., 2020).
- **Trafik Durumu:** Trafik sıklığının seviyesi, sürüş desenini güçlü bir şekilde belirlemektedir. Yoğun trafik, sürekli hızlanma ve yavaşlamalarla karakterize edilen sık dur-kalk'lı bir sürüşe neden olmaktadır. Rejeneratif frenleme bu senaryoda kaybedilen enerjinin bir kısmını geri kazansa da tekrarlanan ivmelenme döngüleri ve aracın rölantide uzun süre çalışması verimliliği düşürmektedir (Chen vd., 2024; Jiang vd., 2023). Ortalama araç hızı büyük ölçüde trafik akışına bağlı olduğundan, gerçek zamanlı trafik verilerinin enerji modellerine entegre edilmesi konusu tahmin doğruluğunu artırmak için üzerinde çalışılan güncel araştırma konularındandır (Li vd., 2024).
- **Yolcu Yüğü (özellikle otobüsler için):** Daha önce de belirtildiği gibi, yolcu yüğü doğrudan aracın toplam kütesini değiştirmektedir. Özellikle EB'lerde bu faktör oldukça dinamik bir yapıya sahip olup, çekiş için gereken enerji miktarını anlık olarak etkilemektedir (Liu ve Liang, 2022; Würtz vd., 2024; Zhang vd., 2024).

- **Durak Sayısı ve Süresi:** Özellikle kentsel otobüs güzergâhlarının belirleyici bir özelliği olan sık duraklar, duruştan tekrar harekete geçmek için enerji-yoğun hızlanmalar gerektirmektedir (Abdelaty ve Mohamed, 2021). Ayrıca duraklama süresi boyunca kapıların açık kalması, HVAC gibi yardımcı sistemlerin enerji sarfiyatını da etkileyebilmektedir.
- **Seyahat Mesafesi ve Süresi:** Toplam enerji tüketimi doğal olarak seyahat mesafesiyle artsa da, birim mesafe başına tüketilen enerji (kWh/km) daha önce sıralanan tüm faktörlerin karmaşık etkileşimine bağlı olarak önemli ölçüde değişkenlik göstermektedir (Pamuła ve Pamuła, 2022). Seyahat süresi ise özellikle zamana bağlı çalışan yardımcı sistemlerin tüketimini etkileyerek toplam enerji kullanımında rol oynamaktadır.

Sıralanan tüm bu faktörlerin hiçbiri bir birinden bağımsız değildir; aksine, birbirini etkileyen katmanlı ve karmaşık bir sistemin parçalarını oluştururlar. Örneğin çevresel bir faktör olan ortam sıcaklığı, bir yandan HVAC sisteminin (yardımcı sistem) enerji talebini tetiklerken, diğer yandan bataryanın (araç özelliği) anlık verimliliğini değiştirmektedir. Benzer şekilde, operasyonel bir koşul olan trafik yoğunluğu, aracın hız profilini ve hızlanma-yavaşlama desenlerini (sürüş davranışı) doğrudan belirler. Bu sürüş davranışları ise yol türü ve eğim gibi rota özellikleriyle sınırlanmaktadır. Bu noktada sürücü, tüm bu çevresel, operasyonel ve topoğrafik koşullara tepki olarak gaz ve fren komutlarını ayarlayarak kritik bir aracı rolü üstlenmektedir.

Örneğin, soğuk bir kış gününde yoğun trafikte çalışan bir EB'yi ele alalım. Düşük ortam sıcaklığı hem HVAC enerji ihtiyacını artıracak hem de batarya verimliliğini düşürecektir. Yoğun trafik, sık dur-kalk manevralarını zorunlu kılacaktır. Sürücü, muhtemelen kaygan yol koşulları nedeniyle daha temkinli ve yumuşak hızlanmalar yapacaktır. Aynı zamanda yolcu yükünün de fazla olması, toplam kütleyi artıracaktır. Bu senaryoda çevresel (sıcaklık), operasyonel (trafik, yolcu yükü), araca özgü (HVAC, batarya) ve sürücüye bağlı faktörlerin tümü, toplam enerji tüketimini belirlemek üzere dinamik olarak etkileşime girmektedir.

Bu nedenle, faktörleri birbirinden bağımsız ele alan basit bir toplamsal modelin, gerçek dünyadaki enerji kullanımını doğru bir şekilde yansıtması mümkün değildir. Bu iç içe geçmiş yapı, bu karmaşık bağımlılıkları veriden öğrenebilen doğrusal olmayan ML algoritmaları gibi daha gelişmiş yaklaşımların gücünü ortaya koymaktadır. Fizik tabanlı modeller ise tüm bu etkileşimleri açık matematiksel denklemlerle ifade etmekte zorlanmaktadır. Bu durum, veri odaklı yöntemlerin önemini ve bu karmaşık bağımlılıkları

etkili bir şekilde temsil edebilecek yeni özellikler (özellik mühendisliği) ile uygun model mimarileri geliştirme ihtiyacını vurgulamaktadır. EB'ler söz konusu olduğunda, özellikle HVAC talebi, yolcu yükü ve durak sıklığı arasındaki etkileşim, güvenilir bir enerji tahmin modeli oluşturmak için kritik bir kombinasyon olarak öne çıkmaktadır. Çizelge 2.2, literatürde EV enerji tüketimini etkileyen temel faktörleri sistematik olarak sınıflandırmaktadır.

Çizelge 2.2. EV enerji tüketimini etkileyen temel faktörler

Kategori	Faktör	Etki Özeti	Örnek Çalışmalar (Yazar, Yıl)
Araç Özellikleri	Araç Kütlesi	Yük arttıkça ivme, yuvarlanma ve yokuş dirençleri için daha fazla enerji gerekir.	• (Würtz vd., 2024) • (Liu ve Liang, 2022) • (Abdelaty ve Mohamed, 2021)
	Aerodinamik Tasarım	Yüksek sürtünme, özellikle yüksek hızlarda enerji tüketimini artırır.	• (Li vd., 2021) • (Abdelaty ve Mohamed, 2021)
	Batarya Özellikleri	SOC, sıcaklık ve yaşlanma verimi etkiler; aşırı sıcak/soğuk performansı düşürür.	• (Qin vd., 2021) • (Zhang vd., 2020)
	Güç Aktarma Organı Verimliliği	Motor ve invertör verimi, yük ve hız gibi koşullara bağlı olarak değişir.	• (Shen vd., 2022) • (Jiang vd., 2023)
	Yardımcı Sistemler	Klima ve diğer donanımlar enerji tüketimini önemli ölçüde artırır.	• (Würtz vd., 2024) • (Abdelaty vd., 2021)
Sürüş Dinamikleri	Hız Profili	Yüksek hız sürtünmeyi, çok düşük hız yardımcı sistem kullanımını artırır.	• (Zhang vd., 2020) • (Jiang vd., 2023)
	Hızlanma ve Yavaşlama Desenleri	Ani hızlanma yüksek enerji tüketir; sert frenleme ise enerji geri kazanımını azaltır.	• (Chen vd., 2024) • (Maity ve Sarkar, 2023)
	Rejeneratif Frenleme	Yavaşlamada enerji geri kazanımı sağlar; verimliliği koşullara göre değişir.	• (Miri vd., 2021) • (Zhang vd., 2020)
Çevresel Faktörler	Ortam Sıcaklığı	Batarya ve klima performansını doğrudan etkiler; aşırı değerler tüketimi artırır.	• (Zhang vd., 2020) • (Würtz vd., 2024)
	Hava Koşulları	Yağış ve buz yuvarlanma direncini ve sürüş tarzını etkiler.	• (Lin vd., 2023) • (De Cauwer vd., 2017)

Rota ve İşletim Koşulları	Yol Eğimi	Yokuş yukarı sürüş yüksek enerji gerektirir; inişte ise rejeneratif frenleme enerjisi geri kazanır.	• (De Cauwer vd., 2017) • (Wang vd., 2024)
	Yüzey ve Yol Türü	Asfalt, çakıl vb. direnci; otoyol veya şehir içi hız ve duraklama düzenini belirler.	• (Chen vd., 2022) • (Yufang vd., 2020)
	Trafik Durumu	Yoğun trafik dur-kalk sürüşü ve düşük ortalama hızları zorunlu kılar.	• (Chen vd., 2024) • (Jiang vd., 2023)
	Yolcu Yüğü	Özellikle otobüslerde kütleli değiştirerek tüketimi doğrudan etkiler.	• (Würtz vd., 2024) • (Liu ve Liang, 2022)
	Durak Sayısı ve Süresi	Sık duraklar sık ivmelenme gerektirir; bekleme süresi yardımcı sistemleri etkiler.	• (Abdelaty ve Mohamed, 2021)
	Mesafe ve Süre	Toplam enerji tüketimi mesafeye göre artar; km başına tüketim koşullara göre değişir.	• (Pamula ve Pamula, 2022)

2.3. EB Enerji Tüketimi Modellemesinde Güncel Durum ve Araştırma Eğilimleri

EB'lerin kendilerine has operasyonel döngüleri ve fiziksel nitelikleri, genel amaçlı EV modellerinin ötesinde, bu araçlara özel enerji tüketim modellerinin geliştirilmesini zorunlu kılmaktadır. Bu alandaki bilimsel çalışmalar, özellikle filo yönetimi, şarj altyapısı planlaması ve işletme optimizasyonu gibi pratik uygulamalarda tahmin doğruluğunu artırmak amacıyla, otobüslere özgü bu dinamiklere odaklanmaktadır.

2.3.1. Otobüslere özgü belirlenen faktörler

EB'lerin enerji tüketimi üzerine yapılan araştırmalar, binek araçlara kıyasla daha belirgin bir etkiye sahip olan veya tamamen otobüslere özgü olan faktörler üzerine yoğunlaşmaktadır:

- **Yolcu Yüğü:** Bir otobüsün hizmet verdiği süreçte yolcu sayısında yaşanması muhtemel değişimler, aracın toplam kütlede ciddi dalgalanmalara neden olmaktadır. Kütledeki bu değişim, ivmelenme için gereken enerjiyi, yuvarlanma direncini ve eğim kuvvetlerini de doğrudan etkilemektedir (Abdelaty vd., 2021; Liu ve Liang, 2022; Würtz vd., 2024). Bu nedenle, doğru enerji tahmini için gerçek zamanlı veya güvenilir şekilde tahmin edilmiş yolcu yükü verilerinin modellere entegre edilmesi menzile belirleme algoritmalarının geliştirilmesinde kritik öneme sahiptirler.
- **HVAC Sistemleri:** Geniş iç hacimleri ve yolcu konforunu her türlü hava koşulunda sağlama zorunluluğu nedeniyle, otobüslerdeki HVAC sistemleri önemli bir

enerji tüketim kalemi olarak karşımıza çıkmaktadır. HVAC sisteminin enerji sarfıyatı; ortam sıcaklığı, güneş ışıınımı, nem seviyesi gibi çevresel faktörlerin yanı sıra, duraklarda kapıların açılma sıklığı ve süresi gibi operasyonel etkenlerden de kuvvetli bir şekilde etkilenmektedir (Abdelaty ve Mohamed, 2021; Widmer vd., 2023; Würtz vd., 2024).

- **Rota Özellikleri ve İşletim Düzenleri:** EB'lerin genellikle önceden belirlenmiş sabit güzergâhlarda çalışması, yol eğimi ve durak konumları gibi detaylı topografik verilerin modellemeye dahil edilmesine olanak tanımaktadır (Li vd., 2021; Pamula ve Pamula, 2022). Bu araçların operasyonel düzenleri, binek araçlardan farklı olarak sık dur-kalk manevraları, değişken duraklama süreleri ve belirli bir tarife uyma zorunluluğu gibi unsurlarla karakterize edilir ve bu durum, kendilerine özgü bir sürüş döngüsü yaratmaktadır (Abdelaty ve Mohamed, 2021; Jiang vd., 2023).

- **Hız Profili Oluşturma:** Özellikle sefer bazında yapılan enerji tahminlerindeki doğruluk bir sonraki hız profilinin doğru bir şekilde öngörülmesi için hayati önem taşımaktadır. Gerçek zamanlı ve yüksek çözünürlüklü geleceğe yönelik hız verileri genellikle mevcut olmadığından, bazı çalışmalarda bu veri eksikliğinin etkisini azaltmak için bilimsel çalışmalarda istatistiksel yöntemlere başvurmaktadır. Bu yöntemler arasında, benzer rota ve zaman dilimlerine ait geçmiş verilerden yararlanmak ve hız geçişlerini modellemek için Markov zincirleri kullanılması veya rota ve trafik bilgilerine dayalı sentetik profiller oluşturması yöntemlerine başvurulmaktadır (Jiang vd., 2023; Li vd., 2021; Wang vd., 2024).

2.3.2. EV araştırmalarında kullanılan veri setleri

EB'lere yönelik enerji tüketim modelleri, farklı niteliklere sahip çeşitli veri kaynaklarından beslenmektedir. Son yıllarda araştırmacılar, hizmette olan EV filolarından toplanan gerçek zamanlı operasyonel verileri kullanmaya giderek daha fazla yönelmektedir. Bu tür veri setlerine örnek olarak Zürih'teki ZTBus (Widmer vd., 2023), ABD'nin Chattanooga kentindeki CARTA filosu (Ayman vd., 2020; Chen vd., 2021), Pekin (Xu vd., 2023b; Zhang vd., 2024) ve Almanya'nın Göttingen (Würtz vd., 2024) kentlerindeki operasyonel kayıtlar gösterilebilir. Bu veri setleri genellikle hız, güç, SOC ve GPS gibi temel parametrelerin yanı sıra yolcu sayısı ve yardımcı sistem kullanımı gibi otobüslere özgü verilerin de yüksek frekanslı (örneğin 1 Hz) ölçümlerini içermektedir.

Gerçek dünya verilerinin yanı sıra, Simulink veya ADVISOR gibi fizik tabanlı simülasyon araçlarıyla üretilen veriler de, özellikle kontrollü deneyler yapmak veya yeterli operasyonel verinin bulunmadığı durumları modellemek için kullanılmaktadır

(Abdelaty vd., 2021; Abdelaty ve Mohamed, 2022). Kullanılan verinin çözünürlüğü (anlık saniye bazlı değerler mi yoksa sefer veya segment bazlı ortalamalar mı olduğu) geliştirilecek modelin türünü ve karmaşıklığını doğrudan etkilemektedir (Jiang vd., 2023; Xu vd., 2023). Toplu taşıma işletmecilerinin genellikle daha kolay erişebildiği düşük frekanslı verilerle çalışmak ise kendine özgü modelleme zorluklarını beraberinde getirmektedir (Abdelaty ve Mohamed, 2022; Jiang vd., 2023).

2.3.3. EB’lerde menzil belirlemede yapılan çalışmalardan elde edilen bulgular ve doğruluk oranları

EB enerji tüketimi modellerinin tahmin performansı, kullanılan modelleme yaklaşımına, veri kalitesine ve dahil edilen faktörlere bağlı olarak büyük ölçüde değişkenlik göstermektedir. Sıkça kullanılan bir metrik olan Ortalama Mutlak Yüzde Hata (MAPE) değeri, sefer veya segment düzeyindeki tahminler için genellikle %3 ile %10 arasında rapor edilmektedir. Örneğin, Chen ve arkadaşları (2021) LSTM modeliyle %3 ve ANN modeliyle %5 MAPE raporlarken, Zhang ve arkadaşları (2024) ise LightGBM modeliyle sefer düzeyinde %4,404’lük bir MAPE değeri elde etmiştir. Benzer şekilde, Jiang ve arkadaşları (2023) da farklı hibrit modellerle sefer bazında %2-%3 aralığında hata oranları bildirmiştir. Buna karşılık, daha düşük çözünürlüklü verilerle veya daha basit modellerle yapılan çalışmalarda daha yüksek hata oranları görülmektedir; nitekim Abdelaty ve Mohamed (2022) düşük çözünürlüklü açık kaynak verilerle %10–18 aralığında MAPE bulmuştur. Literatürde genel olarak yol eğimi, HVAC kullanımı, batarya SOC değeri, yolcu yükü ve sürüş davranışları gibi faktörlerin, tahmin doğruluğunu belirleyen en kritik etkenler olduğu tutarlı bir şekilde vurgulanmaktadır (Abdelaty vd., 2021; Würtz vd., 2024).

Literatürde dikkat çeken önemli bir husus, kullanılan verinin zenginliği ile geliştirilen modelin genellenebilirliği arasındaki karmaşık ilişki olmasıdır. Yüksek doğruluk oranları rapor eden EB enerji modelleri, genellikle ZTBus (Widmer vd., 2023) veya CARTA filosu (Ayman vd., 2020; Chen vd., 2021) gibi ayrıntılı, yüksek frekanslı ve araca özgü veri setlerine dayanmaktadır. Bu zengin veri setleri, anlık güç çekişi, hassas araç dinamikleri ve sürekli yolcu sayımı gibi detaylı parametreler içermektedir. Bu durum, veri içindeki ince etkileşimleri yakalayabilen karmaşık modellerin geliştirilmesine olanak tanıyarak, ilgili filo veya veri seti özelinde üstün bir tahmin performansı elde edilmesini sağlamaktadır.

Ancak, bu tür zengin veri setleri kullanılarak geliştirilen modellerin farklı işletim koşullarına kullanıldığı diğer filolara veya daha sınırlı veriye sahip olanlara göre genellenebilirliği, literatürde halen önemli bir zorluk olarak karşımıza çıkmaktadır (Abdelaty ve Mohamed, 2022; Jiang vd., 2023). Birçok toplu taşıma işletmecisi, yüksek frekanslı sensör verileri yerine yalnızca sefer bazlı özetlere veya düşük frekanslı kayıtlara erişebilmektedir. Sonuç olarak, Zürih'in kapsamlı ZTBus verisiyle eğitilmiş bir modelin, farklı bir şehirde veya farklı bir otobüs modelinde aynı performansı sergilemesi beklenmemektedir. Bu durum, veri kalitesindeki farklılıklara karşı daha dayanıklı veya farklı alanlara uyarlanabilen (transfer öğrenme gibi) esnek modeller geliştirme ihtiyacını ortaya koymaktadır. ZTBus bu noktada EB'lerde menzil belirlemede daha az girdiyle bile daha yüksek performans göstererek genellenebilir özellikler imkanı sunarak daha güvenilir modelleme yaklaşımlarının geliştirilmesi ve test edilmesi için önemli bir referans noktası (benchmark) oluşturmaktadır. Çizelge 2.3, Literatürde EB'lerin enerji tüketimi modellemesine yönelik yapılmış olan önemli çalışmaları özetlenmiştir.

Çizelge 2.3. EV enerji tüketimi modellemesine dair önemli çalışmaların özeti

Çalışma (Yazar, Yıl)	Temel Model Yaklaşımı	Otobüse Özgü Temel Girdiler	Veri Seti Özellikleri	Temel Bulgular / Doğruluk	Yazarların Belirttiği Sınırlamalar
(Jiang vd., 2023)	Hibrit (Markov Tabanlı Hız Profili + GPR)	- Yardımcı Güç (Sıcaklık Etkisi) - Araç Parametreleri - Hız Profili - Dış Etkenler	Çin Ulusal Enerji Araçları Yarışması'ndan Gerçek Dünya, Düşük Çözünürlük Verisi	- Ecmotor Hata <3 % - Toplam Tüketim Hatası <2 % (MAPE)	- Çoğu Modelde Tüm Faktörler Yer Almıyor - Güç Aktarma Verimi Etkisi İncelenmemiş
(Würtz vd., 2024)	Veri Odaklı (Sabit, Aylık, Sıcaklık, NN)	- Rakım - Sıcaklık - Yolcu Yüku - Hız Profili - Duraklar	Göttingen, Almanya – 16 Adet E-Citaro, 13 Ay, Yüksek Çözünürlük	Kombine MAPE = 7,9 %	- Farklı Koşullarda Sistemik Karşılaştırma Eksik - Yeni Veri Kaynakları Entegrasyonu Gerek
(Chen vd., 2021)	Veri Odaklı (LSTM, ANN)	- Hız - İvmeleme - Yol Eğimi - HVAC Durumu (Tahmini) - Yolcu Sayısı (Tahmini)	Carta Filosu, Chattanooga, Tn, ABD – 1 Hz Gerçek Zamanlı Veri	- LSTM MAPE = 3 % - ANN MAPE = 5 %	Güzergâha Özgü Bozulmalar Eklenerek İyileştirme İhtiyacı

(Xu vd., 2023)	Veri Odaklı (Güç Tabanlı, LSTM, XGBoost)	- Hız - İvme - Rakım - Rejenerasyon Sınıflayıcısı	Nanjing, Çin – Elektrikli ve Dizel Otobüsler, Anlık Veri	E-Otobüs Modelleri $R^2 > 0,80$; Güç Tabanlı Model Öneriliyor	- Rejenerasyonda LSTM Zorluk Yaşamış - Dizeller İçin Daha Fazla Motor Verisi Gerekli
(Abdelaty vd., 2021)	Veri Odaklı (MLR, RBF, DT, GBDT, SVM, MLP-NN, RB-NN)	- Soc - Kütle - Cr - Cd - İvme/D - Hız - Duraklar - Uzunluk - Eğim - Sıcaklık - HVAC - Yardımcı Yükler	Simülasyon (Simulink + Altoona Doğrulaması), Tam Faktöriyel Tasarım (907199 Senaryo)	- Tüm Modeller $R^2 > 0,90$ - MLP-NN $R^2 > 0,99$ - Eğim ve SOC en Etkili	- Sadece Simülasyon Verisi - Gerçek Dünya Uygulaması Test Edilmeli
(Pamula ve Pamula, 2022)	Veri Odaklı (Otomatik Kodlayıcı Tabanlı Derin Öğrenme Ağı (Deep Learning Network With Autoencoders-DLNA) + MLP	- Duraklar Arası Seyahat Süresi - Rakım Farkı - Yoğun Saat - Hava Mesafe	Pkm Jaworzno, Polonya – GPS ve Varış Zamanları ile Gerçek Dünya Verisi	- DLNA MAPE = 6,7 % - RMSE = 0,052 KWH	Zaman Serisi Verisi Sıkı Olmayınca LSTM Beklenen Performansı Gösterememiş
(Zhang vd., 2024)	Veri Odaklı (RF, XGBoost, LightGBM, Catboost, SVR, ANN)	- Trafik - Ortam (Sıcaklık, Nem, Rüzgâr) - Araç (SOC, Batarya Sıcaklığı, HVAC Durumu, Kapı Durumu, Yolcu Sayısı) - Sürüş (Hız, İvme, Duraklar)	Pekin, Çin – 3 Hat, Düşük Frekans (10 S Aralık) Sefer Düzeyi Verisi	XGBoost MAPE = 4,404 % (10-Kat CV)	- Birçok Çalışma Az Sayıda Faktöre Odaklanıyor - Kapsamlı Analiz Eksik
(Xu vd., 2023)	Veri Odaklı (Elastic Net, GPR, LightGBM)	Yol-Sürücü-Araç-Çevre Faktörleri (Sıcaklık, Sürüş/Şarj Süreleri, Kat Edilen Yol, SOC, Hız, İvme, HVAC Durumu, Yolcu)	Pekin, Çin – 10 Otobüs, 1 Yıl, Düşük Frekanslı Veri	LightGBM MAPE = 7,7 %	- Düşük Frekans Verisi Nedeniyle Bilgi Süreksiz - Uygun Nicel Yöntem Gerek

(Liu ve Liang, 2022)	Veri Odaklı (Fizik-Tabanlı Kalman Filtresi + RF)	- Hız - İvme (Tahmini) - Yolcu Sayısı (Stokastik) - Rota Bilgisi	St. Albert Transit, Ab, Kanada – Android Tabanlı GPS Verisi	Önerilen Yöntem Gerçek Zamanlı Tahminde Daha Doğru; RF ile Özellik Önemi Belirlendi	- GPS Gürültüsü - Yolcu Sayısı Stokastik - İvme Tahmini Zorluğu
(Abdelaty ve Mohamed, 2022)	Veri Odaklı (Açık Kaynak Düşük Çözünürlüklü Veri Çerçevesi)	- Araçsal Operasyonel - Topolojik - Dış Parametreler	Açık Kaynak Düşük Çözünürlük Verisi	- Tahmin Hatası $\approx 20\%$ (MAPE) $R^2 > 0,90$	- Açık Kaynak Düşük Frekans Verisi Kullanan Modeller Az; - Yaygın Doğrulama Gerek
(Li vd., 2021)	Veri Odaklı (Stokastik RF, KNN RF + Stokastik Hız Profili)	- Kinematik Faktörler (Üretilen Profilden) - Dış Faktörler (Kalkış Zamanı, Gün, Sıcaklık, Islak-Kuru)	Shenzhen, Çin – Şehir Ölçeğinde EB Ağı, Düşük-Çözünürlüklü Büyük Veri	- Stokastik RF: MAE, MAPE, RMSE Gelişimi - KNN RF Hesaplama Süresini Kısaltır	Mevcut Modeller Kapsamlı Faktör Eksikliği veya Veriden Hız Profili Üretme Zorluğu
(Widmer vd., 2023)	Veri Seti Tanımlayıcı (ZTBus)	- Güç Talebi - Tahrik Odometrisi - GPS Sıcaklık - Kapı Açılmaları, - Yolcu Sayısı - Hareket Düzeni	Zürich, İsviçre – 2 Trolleybüs, Nis 2019–Ara 2022, 1 Hz Senkronize Veri	Veri Seti ile ilgili makale; Model Değil	Kamusal Erişimli Zaman Serisi Veri Eksikliği

2.4. EV’lerde Menzil Belirlemede kullanılan Algoritmalarındaki Araştırma

Boşlukları

EV enerji tüketimi üzerine yapılan çalışmalar, özellikle gelişmiş ML algoritmalarını kullanan veri odaklı yaklaşımların öncülük ettiği, dinamik ve hızla ilerleyen zengin bir literatür alanı oluşturmaktadır. Bu yöntemler, yüksek tahmin doğruluğu sunma potansiyeli göstermekle birlikte, veri erişimine ulaşım ve gereken sayıda verinin elde edilmesi vb. bazı zorluklar içermektedir. Ayrıca, birçok karmaşık modelin “kara kutu” yapısı yorumlanabilirliği zorlaştırmakta ve geliştirilen modellerin farklı araçlar ile işletme koşullarında genellenebilirliğinin sağlanması halen önemli bir ihtiyaç olarak görülmektedir.

EB'lere ilişkin arařtırmalar da aktif bir řekilde devam etse de, bu alanda daha kapsamlı ve saęlam modelleme yaklařımlarına ihtiya duyulduęu ařıkardır. EB alıřmaları, binek aralardan bariz řekilde farklı kendine zg bir takım karakteristik dinamik zellikleri bnyesinde bu tez alıřmasının deęiřik konu bařlıkları altında verilmiř olan zellikleri bnyesinde barındırmaktadır. rneęin, yolcu yk ara ktlesini srekli olarak deęiřmesi, byk bir alanın sık sık aılıp kapanan kapıların sistemi bozucu etkisi altında ortamın arzu edilen sıcaklıkta tutulmasında kullanılan HVAC sistemlerinin enerji talebini artırması ve rota zerinde talep edilen noktadan noktaya ulařım talebinin karřılanması iin oluřması muhtemel yoęun dur-kalk dngleri ister istemez araların enerji tketime olan taleplerini artırmaktadır. Bu faktrlerin EB'lerde menzil hesaplamada kullanılan algoritmalara olan etkisi literatrde genel olarak bilinmesine raęmen halen daha bu etkilerin algoritmalara nasıl ve ne řekilde yansıtılması gerektięi konusu zerine yapılacak alıřmalara ihtiya vardır. Halen menzil belirlemede kullanılan algoritmalara bu durum yansıtılamamıř ve kullanılan tahmin modellerine gereken doęrulukta entegre edilmedięi anlařılmaktadır. Bu baęlamda literatrde ne ıkan temel arařtırma bořlukları ařaęıdaki řekilde sıralanabilir.

- Etkileřimli Faktrlerin Btnsel Entegrasyonu,
- Veri Deęiřkenlięine Karřı Model Dayanıklılıęı,
- Dinamik Hız Profili Oluřturma,
- Model Yorumlanabilirlięi ve Doęruluk Arasındaki Denge,
- eřitli Yardımcı Sistemlerin Etki Deęerlendirmesi,

Bu kısımda belirtilen arařtırma bořlukları ařaęıda sırasıyla aıklanmıřtır.

Etkileřimli Faktrlerin Btnsel Entegrasyonu: Birok EB enerji tketim modeli eřitli faktrleri ierse de, bu faktrlerin birbiriyle olan etkileřimini btnsel olarak yakalayan yaklařımlara halen ihtiya duyulmaktadır. EB'lerdeki enerji kullanımı; ara parametreleri (yolcu ykyle deęiřen ktle), iřletim kořulları (yolcu iniř-biniřinin HVAC ykn etkilemesi), srř davranıřları ve evresel kořulların (sıcaklıęın hem batarya hem HVAC performansını etkilemesi) basit bir řekilde toplamı řeklinde olmayıp, aralarında karmařık bir etkileřim sonucu tahmine dayalı algoritmalar geliřtirilerek elde edilebilmektedir. Bu etkileřimli unsurları kendi aralarında btnsel olarak modellemek, tahmin doęruluęunu artırmak iin ok derece nemli bir deęere sahiptir (Jiang vd., 2023; Wrtz vd., 2024).

Veri Deęiřkenlięine Karřı Model Dayanıklılıęı: Yksek doęruluklu EB enerji modelleri, genellikle ZTBus (Widmer vd., 2023) gibi zengin ve yksek znrlkl veri

setlerine dayanmaktadır. Ancak birçok toplu taşıma işletmecisinin sahip olduğu veriler kalite, çözünürlük ve eksiksizlik açısından büyük farklılıklar göstermektedir. Bu veri çeşitliliğiyle başa çıkabilen veya yeni veri ortamlarına kolayca uyum sağlayabilen esnek modellere ihtiyaç duyulmaktadır.

Dinamik Hız Profili Oluşturma: Bir sefer başlamadan önce, değişen trafik ve işletim koşulları altında gelecekteki hız profilini doğru bir şekilde tahmin etmek, öngörüselleştirme için halen önemli bir zorluktur. Markov modelleri (Jiang vd., 2023; Li vd., 2021) ve ML tabanlı hız tahmincileri gibi mevcut yaklaşımların, otobüs işletmelerinin gerçek dinamiklerini yansıtmaları için yolcu yükü ve HVAC gibi EB'ye özgü diğer faktörlerin menzile belirlemede kullanılacak algoritmalara daha iyi entegre edilmesine ihtiyaç vardır.

Model Yorumlanabilirliği ve Doğruluk Arasındaki Denge: Derin öğrenme gibi veri odaklı modeller yüksek doğruluk sunsa da, "kara kutu" yapıları nedeniyle toplu taşıma işletmecileri için pratik anlamda kullanımları zorlaşmaktadır. İşletmeciler, stratejilerini yönlendirmek için hangi faktörlerin enerji kullanımını ne ölçüde etkilediğini anlamak istemektedir (Wang vd., 2021). Bu noktada, doğruluk ve yorumlanabilirlik arasında denge kuran modellerin veya karmaşık modellerin tahminlerini açıklayan yöntemlerin (SHAP değerleri olarak Pokharel ve ark. (2021); Acharyaviriy ve ark. (2023); tarafından bahsedilen) geliştirilmesi için halen daha fazla araştırmaya ihtiyaç vardır.

Çeşitli Yardımcı Sistemlerin Etki Değerlendirmesi: HVAC sistemlerinin yüksek enerji tüketimi yoğun bir şekilde incelenmiş olsa da, kapı mekanizmaları, iç-dış aydınlatma ve yolcu bilgilendirme sistemleri gibi diğer yardımcı sistemlerin toplu etkileri üzerine daha halen yapılması gereken bilimsel çalışmalara ihtiyaç duyulmaktadır. ZTBus (Widmer vd., 2023) gibi zengin veri setleri, araştırmacılara bu sistemlerin farklı operasyonel aşamalarındaki (sürüş, durakta bekleme vb.) enerji tüketimini analiz etmek için çok önemli fırsatlar sunmaktadır.

Yapılan bu tez çalışmasında literatürde ihtiyaç duyulan bu araştırma boşlukları göz önüne alınarak, bu tez çalışmasının ana omurgasını oluşturulmuş ve çalışmanın amacını belirlemede yol gösterecek ve çalışmayı motive edecek unsurların belirlenmesi hedef alınmıştır. Bu tez çalışmasında literatürde eksikliği hissedilen burada bahsedilen araştırma boşluklarına pozitif katkı sağlayabilmek hedeflenmiştir. Yapılan çalışmada etkileşimli faktörleri yakalayabilen kapsamlı bir özellik entegrasyonuna odaklanılmış ve uygulamada karşılaşılan işletme karmaşıklıklarını dikkate alabilecek bir ileri modelleme

tekniklerinin geliştirilmesi araştırılmıştır. Yapılan çalışma sonucunda burada belirtilen amaçlara uygun düşen ZTBus gibi zengin veri setlerinden yararlanılarak yapılan çalışma sonucunda elde edilecek yeni yaklaşımın doğrulanması ve bu alanda yapılacak çalışmalara yeni ön görüler sunulması amaçlanmıştır.

İncelenen literatür, statik öngörü modellerinin geliştirilmesinin ötesinde, EB enerji yönetiminin geleceğinin adaptif ve operasyonel olarak bütünleşik sistemlere doğru evrildiğini göstermektedir. Mevcut araştırmalar, genellikle geçmiş verilere veya önceden tanımlanmış seyahat koşullarına dayalı olarak doğru tahminler sunan modeller geliştirmeye odaklanmıştır (ör. Jiang ve ark. (2023); Chen ve ark. (2021)). Ancak gerçek otobüs işletmeleri, beklenmedik trafik sıkışıklığı, ani yolcu talebi değişiklikleri veya hızlı hava koşulu değişimleri gibi öngörülemeyen olaylara maruz kalan dinamik sistemlerdir. Bu nedenle, başlangıçta ne kadar doğru olursa olsun, statik bir yol öncesi enerji tahmini, sefer sırasında operasyonel koşullar önemli ölçüde saptığında geçerliliğini yitirebilmektedir.

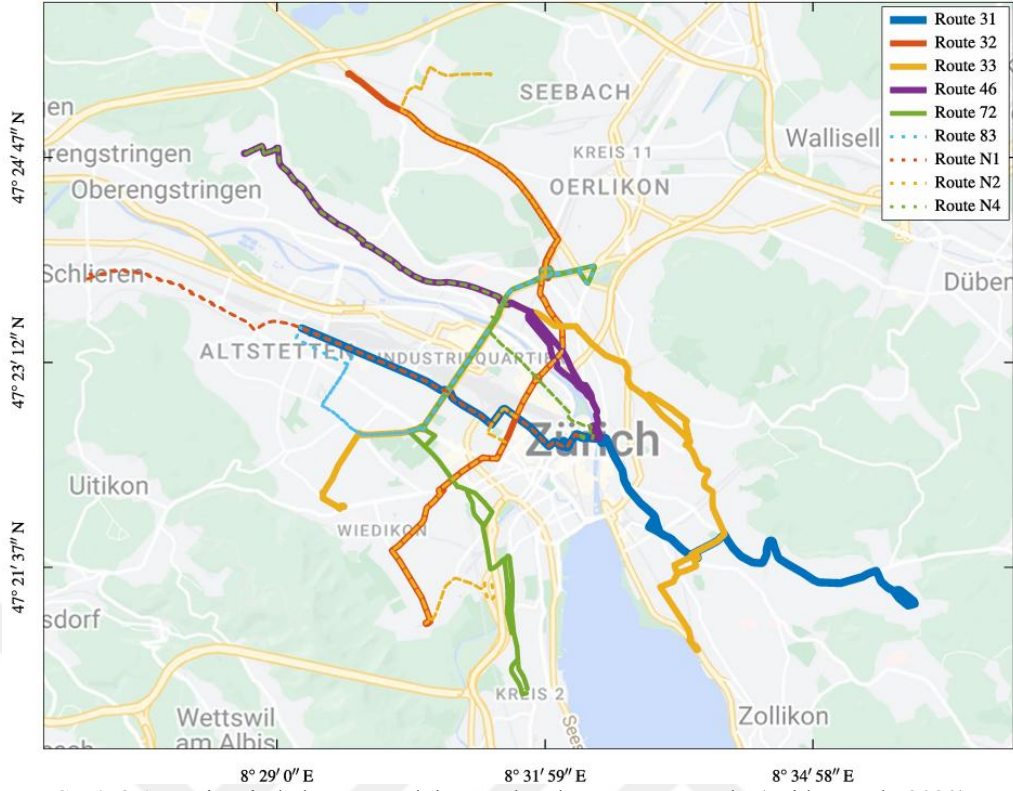
Daha geniş EV literatüründe çevrimiçi model uyarlaması (Zhu vd., 2025) gibi kavramların ortaya çıkması ve gerçek zamanlı tahmin yeteneklerine yönelik çağrılar (Liu ve Liang, 2022; Modi vd., 2020), bu alandaki önemli bir gelişim yönünü işaret etmektedir. Sıkı operasyon programlarına ve kritik enerji kısıtlarına sahip EB'ler için, sefer sırasında enerji tüketimi tahminlerini dinamik olarak güncelleme yeteneği önemli bir ilerleme olacaktır. Bu tür öngörülerin pratik değeri ise, filo yönetimi ve karar destek sistemlerine entegre edildiğinde tam olarak ortaya çıkacaktır. Bu entegrasyon, tahmin edilen enerji mevcudiyetine dayalı dinamik yönlendirmeler veya optimize edilmiş şarj programları gibi proaktif kararlar alınmasına olanak tanıyacaktır. Bu alandaki yaklaşımlara örnek olarak, Jiang ve arkadaşlarının (2023) şarj optimizasyonu üzerine yaptığı çalışmalar ve yine aynı kaynakta referans gösterilen Bie ve arkadaşlarının (2021) güzergâh planlaması üzerine olan çalışması gösterilebilir. Bu tezin kapsamı temel olarak öngörü modelini oluşturmaya odaklanmış olsa da, bu gelecek vizyonunu kabul etmek, yapılan çalışmanın daha akıllı ve verimli EB filo yönetimine doğru atılmış hayati bir adım olarak çerçevelenmesini sağlamaktadır. ZTBus (Widmer vd., 2023) gibi kapsamlı, zaman çözünürlüklü veri setleri, bu tür adaptif ve dinamik öngörü yeteneklerinin geliştirilmesi ve titiz test edilmesi için ideal olarak uygundur.

3. MATERYAL VE YÖNTEM

3.1. Veri Seti ve Özellikleri

Bu çalışma, kentsel elektrikli otobüslerin anlık enerji tüketimini yüksek doğrulukla modellemeyi amaçlamakta ve bu doğrultuda, kamuya açık ve yüksek çözünürlüklü ZTBus veri setini temel almaktadır. Alanında en kapsamlı çalışmalardan biri olarak kabul edilen bu veri seti, Widmer ve arkadaşları (2023) tarafından akademik kullanıma sunulmuştur. Veri seti, her bir seferin (görev) saniyelik verilerini içeren ve toplamda birkaç gigabayt büyüklüğe ulaşan Virgülle Ayrılmış Değerler (CSV) formatındaki dosyalardan oluşmaktadır. Bu veri setinin tercih edilmesindeki temel nedenler; verilerin 1 Hz. frekansında (saniyelik) kaydedilmiş olması, araç dinamiği, operasyonel durum ve çevresel koşulları kapsayan çok sayıda değişkeni senkronize bir şekilde sunması ve kamuya açık erişimli olmasıdır. Bu yüksek frekanslı veriler, sürüş sırasındaki ani hızlanma veya frenleme gibi anlık değişimlerin enerji tüketimi üzerindeki etkilerini mikroskobik düzeyde analiz etme olanağı sağladığı için bu çalışma açısından kritik bir öneme sahiptir.

Veriler, İsviçre'nin Zürih kentinde, Zürih Toplu Taşıma (VBZ) filosunda hizmet veren iki adet HESS lighTram® 19 DC model körüklü elektrikli trolleybüsten toplanmıştır. Nisan 2019 ile Aralık 2022 tarihleri arasında 3,5 yılı aşkın bir süreyi kapsayan bu zengin veri seti, toplamda 1409 tamamlanmış sefer içermektedir. Bu süre zarfında kaydedilen saniyelik verilerin toplamı, yalnızca bu çalışmada kullanılan örneklemde dahi 5,5 milyondan fazla satıra ulaşmakta, bu da veri setinin yoğunluğunu ve zenginliğini ortaya koymaktadır. Veri seti; aracın Kontrolör Alan Ağı (CAN bus) sinyallerinin, yüksek hassasiyetli Küresel Navigasyon Uydu Sistemi (GNSS) ve kızılötesi yolcu sayım sensörlerinden gelen verilerle birleştirilmesiyle oluşturulmuştur. Araştırmanın yapıldığı Zürih şehri, farklı topografik özelliklere (ör. yokuşlar, düz yollar) ve değişken trafik yoğunluklarına sahip güzergahları barındırması nedeniyle, geliştirilecek modelin farklı sürüş koşulları altındaki genelleme kabiliyetini test etmek için ideal bir operasyonel ortam sunmaktadır (Widmer ve ark., 2023). Çalışmada incelenen temel otobüs hatlarının güzergahları Şekil 3.1'de sunulmuştur.



Şekil 3.1. Veri setinde kapsanan dokuz otobüs hattının güzergahı (Widmer vd., 2023)

ZTBus veri seti, yapısal olarak iki ana bileşenden oluşmaktadır: her bir sefere ait genel özellikleri (ör. sefer kimliği, başlangıç zamanı, güzergâh) içeren merkezi bir meta veri dosyası (metaData.csv) ve her seferin ayrıntılı zaman serisi verilerini barındıran çok sayıda sefer dosyası. Bir dizin işlevi gören meta veri dosyası, her bir sefer için özet bilgiler sunar. Her bir sefer dosyası ise “name” sütunu aracılığıyla meta veri dosyasındaki ilgili satıra bağlanarak o sefere ait saniyelik operasyonel verileri (hız, güç talebi vb.) içerir. Bu yapı, araştırma sürecinde öncelikle meta veri üzerinden ilgili seferlerin seçilmesine, ardından bu seferlere ait dosyaların detaylı analiz için çağrılmasına olanak tanımıştır.

Bu çalışmada, enerji tüketimini en iyi şekilde temsil edeceği öngörülen özellikler, literatürdeki benzer çalışmalar ve aracın temel fiziksel dinamikleri dikkate alınarak seçilmiştir. Bu kapsamda 11 temel girdi ve bir hedef özelliği belirlenmiştir. ZTBus veri seti, bu çalışmada kullanılmayan çok sayıda ek değişken (ör. her tekerleğin ayrı hızı veya detaylı GNSS sinyalleri) içermekle birlikte, bu araştırma model karmaşıklığını artırmadan en yüksek ilgililiğe sahip olduğu düşünülen değişkenlere odaklanmıştır. Hedef değişken, aracın bataryadan veya elektrik hattından çektiği anlık toplam gücü Watt cinsinden ifade eden *electric_powerDemand* (anlık toplam güç talebi) olarak tanımlanmıştır.

Seçilen girdi özellikleri, aşağıda listelenen üç ana kategori altında gruplandırılmıştır. Bunlar;

- Araç Dinamiği Değişkenleri,
- Operasyonel Durum Değişkenleri,
- Çevresel ve Zamansal Değişkenler.

- **Araç Dinamiği Değişkenleri:**

- odometry_vehicleSpeed: Aracın anlık hızı (m/s).
- acceleration: Aracın anlık ivmesi (m/s²).
- traction_tractionForce: Motor tarafından tekerleklerle uygulanan anlık çekiş kuvveti (N).
- traction_brakePressure: Fren sistemindeki anlık basınç (Pa).

- **Operasyonel Durum Değişkenleri:**

- itcs_numberOfPassengers: Araç içindeki anlık yolcu sayısı.
- status_doorIsOpen: Kapıların açık (1) veya kapalı (0) olma durumu.
- status_gridIsAvailable: Aracın elektrik hattına bağlı (1) veya değil (0) olma durumu.

- **Çevresel ve Zamansal Değişkenler:**

- gnss_altitude: Deniz seviyesinden yükseklik (m).
- temperature_ambient: Ortamın dış hava sıcaklığı (K).
- hour_of_day: Günün saati (0-23).
- day_of_week: Haftanın günü (Pazartesi=0, Pazar=6).

Bu değişkenlerin istatistiksel bir özeti Çizelge 3.1'de sunulmuştur.

Çizelge 3.1. Modellemede kullanılan değişkenlerin istatistiksel özeti

	Adet	Ortalama	Standart Sapma	Minimum	Maksimum	Varyans
<i>odometry_vehicleSpeed</i>	5513259	4,25	4,20	-0,33	18,16	17,60
<i>acceleration</i>	5513259	0,00	0,46	-3,55	1,72	0,22
<i>road_grade</i>	5513259	0,00	0,05	-0,25	0,25	0,00
<i>itcs_numberOfPassengers</i>	5513259	16,99	14,03	0,00	122,00	196,96
<i>temperature_ambient</i>	5513259	289,30	8,00	273,15	311,15	64,05
<i>traction_tractionForce</i>	5513259	1,96E+03	1,06E+04	-4,72E+04	5,03E+04	1,13E+08
<i>traction_brakePressure</i>	5513259	8,18E+04	1,14E+05	0,00	9,33E+05	1,29E+10
<i>status_doorIsOpen</i>	5513259	0,21	0,41	0,00	1,00	0,17
<i>status_gridIsAvailable</i>	5513259	0,66	0,47	0,00	1,00	0,22
<i>hour_of_day</i>	5513259	11,80	4,98	0,00	23,00	24,77
<i>day_of_week</i>	5513259	2,70	1,86	0,00	6,00	3,47
<i>electric_powerDemand</i>	5513259	2,52E+04	6,80E+04	-3,12E+05	3,65E+05	4,62E+09

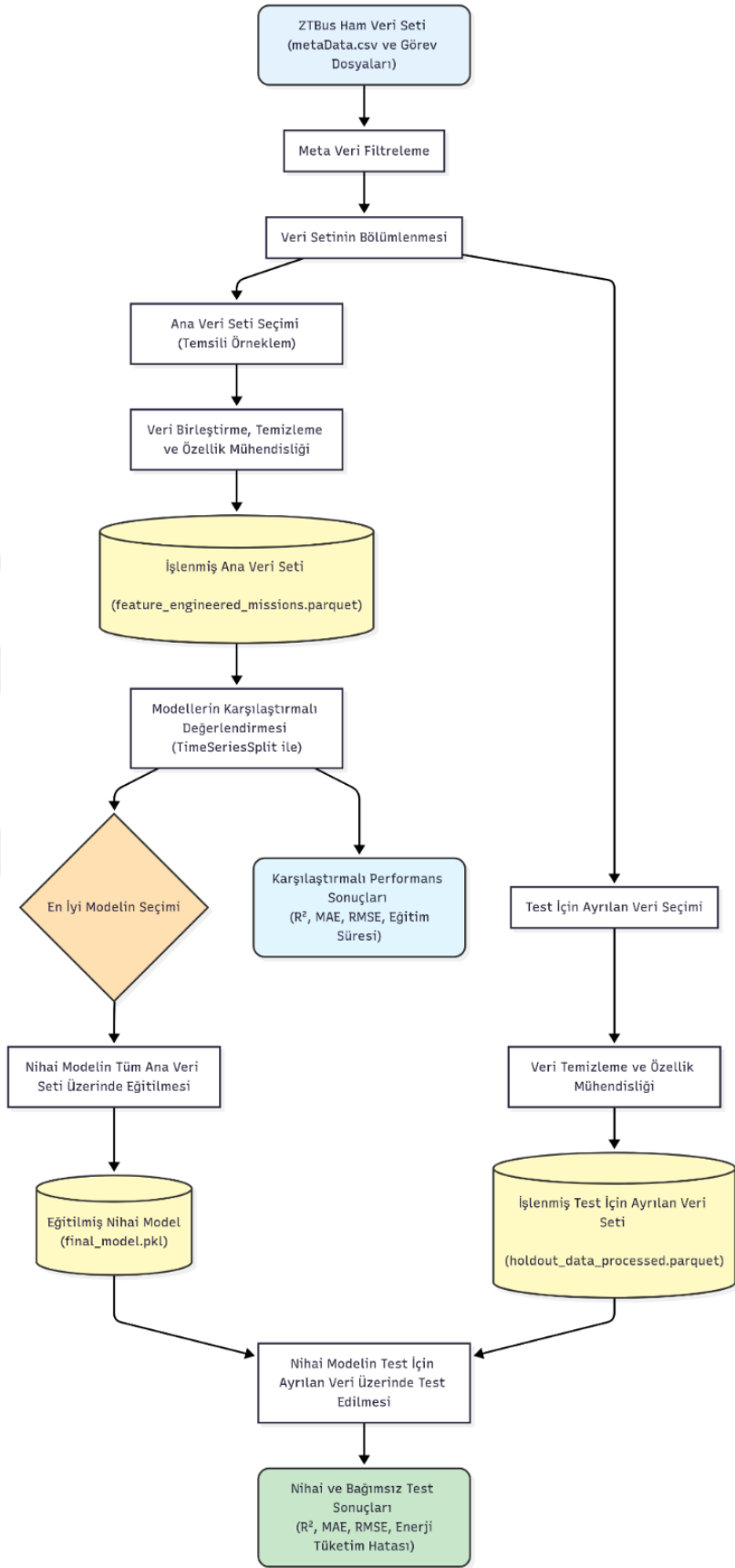
Çizelge 3.1'de sunulan istatistiksel özet, modellemede kullanılan değişkenlerin doğasını ve ölçeğini ortaya koymaktadır. *electric_powerDemand* (anlık güç talebi) değişkenindeki negatif değerler, aracın rejeneratif frenleme esnasında gerçekleştirdiği enerji geri kazanımını temsil etmektedir. *odometry_vehicleSpeed* (araç hızı) değişkeninde gözlemlenen nadir negatif değerler ise aracın kısa süreli geri manevralarından veya düşük hızlardaki sensör gürültüsünden kaynaklanmaktadır. *traction_tractionForce* gibi bazı değişkenlerde görülen yüksek varyans, bu değişkenlerin doğal olarak geniş bir değer aralığına sahip olmasından ileri gelmektedir; bu durum, modelleme öncesinde uygulanacak olan veri ölçeklendirme adımında ele alınacaktır. *status_doorIsOpen* (kapı durumu) gibi kategorik değişkenler ise beklendiği üzere yalnızca 0 ve 1 değerlerini

almaktadır. Değişkenlerdeki bu istatistiksel çeşitlilik, geliştirilecek modelin farklı operasyonel senaryoları öğrenebilmesi için zengin bir temel sağlamaktadır.

3.2. Araştırma Metodolojisi Akış Şeması

Bu tez çalışmasında izlenen metodoloji, veri ön işlemeden nihai modelin değerlendirilmesine kadar uzanan bir dizi sistematik adımdan oluşmaktadır. Araştırma sürecinin bütünü, adımlar arasındaki mantıksal ilişkiyi ve veri akışını şematik olarak özetleyen metodoloji akış şeması, Şekil 3.2'de sunulmuştur. Bu şema, projenin temel yapı taşlarını ve veri akışını görselleştirerek uygulanan yöntemin anlaşılabilirliğini artırmaktadır.





Şekil 3.2. Araştırma metodolojisi akış şeması

Şekil 3.2'de de gösterildiği gibi, araştırma süreci ham meta verinin filtrelenmesiyle başlamaktadır. Ardından, filtrelenmiş görev listesinden temsili bir örneklem alınarak model geliştirme için kullanılacak ana veri seti oluşturulur. Geriye kalan görevler arasından ise nihai değerlendirme için bağımsız bir test (holdout) veri seti ayrılır.

Bu iki ayrı görev listesi tanımlandıktan sonra, her biri için verilerin birleştirilmesi, temizlenmesi ve özellik mühendisliği gibi sıralı ön işleme adımları uygulanarak iki ayrı işlenmiş veri seti elde edilir. Ana veri seti, çeşitli makine öğrenmesi modellerinin karşılaştırmalı olarak değerlendirilmesi amacıyla kullanılır ve bu aşamanın çıktısı olarak, modellerin performansını özetleyen bir sonuç tablosu üretilir.

Karşılaştırmalı analiz sonucunda en iyi performansı gösteren model belirlenir ve bu model, tüm ana veri seti üzerinde yeniden eğitilerek nihai model haline getirilir. Son aşamada ise bu nihai modelin performansı, sürecin önceki aşamalarında hiç kullanılmamış olan bağımsız test (holdout) verileri üzerinde ölçülerek çalışmanın yansız ve kesin sonuçları raporlanır. Her adımın çıktısının bir sonraki adım için girdi teşkil etmesi, metodolojinin tekrarlanabilir ve sistematik bir yapıya sahip olmasını güvence altına almaktadır.

3.3. Veri Ön İşleme Süreci

Ham verilerin doğrudan makine öğrenmesi modellerinde kullanılması genellikle hatalı ve güvenilmez sonuçlara yol açar. Bu nedenle modelleme aşamasından önce, veri setinin tutarlı, temiz ve anlamlı bir formata dönüştürülmesi kritik öneme sahiptir. Bu bölümde, ham ZTBus verilerini, nihai modelin girdisi olarak kullanılacak işlenmiş veri setlerine dönüştüren çok adımlı ön işleme süreci detaylandırılmaktadır.

Bu süreç; veri setinin ana ve test (holdout) gruplarına ayrılması, ana veri setinden temsili bir örneklem seçilmesi, görev verilerinin birleştirilmesi, eksik ve hatalı verilerin ayıklanması ve son olarak modelin tahmin gücünü artırmak amacıyla özellik mühendisliği adımlarını kapsamaktadır. Her bir adım, bir sonraki aşama için sağlam bir temel oluşturacak şekilde titizlikle tasarlanmıştır.

3.3.1. Meta veri filtreleme ve geçerli görev listesinin oluşturulması

Veri hazırlama sürecinin ilk adımı, ham meta veri dosyasının (metaData.csv) filtrelenerek analiz için uygun ve geçerli görevlerin belirlenmesidir. Orijinal meta veri dosyası 1409 görev içermesine rağmen, bu görevlerin tümü çalışmanın odaklandığı standart gündüz operasyonlarını temsil etmemekteydi. Bu nedenle, sonuçların tutarlılığını

sağlamak ve modelin doğru verilerle eğitildiğinden emin olmak amacıyla iki temel kritere dayalı bir filtreleme işlemi uygulanmıştır. Bunlar;

- Gece Görevlerinin Çıkarılması,
 - Geçersiz Görevlerin Çıkarılması.
- **Gece Görevlerinin Çıkarılması:** *busRoute* (güzergâh kimliği) sütununda 'N' ön ekine sahip olan görevler, gece operasyonlarına ait oldukları için veri setinden çıkarılmıştır. Gece görevleri; trafik yoğunluğu, yolcu sayısı ve sürüş davranışları açısından gündüz operasyonlarından önemli farklılıklar gösterdiğinden, bu görevlerin analize dahil edilmesinin model performansını olumsuz etkileyeceği öngörülmüştür.
 - **Geçersiz Görevlerin Çıkarılması:** *busRoute* sütununda '-' gibi geçersiz veya tanımsız bir kimlik barındıran görevler de veri setinden ayıklanmıştır. Bu görevlerin belirli bir güzergâhla ilişkilendirilememesi, onları güzergâh bazlı analizler için elverişsiz kılmaktadır.

Bu iki filtreleme adımının uygulanması sonucunda, 1409 görevlik başlangıç havuzundan, 1325 adet geçerli gündüz görevini içeren bir liste elde edilmiştir. Bu filtrelenmiş görev listesi, hem modelin geliştirileceği ana veri setinin hem de nihai test için kullanılacak bağımsız test (holdout) veri setinin seçileceği temel havuzu teşkil etmektedir.

3.3.2. Ana veri seti için temsili örneklem seçimi

Model geliştirme ve değerlendirme süreçlerinde hesaplama verimliliğini artırmak amacıyla, 1325 görevlik geçerli görev havuzundan daha küçük, ancak istatistiksel olarak anlamlı bir alt küme oluşturulmuştur. Bu çalışmanın ana veri setini teşkil etmek üzere, toplam popülasyonun yaklaşık %12'sine karşılık gelen 162 görev seçilmiştir. Bu örneklem büyüklüğü, veri çeşitliliğini temsil etme kabiliyeti ile model eğitim süreçlerinin makul sürelerde tamamlanması arasındaki denge gözetilerek belirlenmiştir.

Veri setinin farklı güzergâh ve mevsim koşullarını içeren heterojen yapısı nedeniyle, basit rastgele örnekleme yönteminin popülasyon dağılımını doğru yansıtmama riski mevcuttu. Bu riski ortadan kaldırmak ve örneklemin, popülasyonun temel özelliklerini yüksek doğrulukla temsil etmesini sağlamak amacıyla Tabakalı Oransal Örnekleme (Stratified Proportional Sampling) yöntemi kullanılmıştır. Bu yaklaşım, popülasyonu anlamlı alt gruplara (tabakalara) ayırmayı ve ardından her tabakadan, o tabakanın popülasyon içindeki oranıyla orantılı sayıda örnek seçmeyi esas alır.

Bu çalışmada, görevlerin operasyonel çeşitliliğini en iyi yansıtacağı düşünülen iki temel değişken, tabakalama kriteri olarak kullanılmıştır. Bunlar;

- Güzergâh Kimliği (*busRoute*),
 - Mevsim (*season*).
- **Güzergâh Kimliği (*busRoute*):** Her güzergahın kendine özgü topografik ve trafik profiline sahip olması nedeniyle.
 - **Mevsim (*season*):** Hava koşullarının ve yolcu yoğunluğunun mevsime bağlı olarak enerji tüketimi üzerindeki güçlü etkileri nedeniyle.

Bu süreç sonunda, 162 görevden oluşan temsili bir ana veri seti oluşturulmuştur. Örneklemin orijinal popülasyon dağılımını ne kadar başarılı temsil ettiğini doğrulamak amacıyla, görevlerin mevsimsel ve güzergâh bazlı dağılımları, popülasyon ve örneklem için karşılaştırmalı olarak Çizelge 3.2 ve Çizelge 3.3'te sunulmuştur. Tablolarda görüldüğü üzere, örneklemdeki dağılımlar ile popülasyon dağılımları arasında belirgin bir sapma olmaması, seçilen örneklemin geçerliliğini ve popülasyonu yüksek doğrulukla temsil ettiğini istatistiksel olarak teyit etmektedir.

Son olarak, ana veri setini oluşturan 162 görevin listesi, başlangıç zamanlarına (*startTime_iso*) göre kronolojik olarak sıralanmıştır. Bu sıralama, TimeSeriesSplit gibi zamana duyarlı çapraz doğrulama yöntemlerinin doğru çalışması ve zaman serisi verilerinde model öğrenme tutarlılığının sağlanması için kritik öneme sahiptir. Sıralanmış görev listesi, bir sonraki veri birleştirme adımı için temel girdi olarak kullanılmıştır.

Çizelge 3.2. Ana veri seti ve temsili örneklemdaki görevlerin mevsimsel dağılımının karşılaştırılması

Mevsim	Ana Veri Seti Dağılımı (%)	Örneklem Dağılımı (%)
Sonbahar	30,79	30,25
Bahar	26,64	27,16
Yaz	30,49	30,25
Kış	12,08	12,35

Çizelge 3.3. Ana veri seti ve temsili örneklemedeki görevlerin güzergâh dağılımının karşılaştırılması

Güzergâh Kimliği	Ana Veri Seti Dağılımı (%)	Örneklem Dağılımı (%)
31	1,28	2,47
32	1,96	2,47
33	11,7	11,73
46	9,28	9,26
72	11,92	11,73
83	63,85	62,35

3.3.3. Verilerin ön işlenmesi ve özellik mühendisliği

Ana ve bağımsız test (holdout) veri setlerine ait görev listeleri tanımlandıktan sonraki aşamada, bu görevlerin ham zaman serisi verileri işlenmiştir. Ham veriler, makine öğrenmesi modellerine doğrudan girdi olarak sunulmaya elverişli olmadığından, kapsamlı bir ön işleme ve özellik mühendisliği sürecine tabi tutulmuştur. Bu sürecin temel amacı; ham verileri bütünleştirmek, temizlemek ve model performansı için kritik öneme sahip yeni ve bilgilendirici özellikler türeterek güvenilir bir modelleme altyapısı hazırlamaktır.

Sonuçların tutarlılığını ve karşılaştırılabilirliğini güvence altına almak amacıyla, bu işlemler hem ana veri setini oluşturan 162 görev hem de holdout veri setindeki görevler için birebir aynı şekilde uygulanmıştır. Bu çok adımlı süreç, aşağıdaki alt bölümlerde ayrıntılı olarak açıklanmaktadır.

3.3.3.1. Verilerin birleştirilmesi ve ilk hazırlık

Ön işleme sürecinin ilk adımı, seçilen görevlere ait zaman serisi verilerinin ilgili CSV dosyalarından okunarak tek bir ana veri çerçevesinde (DataFrame) birleştirilmesidir. Grup bazlı operasyonların (örneğin ivme hesaplaması) her görev özelinde doğru yapılmasını sağlamak ve farklı görevlere ait verilerin karışmasını önlemek amacıyla, bu aşamada her bir veri çerçevesine o görevi tanımlayan *mission_id* adlı bir kimlik sütunu eklenmiştir. Son olarak, bu bireysel veri çerçeveleri Pandas kütüphanesi aracılığıyla birleştirilerek, sonraki tüm işlemlerin temelini oluşturan 5,5 milyondan fazla satırlık ana veri çerçevesi oluşturulmuştur.

3.3.3.2. Veri temizliği ve eksik veri yönetimi

Birleştirilmiş ana veri çerçevesi üzerinde, eksik (NaN) veya geçersiz değerleri yönetmek ve veri kalitesini artırmak amacıyla kapsamlı bir veri temizliği işlemi

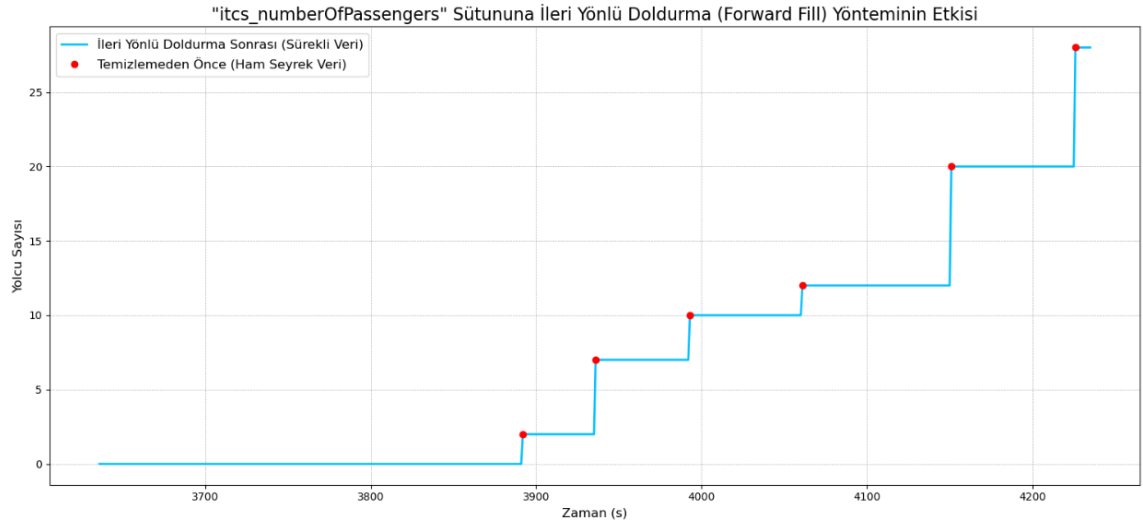
uygulanmıştır. Bu adımda, her özellik için en uygun yöntemler belirlenerek, görev bazında (groupby('mission_id')) aşağıdaki düzeltmeler yapılmıştır. Bunlar;

- *itcs_busRoute* (Güzergâh Kimliği),
 - *itcs_numberOfPassengers* (Yolcu Sayısı),
 - *gnss_altitude* (Yükseklik),
 - Diğer Operasyonel Değişkenler.
-
- ***itcs_busRoute* (Güzergâh Kimliği):** Veri setinde, bazı anlık kayıtlarda güzergâh kimliğinin geçersiz bir karakter olan '-' ile kaydedildiği gözlemlenmiştir. Her görevin tek bir sabit güzergâhta gerçekleştiği varsayımına dayanarak, bu tutarsızlığı gidermek amacıyla her görev grubu içindeki geçersiz kayıtlar, o göreve ait geçerli güzergâh kimliği ile doldurulmuştur. Bu işlem, veri setinin bütünlüğünü ve tutarlılığını sağlamıştır.
 - ***itcs_numberOfPassengers* (Yolcu Sayısı):** Yolcu sayısını ölçen sensörler, doğası gereği olaya dayalı çalıştığından, veriyi sadece duraklarda yolcu sayısında bir değişiklik olduğunda kaydetmektedir. Bu durum, duraklar arasındaki zaman dilimlerinde verinin eksik (NaN) olmasına yol açmaktadır. Bu seyrek yapıyı makine öğrenmesi modellerinin işleyebileceği sürekli bir zaman serisine dönüştürmek için, her görev özelinde ileri yönlü doldurma (ffill) yöntemi uygulanmıştır. Bu yöntem, her eksik değeri aynı görevdeki bir önceki geçerli değerle doldurarak, yolcu sayısının iki durak arasında sabit kaldığı mantıksal varsayımını temel almaktadır. Görevlerin başlangıcındaki olası eksik değerler ise sıfır (0) olarak atanmıştır. Bu sürecin seyrek veriyi sürekli bir zaman serisine dönüştürmedeki etkisi Şekil 3.3'te görselleştirilmiştir.
 - ***gnss_altitude* (Yükseklik):** Yol eğimi (*road_grade*) gibi kritik bir özelliğin türetilmesi, doğrudan yükseklik verisinin kalitesine bağlıdır. Ham GNSS verilerindeki sinyal kaybı kaynaklı anlık boşluklar, eğim hesaplamalarında ciddi hatalara yol açabileceğinden, bu riski ortadan kaldırmak ve sürekli bir yükseklik profili oluşturmak amacıyla *gnss_altitude* sütununa üç aşamalı ve sağlam bir doldurma yöntemi uygulanmıştır:
 1. **Doğrusal Enterpolasyon (interpolate):** Ara noktadaki boşluklar, iki bilinen nokta arasındaki doğrusal eğilim varsayımıyla mantıksal olarak doldurulmuştur.
 2. **Geri Yönlü Doldurma (bfill):** Görevlerin başlangıcındaki olası eksik değerler, o göreve ait ilk geçerli değerle doldurulmuştur.

3. **İleri Yönlü Doldurma (ffill):** Kalan tüm olası boşluklar, bir önceki geçerli değerle doldurularak sütunun tamamen temizlenmesi garanti altına alınmıştır.

Bu çok adımlı süreç, yol eğimi hesaplaması için güvenilir bir temel oluşturmuştur.

- **Diğer Operasyonel Değişkenler:** *odometry_vehicleSpeed*, *temperature_ambient* ve *traction_tractionForce* gibi diğer kritik değişkenlerde de sinyal kaybı nedeniyle nadiren eksik değerler bulunabilmekteydi. Modelin eğitim sürecinde herhangi bir hatayla karşılaşmasını önlemek amacıyla bu sütunlardaki olası eksik değerler de ffill ve bfill yöntemleri kullanılarak giderilmiş ve böylece veri setinin modellemeye tamamen hazır olması sağlanmıştır.



Şekil 3.3. itcs_numberOfPassengers sütununa ileri yönlü doldurma (forward fill) yönteminin uygulanmasının etkisi: ham veriler (kırmızı noktalar) ve temizlenmiş sürekli veriler (mavi çizgi)

3.3.3.3. Yeni özelliklerin mühendisliği

Veri temizliği adımı modelleme için tutarlı bir temel oluştursa da, ham sensör verileri genellikle enerji tüketimini yönlendiren temel fiziksel ve operasyonel dinamikleri yansıtmakta yetersiz kalmaktadır. Bu nedenle, modelin tahmin gücünü artırmak ve ham verilerde örtük olarak bulunan bilgileri açık hale getirmek amacıyla bir özellik mühendisliği süreci uygulanmıştır. Bu süreçte, temel değişkenlerden yola çıkılarak enerji tüketimiyle daha güçlü ve anlamlı bir ilişkiye sahip olduğu düşünülen yeni değişkenler türetilmiştir. Bunlar;

- Zamansal Özellikler (*hour_of_day*, *day_of_week*),
- İvme (*acceleration*),
- Yol Eğimi (*road_grade*).

- **Zamansal Özellikler:** Makine öğrenmesi modelleri, standart bir zaman damgasından yola çıkarak trafik yoğunluğundaki günlük pik saatler veya hafta içi/hafta sonu operasyonel farklılıkları gibi döngüsel kalıpları doğrudan yorumlayamaz. Bu zamansal bilgiyi modele açıkça sunmak amacıyla, *time_iso* sütunundan *hour_of_day* (günün saati) ve *day_of_week* (haftanın günü) olmak üzere iki yeni sayısal özellik türetilmiştir. Bu özellikler, modelin günün farklı saatlerindeki ve haftanın farklı günlerindeki tipik enerji tüketim desenlerini öğrenmesine olanak tanır.
- **İvme:** Enerji tüketimi, yalnızca aracın hızıyla değil, aynı zamanda hızındaki değişim oranıyla da yakından ilişkilidir. Anlık ivme, sürücü davranışının (örneğin agresif veya yumuşak hızlanma) ve araca etki eden atalet kuvvetlerinin doğrudan bir göstergesidir. Bu kritik bilgiyi modellemek amacıyla *acceleration* özelliği, *odometry_vehicleSpeed* (araç hızı) verisinin zamana göre türevi alınarak (her saniyedeki değişimi) hesaplanmıştır.
- **Yol Eğimi:** Yer çekimi kuvveti, özellikle engebeli arazilerde enerji tüketimini belirleyen en önemli faktörlerden biridir. Yokuş yukarı hareket bu kuvvete karşı ek enerji gerektirirken, yokuş aşağı hareket rejeneratif frenleme ile enerji geri kazanımı için bir fırsat sunar. Bu temel fiziksel etkiyi modellemek amacıyla *road_grade* özelliği, *gnss_altitude* (yükseklik) ve *odometry_vehicleSpeed* verileri kullanılarak yolun anlık eğimi olarak hesaplanmıştır. Bu özelliğin kalitesini artırmak için, GPS sensörlerinin doğal gürültüsünden kaynaklanabilecek aşırı değerler bir "kırpma" (clipping) fonksiyonu ile sınırlandırılmış ve ardından daha kararlı bir eğim profili elde etmek amacıyla kayan ortalama (rolling average) filtresi ile yumuşatılmıştır.

Bu süreçte türetilen yeni özellikler, açıklamaları ve kaynaklandıkları orijinal değişkenler Çizelge 3.4'te özetlenmiştir.

Çizelge 3.4. Özellik mühendisliği ile türetilen yeni değişkenler

Kaynak Değişkenler	Açıklama	Özellik Adı
<i>time_iso</i>	Günün saati (0-23 aralığı)	<i>hour_of_day</i>
<i>time_iso</i>	Haftanın günü (Pazartesi=0, Pazar=6)	<i>day_of_week</i>
<i>odometry_vehicleSpeed</i>	Aracın anlık ivmesi (m/s^2)	<i>acceleration</i>
<i>gnss_altitude, odometry_vehicleSpeed</i>	Tahmini anlık yol eğimi (oran olarak)	<i>road_grade</i>

3.4. Modelleme ve Değerlendirme Yaklaşımı

Bu bölümde, anlık enerji tüketimini en yüksek doğrulukla tahmin edecek makine öğrenmesi modelini seçmek amacıyla izlenen sistematik yaklaşım detaylandırılmaktadır. Bu yaklaşım, ön işleme ve özellik mühendisliği adımlarından geçen ana veri setini kullanarak model geliştirme, değerlendirme ve seçim süreçlerini kapsar. Süreç boyunca farklı algoritmalar adil bir şekilde karşılaştırılmış, performansları bilimsel olarak geçerli yöntemlerle ölçülmüş ve nihai model bu bulgulara dayanılarak belirlenmiştir.

3.4.1. Değerlendirilen makine öğrenmesi algoritmaları

Problemin hem doğrusal hem de karmaşık, doğrusal olmayan yapılarını analiz edebilmek amacıyla, farklı yaklaşımları temsil eden dört denetimli öğrenme algoritması değerlendirilmeye alınmıştır. Bunlar;

- Doğrusal Regresyon (MLR),
 - XGBoost (eXtreme Gradient Boosting),
 - LightGBM (Light Gradient Boosting Machine),
 - HistGradientBoostingRegressor.
-
- **Doğrusal Regresyon (MLR):** Girdi özellikleri ile hedef değişken arasında doğrusal bir ilişki varsayan bu temel model, yorumlanabilirliğinin yüksek ve hesaplama maliyetinin düşük olması nedeniyle, daha karmaşık modellerin performansını ölçmek için bir temel seviye modeli olarak kullanılmıştır.
 - **XGBoost (eXtreme Gradient Boosting):** Yüksek tahmin doğruluğu, hızı ve ölçeklenebilirliği ile bilinen, Gradyan Artırma (Gradient Boosting) ailesinin endüstri standardı haline gelmiş bir uygulamasıdır.
 - **LightGBM (Light Gradient Boosting Machine):** Microsoft tarafından geliştirilen ve Gradyan Artırma tekniğini kullanan bu algoritma, özellikle büyük veri setleri üzerinde XGBoost'a kıyasla daha hızlı olması ve daha az bellek tüketmesiyle öne çıkan oldukça verimli bir alternatiftir.
 - **HistGradientBoostingRegressor:** Scikit-learn kütüphanesinin modern bir Gradyan Artırma uygulamasıdır. LightGBM'den ilham alan histogram tabanlı yapısı sayesinde büyük veri setlerinde yüksek hız ve verimlilik sunmaktadır.

Başlangıçta, Random Forest (RF) ve Destek Vektör Regresyonu (SVR) gibi diğer yaygın algoritmalar da değerlendirme için düşünülmüştü. Ancak, veri setinin 5,5 milyondan fazla satır içermesi nedeniyle yapılan ilk testlerde, bu iki modelin eğitim sürelerinin kabul edilemez derecede uzun olduğu ve özellikle Random Forest modelinin yüksek bellek gereksinimi nedeniyle MemoryError hatasına yol açtığı gözlemlenmiştir. Bu nedenle, bu iki model yerine hesaplama açısından çok daha verimli olan LightGBM ve HistGradientBoosting algoritmalarının kullanılmasına karar verilmiştir.

3.4.2. Model değerlendirme stratejisi

Geliştirilen makine öğrenmesi modellerinin genelleme performansını güvenilir, yansız ve bilimsel olarak geçerli bir şekilde ölçebilmek amacıyla, bu çalışmada iki temel stratejiyi birleştiren bütüncül bir yaklaşım benimsenmiştir. Bunlar;

- Veri sızıntısını önlemek için *Pipeline* Mimarisi
 - Zaman serisi verilerinin doğasına uygun *TimeSeriesSplit* ile Çapraz Doğrulama
- **Veri Sızıntısını Önleme: Pipeline Mimarisi ve Veri Ölçeklendirme:** Makine öğrenmesi modellerinin eğitiminde karşılaşılan en kritik risklerden biri veri sızıntısıdır. Bu durum, modelin eğitim aşamasında, test aşamasında görmemesi gereken bilgileri dolaylı yoldan öğrenmesi ve bu nedenle performansının yapay olarak yüksek çıkmasıdır. Bu riski ortadan kaldırmak için *scikit-learn* kütüphanesinin Pipeline mimarisi kullanılmıştır. Bir Pipeline, veri ön işleme adımlarını (bu çalışmada veri ölçeklendirme) ve modelin kendisini tek bir birleşik nesne içinde zincirler. Bu yapının temel avantajı, çapraz doğrulama sırasında her bir adımın doğru veri üzerinde uygulanmasını garanti altına almasıdır.

Bu çalışmada kullanılan MinMaxScaler, tüm girdi özelliklerini standart bir aralık olan [0, 1] arasına dönüştüren bir veri ölçeklendirme tekniğidir. Bu işlem, *traction_tractionForce* (on binlerce Newton) gibi geniş bir ölçeğe sahip özelliklerin, *road_grade* (oran olarak ifade edilen küçük bir sayı) gibi daha küçük ölçekli özelliklere baskın çıkmasını önler. Eğer ölçeklendirme yapılmazsa, birçok algoritma sadece sayısal büyüklükleri nedeniyle geniş ölçekli özelliklere daha fazla önem atfeder. MinMaxScaler, her özelliğin minimum değerini 0'a, maksimum değerini 1'e ve aradaki diğer tüm değerleri bu aralıkta orantılı olarak eşleyerek bu sorunu çözer.

Pipeline mimarisi, bu ölçeklendirme işlemini çapraz doğrulamanın her bir döngüsü (fold) içinde ve sadece o anki eğitim verisi üzerinde gerçekleştirir. Yani, model

bir veri bölümü üzerinde eğitilirken, ölçekleyici de sadece o bölümün minimum ve maksimum değerlerini öğrenir. Test verisi ise, eğitim verisinden öğrenilen bu ölçek ile dönüştürülür. Bu sayede, test verisinden eğitim verisine hiçbir bilgi sızmaz ve modelin performansı yansız bir şekilde ölçülür.

- **Çapraz Doğrulama: Standart Yöntemler ve Zaman Serisi Verilerinin Zorlukları:** Bir modelin genelleme yeteneğini ölçmenin en güvenilir yollarından biri çapraz doğrulamadır (Cross-Validation). Bu yöntemin temel amacı, modeli veri setinin farklı alt kümeleri üzerinde tekrar tekrar eğitip test ederek, performansının tek bir rastgele veri bölünmesine bağlı olmadığından emin olmaktır. K-Fold gibi klasik çapraz doğrulama yöntemleri, veri setini rastgele bir şekilde K adet eşit parçaya böler ve her seferinde bir parçayı test, geri kalanını ise eğitim için kullanır.

Ancak bu klasik yöntemler, veriler arasında zamansal bir bağımlılık bulunan zaman serisi verileri için tamamen uygunsuzdur. Veriyi rastgele karıştırmak, modelin gelecekteki verileri (örneğin, salı gününe ait veriler) kullanarak geçmişteki verileri (örneğin, pazartesi gününe ait veriler) "tahmin etmesine" yol açar. Bu, gerçek dünya senaryolarıyla tutarlı olmayan ve modelin performansının yanıltıcı derecede yüksek çıkmasına neden olan ciddi bir veri sızıntısı türüdür.

3.4.2.1. Zaman serisi çapraz doğrulama: timeseriessplit yaklaşımı

Bu çalışmada, zaman serisi verilerinin kronolojik yapısını korumak ve gerçek dünya tahmin senaryosunu doğru bir şekilde simüle etmek amacıyla scikit-learn kütüphanesinin TimeSeriesSplit yöntemi kullanılmıştır. Bu yaklaşım, veriyi asla rastgele karıştırmaz ve "ileri yürüme" (walk-forward) olarak bilinen bir mantıkla çalışır.

`n_splits=5` olarak ayarlandığında, TimeSeriesSplit veri setini ardışık zaman bloklarına bölerek aşağıdaki gibi 5 adet eğitim/test çifti oluşturur.

- **1. Döngü:** Model, Blok 1 üzerinde eğitilir ve Blok 2 üzerinde test edilir.
- **2. Döngü:** Model, Blok 1 ve 2 üzerinde eğitilir ve Blok 3 üzerinde test edilir.
- **3. Döngü:** Model, Blok 1, 2 ve 3 üzerinde eğitilir ve Blok 4 üzerinde test edilir.
- Bu süreç, tüm veri kullanılana kadar bu şekilde devam eder.

Bu yöntem, modelin her zaman geçmiş verilerle eğitilip, hiç görmediği gelecek veriler üzerinde test edilmesini garanti altına alır. Bu yaklaşım, bir modelin gerçek

operasyonel koşullarda nasıl performans göstereceğini ölçmek için en doğru ve bilimsel olarak en geçerli yöntemdir.

3.4.3. Performans değerlendirme metrikleri

Geliştirilen modellerin performansını nicel olarak değerlendirmek ve farklı yönlerini (ortalama hata, aykırı değerlere duyarlılık, açıklama gücü vb.) analiz etmek amacıyla, literatürde yaygın olarak kabul görmüş ve bu çalışmanın hedeflerine en uygun metrikler seçilmiştir. Bu metrikler, modellerin hem tahmin doğruluğunu hem de hesaplama verimliliğini kapsamlı bir şekilde analiz etmek için kullanılmıştır.

- Belirlilik Katsayısı (R^2),
- Ortalama Mutlak Hata (MAE),
- Hataların Ortalama Karekökü (RMSE),
- Elenen Metrik (MAPE),
- Verimlilik Metriği: Toplam Eğitim Süresi.

- **Belirlilik Katsayısı (R^2)**

MAE ve RMSE'nin aksine R^2 , modelin hatasından ziyade "uyum iyiliğini" (goodness of fit) ölçmektedir. Bu metrik, hedef değişkendeki varyansın, modelin girdi özellikleri tarafından ne kadarının açıklanabildiğini gösterir. Değeri genellikle 0 ile 1 arasında değişen R^2 için daha yüksek bir değer, daha iyi bir model performansına işaret etmektedir. Belirlilik katsayısı aşağıda verilen formül 3.1'den hesaplanabilir. Formül 3.1'de n :denek sayısını, $\hat{y}_i$: test örneğindeki tahmin edilen değerleri y_i : ise gerçek değerleri $\underline{y}$, ise hedef değişkenin gerçek değerlerinin ortalamasını temsil etmektedir.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \underline{y})^2} \quad (3.1)$$

- **Ortalama Mutlak Hata (MAE)**

Bu metrik, formül 3.2 'den tüm n (denek sayısı) test örneğindeki tahmin edilen değerler ($\hat{y}_i$) ile gerçek değerler (y_i) arasındaki farkın mutlak değerlerinin ortalamasından hesaplamaktadır. Hatanın doğrudan ve sezgisel bir yorumunu sunan MAE'nin birimi, hedef değişkenin birimiyle (bu çalışmada Watt) aynıdır.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3.2)$$

- **Hataların Ortalama Karekökü (RMSE)**

Bu metrik de modelin ortalama hatasını ölçer, ancak temel farkı, büyük hatalara çok daha fazla ağırlık vermesidir. RMSE değeri formül 3.2'den gerçek ve tahmin edilen değerler arasındaki farkların karelerini alınarak hesaplanır, Karesi alınarak oluşması muhtemel bariz hatalı tahminlere daha ağır bir ceza uygulanır. Formül 3.3'de n :denek sayısını, $\hat{y}_i$: test örneğindeki tahmin edilen değerleri y_i : ise gerçek değerleri temsil etmektedir. Bu metriğin birimi de hedef değişkenin birimiyle (Watt) aynıdır.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3.3)$$

- **Elenen Metrik: Ortalama Mutlak Yüzde Hata (MAPE)**

Metrik seçim sürecinde, yüzde olarak hata sunması nedeniyle yorumlama kolaylığı sağlayan MAPE de değerlendirilmiştir. Bunun değeri ise formül 3.4 den hesaplanabilir. Formül 3.4'de n :denek sayısını, $\hat{y}_i$: test örneğindeki tahmin edilen değerleri y_i : ise gerçek değerleri temsil etmektedir.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3.4)$$

Formül 3.4'de MAPE'nin paydasında gerçek değer (y_i) yer aldığı görülmektedir. Bu çalışmadaki hedef değişken olan *electric_powerDemand*, aracın durduğu veya enerji geri kazanımı yaptığı anlarda sıfır, sıfıra yakın ve hatta negatif değerler alabilmektedir. Bu durum, MAPE hesaplamasının tanımsız (sonsuz) veya yanıltıcı derecede büyük hata yüzdelere yol açmasına neden olmaktadır. Dolayısıyla, bu probleme özgü kararsızlığı nedeniyle MAPE'nin kullanımından vazgeçilmiş ve analizler de MAE, RMSE ve R^2 gibi bu çalışma için daha uygun olduğu düşünülen metrikler üzerine odaklanılarak gerçekleştirilmiştir.

- **Verimlilik Metriği: Toplam Eğitim Süresi**

Tahmin doğruluğunun yanı sıra, bir modelin pratik uygulanabilirliği hesaplama verimliliğine de bağlıdır. Bu nedenle, her modelin TimeSeriesSplit ile yapılan 5 katlı çapraz doğrulama sürecinin toplam tamamlanma süresi (saniye cinsinden) da bir

performans metriği olarak kaydedilmiş ve modellerin verimliliği bu açıdan da karşılaştırılmıştır.

3.5. Nihai Modelin Seçimi ve Eğitimi

Önceki aşamalarda detaylandırılan sistematik değerlendirme sürecinin ardından bu bölümde, en iyi performansı gösteren makine öğrenmesi modeli seçilmektedir. Ayrıca, bu seçimin gerekçeleri sunulmakta ve modelin nihai kullanım için eğitilip saklanması adımları açıklanmaktadır.

3.5.1. Modellerin karşılaştırmalı performans analizi

Değerlendirilen dört makine öğrenmesi algoritmasının (Linear Regression, XGBoost, LightGBM ve HistGradientBoosting) zaman serisi çapraz doğrulama ile elde edilen performans sonuçları Çizelge 3.5'da özetlenmiştir. Bu tablo, modellerin tahmin doğruluğunu (R^2), ortalama hata miktarını (MAE ve RMSE) ve hesaplama verimliliğini (Eğitim Süresi) bir arada sunarak, bütüncül bir karşılaştırma imkânı sağlamaktadır.

Çizelge 3.5. Modellerin zaman serisi çapraz doğrulama ile karşılaştırmalı performans sonuçları

<i>model</i>	<i>Eğitim Süresi (s)</i>	<i>R² (Mean)</i>	<i>R² (Std)</i>	<i>MAE (Mean)</i>	<i>MAE (Std)</i>	<i>RMSE (Mean)</i>	<i>RMSE (Std)</i>
<i>LightGBM</i>	55,0299	0,97708	0,00628	7415,63	1772,39	10091,6	1490,93
		7	7	3	1	7	5
<i>HistGradientBoosting</i>	99,1151	0,97685	0,00631	7446,53	1768,83	10144,7	1488,54
	3	3	3	5	3	3	8
<i>XGBoost</i>	57,1336	0,97547	0,00651	7635,73	1720,39	10449,0	1488,00
	7	8	9	9	9	2	2
<i>Linear Regression</i>	9,48175	0,72146	0,00283	23207,7	655,910	35559,1	205,166
	8	2	3	8	8		5

Analiz sonuçları, Gradyan Artırma (Gradient Boosting) ailesine ait modellerin (XGBoost, LightGBM, HistGradientBoosting), temel doğrusal regresyon modeline kıyasla belirgin bir şekilde daha iyi performans gösterdiğini ortaya koymaktadır. Bu durum, elektrikli otobüslerin enerji tüketim dinamiklerinin, doğrusal varsayımlarla tam olarak yakalanamayan karmaşık ve doğrusal olmayan bir yapıya sahip olduğu hipotezini güçlü bir şekilde doğrulamaktadır. Ayrıca, üç gelişmiş modelin de birbirine oldukça yakın ve yüksek doğruluk oranları sergilemesi, elde edilen sonuçların tek bir algoritmaya özgü olmadığını ve genel olarak bu model ailesinin probleme uygunluğunu göstermektedir.

3.5.2. En iyi modelin seçimi ve gerekçesi

Çizelge 3.5'daki sonuçların detaylı incelenmesi sonucunda, bu çalışmanın nihai modeli olarak LightGBM algoritması seçilmiştir. Bu seçimin temel gerekçeleri çok yönlüdür. Bular;

- En Yüksek Tahmin Doğruluğu,
 - En Düşük Hata Oranları,
 - Hesaplama Verimliliği.
- **En Yüksek Tahmin Doğruluğu:** LightGBM, 0,977'lik ortalama R^2 skoru ile test edilen modeller arasında en yüksek tahmin doğruluğuna ulaşmıştır. Bu, hedef değişkendeki varyansın en büyük kısmını açıklama kapasitesine sahip olduğu anlamına gelir.
 - **En Düşük Hata Oranları:** Buna ek olarak, hem Ortalama Mutlak Hata (MAE) hem de Kök Ortalama Kare Hata (RMSE) metriklerinde en düşük değerleri kaydederek, tahminlerinin gerçek değerlere ortalamada en yakın olan model olduğunu kanıtlamıştır.
 - **Hesaplama Verimliliği:** Eğitim süresi açısından XGBoost ile rekabetçi bir sürede eğitimi tamamlaması, bu modeli hem doğruluk hem de verimlilik arasında en iyi dengeyi sunan aday olarak öne çıkarmıştır.

3.5.3. Nihai modelin tüm veri seti üzerinde eğitilmesi ve saklanması

Model seçiminin ardından, seçilen LightGBM algoritmasının genelleme kabiliyetini en üst düzeye çıkarmak amacıyla, model son bir kez tüm ana veri seti (X_{main} ve y_{main}) üzerinde eğitilmiştir. Bu yaklaşım, modelin mevcut tüm bilgi ve kalıplardan öğrenmesini sağlayarak, daha önce hiç görmediği bağımsız test (holdout) verileri üzerinde mümkün olan en yüksek performansı göstermesini hedefler.

Eğitim süreci, veri ön işleme ve modellemeyi entegre eden bir Pipeline mimarisi kullanılarak gerçekleştirilmiştir. Bu Pipeline, ilk adımda verileri MinMaxScaler ile [0, 1] aralığına ölçeklendirir ve ikinci adımda bu ölçeklenmiş veriyi LGBMRegressor modeline sunar. Son olarak, bu eğitilmiş nihai Pipeline, tüm öğrenilmiş parametreleriyle birlikte, gelecekteki kullanım ve yansız değerlendirme için joblib kütüphanesi aracılığıyla *final_model.pkl* adıyla bir dosyaya kaydedilmiştir. Bu serileştirme işlemi, zaman alıcı

eđitim s¼recini tekrar etmeksizin, eđitilmiř modelin dođrudan y¼klenip yeniden kullanılabilmesini m¼mk¼n kılar.



4. ARAŞTIRMA BULGULARI VE TARTIŞMA

Bu bölümde, yürütülen modelleme çalışmalarından elde edilen ampirik bulgular sunulmakta ve bu bulguların çıkarımları, araştırmanın temel hipotezleri ve ilgili literatür bağlamında tartışılmaktadır. Bölümde, modelleme sürecine temel teşkil eden değişkenler arasındaki ilişkileri ortaya koyan keşifsel bir veri analizi ile başlanmaktadır.

4.1. Keşifsel Veri Analizi ve Değişken İlişkileri

Makine öğrenmesi modelleri geliştirilmeden önce, girdi değişkenleri ile hedef değişken arasındaki ve girdi değişkenlerinin kendi aralarındaki ilişkilerin incelenmesi, model davranışını anlamak ve sonuçları doğru yorumlamak için kritik bir adımdır. Bu amaçla, çalışmada kullanılan 11 temel girdi ve 1 hedef değişken arasındaki doğrusal ilişkiyi ölçmek için bir korelasyon matrisi hesaplanmıştır.

Bilindiği üzere, korelasyon katsayısı -1 ile +1 arasında bir değer alır; +1 mükemmel pozitif, -1 mükemmel negatif doğrusal ilişkiyi, 0 ise doğrusal bir ilişkinin yokluğunu ifade eder. Bu analizin sonuçları, Çizelge 4.1'de sunulan korelasyon matrisinde özetlenmiştir.

Çizelge 4.1. Modellemede kullanılan değişkenler arasındaki pearson korelasyon matrisi

	<i>odometry_vehicle Speed</i>	<i>accelerat ion</i>	<i>road_gr ade</i>	<i>itscs_numberOfPasse ngers</i>	<i>temperature_am bient</i>	<i>traction_traction Force</i>	<i>traction_brakePre ssure</i>	<i>status_doorIs Open</i>	<i>status_gridIsAvai lable</i>	<i>hour_of_ day</i>	<i>day_of_w eek</i>	<i>electric_powerDe mand</i>
<i>odometry_vehicleSp eed</i>	1	0,055	-0,039	0,064	-0,017	0,063	-0,723	-0,53	0,01	0,009	0,014	0,188
<i>acceleration</i>	0,055	1	-0,018	0	0,002	0,858	-0,042	-0,002	0	-0,001	0	0,677
<i>road_grade</i>	-0,039	-0,018	1	0,003	0,002	0,122	0,022	0,013	0,007	-0,003	-0,004	0,129
<i>itscs_numberOfPasse ngers</i>	0,064	0	0,003	1	0,004	0,022	-0,074	-0,015	0,054	0,07	-0,042	0,026
<i>temperature_ambien t</i>	-0,017	0,002	0,002	0,004	1	-0,008	0,013	0,036	-0,053	0,045	-0,049	-0,031
<i>traction_tractionFor ce</i>	0,063	0,858	0,122	0,022	-0,008	1	-0,136	-0,096	0,012	-0,001	-0,001	0,831
<i>traction_brakePress ure</i>	-0,723	-0,042	0,022	-0,074	0,013	-0,136	1	0,743	0,042	-0,02	-0,004	-0,158
<i>status_doorIsOpen</i>	-0,53	-0,002	0,013	-0,015	0,036	-0,096	0,743	1	0,025	-0,008	-0,016	-0,111
<i>status_gridIsAvailab le</i>	0,01	0	0,007	0,054	-0,053	0,012	0,042	0,025	1	0,028	0,164	0,026
<i>hour_of_day</i>	0,009	-0,001	-0,003	0,07	0,045	-0,001	-0,02	-0,008	0,028	1	0,124	-0,001
<i>day_of_week</i>	0,014	0	-0,004	-0,042	-0,049	-0,001	-0,004	-0,016	0,164	0,124	1	0,006
<i>electric_powerDem and</i>	0,188	0,677	0,129	0,026	-0,031	0,831	-0,158	-0,111	0,026	-0,001	0,006	1

Bu matrisin incelenmesi, enerji tüketimi tahmini için önemli ipuçları sunmaktadır:

- **Hedef Değişken ile En Güçlü İlişkiler:** Beklendiği üzere, hedef değişken olan *electric_powerDemand* (anlık güç talebi), en güçlü pozitif korelasyonu *traction_tractionForce* (çekiş kuvveti) ($r = 0,83$) ve *acceleration* (ivme) ($r = 0,68$) değişkenleri ile göstermektedir. Bu bulgu, temel fizik prensipleriyle tam uyum içindedir; bir aracı hızlandırmak için gereken anlık mekanik güç, doğrudan uygulanan kuvvet ve ivme ile ilişkilidir. Bu yüksek korelasyon, söz konusu iki değişkenin, modelin tahmin doğruluğu açısından en belirleyici faktörler olacağına dair güçlü bir gösterge sunmaktadır.
- **Diğer Önemli İlişkiler:** *odometry_vehicleSpeed* (araç hızı), *electric_powerDemand* ile daha zayıf ancak pozitif bir korelasyona ($r = 0,19$) sahiptir. Diğer yandan, *traction_brakePressure* (fren basıncı) ile *electric_powerDemand* arasındaki negatif korelasyon ($r = -0,10$), frenleme anlarında enerji tüketiminin azaldığını veya rejeneratif frenleme ile enerji geri kazanımı (negatif güç talebi) yaşandığını doğrulamaktadır.
- **Beklenenden Düşük Korelasyonlar:** Dikkat çekici bir bulgu, literatürde genellikle önemli kabul edilen *itcs_numberOfPassengers* (yolcu sayısı) *electric_powerDemand* ile neredeyse sıfıra yakın bir doğrusal ilişkiye sahip olmasıdır. Bu durum, bu değişkenin önemsiz olduğu anlamına gelmez. Aksine, bu düşük korelasyon, bu değişkenlerin etkisinin (kütle artışı gibi), daha kapsayıcı ve doğrudan ölçülen bir değişken olan *traction_tractionForce* sinyali içerisinde dolaylı olarak zaten kapsandığı hipotezini güçlendirmektedir. Başka bir deyişle, modelin, yolcu yükü nedeniyle artan direnci, doğrudan çekiş kuvvetindeki artıştan öğrenmesi beklenir.

Bu keşifsel analiz, modelin hangi değişkenlere daha fazla odaklanabileceğine dair değerli bir öngörü sağlamış ve sonraki bölümlerde sunulacak olan model performans analizinin temelini oluşturmuştur.

4.2. Modellerin Karşılaştırmalı Performans Analizi

Bu bölümde, Bölüm 3'te metodolojisi detaylandırılan dört farklı makine öğrenmesi modelinin (Linear Regression, XGBoost, LightGBM ve HistGradientBoosting), daha önce hiç görmedikleri test verileri üzerindeki nicel performansı kapsamlı bir şekilde karşılaştırılmaktadır. Bu analizin temel amacı, tahmin

doğruluğu, hata oranları ve hesaplama verimliliği gibi çok yönlü kriterler üzerinden en uygun ve en güçlü modeli belirlemektir.

4.2.1. Tahmin doğruluğu ve hata oranlarının değerlendirilmesi

Modellerin performans sonuçları, Bölüm 3'te sunulan Çizelge 3.5 temel alınarak bu bölümde yorumlanmaktadır. Bu tablo, modellerin tahmin doğruluğunu, hata oranlarını ve bu metriklerin kararlılığını (standart sapma ile) bir arada sunarak bütüncül bir karşılaştırma imkânı sağlamaktadır.

Belirlilik Katsayısı (R^2) açısından incelendiğinde, Gradyan Artırma (Gradient Boosting) ailesine ait üç modelin (LightGBM, HistGradientBoosting ve XGBoost), temel doğrusal regresyon modeline kıyasla belirgin ölçüde daha üstün bir performans sergilediği görülmektedir. Bu üç modelin ortalama R^2 skorlarının %97,5'in üzerinde olması, hedef değişkendeki varyansın çok büyük bir kısmını başarıyla açıkladıklarını göstermektedir. Bu modeller arasında LightGBM, 0,9771'lik ortalama R^2 skoru ile en yüksek tahmin doğruluğuna ulaşmıştır. R^2 (Std) değerlerinin tüm gelişmiş modeller için oldukça düşük ($\sim 0,006$) olması, bu yüksek performansın veri setinin farklı zaman dilimlerinde de kararlı ve tutarlı olduğunu teyit etmektedir. Buna karşılık, Linear Regression modelinin %72'lik R^2 skoru, problemin doğasının basit doğrusal varsayımlarla tam olarak yakalanamayacak kadar karmaşık olduğunu kanıtlamaktadır.

Hata metrikleri (MAE ve RMSE) incelendiğinde, en iyi performans gösteren modeller ile temel model arasındaki fark daha da belirginleşmektedir. LightGBM modeli, 7415 Watt'lık Ortalama Mutlak Hata (MAE) ve 10091 Watt'lık Kök Ortalama Kare Hata (RMSE) ile en düşük hata oranlarına sahiptir. Bu, tahminlerinin gerçek değerlere ortalamada en yakın olan model olduğunu kanıtlamaktadır. HistGradientBoosting ve XGBoost modelleri de oldukça yakın ve başarılı sonuçlar elde etmiştir. Bu bulgu, Gradyan Artırma algoritmalarının, araç dinamiği ve operasyonel koşullar arasındaki karmaşık ve doğrusal olmayan ilişkileri öğrenme konusundaki üstün yeteneğini vurgulamaktadır.

Sonuç olarak, hem doğruluk hem de hata metrikleri açısından LightGBM en başarılı model olarak öne çıkarken, Linear Regression beklendiği gibi en zayıf performansı göstermiştir.

4.2.2. Eğitim süresi ve hesaplama verimliliği açısından karşılaştırma

Bir modelin pratik uygulanabilirliği, sadece tahmin doğruluğuna değil, aynı zamanda hesaplama verimliliğine de bağlıdır. Çizelge 3.5'deki "Eğitim Süresi (s)" sütunu, her bir modelin 5 katlı zaman serisi çapraz doğrulama sürecini tamamlama süresini saniye cinsinden göstermektedir.

Bu analizde, Linear Regression modeli beklendiği gibi yaklaşık 9,5 saniye ile en hızlı eğitimi tamamlarken, bu hızın bedeli düşük doğruluk olmuştur. Karmaşık modeller arasında ise, LightGBM yaklaşık 55 saniyelik eğitim süresi ile en verimli modellerden biri olarak dikkat çekmektedir. Histogram tabanlı yapıları sayesinde LightGBM ve HistGradientBoosting, büyük veri setleri üzerinde oldukça verimli çalışmaktadır.

Nihayetinde bu çok yönlü analiz, bir modelin seçiminde sadece en yüksek doğruluğun değil, aynı zamanda en düşük hatanın ve en verimli hesaplama süresinin de göz önünde bulundurulması gerektiğini göstermektedir. LightGBM modeli, en yüksek doğruluğu en düşük hata oranları ve oldukça rekabetçi bir eğitim süresi ile birleştirdiği için, bu projenin hedeflerine en uygun model olarak belirlenmiştir.

4.3. En İyi Modelin Seçimi ve Özellik Önem Analizi

Önceki bölümde sunulan karşılaştırmalı performans analizi, en uygun modelin seçilmesi için gerekli olan ampirik verileri sağlamıştır. Bu bölümde, bu veriler ışığında en iyi modelin seçimi gerekelecek ve seçilen modelin içsel davranışını anlamak amacıyla, girdi özelliklerinin tahmin üzerindeki etkileri detaylı bir şekilde analiz edilecektir.

4.3.1. En iyi modelin seçimi ve gerekçesi

Önceki bölümde (Çizelge 3.5) sunulan karşılaştırmalı analizin de gösterdiği gibi, LightGBM modeli en iyi genel performansı sergilemiştir. En yüksek tahmin doğruluğuna (R^2) ve en düşük hata oranlarına (MAE, RMSE) sahip olmasının yanı sıra, hesaplama açısından da verimli olması, bu modeli doğruluk ve verimlilik arasında en iyi dengeyi sunan aday yapmıştır. Bu gerekçeler doğrultusunda, bu çalışmanın nihai modeli olarak LightGBM algoritması seçilmiştir.

4.3.2. Nihai modelin özellik önem analizi

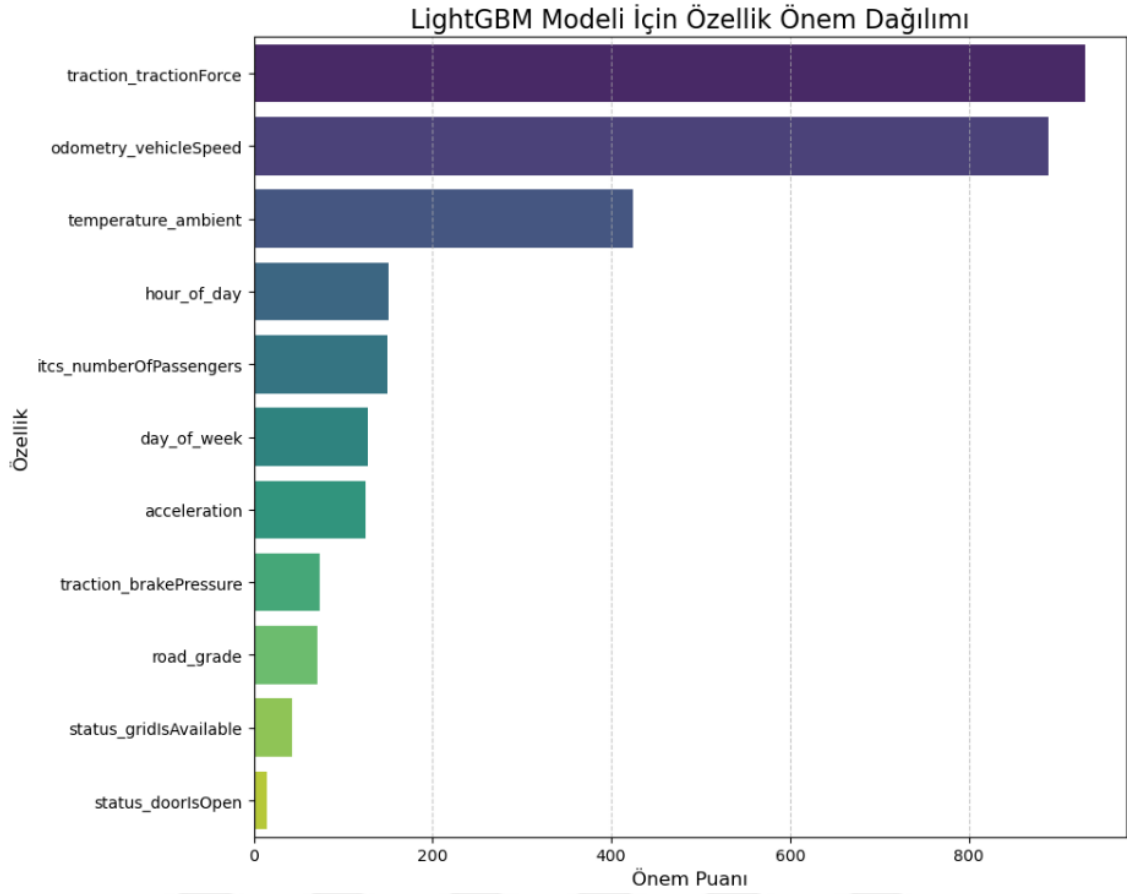
Bir modelin sadece ne kadar doğru tahmin yaptığını bilmek yeterli değildir; aynı zamanda bu tahminlere nasıl ulaştığını ve hangi faktörlerin kararlarında daha etkili

olduğunu anlamak da kritik öneme sahiptir. Bu amaçla, eğitilmiş nihai LightGBM modelinin özellik önem analizi gerçekleştirilmiştir. Bu analiz, her bir girdi özelliğinin modelin tahmin doğruluğuna olan toplam katkısını ölçer. LightGBM, bu önemi bir özelliğin karar ağaçları boyunca ne sıklıkta bir bölünme (split) noktası olarak kullanıldığına göre hesaplar. Bir özelliğin önem puanının yüksek olması, modelin o özelliği enerji tüketimindeki değişimleri ayırt etmek için daha sık ve etkili bir şekilde kullandığı anlamına gelir.

Bu analizin sonuçları Şekil 4.1 ve Çizelge 4.2'te özetlenmiştir

Çizelge 4.2. Nihai lightgbm modeli için özellik önem puanları

Özellik	Önemi
<i>traction tractionForce</i>	930
<i>odometry vehicleSpeed</i>	889
<i>temperature ambient</i>	424
<i>hour of day</i>	151
<i>itcs numberOfPassengers</i>	150
<i>day_of_week</i>	128
<i>acceleration</i>	125
<i>traction brakePressure</i>	74
<i>road_grade</i>	71
<i>status_gridIsAvailable</i>	43
<i>status_doorIsOpen</i>	15



Şekil 4.1. Nihai lightgbm modeli için özellik önem dağılımı

Analiz sonuçları, özellikleri üç ana etki katmanına ayırmaktadır. Bunlar;

- Birincil Derecede Önemli Özellikler,
 - İkincil Derecede Önemli Özellikler,
 - Üçüncül Derecede Önemli Özellikler.
-
- **Birincil Derecede Önemli Özellikler:** *traction_tractionForce* (çekiş kuvveti) ve *odometry_vehicleSpeed* (araç hızı), beklendiği gibi modelin karar verme sürecinde en baskın iki özelliktir. Bu iki değişken, aracın anlık mekanik gücünün temel bileşenleridir ve modelin, bu temel fiziksel ilişkiyi enerji tüketimini tahmin etmek için birincil dayanak olarak kullandığını göstermektedir.
 - **İkincil Derecede Önemli Özellikler:** *temperature_ambient* (ortam sıcaklığı), *hour_of_day* (günün saati), *itcs_numberOfPassengers* (yolcu sayısı) ve *day_of_week* (haftanın günü) gibi değişkenler, ikinci etki katmanını oluşturur. *temperature_ambient* özelliğinin yüksek önemi, HVAC sistemlerinin enerji tüketimi üzerindeki dolaylı ancak

güçlü etkisini teyit etmektedir. *hour_of_day* ve *day_of_week* ise trafik yoğunluğu gibi operasyonel kalıpları yakalayarak modele önemli bağlamsal bilgiler sunar.

- **Üçüncül Derecede Önemli Özellikler:** *acceleration* (ivme), *traction_brakePressure* (fren basıncı) ve *road_grade* (yol eğimi) gibi fiziksel olarak anlamlı değişkenlerin bu katmanda yer alması dikkat çekicidir. Bu durum, bu değişkenlerin önemsiz olduğu anlamına gelmez; aksine, bu değişkenlerin içerdiği bilginin büyük bir kısmının, daha baskın olan *traction_tractionForce* ve *odometry_vehicleSpeed* tarafından dolaylı olarak zaten yakalandığını gösterir. Örneğin, yüksek ivmelenme veya yokuş tırmanma, doğrudan çekiş kuvvetinde bir artışa neden olur ve model bu birincil sinyali kullanmayı tercih eder.

Bu analiz, modelin sadece bir "kara kutu" olmadığını, aksine tahminlerini aracın temel fiziksel ve operasyonel dinamiklerine dayandığını göstermektedir. Bu bulgular, elektrikli ulaşım sistemlerinde enerji tüketimini etkileyen faktörlerin hiyerarşisini anlamak için değerli bilgiler sunmaktadır.

4.4. Nihai Modelin Bağımsız Veri Seti (Holdout) Üzerinde Test Edilmesi

Önceki bölümlerde, çeşitli makine öğrenmesi modelleri karşılaştırmalı olarak değerlendirilmiş ve aralarından en yüksek performansı gösteren nihai model seçilerek tüm ana veri seti üzerinde eğitilmiştir. Ancak, bir modelin gerçek dünyadaki genelleme kabiliyetinin en güvenilir ölçüsü, modelin geliştirme sürecinin hiçbir aşamasında karşılaşmadığı, tamamen bağımsız bir veri seti üzerindeki performansdır. Bu bölümün temel amacı, eğitilmiş nihai modelin bu bağımsız test (holdout) veri seti üzerindeki performansını yansız bir şekilde değerlendirerek, modelin gerçek operasyonel koşullar altındaki başarısını ortaya koymaktır.

4.4.1. Anlık güç talebi tahmin performansı

Nihai LightGBM modelinin, model geliştirme sürecinde hiç karşılaşmadığı bağımsız test (holdout) veri seti üzerindeki performansı bu bölümde analiz edilmektedir. Bu son değerlendirme, modelin gerçek dünya koşullarındaki genelleme kabiliyetini ve güvenilirliğini ölçmek için en önemli adımdır. Modelin anlık güç talebi tahmin performansı, daha önce tanımlanan R^2 , MAE ve RMSE metrikleri kullanılarak ölçülmüş ve sonuçlar Çizelge 4.3'te özetlenmiştir.

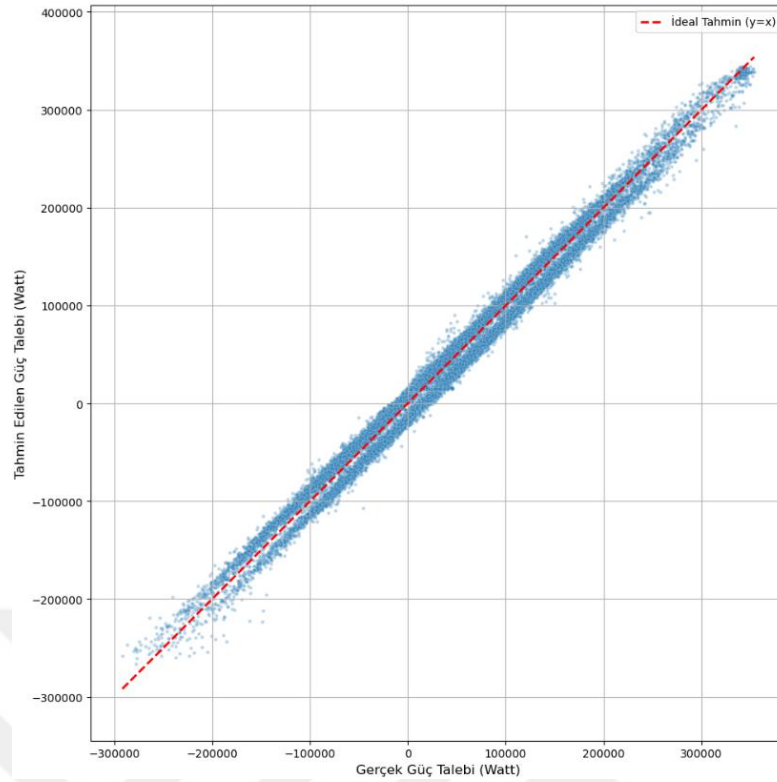
Çizelge 4.3. Nihai modelin bağımsız test (holdout) veri seti üzerindeki performans metrikleri

	METRİK	DEĞER
	R^2	0,9759
MAE (ORTALAMA MUTLAK HATA)		11456,06 Watt
RMSE (HATALARIN ORTALAMA KAREKÖKÜ)		12360,03 Watt

Elde edilen 0,9759'luk R^2 skoru, modelin holdout veri setindeki varyansın yaklaşık %97,6'sını başarıyla açıklayabildiğini göstermektedir. Bu sonuç, çapraz doğrulama aşamasında elde edilen ortalama R^2 skoruyla (0,9771) oldukça tutarlıdır ve modelin aşırı öğrenmeye (overfitting) maruz kalmadığını, aksine verideki temel kalıpları başarıyla öğrenip yeni verilere genelleme yapabildiğini güçlü bir şekilde teyit etmektedir.

Hata metrikleri incelendiğinde, modelin Ortalama Mutlak Hatasının (MAE) yaklaşık 11,5 kW olduğu görülmektedir. Bu değer, modelin anlık güç talebi tahminlerinin, gerçek değerlerden ortalama olarak ne kadar saptığını ifade eder. RMSE değerinin (yaklaşık 12,4 kW) MAE'den biraz daha yüksek olması, modelin bazı anlarda (örneğin ani hızlanma gibi dinamiklerin yüksek olduğu anlarda) daha büyük hatalar yaptığını, ancak bu hataların genel performansı önemli ölçüde düşürmediğini göstermektedir. Bu performans seviyesi, aracın anlık güç talebinin -300 kW ile +300 kW arasında değişebildiği geniş bir operasyonel aralık göz önüne alındığında, modelin pratikteki güvenilirliğinin ve uygulanabilirliğinin altını çizmektedir.

Modelin tahmin doğruluğunu görsel olarak incelemek amacıyla, holdout veri setindeki gerçek değerler ile modelin tahmin ettiği değerler arasındaki ilişki bir saçılım grafiği üzerinde Şekil 4.2'da gösterilmiştir. Noktaların, ideal tahmini temsil eden $y=x$ doğrusu etrafında yoğun bir şekilde toplanması, modelin yüksek performansını görsel olarak da doğrulamaktadır.



Şekil 4.2. Holdout veri seti için gerçek ve tahmin edilen değerlerin karşılaştırılması

4.4.2. Toplam görev enerjisi tüketimi tahmin sonuçları

Modelin anlık güç talebi tahminindeki yüksek doğruluğunun yanı sıra, bu çalışmanın temel hedeflerinden biri de bu anlık tahminlerin, bütün bir görev (sefer) boyunca toplam enerji tüketimini ne kadar başarılı bir şekilde yansıttığını ölçmektir. Bu analiz, modelin pratik uygulanabilirliğini ve operasyonel planlama (örneğin şarj stratejileri) gibi alanlardaki potansiyelini değerlendirmek için kritik öneme sahiptir.

Bu amaçla, holdout veri setini oluşturan görevin hem gerçek hem de model tarafından tahmin edilen anlık güç talebi değerleri, tüm görev süresi boyunca entegre edilerek (toplanarak) toplam enerji tüketimi hesaplanmıştır. Elde edilen sonuçlar Çizelge 4.4'te karşılaştırmalı olarak sunulmuştur.

Çizelge 4.4. Nihai Modelin Holdout Görevi Üzerindeki Toplam Enerji Tüketimi Tahmin Performansı

Metrik	Değer
Gerçek Tüketilen Enerji	629,22 kWh
Tahmin Edilen Enerji	629,71 kWh
Nihai Tahmin Hatası (%)	0,08 %

Sonuçlar, geliştirilen nihai LightGBM modelinin, daha önce hiç görmediği bir görevin toplam enerji tüketimini %0,08 gibi son derece düşük bir hata payıyla tahmin

edebildiğini göstermektedir. Bu, modelin sadece anlık dinamikleri değil, aynı zamanda bu dinamiklerin birikimli etkisini de büyük bir doğrulukla yakalayabildiğini kanıtlamaktadır. Bu bulgu, modelin elektrikli otobüs filolarının operasyonel verimliliğini artırmak için güçlü bir araç olarak kullanılabileceğine dair somut bir kanıt sunmaktadır.

4.5. Tartışma

Bu bölümde, araştırma kapsamında elde edilen ampirik bulgular, tezin başlangıcında sunulan literatür çerçevesinde etraflıca yorumlanmakta; çalışmanın bilime ve uygulamaya olan katkıları, metodolojik üstünlükleri, sınırlılıkları ve gelecek araştırmalar için potansiyel yönleri kapsamlı bir şekilde tartışılmaktadır.

4.5.1. Bulguların genel değerlendirmesi ve literatürdeki yeri

Bu çalışmada geliştirilen nihai LightGBM modelinin, bağımsız test (holdout) veri seti üzerinde %97,59'luk bir R^2 skoru elde etmesi ve bir tam görevin toplam enerji tüketimini %0,08 gibi son derece düşük bir hata payıyla tahmin etmesi, veri-odaklı yaklaşımların elektrikli otobüslerin enerji tüketim dinamiklerini modellemedeki yüksek potansiyelini güçlü bir şekilde ortaya koymaktadır. Elde edilen bu yüksek doğruluk oranı, literatürdeki benzer çalışmalarla karşılaştırıldığında oldukça rekabetçi ve dikkat çekici bir sonuçtur.

Literatürdeki modelleme yaklaşımları genel olarak fizik-temelli, veri-odaklı ve hibrit olarak üç ana kategoriye ayrılmaktadır. Fiori ve ark. (2016) gibi fizik-temelli (beyaz kutu) modeller, aracın temel dinamiklerini Newton mekaniği ve termodinamik yasalarıyla başarılı bir şekilde formüle etseler de, bu modeller genellikle sürüş davranışı, anlık trafik yoğunluğu ve anlık çevresel koşullar gibi stokastik ve dinamik faktörleri yakalamada yetersiz kalmaktadır. Bu çalışmanın benimsediği veri-odaklı yaklaşım ise, özellikle kullanılan yüksek frekanslı (1 Hz) veriler sayesinde, bu anlık ve karmaşık kalıpları daha detaylı öğrenme kapasitesine sahip olduğunu kanıtlamıştır. Bu durum, özellikle Wu ve ark. (2023) tarafından da belirtildiği gibi, atalet kuvvetlerinin (ivmelenme ve frenleme) anlık etkilerinin doğru bir şekilde modellenebilmesi için kritik öneme sahiptir.

4.5.2. Model performansının ve algoritmik seçimin literatürle karşılaştırılması

Bu çalışmada, Gradyan Artırma ailesine ait modellerin temel doğrusal regresyon modeline kıyasla belirgin ölçüde üstün performans göstermesi, literatürdeki genel eğilimle tam bir tutarlılık göstermektedir. EV'lerin enerji tüketiminin karmaşık ve

doğrusal olmayan doğası, De Cauwer ve ark. (2017) tarafından kullanılan Yapay Sinir Ağları (ANN) ve Zhang ve ark. (2020) tarafından kullanılan Destek Vektör Regresyonu (SVR) gibi modellerin başarısıyla da kanıtlanmıştır.

Ancak bu çalışmada, özellikle 5,5 milyondan fazla veri satırı içeren büyük bir veri seti üzerinde, LightGBM gibi ağaç tabanlı topluluk modellerinin, SVR gibi çekirdek (kernel) tabanlı yöntemlere kıyasla hesaplama verimliliği açısından çok daha ölçeklenebilir olduğu görülmüştür. Başlangıçta SVR ve Random Forest (RF) modelleri de değerlendirme için düşünülmüş, ancak yüksek hesaplama maliyetleri ve bellek gereksinimleri nedeniyle pratik olarak uygulanamamıştır. Bu durum, büyük veri setleri için algoritma seçiminde verimliliğin en az doğruluk kadar önemli bir kriter olduğunu göstermektedir. Benzer şekilde, ANN modelleri de yüksek doğruluk potansiyeline sahip olmalarına rağmen, genellikle daha fazla hiperparametre ayarı gerektirir ve Gradyan Artırma modellerine göre daha "kara kutu" bir yapıya sahiptirler. Bu çalışmada elde edilen %97'nin üzerindeki R^2 skoru, Gradyan Artırma modellerinin, doğru özellik mühendisliği ile birleştirildiğinde, daha karmaşık ve ayarlanması zor olan derin öğrenme modelleriyle rekabet edebilecek bir "altın standart" oluşturduğunu ortaya koymaktadır.

4.5.3. Metodolojik yaklaşımın ve özellik öneminin derinlemesine analizi

Bu çalışmanın literatüre en önemli katkılarından biri, metodolojik yaklaşımının sağlamlığında ve yüksek çözünürlüklü verilerin getirdiği yeni anlayışlarda yatmaktadır.

- **Veri Frekansının Etkisi:** Yüksek frekanslı (1 Hz) ZTBus veri setinin kullanılması, enerji tüketimini etkileyen faktörlerin hiyerarşisini, literatürdeki düşük frekanslı veri kullanan çalışmalardan farklı bir şekilde ortaya çıkarmıştır. Örneğin, Wang ve ark. (2023) düşük frekanslı (0,1 Hz) verilerle yaptığı çalışmada, *road_grade* (yol eğimi) gibi dolaylı değişkenlerin önemi vurgulanırken, bu çalışmadaki özellik önem analizi, en önemli değişkenlerin *traction_tractionForce* (çekiş kuvveti) ve *odometry_vehicleSpeed* (araç hızı) gibi doğrudan araç mekaniğiyle ilişkili sinyaller olduğunu göstermiştir. Bu bulgu, yüksek çözünürlüklü CAN verilerinin, enerji tüketimini etkileyen birçok fiziksel faktörü (yokuş, yolcu yükü vb.) tek bir güçlü sinyalde birleştirerek modelleme sürecini basitleştirme ve doğruluğu artırma potansiyeline sahip olduğunu göstermektedir. Bu, literatürde sıklıkla ayrı ayrı modellenen yuvarlanma direnci ve aerodinamik sürüklenme gibi etkilerin (Wu vd., 2023), *traction_tractionForce* içinde dolaylı olarak zaten mevcut olduğu anlamına gelmektedir.

- **Değerlendirme Metodolojisinin Gücü:** Model performansının değerlendirilmesinde TimeSeriesSplit ile çapraz doğrulama ve bağımsız bir holdout veri setinin bir arada kullanılması, bu çalışmanın metodolojik gücünü önemli ölçüde artırmaktadır. Literatürdeki birçok çalışmanın (örneğin, (Yufang vd., 2020)) basit bir eğitim-test ayırımına dayanmasının aksine, bu iki aşamalı doğrulama süreci, modelin sadece eğitim verisini ezberlemediğini (overfitting), aynı zamanda daha önce hiç görmediği veriler üzerinde de tutarlı ve güvenilir bir performans sergilediğini bilimsel olarak kanıtlamaktadır.



5. SONUÇLAR VE ÖNERİLER

Bu bölümde, araştırma süresince yürütülen analizler sonucunda elde edilen temel bulgular özetlenecek, çalışmanın sınırlılıkları tartışılacak ve gelecek araştırmalar için potansiyel yönler sunulacaktır.

5.1. Sonuçlar

Bu tez çalışmasında, yüksek frekanslı operasyonel veriler kullanılarak elektrikli bir otobüsün anlık ve toplam enerji tüketimini tahmin etmek amacıyla kapsamlı bir makine öğrenmesi metodolojisi geliştirilmiş ve uygulanmıştır. Yürütülen analizler sonucunda elde edilen temel bulgular aşağıda özetlenmiştir. Bunlar;

- Model Performansı ve Algoritma Seçimi,
 - Özellik Önem Analizi,
 - Nihai Modelin Genelleme Performansı,
 - Toplam Enerji Tüketimi Tahmini.
- **Model Performansı ve Algoritma Seçimi:** Dört farklı makine öğrenmesi algoritması (Linear Regression, XGBoost, LightGBM, HistGradientBoosting) karşılaştırmalı olarak değerlendirilmiştir. Gradyan Artırma (Gradient Boosting) tabanlı modellerin, temel doğrusal regresyon modeline kıyasla belirgin ölçüde daha üstün bir performans sergilediği tespit edilmiştir. Bu modeller arasında LightGBM, hem en yüksek tahmin doğruluğuna (ortalama $R^2 = 0,9771$) hem de en düşük hata oranlarına (ortalama MAE = 7415,63 Watt) ulaşarak en başarılı model olarak seçilmiştir.
 - **Özellik Önem Analizi:** Nihai LightGBM modeli üzerinde yapılan özellik önem analizi, enerji tüketimini etkileyen faktörlerin hiyerarşisini ortaya koymuştur. *traction_tractionForce* (çekiş kuvveti) ve *odometry_vehicleSpeed* (araç hızı) gibi doğrudan araç mekaniğiyle ilişkili değişkenlerin, modelin tahmin gücünün büyük bir kısmını oluşturduğu belirlenmiştir. Bu bulgu, yüksek çözünürlüklü CAN verilerinin, enerji tüketimini etkileyen birçok fiziksel faktörü tek bir güçlü sinyalde birleştirme potansiyelini göstermektedir.
 - **Nihai Modelin Genelleme Performansı:** Modelin gerçek dünya performansını ölçmek amacıyla, eğitim sürecinde hiç karşılaşmadığı bağımsız bir test (holdout) veri seti üzerinde yapılan nihai değerlendirme, modelin yüksek genelleme kabiliyetini

doğrulamıştır. Model, anlık güç talebini %97,59'luk bir R^2 skoru ile tahmin etmeyi başarmıştır.

- **Toplam Enerji Tüketimi Tahmini:** Çalışmanın en önemli pratik çıktısı olarak, nihai modelin bir tam görevin toplam enerji tüketimini %0,08 gibi son derece düşük bir hata payıyla tahmin edebildiği kanıtlanmıştır. Bu sonuç, geliştirilen modelin, filo yönetimi, şarj optimizasyonu ve güzergâh planlaması gibi operasyonel karar destek sistemleri için güvenilir bir temel oluşturma potansiyelini güçlü bir şekilde ortaya koymaktadır.

Bu sonuçlar, veri-odaklı yaklaşımların ve modern makine öğrenmesi algoritmalarının, elektrikli ulaşım sistemlerinin verimliliğini artırmada ne denli etkili olabileceğini göstermektedir.

5.2. Araştırmanın Kısıtları

Bu çalışmada yüksek doğruluk oranlarına sahip bir makine öğrenmesi modeli geliştirilmiş olmasına rağmen, elde edilen sonuçların yorumlanmasında dikkate alınması gereken bazı metodolojik ve veri kaynaklı sınırlılıklar mevcuttur. Bu sınırlılıkların farkında olmak, hem bu çalışmanın kapsamını doğru bir şekilde anlamak hem de gelecek araştırmalar için yön belirlemek açısından önemlidir. Bunlar;

- Modelin Genellenebilirliği ve Dış Geçerliliği,
 - Veri Setine Bağımlılık ve Sinyal Ayırıştırma Sorunu,
 - Hiperparametre Optimizasyonunun Kapsamı,
 - Statik Model Yapısı.
-
- **Modelin Genellenebilirliği ve Dış Geçerliliği:** Bu çalışmanın en temel sınırlılığı, geliştirilen modelin coğrafi ve operasyonel bağlama özgü olmasıdır. Model, Zürih şehrinin özgün topografik yapısı, trafik desenleri ve iklim koşulları altında, spesifik bir otobüs modeli (HESS lighTram® 19 DC) üzerinden toplanan verilerle eğitilmiştir. Bu nedenle, modelin farklı coğrafi konumlarda, iklim koşullarında veya araç modelleriyle yeniden kalibrasyon yapılmaksızın aynı performansı göstermesi beklenmemelidir. Bu durum, modelin dış geçerliliği konusunda bir sınırlama teşkil etmektedir.
 - **Veri Setine Bağımlılık ve Sinyal Ayırıştırma Sorunu:** Modelin performansı, ZTBus veri setinin sunduğu yüksek çözünürlüklü verilerinin mevcudiyetine ve kalitesine

doğrudan bağlıdır. Önemli bir sınırlılık, veri setinin yardımcı sistemlerin (örneğin HVAC) enerji tüketimini, ana güç talebinden (*electric_powerDemand*) ayrı bir değişken olarak sunmamasıdır. Bu durum, HVAC sistemlerinin enerji tüketimi üzerindeki etkisinin doğrudan ve izole bir şekilde analiz edilmesini engellemiştir. Her ne kadar *temperature_ambient* değişkeni bu etkiyi dolaylı olarak modellemeye yardımcı olsa da, HVAC tüketiminin ayrı bir sinyal olarak mevcut olmaması bu önemli enerji tüketicisinin rolünü daha derinlemesine anlama fırsatını kısıtlamıştır.

- **Hiperparametre Optimizasyonunun Kapsamı:** Bu çalışmada, farklı algoritma ailelerinin temel performansını adil bir şekilde karşılaştırmak amacıyla, modellerin literatürde kabul görmüş varsayılan hiperparametreleri ile değerlendirilmesi tercih edilmiştir. Bu yaklaşım, algoritmaların temel potansiyelini ortaya koysa da, LightGBM gibi karmaşık modellerin hiperparametrelerinin sistematik bir optimizasyon sürecinden geçirilmemesi, modelin ulaşabileceği maksimum potansiyel performansın kilidini açmamış olabileceği anlamına gelir.
- **Statik Model Yapısı:** Geliştirilen nihai model, mevcut veri seti üzerinde bir kez eğitilmiş statik bir modeldir. Ancak, gerçek dünya koşulları dinamiktir; araçların zamanla yaşlanması, batarya sağlığının değişmesi veya trafik desenlerinin değişmesi gibi durumlar, modelin performansının zamanla düşmesine neden olabilir. Bu çalışma, "çevrimiçi model uyarlaması" olarak bilinen dinamik güncelleme mekanizmalarını içermemektedir.

5.3. Gelecek Çalışmalar İçin Öneriler

Bu tez çalışması, yüksek frekanslı veriler kullanarak elektrikli otobüslerin enerji tüketimini yüksek bir doğrulukla tahmin etmek için sağlam bir metodolojik temel oluşturmuştur. Elde edilen bulgular ve tespit edilen sınırlılıklar ışığında, gelecekteki araştırmalar için odaklanılabilecek bazı önemli alanlar aşağıda sunulmaktadır. Bunlar;

- Transfer Öğrenimi,
- Hiperparametre Optimizasyonu,
- Düşük Çözünürlüklü Verilerle Modelleme,
- Gerçek Zamanlı Karar Destek Sistemlerinin Geliştirilmesi,
- Çevrimiçi Model Uyarlaması.

- **Transfer Öğrenimi (domain adaptation):** Bu çalışmanın en temel sınırlılığı, modelin tek bir şehir (Zürih) ve tek bir otobüs modeline özgü olmasıdır. Gelecek çalışmalarda, modelin genellenebilirliğini artırmak için transfer öğrenimi tekniklerinden yararlanılabilir. Bu yaklaşımla, Zürich verileri üzerinde eğitilmiş olan bu modelin temel bilgileri, daha az veriye sahip farklı bir şehrin veri seti üzerinde ince ayar (fine-tuning) yapılarak yeni koşullara uyarlanabilir. Bu yöntem, yeni bir şehre özel, yüksek doğruluklu bir model geliştirmek için gereken veri ve zaman miktarını önemli ölçüde azaltma potansiyeline sahiptir.
- **Hiperparametre Optimizasyonu:** Bu çalışmada modeller, farklı algoritma ailelerinin temel performansını adil bir şekilde karşılaştırmak amacıyla varsayılan hiperparametrelerle değerlendirilmiştir. Gelecekte, GridSearchCV veya daha gelişmiş **Bayesyen Optimizasyon** gibi yöntemlerle LightGBM modelinin hiperparametreleri (ör. ağaç sayısı, öğrenme oranı, ağaç derinliği vb.) sistematik olarak optimize edilebilir. Bu süreç, modelin mevcut potansiyelinin en üst sınırına ulaşmasını sağlayarak, tahmin doğruluğunu daha da artırma imkânı sunar.
- **Düşük Çözünürlüklü Verilerle Modelleme:** Bu çalışmada kullanılan 1 Hz'lik yüksek frekanslı veriler, her filo operatörü için mevcut olmayabilir. Pratik uygulanabilirlik açısından kritik bir araştırma yönü, modelin daha düşük çözünürlüklü (örneğin 10 saniyede bir veya dakikada bir kaydedilen) verilerle ne kadar başarılı bir şekilde eğitilebileceğini araştırmaktır. Bu, geliştirilen modellerin daha geniş bir endüstriyel uygulama alanına sahip olmasını sağlayacaktır.
- **Gerçek Zamanlı Karar Destek Sistemlerinin Geliştirilmesi:** Bu tezin kapsamı öngörü modelini oluşturmaya odaklansa da, bu modelin pratik değeri, filo yönetimi ve karar destek sistemlerine entegre edildiğinde tam olarak ortaya çıkacaktır. Tahmin edilen anlık ve toplam enerji tüketimi, aşağıdaki gibi pratik uygulamalar için temel bir girdi olarak kullanılabilir:
 - **Dinamik Güzergâh Optimizasyonu:** Anlık trafik ve hava koşullarına göre en verimli rotanın belirlenmesi.
 - **Optimize Edilmiş Şarj Programları:** Aracın kalan menzilini ve bir sonraki şarj istasyonuna ulaşmak için gereken enerjiyi daha doğru tahmin ederek "menzil endişesini" azaltmak.

- **Eko-Sürüş (Eco-Driving) Geri Bildirim Sistemleri:** Sürücülere anlık olarak enerji verimliliklerini artıracak geri bildirimler sunmak.
- **Çevrimiçi Model Uyarlaması:** Geliştirilen statik model, zamanla araçların yaşlanması veya operasyonel koşulların değişmesi gibi nedenlerle performansını yitirebilir. Literatürde Zhu ve ark. (2025) tarafından önerilen çevrimiçi model uyarlaması gibi kavramlar bu soruna bir çözüm sunmaktadır. Gelecek çalışmalarda, bu modelin periyodik olarak yeni verilerle güncellendiği veya anlık hatalara göre kendisini dinamik olarak uyarladığı bir çevrimiçi öğrenme (online learning) mimarisi geliştirilebilir.



6. KAYNAKLAR

- Abdelaty, H., Al-Obaidi, A., Mohamed, M., ve Farag, H. E. Z. (2021). Machine learning prediction models for battery-electric bus energy consumption in transit. *Transportation Research Part D: Transport and Environment*, 96, 102868. <https://doi.org/10.1016/j.trd.2021.102868>
- Abdelaty, H., ve Mohamed, M. (2021). A Prediction Model for Battery Electric Bus Energy Consumption in Transit. *Energies*, 14(10), Article 10. <https://doi.org/10.3390/en14102824>
- Abdelaty, H., ve Mohamed, M. (2022). A framework for BEB energy prediction using low-resolution open-source data-driven model. *Transportation Research Part D: Transport and Environment*, 103, 103170. <https://doi.org/10.1016/j.trd.2022.103170>
- Acharyaviriya, W., Wongsapai, W., Janpoom, K., Katongtung, T., Mona, Y., Tippayawong, N., ve Suttakul, P. (2023). Estimating Energy Consumption of Battery Electric Vehicles Using Vehicle Sensor Data and Machine Learning Approaches. *Energies*, 16(17), Article 17. <https://doi.org/10.3390/en16176351>
- Ayman, A., Wilbur, M., Sivagnanam, A., Pugliese, P., Dubey, A., ve Laszka, A. (2020). *Data-Driven Prediction of Route-Level Energy Use for Mixed-Vehicle Transit Fleets* (No. arXiv:2004.06043). arXiv. <https://doi.org/10.48550/arXiv.2004.06043>
- Bie, Y., Ji, J., Wang, X., ve Qu, X. (2021). Optimization of electric bus scheduling considering stochastic volatilities in trip travel time and energy consumption. *Computer-Aided Civil and Infrastructure Engineering*, 36(12), 1530-1548. <https://doi.org/10.1111/mice.12684>
- Chen, J., Li, B., Zhang, Z., Liu, M., Piao, C., ve Huang, M. (2024). *Energy Consumption Estimation of Electric Vehicles for Future Driving Process Based*

- on Driving Conditions Recognition* (SSRN Scholarly Paper No. 4931664).
Social Science Research Network. <https://papers.ssrn.com/abstract=4931664>
- Chen, X., Lei, Z., ve Ukkusuri, S. V. (2022). Prediction of Road-level Energy Consumption of Battery Electric Vehicles. *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2550-2555.
<https://doi.org/10.1109/ITSC55140.2022.9921975>
- Chen, Y., Zhang, Y., ve Sun, R. (2021). Data-driven estimation of energy consumption for electric bus under real-world driving conditions. *Transportation Research Part D: Transport and Environment*, 98, 102969.
<https://doi.org/10.1016/j.trd.2021.102969>
- De Cauwer, C., Verbeke, W., Coosemans, T., Faid, S., ve Van Mierlo, J. (2017). A Data-Driven Method for Energy Consumption Prediction and Energy-Efficient Routing of Electric Vehicles in Real-World Conditions. *Energies*, 10(5), Article 5. <https://doi.org/10.3390/en10050608>
- Fiori, C., Ahn, K., ve Rakha, H. A. (2016). Power-based electric vehicle energy consumption model: Model development and validation. *Applied Energy*, 168, 257-268. <https://doi.org/10.1016/j.apenergy.2016.01.097>
- Foiadelli, F., Longo, M., ve Miraftebzadeh, S. (2018). Energy Consumption Prediction of Electric Vehicles Based on Big Data Approach. *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, 1-6. <https://doi.org/10.1109/EEEIC.2018.8494573>
- Guo, J., Jiang, Y., Yu, Y., ve Liu, W. (2020). A novel energy consumption prediction model with combination of road information and driving style of BEVs.

Sustainable Energy Technologies and Assessments, 42, 100826.

<https://doi.org/10.1016/j.seta.2020.100826>

Huang, Y.-W., Prehofer, C., Lindskog, W., Puts, R., Mosca, P., ve Kauermann, G.

(2023). Predictive Energy Management for Battery Electric Vehicles with

Hybrid Models. İçinde A. L. Martins, J. C. Ferreira, A. Kocian, & U.

Tokkozhina (Ed.), *Intelligent Transport Systems* (ss. 182-196). Springer Nature

Switzerland. https://doi.org/10.1007/978-3-031-30855-0_13

Jiang, J., Yu, Y., Min, H., Cao, Q., Sun, W., Zhang, Z., ve Luo, C. (2023). Trip-level

energy consumption prediction model for electric bus combining Markov-based

speed profile generation and Gaussian processing regression. *Energy*, 263,

125866. <https://doi.org/10.1016/j.energy.2022.125866>

Karadag, T., Ekici, Y. E., Akdağ, O., ve Aydin, A. A. (2023). *Electric City Buses*

Enhanced Energy Consumption Model Using Real-Time Big-Data (SSRN

Scholarly Paper No. 4613853). Social Science Research Network.

<https://doi.org/10.2139/ssrn.4613853>

Kaya, İ., Uzun, E., Ergun, B., Akyıldız, A., Çakıroğlu, Y., Zehir, A., Küçüktezcan, C.

F., Kocaarslan, İ., ve Gülbahçe, M. O. (2024). Development of a Real-Time

Data-Supported Energy Consumption Model for Electric Buses. *2024 15th*

National Conference on Electrical and Electronics Engineering (ELECO), 1-5.

<https://doi.org/10.1109/ELECO64362.2024.10847236>

Li, J., Li, Z., Zhuang, W., He, Z., ve Fu, J. (2022). Data-driven Energy Consumption

Prediction of Electric Bus on a Given Route. *2022 IEEE 5th International*

Electrical and Energy Conference (CIEEC), 1698-1703.

<https://doi.org/10.1109/CIEEC54735.2022.9846407>

- Li, P., Zhang, Y., Zhang, Y., Zhang, Y., ve Zhang, K. (2021). Prediction of electric bus energy consumption with stochastic speed profile generation modelling and data driven method based on real-world big data. *Applied Energy*, 298, 117204. <https://doi.org/10.1016/j.apenergy.2021.117204>
- Li, X., Wang, T., Li, J., Tian, Y., ve Tian, J. (2022). Energy Consumption Estimation for Electric Buses Based on a Physical and Data-Driven Fusion Model. *Energies*, 15(11). <https://doi.org/10.3390/en15114160>
- Li, Y., Jiang, Y., Guo, J., ve Xie, D. (2024). A Novel Energy Consumption Prediction Model Integrating Real-Time Traffic State Recognition and Velocity Prediction of BEVs. *IEEE Access*, 12, 196178-196194. <https://doi.org/10.1109/ACCESS.2024.3509456>
- Lin, M., Chen, S., Wang, W., ve Wu, J. (2023). Multi-Feature Fusion-Based Instantaneous Energy Consumption Estimation for Electric Buses. *IEEE/CAA Journal of Automatica Sinica*, 10(10), 2035-2037. <https://doi.org/10.1109/JAS.2022.106010>
- Liu, K., Wang, J., Yamamoto, T., ve Morikawa, T. (2018). Exploring the interactive effects of ambient temperature and vehicle auxiliary loads on electric vehicle energy consumption. *Applied Energy*, 227, 324-331. <https://doi.org/10.1016/j.apenergy.2017.08.074>
- Liu, Y., ve Liang, H. (2022). A Data-Driven Approach for Electric Bus Energy Consumption Estimation. *IEEE Transactions on Intelligent Transportation Systems*, 23(10), 17027-17038. <https://doi.org/10.1109/TITS.2022.3152679>
- Liu, Y., Zhang, Q., Lyu, C., ve Liu, Z. (2021). Modelling the energy consumption of electric vehicles under uncertain and small data conditions. *Transportation*

Research Part A: Policy and Practice, 154, 313-328.

<https://doi.org/10.1016/j.tra.2021.10.009>

Maity, A., ve Sarkar, S. (2023). *Data-Driven Probabilistic Energy Consumption Estimation for Battery Electric Vehicles with Model Uncertainty*. arXiv.Org.
<https://arxiv.org/abs/2307.00469v1>

Miri, I., Fotouhi, A., ve Ewin, N. (2021). Electric vehicle energy consumption modelling and estimation—A case study. *International Journal of Energy Research*, 45(1), 501-520. <https://doi.org/10.1002/er.5700>

Modi, S., ve Bhattacharya, J. (2022). A system for electric vehicle's energy-aware routing in a transportation network through real-time prediction of energy consumption. *Complex & Intelligent Systems*, 8(6), 4727-4751.
<https://doi.org/10.1007/s40747-022-00727-4>

Modi, S., Bhattacharya, J., ve Basak, P. (2020). Estimation of energy consumption of electric vehicles using Deep Convolutional Neural Network to reduce driver's range anxiety. *ISA Transactions*, 98, 454-470.
<https://doi.org/10.1016/j.isatra.2019.08.055>

Pamuła, T., ve Pamuła, D. (2022). Prediction of Electric Buses Energy Consumption from Trip Parameters Using Deep Learning. *Energies*, 15(5).
<https://doi.org/10.3390/en15051747>

Petersen, P., Rudolf, T., ve Sax, E. (2022). A Data-driven Energy Estimation based on the Mixture of Experts Method for Battery Electric Vehicles. *Proceedings of the 8th International Conference on Vehicle Technology and Intelligent Transport Systems - VEHITS*, 384-390. <https://doi.org/10.5220/0011081000003191>

Pokharel, S., Sah, P., ve Ganta, D. (2021). Improved Prediction of Total Energy Consumption and Feature Analysis in Electric Vehicles Using Machine

- Learning and Shapley Additive Explanations Method. *World Electric Vehicle Journal*, 12(3). <https://doi.org/10.3390/wevj12030094>
- Qin, W., Wang, L., Liu, Y., ve Xu, C. (2021). Energy Consumption Estimation of the Electric Bus Based on Grey Wolf Optimization Algorithm and Support Vector Machine Regression. *Sustainability*, 13(9). <https://doi.org/10.3390/su13094689>
- Shen, H., Wang, Z., Zhou, X., Lamantia, M., Yang, K., Chen, P., ve Wang, J. (2022). Electric Vehicle Velocity and Energy Consumption Predictions Using Transformer and Markov-Chain Monte Carlo. *IEEE Transactions on Transportation Electrification*, 8(3), 3836-3847. <https://doi.org/10.1109/TTE.2022.3157652>
- Shen, H., Zhou, X., Yu, A., Lamantia, M., Chen, P., ve Wang, J. (2023). An Experimental Comparison of Physics-based and Machine-Learning-based Electric Vehicle Energy Consumption Estimation Methods. *IFAC-PapersOnLine*, 56(2), 8672-8677. <https://doi.org/10.1016/j.ifacol.2023.10.045>
- Wang, C., Wang, Y., Tan, J., Liu, F., Jiang, Y., ve Wang, Z. (2023). Regionally Differentiated Real-Time Energy Consumption Prediction of Electric Vehicles Oriented to Travel Characteristics. İçinde W. Wang, J. Wu, X. Jiang, R. Li, & H. Zhang (Ed.), *Green Transportation and Low Carbon Mobility Safety* (ss. 631-650). Springer Nature Singapore.
- Wang, L., Yang, Y., Zhang, K., Liu, Y., Zhu, J., ve Dang, D. (2024). Enhancing Electric Vehicle Energy Consumption Prediction: Integrating Elevation into Machine Learning Model. *2024 IEEE Intelligent Vehicles Symposium (IV)*, 2936-2941. <https://doi.org/10.1109/IV55156.2024.10588445>
- Wang, P., Liu, Q., Xu, N., Ou, Y., Wang, Y., Meng, Z., Liu, N., Fu, J., ve Li, J. (2024b). Energy Consumption Estimation Method of Battery Electric Buses

- Based on Real-World Driving Data. *World Electric Vehicle Journal*, 15(7).
<https://doi.org/10.3390/wevj15070314>
- Wang, Y., Chenlong, W., ve Hongwen, H. (2021). Estimating the Energy Consumption and Driving Range of Electric Vehicles with Machine Learning. *Journal of Physics: Conference Series*, 2005(1), 012131. <https://doi.org/10.1088/1742-6596/2005/1/012131>
- Widmer, F., Ritter, A., ve Onder, C. H. (2023). ZTBus: A Large Dataset of Time-Resolved City Bus Driving Missions. *Scientific Data*, 10(1), 687.
<https://doi.org/10.1038/s41597-023-02600-6>
- Wu, Z., Chen, X., ve Jiang, Y. (2023). Prediction of Electric Vehicle Energy Consumption by Combining Real Vehicle Data and Machine Learning Methods. *Medicon Engineering Themes*. <https://doi.org/10.55162/MCET.04.115>
- Würtz, S., Bogenberger, K., Göhner, U., ve Rupp, A. (2024). Towards Efficient Battery Electric Bus Operations: A Novel Energy Forecasting Framework. *World Electric Vehicle Journal*, 15(1), Article 1. <https://doi.org/10.3390/wevj15010027>
- Xu, H., Tu, R., Li, T., ve Chen, H. (2023). Interpretable bus energy consumption model with minimal input variables considering powertrain types. *Transportation Research Part D: Transport and Environment*, 119, 103742.
<https://doi.org/10.1016/j.trd.2023.103742>
- Xu, Z., Wang, J., Lund, P. D., ve Zhang, Y. (2023). Analysis of energy consumption for electric buses based on low-frequency real-world data. *Transportation Research Part D: Transport and Environment*, 122, 103857.
<https://doi.org/10.1016/j.trd.2023.103857>
- Yufang, L., Jun, Z., Chen, R., ve Xiaoding, L. (2020). Prediction of vehicle energy consumption on a planned route based on speed features forecasting. *IET*

Intelligent Transport Systems, 14(6), 511-522. <https://doi.org/10.1049/iet-its.2019.0538>

Zhang, J., Liu, W., Wang, Z., ve Fan, R. (2024). Electric Vehicle Power Consumption Modelling Method Based on Improved Ant Colony Optimization-Support Vector Regression. *Energies*, 17(17). <https://doi.org/10.3390/en17174339>

Zhang, J., Wang, Z., Liu, P., ve Zhang, Z. (2020). Energy consumption analysis and prediction of electric vehicles based on real-world driving data. *Applied Energy*, 275, 115408. <https://doi.org/10.1016/j.apenergy.2020.115408>

Zhang, Z., Ye, B., Wang, S., ve Ma, Y. (2024). Analysis and estimation of energy consumption of electric buses using real-world data. *Transportation Research Part D: Transport and Environment*, 126, 104017. <https://doi.org/10.1016/j.trd.2023.104017>

Zhu, Q., Huang, Y., Feng Lee, C., Liu, P., Zhang, J., ve Wik, T. (2025). Predicting Electric Vehicle Energy Consumption From Field Data Using Machine Learning. *IEEE Transactions on Transportation Electrification*, 11(1), 2120-2132. <https://doi.org/10.1109/TTE.2024.3416532>