

Raylı Sistem Kullanım Oranlarının Açıklanabilir Yapay Zeka ile Değerlendirilmesi: Kentsel Hareketlilik Perspektifi

Evaluation of Rail System Usage Rates with Explainable Artificial Intelligence: An Urban Mobility Perspective

Mert Erkan SÖZEN, İzmir Metro A.Ş., Türkiye, merterkansozen@gmail.com

Orcid No: 0000-0002-7965-6461

Öz: Raylı sistemlerin kullanım tahminleri, özellikle nüfus yoğunluğu yüksek kentsel alanlarda sürdürülebilir ulaşım planlaması için büyük önem taşımaktadır. Bu tahminler, makine öğrenimi (ML) modelleriyle gerçekleştirilebilmekte ancak bu modellerin "kara kutu" yapısı, açıklanabilirlik sorununu beraberinde getirmektedir. Bu sorunu aşmak için, Açıklanabilir Yapay Zekâ (XAI) tekniklerinden biri olan SHAP (SHapley Additive exPlanations) yöntemi, model çıktılarının şeffaf ve yorumlanabilir olmasını sağlamaktadır. Bu çalışmada, raylı sistem kullanım oranlarını tahmin etmek amacıyla ağaç tabanlı bir topluluk algoritması geliştirilmiş ve modelin çıktıları SHAP ile analiz edilmiştir. Analizde, haftanın günü, özel günler ve yolcu biniş türlerinin günlük dağılımı gibi değişkenler kullanılmıştır. İzmir şehrine ait açık veri seti üzerinde test edilen model, yüksek doğruluk oranları elde etmiştir. SHAP analizi, veri seti içindeki "DİĞER_O" ve "TAM_O" inputlarının modelin tahminlerinde en belirleyici faktörler olduğunu ortaya koyarken, özelliklerin çıktı değişkeni olan "TOPLAM" ile arasındaki etkileşimler de dikkat çekmiştir. Bu sonuçlar, toplu taşıma sistemlerinde "Diğer" kullanıcı gruplarının davranışsal dinamiklerinin daha kapsamlı bir şekilde analiz edilmesi gerektiğini vurgulamaktadır. Bu bulgular, kentsel hareketlilik perspektifine bağlı ulaşım politikalarının geliştirilmesine önemli katkılar sunmaktadır.

Anahtar Sözcükler: Raylı Sistemler, Makine Öğrenmesi, Açıklanabilir Yapay Zeka (XAI), Yolcu Talebi Tahmini, SHAP (SHapley Additive exPlanations).

JEL Sınıflandırması: C33, E37, L91.

Abstract: The prediction of rail system usage holds critical importance for sustainable transportation planning, particularly in densely populated urban areas. While machine learning (ML) models can generate these predictions, their "black-box" nature raises interpretability challenges. To address this, the SHAP (SHapley Additive exPlanations) method, an Explainable Artificial Intelligence (XAI) technique, ensures transparency and interpretability of model outputs. In this study, a tree-based ensemble algorithm was developed to predict rail system usage rates, and the model's outputs were analyzed using SHAP. Variables such as day of the week, special events, and daily distribution of passenger boarding types were incorporated into the analysis. Tested on an open dataset from the city of İzmir, the model achieved high accuracy rates. SHAP analysis revealed that the "DİĞER_O" (OTHER_P) and "TAM_O" (FULL_P) inputs were the most decisive factors in the model's predictions, while interactions between features and the output variable "TOPLAM" (TOTAL) also drew attention. These results underscore the need for more comprehensive analysis of behavioral dynamics among "Diğer" (Other) user groups in public transportation systems. The findings offer significant contributions to the development of urban mobility policies grounded in a transportation-oriented perspective.

Keywords: Rail Systems, Machine Learning (ML), Explainable Artificial Intelligence (XAI), Passenger Demand Prediction, SHAP (SHapley Additive exPlanations).

JEL Classification: C33, E37, L91

Makale Geçmişi / Article History

Başvuru Tarihi / Date of Application : 1 Eylül / September 2025

Kabul Tarihi / Acceptance Date : 22 Ekim / October 2025

© 2025 Journal of Yaşar University. Published by Yaşar University. Journal of Yaşar University is an open access journal.

1. Giriş ve Çalışmanın Amacı

Kentsel hareketlilik, hızla büyüyen modern şehirlerin ekonomik canlılığı, sosyal erişilebilirliği ve çevresel sürdürülebilirliği açısından temel dinamiklerinden biridir. Bu bağlamda, raylı sistemler gibi yüksek kapasiteli toplu taşıma modlarının kullanım oranlarının doğru bir şekilde tahmin edilmesi, akılcı ulaşım planlaması ve kaynak optimizasyonu için kritik bir gerekliliktir. Geleneksel istatistiksel yöntemlerin ötesine geçerek yüksek tahmin doğruluğu sunan makine öğrenmesi (ML) modelleri bu alanda yaygınlaşmıştır. Ancak bu modellerin "kara kutu" doğası, karar verme süreçlerinin anlaşılabilirliğini kısıtlamakta ve politika yapıcılarının tahminlere olan güvenini azaltmaktadır. Bu sorunun üstesinden gelmek ve model kararlarını şeffaflaştırmak, çalışmamızın temel motivasyonunu oluşturmaktadır. Bu amaçla, bu çalışma raylı sistem yolcu talebi tahmin modelini Açıklanabilir Yapay Zekâ (XAI) tekniklerinden SHAP (SHapley Additive exPlanations) yöntemi ile analiz etmeyi hedeflemektedir. Çalışmamızın özgün değeri, yolcu talebini etkileyen karmaşık değişkenler arasındaki etkileşimleri (özellikle farklı biniş türlerinin etkisini) XAI ile niceliksel olarak açıklayarak, kentsel hareketlilik perspektifine dayalı, veriye dayalı ve güvenilir politika geliştirme süreçlerine somut katkılar sunmaktır.

Birleşmiş Milletler Sürdürülebilir Kalkınma Hedefleri (SKH'ler) bağlamında ele alındığında, kentsel hareketliliğin optimize edilmesi; iklim değişikliğiyle mücadele (Hedef 13), kapsayıcı ve güvenli şehirlerin inşası (Hedef 11) ve ekonomik büyümenin desteklenmesi (Hedef 8) gibi çoklu hedeflere sinerjik katkı sağlamaktadır (EUROSAI Project Group, 2023).

Özellikle mega-kentlerde gözlemlenen nüfus yoğunluğu ve demografik çeşitlilik, ulaşım sistemlerine yönelik talebi katlanarak artırmaktadır. Bu bağlamda, altyapının etkin yönetimi, yalnızca trafik sıkışıklığını hafifletmekle sınırlı kalmayan stratejik bir paradigmadır. Akıllı ulaşım sistemleri (AUS), entegre raylı ağlar ve veriye dayalı talep tahmin modelleri gibi yenilikçi yaklaşımlar, kaynak kullanım verimliliğini maksimize ederken karbon ayak izinin minimize edilmesine olanak tanır. Ayrıca, toplu taşıma odaklı politikaların benimsenmesi, fosil yakıt bağımlılığını azaltarak hava kalitesi iyileştirmelerine ve biyoçeşitliliğin korunmasına katkıda bulunur. Ekonomik perspektiften bakıldığında ise optimize edilmiş hareketlilik sistemleri, işgücü verimliliğini artırır, lojistik maliyetlerini düşürür ve turizm ile ticaret hacmini doğrudan destekleyerek şehirlerin küresel rekabet gücünü güçlendirir.

Bu bağlamda İzmir özelinde yapılan projeksiyonlar, 2030 yılında kentteki çalışan nüfusun 2.194.416'ya, öğrenci sayısının 1.193.224'e ve hastane yatak kapasitesinin 19.753'e ulaşacağını öngörmektedir. Bununla birlikte otomobil sayısının 1.444.274 olması beklenirken,

bin kişiye düşen otomobil sayısının da 233 olacağı tahmin edilmiştir (Proje, 2017, s. 17). Bu durum, mevcut ulaşım altyapısına gelecekteki baskıyı kaldırabilmesi için akılcı çözümler ile yaklaşılması gerektirdiğini ortaya koymaktadır.

Günümüzde kentsel hareketliliğin sürdürülebilirliği, çevresel dirençlilik, ekonomik verimlilik ve sosyal eşitlik ekseninde şekillenen üçlü sacayağı üzerine inşa edilmektedir. Bu doğrultuda, toplu taşıma sistemlerinin entegrasyonu, sadece bir seçenek değil, kentlerin geleceğini şekillendiren ve karbon salınımını azaltmada stratejik bir gereklilik olarak öne çıkmaktadır. Nitekim Li, Hao, Xie, Han ve Wang (2023), çalışmalarında Çin'in Karbon Emisyon Eylem Planı dahilinde karbon salınımının azaltılması için araç, ekipman ve raylı sistemlerde uygulayacakları politikalardan bahsederek bu iddiayı somutlaştırmaktadır.

Raylı sistemlerin avantajları yalnızca çevresel boyutla sınırlı değildir. Özellikle büyük şehirlerde, trafikten bağımsız işletim altyapısı ve elektrikli sistemlerle uyumluluğu sayesinde yüksek yolcu taşıma kapasitesine sahip olmaları, onları sürdürülebilir ulaşım planlamasının merkezine yerleştirmektedir.

Nar ve Araslankaya (2022) çalışmalarında, toplu taşımanın stratejik ve detaylı planlamasının başarısının büyük ölçüde doğru talep bilgisi verilerine bağlı olduğunu belirtmiştir. Raylı ulaşım sistemlerinde yolcu talebi tahmini, doğru sefer planlaması açısından büyük önem taşımaktadır. Yolcu talebi doğru tahmin edildiğinde, sefer sıklığı optimize edilir, yolcu memnuniyeti artar ve işletme maliyetleri düşer.

Bu noktada, yolcu talebi tahminlerinin, yolcu davranışlarının tahminlenmesinde de önemli bir rol oynadığı görülmektedir. Kusunthum, Srinavin, Leungbootnak, ve Chaitongrat (2022), çalışmalarında Tayland Bangkok'daki yeraltı tren yolcusunu tahminlemek için makine öğrenmesi tekniklerini kullandıkları çalışmalarında yolcu davranışlarının rastgele oluşmadığını ve yolcu akışının etkili bir şekilde tahmin edildiğini göstermişlerdir.

Yolcu talebi tahminleri bağlamında, toplu taşıma kullanım dinamiklerini şekillendiren ana faktörlerin sistematik olarak tanımlanması ve bu değişkenlerin nicel veriye dayalı tahmin modellerine entegre edilmesi, operasyonel verimlilikten kaynak optimizasyonuna kadar uzanan etkili karar alma süreçlerinin analitik temelini oluşturmaktadır (Sariyer, Mangla, Sozen, Li, & Kazancoglu, 2024).

Bu çalışma, raylı sistem kullanım oranlarının tahmininde makine öğrenmesi modellerinin potansiyelini keşfetmeyi ve bu modellerin karar süreçlerini Açıklanabilir Yapay Zekâ (XAI) teknikleriyle şeffaf hale getirmeyi amaçlamaktadır. Toplu taşıma talebini etkileyen değişkenler arasında haftanın günleri, özel günler ve yolcu biniş türleri (örn. tam ücretli, öğrenci, 60 yaş) oransal dağılımı gibi dinamikler analize dahil edilmiştir. Geliştirilen model,

İzmir’e ait açık veri seti üzerinde test edilerek yüksek doğruluk performansı sergilemiş; SHAP (SHapley Additive exPlanations) analiziyle ise “Tam Ücretli” ve “Diğer” kategorilerindeki yolcu oranlarının tahminler üzerinde en belirleyici rolü oynadığı tespit edilmiştir. Bu bulgular, kentsel planlamada geleneksel olarak göz ardı edilen diğer kullanıcı gruplarının (60 yaş, bilet, personel vb.) dinamiklerinin dikkate alınması gerektiğini vurgulamaktadır. Çalışmanın nihai hedefi, yapay zekâ destekli tahminlerin politika yapıcılara sunduğu nicel ve nitel içgörülerle sürdürülebilir ulaşım stratejilerinin geliştirilmesine katkı sağlamaktır.

2. Çalışma Konusuyla İlgili Kavramsal / Kuramsal Çerçeve

Kentsel hareketlilik başlığı altında yolcu talebinin tahminlemesi ile alakalı yapılan çalışmalar incelendiğinde mevcut çalışmaların çoğu raylı sistem kullanım modunu analiz etmiş ve metro kullanım oranlarını tahmin etmeyi amaçlamıştır. Chen, Zhirui, Wang, ve Xu (2020), çalışmalarında Çin, Nanjing'deki Olimpiyat Spor Merkezi yakınlarındaki iki istasyondan alınan toplu taşıma akıllı kart verilerine dayalı olarak yolcu akışının ortalamasını ve oynaklığını modellemişlerdir. Jing ve Yin (2020), ise çalışmalarında Pekin şehrindeki raylı sistem ağının 1 Haziran 2017 ile 31 Mayıs 2018 tarihleri arasındaki akıllı kart kullanımı verilerini baz alarak yolcu sayısını tahmin etmeye çalışmışlardır. Utku ve Kaya (2023), ise kısa vadeli transit yolcu hacminin tahminlenmesi çalışmalarında LSTM tabanlı derin öğrenme modeli ile rastgele orman (RF) ve destek vektör makineleri gibi makine öğrenmesi modellerini karşılaştırmış, LSTM tabanlı derin öğrenme modelinin her aktarma güzergahı için tahmin açısından diğer modellerden daha iyi performans gösterdiğini belirtmişlerdir. Zhang, Liu, Spurgeon, ve Yu (2021), ise Londra Metrosu trenlerinin yolcu akışının tahminlenmesi üzerine çalışma yapmışlardır. Son olarak yine Liu, Liu, & Jia (2019), metroya gelen/giden yolcu akışının tahminlenmesi üzerine bir çalışma yapmışlardır.

Raylı sistem kullanım oranlarının doğru tahmini, hem operasyonel verimlilik hem de kentsel hareketliliğin belirlenmesi açısından kritik bir role sahiptir. Kentsel hareketlilikte raylı sistemlerin kullanımına yönelik artan ilgi, bu ulaşım modunun daha etkin ve verimli hale getirilmesi için kapsamlı geliştirme stratejilerinin oluşturulmasını zorunlu kılmaktadır. Ancak Borucka, Mazurkiewicz, ve Lagowska (2021), bu stratejilerin doğru bir şekilde tanımlanabilmesi için, demiryolu taşımacılığı hizmetlerine yönelik talebin etkili bir şekilde analiz edilmesi ve güvenilir tahmin modellerine dayalı öngörüler sağlanması gerektiğini belirtmektedir.

Makine öğrenmesi modelleri, yüksek doğruluk oranlarına ulaşmasına rağmen, çoğu zaman kara kutu olarak değerlendirilmekte ve karar verme süreçlerinin anlaşılabilirliğini ve güvenilirliğini sınırlandırmaktadır. Madanu, Abbod, Hsiao, ve Chen (2022), çalışmalarında makine öğrenmesinin klinik tanı ve tedavi önerileri gibi tıbbi karar alma süreçlerinde yaygın olarak kullanıldığını, ancak bu modellerin nasıl karar verdiğinin her zaman kolay anlaşılmadığını belirtmiştir. Bu bağlamda, Açıklanabilir Yapay Zekâ (XAI)'ın, makine öğrenmesi kararlarını bağlamsal açıklamalarla ilişkilendirerek karar alma süreçlerini daha şeffaf hale getirdiğini vurgulamışlardır.

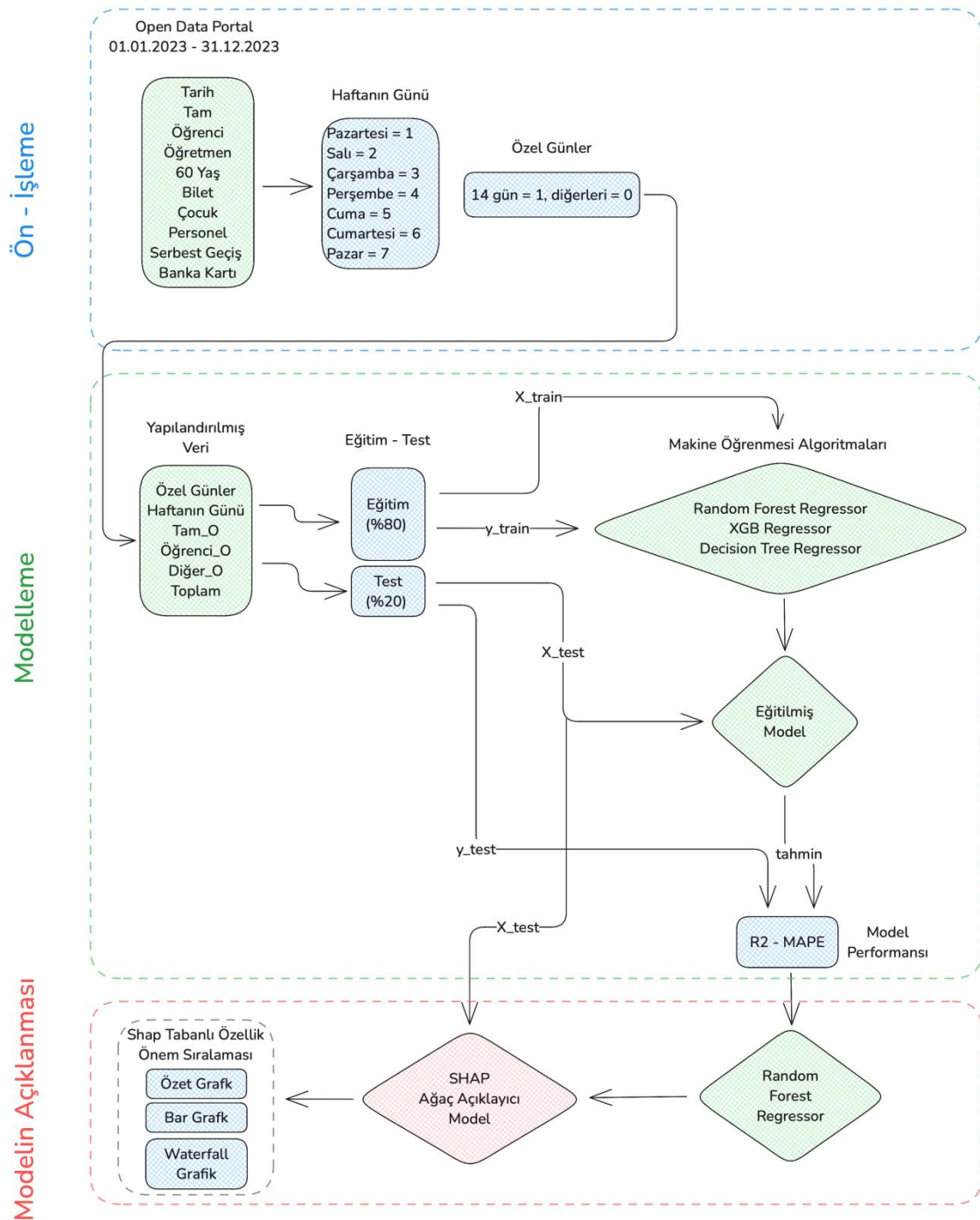
Kentsel hareketlilik alanında yolcu talebi tahminine yönelik mevcut çalışmaların çoğu, raylı sistem kullanımını ve metro yolcu akışını analiz etmeye odaklanmıştır. Bu çalışmalar, akıllı kart verileri üzerinden yolcu akışının modellenmesi veya derin öğrenme ve makine öğrenimi modellerinin karşılaştırılması gibi teknik yaklaşımlar sergilemiştir. Ancak bu modeller, raylı sistem kullanım oranlarının doğru tahmini ve kentsel hareketlilik stratejilerinin oluşturulması açısından kritik öneme sahip olsa da, çoğunlukla "kara kutu" yapısında olup, yüksek doğruluk oranlarına rağmen karar verme süreçlerinin anlaşılabilirliğini ve güvenilirliğini kısıtlamaktadır.

Mevcut çalışma, literatürdeki bu açıklanabilirlik boşluğunu doldurarak önemli bir katkı sağlamaktadır. Çalışma, raylı sistem kullanım oranlarını tahmin etmek için ağaç tabanlı bir topluluk algoritması (Random Forest Regressor) geliştirmiş ve bu "kara kutu" yapısındaki modelin çıktılarını şeffaf ve yorumlanabilir hale getirmek için Açıklanabilir Yapay Zekâ (XAI) tekniklerinden SHAP (SHapley Additive explanations) yöntemini kullanmıştır. Diğer çalışmalar talep tahminine odaklanırken, bu çalışma, SHAP analizi ile model tahminlerinde hangi değişkenlerin (özellikle "DİĞER_O" ve "TAM_O" gibi yolcu biniş türlerinin) en belirleyici faktörler olduğunu ortaya koymuş ve bu sayede toplu taşıma politikalarının geliştirilmesine somut, veriye dayalı içgörüler sunmuştur. Bu bağlamda, mevcut çalışma, sadece tahmin yapmakla kalmayıp, tahminin arkasındaki mekanizmaları şeffaf bir şekilde açıklayarak, kentsel planlamada sürdürülebilir ulaşım stratejilerinin daha etkin bir şekilde tasarlanmasına yönelik bir karar destek çerçevesi sunmasıyla literatürdeki yerini sağlamlaştırmaktadır.

3. Yöntem

Bu çalışmada, raylı sistem kullanımına yönelik talep tahmini için makine öğrenmesi tabanlı bir model geliştirilmiştir ve ilgili model açıklanabilir yapay zeka ile değerlendirilmiştir.

Yöntem, ön işleme, modelleme ve modelin açıklanması olmak üzere üç temel aşamadan oluşmaktadır (Şekil-1).



Şekil 1. Önerilen Modelin Akış Şeması

3.1 Ön-İşleme:

Bu çalışmada kullanılan veri seti, İzmir Büyükşehir Belediyesi'nin Açık Veri Platformunda yer alan toplu taşıma verilerinden alınmış olup, 1 Ocak 2023 ile 31 Aralık 2023 tarihleri arasındaki yıllık bazlı yolcu binış verilerini içermektedir (Tablo-1).

Tablo 1. Örnek Veri Seti

<i>Tarih</i>	<i>Tam</i>	<i>Öğrenci</i>	<i>Öğretmen</i>	<i>60 Yaş</i>	<i>Bilet</i>	<i>Çocuk</i>	<i>Personel</i>	<i>Serbest Geçiş</i>	<i>Banka Kartı</i>
31.12.2023	168.640	138.645	1.201	8.163	3	354	9.753	102.835	18.712
31.12.2023	44.483	36.143	202	2.069	0	150	2.480	27.046	4.038
31.12.2023	16.441	16.230	224	574	3	12	658	5.223	4.080
31.12.2023	235	204	6	1	0	0	0	0	1.296
31.12.2023	74	36	0	0	0	0	0	0	49

İlgili veri setinde tam, öğrenci, öğretmen, 60 yaş, bilet, çocuk, personel, serbest geçiş ve banka kartı olmak üzere toplamda dokuz biniş türü mevcuttur. Yolcu biniş türlerini “tam, öğrenci ve diğer” olmak üzere üç grupta toplanmıştır. Öğretmen, 60 yaş, bilet, çocuk, personel, serbest geçiş ve banka kartı binişlerini ise “diğer” grubunda toplanmıştır. Bu üç biniş türü toplam yolcu sayısına oranlanarak modelin girdileri olarak tanımlanmıştır. Tam biniş sayılarının toplam biniş sayılarına oranı “TAM_O”, öğrenci biniş sayılarının toplam biniş sayılarına oranı “ÖĞRENCİ_O”, diğer ücretli biniş sayılarının toplam biniş sayılarına oranı ise “DİĞER_O” olarak tanımlanmıştır. Ayrıca “HAFTANIN GÜNÜ” ve “ÖZEL GÜNLER” de modelin diğer girdileri olarak tanımlanmıştır. Haftanın her günü artan düzende kategorik olarak tanımlanırken, o günün özel (ulusal veya dini bayram) olup olmadığını ikili bir yapıda temsil etmek için bir değişken eklenmiştir. Son olarak, tüm biniş türlerinin toplam biniş sayısı hedef değişken olarak tanımlanmıştır. Dolayısıyla, 2023’ün 365 günü için veri setimiz 365 satır ve 6 sütun içerecek şekilde hazırlanmıştır (Tablo 2).

Tablo 2. Yapılandırılmış Veri Seti

<i>Özel Günler</i>	<i>Haftanın Günü</i>	<i>Tam_O</i>	<i>Öğrenci_O</i>	<i>Diğer_O</i>	<i>Toplam</i>
1	7	0.51	0.34	0.15	360.887
0	1	0.50	0.29	0.22	615.834
0	2	0.46	0.31	0.23	641.135
0	3	0.49	0.31	0.19	503.317
0	4	0.50	0.31	0.19	537.969

3.2 Modelleme:

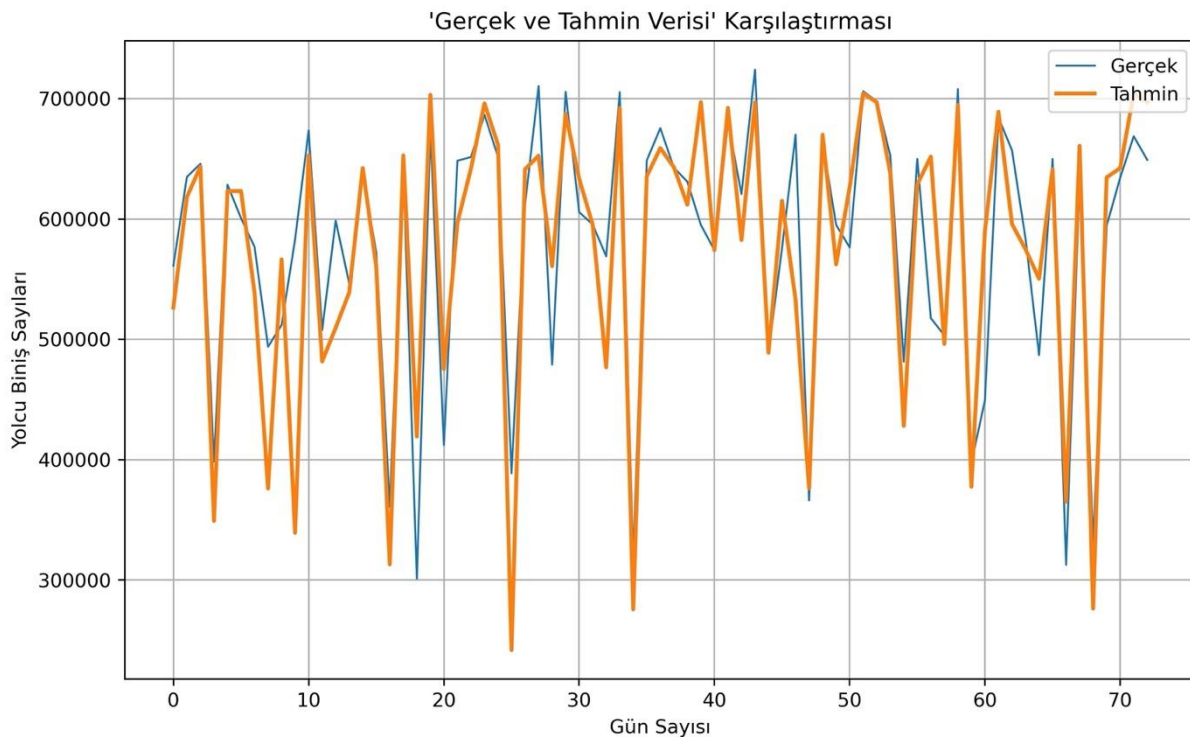
Yapılandırılmış veri setinin %80’i eğitim, %20’si test veri seti olarak ayrılmıştır. Talep tahmini için Random Forest Regressor (RF), XGB Regressor (XGB) ve Decision Tree Regressor (DT) algoritmaları kullanılmıştır. Yolcu biniş tahmini için en yüksek R² ve en düşük hata istatistiklerine sahip en iyi performans gösteren algoritma olan RF seçilmiştir

(Tablo-3). Bu algoritma ile Şekil – 2 de görüleceği üzere günlük toplam yolcu biniş sayılarına ilişkin tahminler üretilmiş, test verilerine bağlı olarak gerçek ve tahmin değerleri karşılaştırılmıştır.

Test veri setinin birçok örneği için gerçek ve tahmin edilen değerler birbirine yakın olduğu gözlemlenmiştir. Model bazı durumlarda gerçek değerleri fazla tahmin ederken, bazı durumlarda az tahmin etmiştir ama genel olarak yüksek performanslı tahminlere sahiptir.

Tablo 3. Algoritmaların Tahmin Performansları

Algoritmalar	R^2	RMSE	MAPE
RF	0,70	59.506,51	0,08
XGB	0,61	68.276,46	0,09
DT	0,55	72.936,13	0,10



Şekil 2. Gerçek ve Tahmin Verilerinin Karşılaştırılması

3.3 Modelin Açıklanması:

Giriş değişkenlerinin modelin günlük toplam yolcu biniş sayılarını tahminleri üzerindeki etkilerinin yorumlamak için SHAP (SHapley Additive exPlanations) değerleri kullanılmıştır. SHAP, modeli açıklamak için Shapley değerlerine dayalı farklı grafikler (Summary, Bar ve WaterFall) sunar. SHAP ile RF algoritmasının değişkenleri nasıl değerlendirdiği ortaya çıkartılarak tahmin üzerinde en etkili değişkenlerin (örneğin, DİĞER_O ve TAM_O) belirlenmesi ve bu tahminlerin daha güvenilir hale gelmesi sağlanmıştır.

4. Bulgular Ve Tartışma

Şekil-3a de gösterilen Özet Grafik, Tablo – 4 de görüleceği üzere her bir giriş değişkeninin modelin çıktısına olan etkisini Shapley değerleri üzerinden niceliksel olarak ve renk skalası ile gösterirken, etki yönünü de pozitif ve negatif olarak ortaya koyar. Niceliksel olarak Shapley değerlerinin yüksek olması ilgili özelliğin modelin tahmini üzerinde etkisinin büyük olduğunu gösterirken, düşük değerler model tahmin üzerinde daha az etkili olduğunu gösterir. Renk olarak kırmızı renk, veri setindeki ilgili özelliğin yüksek değerlerini, mavi renk ise ilgili özelliğin düşük değerlerini göstermektedir.

Tablo 4. SHAP Değer Tanım Tablosu

<i>Shapley Değeri</i>	<i>Renk</i>	<i>Değişken Değeri</i>	<i>Anlamı</i>
<i>Yüksek (+)</i>	Kırmızı	Yüksek	Değişkenin yüksek değeri, tahmini güçlü bir şekilde artırır.
<i>Yüksek (+)</i>	Mavi	Düşük	Değişkenin düşük değeri, tahmini güçlü bir şekilde artırır.
<i>Yüksek (-)</i>	Kırmızı	Yüksek	Değişkenin yüksek değeri, tahmini güçlü bir şekilde azaltır.
<i>Yüksek (-)</i>	Mavi	Düşük	Değişkenin düşük değeri, tahmini güçlü bir şekilde azaltır.
<i>Düşük (+)</i>	Kırmızı	Yüksek	Değişkenin yüksek değeri, tahmini hafif bir şekilde artırır.
<i>Düşük (+)</i>	Mavi	Düşük	Değişkenin düşük değeri, tahmini hafif bir şekilde artırır.
<i>Düşük (-)</i>	Kırmızı	Yüksek	Değişkenin yüksek değeri, tahmini hafif bir şekilde azaltır.
<i>Düşük (-)</i>	Mavi	Düşük	Değişkenin düşük değeri, tahmini hafif bir şekilde azaltır.

DİĞER_O değişkeninin analizinde dikkat çeken iki farklı davranış paterni gözlemlenmiştir:

Pozitif Etki Senaryosu:

- Şekil 4a'da görüldüğü üzere, yüksek değerler (0.21, 0.22, 0.23 gibi) kırmızı renk kodlamasıyla temsil edilmekte ve Şekil 3a'daki SHAP dağılımında belirgin pozitif değerler (+100.000 birim civarı) sergilemektedir.
- Bu durum, değişkenin yüksek değerlerinin model tahminlerini istikrarlı şekilde artırdığını göstermektedir.

Negatif Etki Senaryosu:

- Düşük değerler (0.13, 0.14, 0.15 gibi) mavi renkle işaretlenmiş olup, Şekil 3a'da -200.000 birime varan aşırı negatif SHAP değerleri üretmektedir.

- Bu da değişkenin düşük değerlerinin tahminleri benzer şiddette baskıladığını ortaya koymaktadır.

Değişkenin X eksenindeki SHAP değerlerinin -200.000 ile +100.000 aralığında asimetrik dağılımı, model çıktıları üzerinde çift yönlü aşırı duyarlılık ve diğer değişkenlere kıyasla anormal ölçüde baskın rol sergilediğini görsel olarak doğrulamaktadır.

TAM_O değişkeninin düşük değerleri (Şekil 4b'de görülen 0.38, 0.40 gibi değerler), Şekil 3a'da mavi renk kodlamasıyla temsil edilmiş ve pozitif yönde yüksek SHAP değerleri sergilemiştir. Bu bulgu, TAM_O oranındaki düşüşlerin model tahminlerini beklenmedik şekilde güçlü pozitif yönde etkilediğini göstermektedir.

HAFTANIN GÜNÜ değişkenine ait yüksek değerler (Şekil 4d'de görülen ve 6 - Cumartesi ve 7 - Pazar olarak kodlanan hafta sonu günleri), Şekil 3a'da kırmızı renkle gösterilmiş ve negatif yönde düşük SHAP değerleri üretmiştir. Bu durum, hafta sonu günlerinin model tahminlerini negatif yönde etkilediğini ortaya koymaktadır.

ÖĞRENCİ_O ve ÖZEL GÜNLER değişkenlerinin Şekil 3'teki SHAP dağılımları incelendiğinde:

- Her iki değişken de $\pm 20,000$ birimlik dar bir etki aralığı sergilemektedir.
- Pozitif/negatif yönde istikrarlı ancak marjinal katkılar sağlamaktadır.
- Model tahminleri üzerinde belirleyici bir rol oynamadıkları gözlemlenmiştir.

Şekil 3b'de sunulan SHAP önem sıralaması grafiği, modelin tahminlerine en fazla katkı sağlayan değişkenleri nicel olarak ortaya koymaktadır. Değişkenlerin ortalama mutlak SHAP değerleri temel alındığında, DİĞER_O (+70,190.68) ve TAM_O (+26,621.85) değişkenlerinin diğer faktörlere kıyasla model çıktıları üzerinde istatistiksel olarak anlamlı ölçüde daha belirleyici olduğu görülmektedir. Bu bulgular, söz konusu iki değişkenin modelin karar mekanizmalarında baskın rol oynadığını göstermektedir.

Özellikle dikkat çekici olan, bu değişkenlerin toplam binişler içindeki oransal dağılımı temsil etmesi ve aralarındaki ters orantılı ilişkidir. Şekil 3b'deki verilerle desteklenen bu ilişki şu şekilde özetlenebilir:

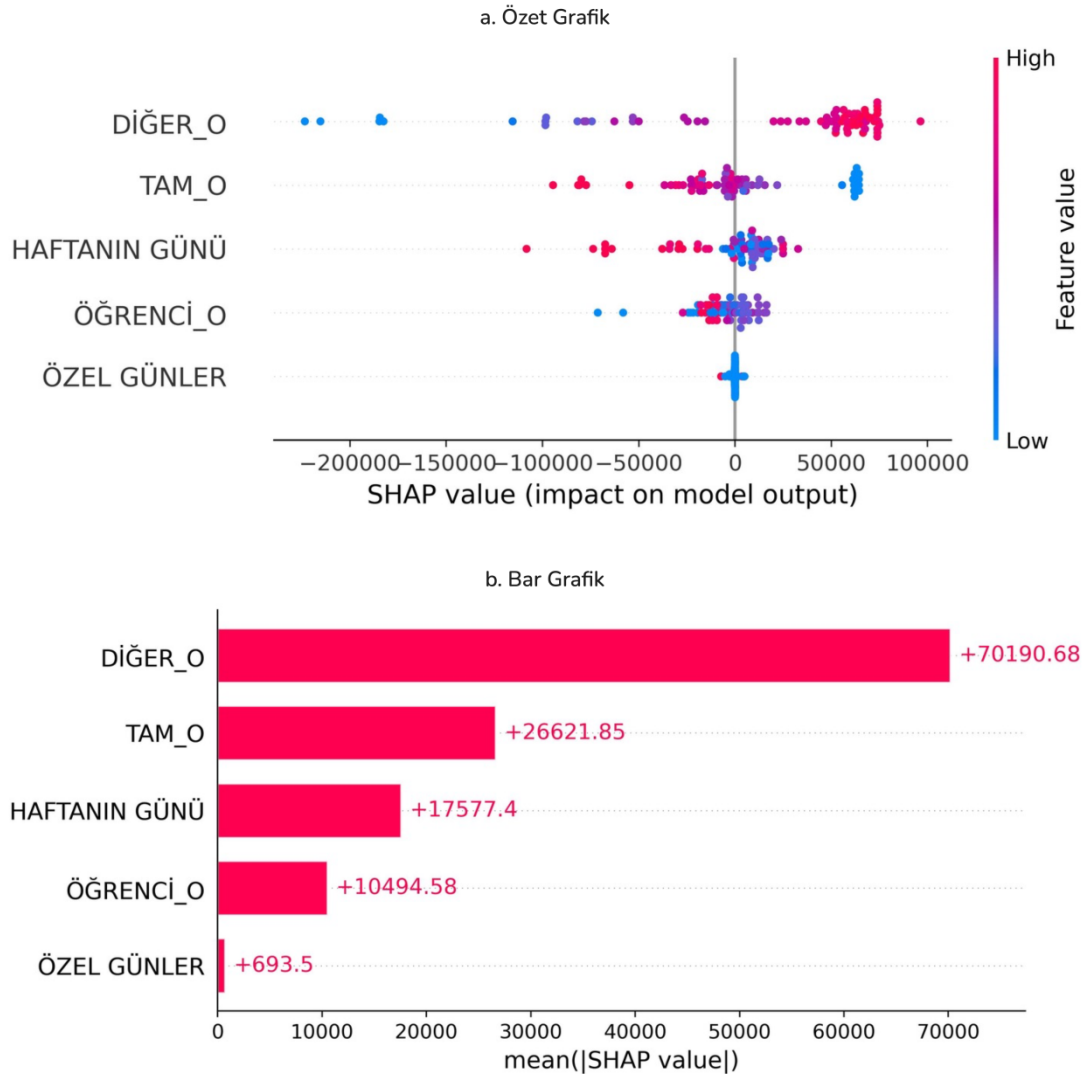
- TAM_O'daki azalma $\rightarrow$ DİĞER_O'da artış
- DİĞER_O'daki artış $\rightarrow$ TAM_O'da azalma

Bu karşılıklı etkileşim, raylı sistem talebi modellemesinde iki kritik sonucu ortaya çıkarmaktadır:

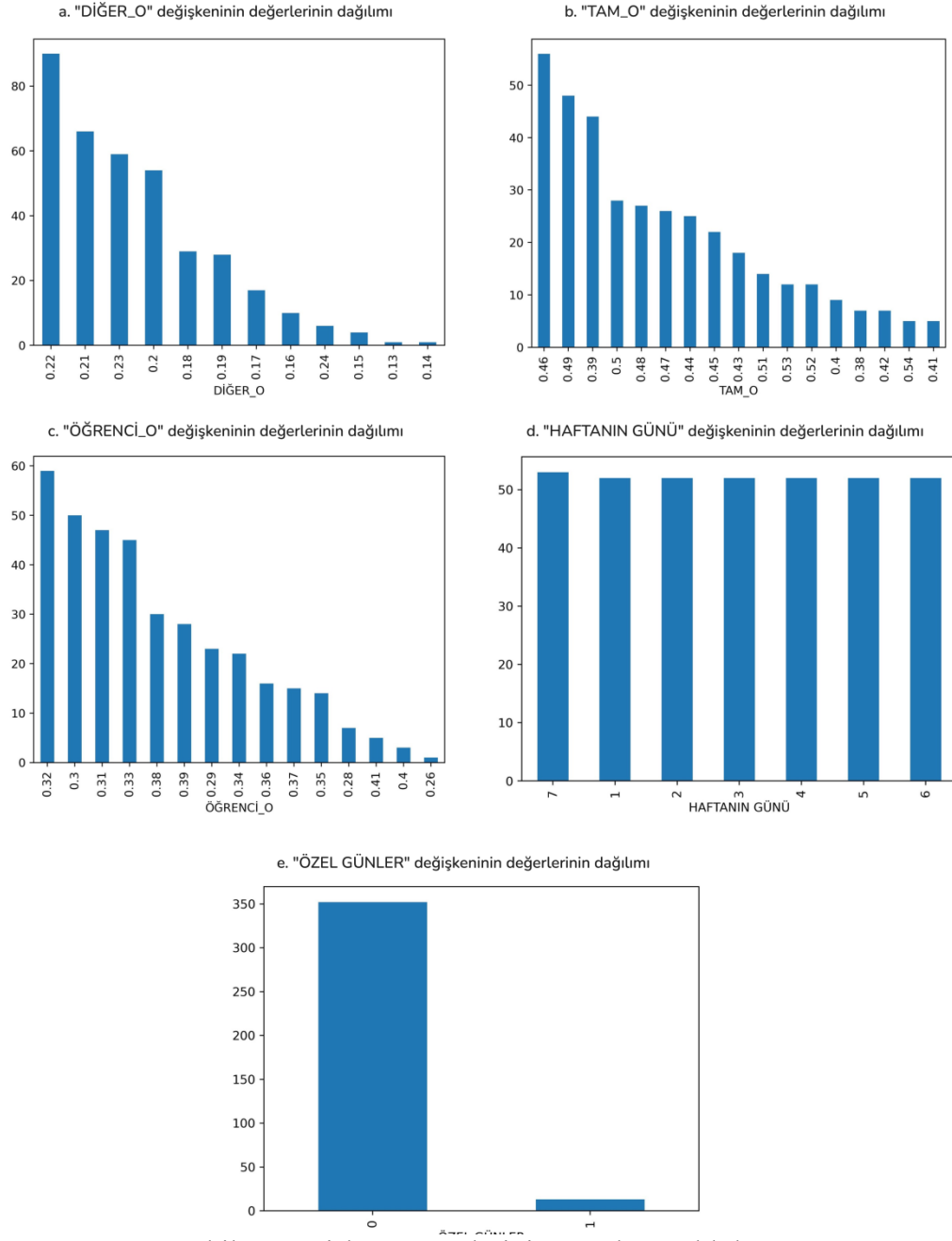
- Toplam biniş hacmindeki değişimler, bu iki değişkenin oransal dengelerine duyarlıdır.

- Modelin tahmin performansı, büyük ölçüde bu iki değişkenin etkileşim dinamiğini doğru yakalayabilme kapasitesine bağlıdır.

TAM_O ve ÖĞRENCİ_O değişkenlerinin ± 50.000 birimlik dar SHAP aralıkları, bu değişkenlerin modele tutarlı ve öngörülebilir etkiler yaptığını gösterir. Buna karşın DİĞER_O değişkeninin SHAP değerlerinin -200.000 ile +100.000 arasında değişmesi, bu değişkenin modele daha dengesiz ve sıra dışı etkiler yapabileceğini ortaya koymaktadır.



Şekil 3. SHAP Tabanlı Özellik Önem Sıralaması



Şekil 4. Değişken Değerlerinin Dağılım Tabloları

SHAP önem sıralamasında belirgin şekilde öne çıkan DİĞER_O değişkeninin model davranışı üzerindeki etkileri, Şekil 5'te sunulan SHAP waterfall grafiği ile daha detaylı analiz edildiğinde, bu değişkenin tahminler üzerinde asimetrik ve bağlama duyarlı bir etki profili sergilediği görülmektedir. Özellikle dikkat çeken bulgular şunlardır:

Düşük Değerler (0.13-0.15):

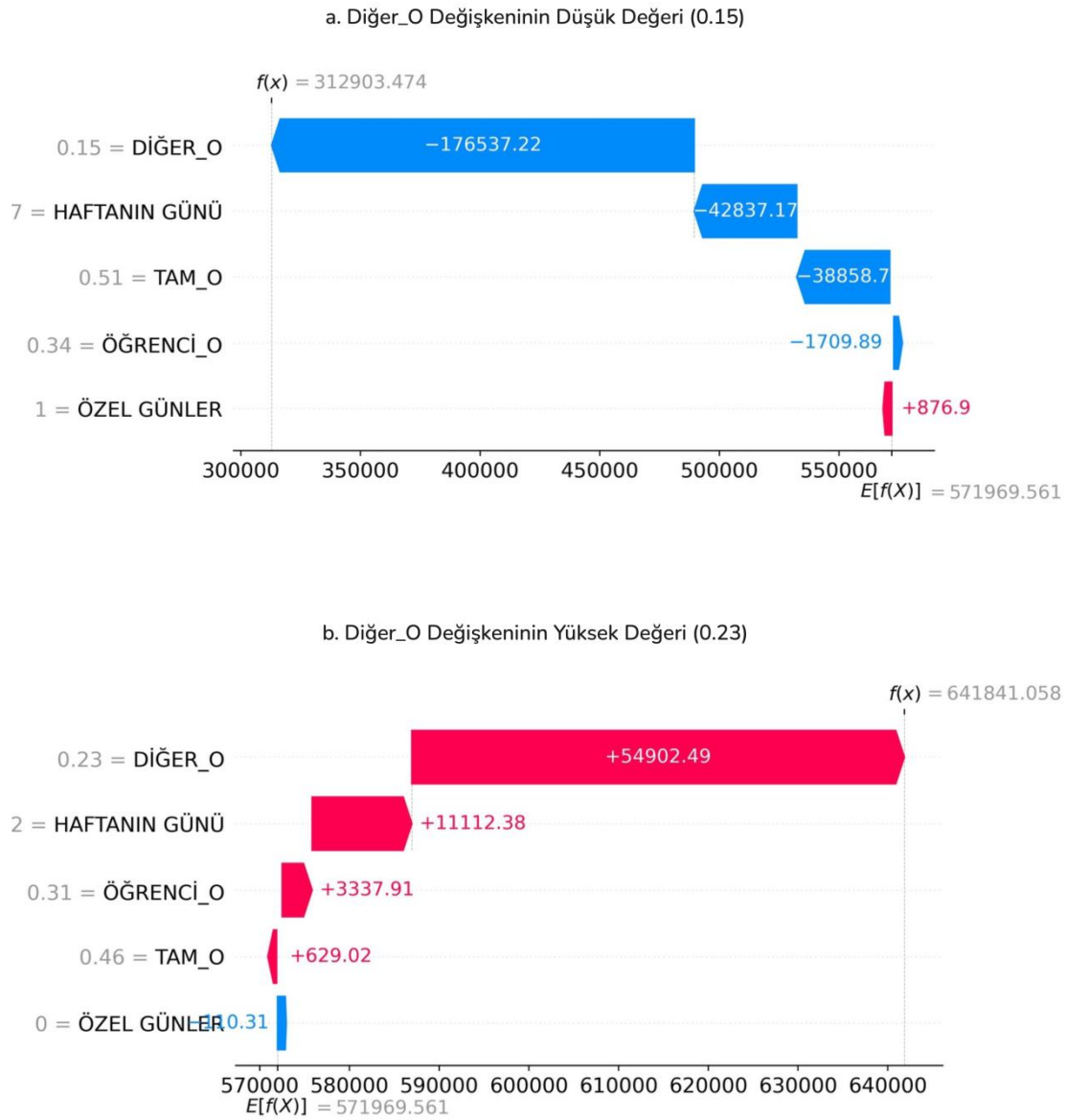
- -176,537.22 birime varan şiddetli negatif etkiye (tahminleri %30.9 baskılama) sebep olmaktadır.
- Modelin ortalama tahmini toplam yolcu sayısı 571,969.56 olup $E[f(X)]$ ile gösterilirken, DİĞER_O değişkeninin düşük değerlerinin (0.15) baz alındığı senaryo için (HAFTANIN GÜNÜ:7, TAM_O:0.51, ÖĞRENCİ_O:0.34 ve ÖZEL GÜNLER:1) modelin tahmin ettiği toplam yolcu sayısı 312,903.47 olup $f(x)$ ile ifade edilmiştir (Şekil 5a). Bu değer ortalama tahminin çok çok altındadır. DİĞER_O'nun SHAP etkisi diğer tüm değişkenlerden kat kat fazla (-176,537.22 birim) modelin tahminine negatif katkı sağlayarak modelin tahminini ciddi oranda düşürmektedir.

Yüksek Değerler (0.21-0.23):

- +54,902.49 birimlik pozitif katkı (tahminleri %9.6 artırma) sağlamaktadır.

Modelin ortalama tahmini toplam yolcu sayısı 571,969.56 iken DİĞER_O değişkeninin yüksek değerlerinin (0.23) baz alındığı senaryo için (HAFTANIN GÜNÜ:2, TAM_O:0.46, ÖĞRENCİ_O:0.31 ve ÖZEL GÜNLER:0) modelin tahmin ettiği toplam yolcu sayısı 641,841.06 birimdir. (Şekil 5b). DİĞER_O değişkeni modelin tahminine diğer değişkenlere göre yine çok daha fazla (+54,902.49 birimlik) katkı sağlamaktadır. HAFTANIN GÜNÜ +11,112.38 birimlik ve ÖĞRENCİ_O değişkeni de +3,337.91 birimlik katkı sağlamaktadır.

DİĞER_O değişkeninin düşük ve yüksek değerlerindeki bu 231,439.71 (-176,537.22 ve +54,902.49) birimlik mutlak fark, değişkenin model tahminlerinde eşik etkisi yaratarak doğrusal olmayan bir katkı sağladığını göstermektedir.



Şekil 5. SHAP Waterfall Grafiği

5. Sonuç ve Öneriler

Bu çalışma, makine öğrenimi modellerinin toplu taşıma sistemlerinde yolcu talebini yüksek doğrulukla tahmin etme potansiyelini ortaya koymuştur. Özellikle haftanın günleri, özel günler, biniş türleri ve toplam biniş gibi bağımlı ve bağımsız değişkenlerin doğru bir şekilde tanımlanmasının model performansı üzerinde belirleyici olduğu gözlemlenmiştir. Öte yandan, bağımlı ve bağımsız değişkenler arasındaki karmaşık etkileşimlerin modellenmesi, sistemin esnekliğini ve güvenilirliğini pekiştirmiş; bu da sürdürülebilir kentsel planlama için veriye dayalı karar destek sistemlerinin önemini bir kez daha vurgulamıştır.

Bu çalışmanın bulguları, makine öğrenimi tabanlı yaklaşımların toplu taşıma araçlarına yönelik yolcu talebini son derece güvenilir bir biçimde öngörebildiğini göstermektedir. Haftanın günleri, özel günler, biniş türleri ve toplam biniş gibi bağımlı ve bağımsız

değişkenlerin doğru bir şekilde tanımlanmasının model performansı üzerinde belirleyici bir etkiye sahip olduğu gözlemlenmiştir. Diğer yandan, bağımlı ve bağımsız değişkenler arasındaki etkileşimin farklı makine öğrenmesi algoritmalarının — özellikle Random Forest Regressor, XGB Regressor ve Decision Tree Regressor — güçlü yönlerinden yararlanacak şekilde modellenmesi, karmaşık veri yapılarıyla daha iyi başa çıkabilen ve çeşitli senaryolarda tutarlı tahminler sunan hibrit bir çerçeve oluşturmuştur; yapılan karşılaştırmalı analizler sonucunda ise Random Forest Regressor'un sağladığı en yüksek doğruluk ve güvenilirlik sebebiyle nihai model olarak tercih edilmiştir.

Kentsel planlama bağlamında ele alındığında ise, veriye dayalı talep tahmin modellerinin önemi daha da netleşmektedir. Şehir içi ulaşım altyapısının etkin kullanımı; yolcu beklentilerinin doğru öngörülmesi, güzergâh optimizasyonu, sefer sıklığı düzenlemeleri ve araç filosu yönetimi gibi pek çok stratejik kararı besler. Özellikle hızla büyüyen kentlerde altyapı yatırımlarının maliyetinin yüksekliği ve toplu taşımaya duyulan talebin belirsizliği göz önüne alındığında, makine öğrenimi destekli analizler hem ekonomik kaynakların verimli dağılımını hem de çevresel sürdürülebilirlik hedeflerini yakalamayı mümkün kılar. Bu çerçevede, geliştirilen modeller yalnızca teknik birer araç olmaktan çıkıp, geleceğin şehirlerini şekillendirecek akıllı karar destek sistemlerinin temel taşlarından biri haline gelmektedir.

İzmir Büyükşehir Belediyesi'nin Açık Veri Platformundan alınan verilerle İzmir ili özelinde yapılmış olan bu çalışma, DİĞER_O değişkeninin (öğretmen, 60 yaş üstü ve serbest geçişli yolcular) raylı sistem talebi üzerinde beklenenden çok daha belirleyici bir role sahip olduğunu ortaya koymaktadır. Analizler, bu değişkenin model üzerinde adeta bir "davranış anahtarı" işlevi gördüğünü ve %8'lik küçük değişimlerin bile (0.15-0.23 aralığında) 230.000 birimlik radikal tahmin kaymalarına yol açabildiğini göstermektedir. Özellikle dikkat çeken, değişkenin düşük değerlerinde (<%15) tahminleri -176,537.22 birim baskılayabilirken, yüksek değerlerinde (>%20) +54,902.49 birime varan pozitif katkı sağlayabilmesidir. Bu aşırı duyarlılık, modelin bu değişkene karşı anormal ölçüde hassas olduğunu ve DİĞER_O'daki küçük dalgalanmaların toplam yolcu tahminlerinde $\pm\%10$ 'luk sapmalara neden olabileceğini ortaya koymaktadır. Buna karşılık, "TAM_O" binişlerindeki stabilite, modelin bu değişkene olan duyarlılığını sınırlandırmıştır. Bu bulgular ışığında, belediye yönetimlerinin kentsel ulaşım politikalarını oluştururken sadece yoğun kullanıcı gruplarını değil, bu özel grupların dinamiklerini de dikkate alması ve modelin bu aşırı tepkiselliğini dengelemek için regularizasyon teknikleri uygulaması önerilmektedir.

Bu çalışmanın sınırlılığı olarak, modelin DİĞER_O'daki değişimlere gösterdiği bu şiddetli tepkinin altında yatan sosyo-ekonomik faktörlerin ve mevsimsel etkilerin henüz tam olarak modellenmemiş olması gösterilebilir. İleriki çalışmalarda DİĞER_O kategorisinin alt gruplara (yaşlı/engelli/öğretmen) ayrıştırılarak analiz edilmesi tavsiye edilmektedir. Ayrıca model sadece bir yıllık veriye dayandığı için (01.01.2023 - 31.12.2023), daha uzun vadeli veri setlerinin kullanımıyla ortaya çıkabilecek önemli trendleri ve mevsimsel dalgalanmaları (örneğin yaz tatili ve okul dönemleri gibi) yansıtmaya kapasitesi sınırlıdır.

Bu çalışmanın bulguları ışığında, açıklanabilir yapay zeka (XAI) yöntemleriyle geliştirilen tahmin modellerinin, kentsel ulaşım yönetiminde kritik karar destek araçları olarak kullanılabileceği ortaya çıkmaktadır. Özellikle SHAP analizleriyle desteklenen bu modeller, DİĞER_O değişkeninin beklenmedik etki büyüklüğünü ortaya çıkararak, karar vericilere üç temel alanda iyileştirme fırsatı sunmaktadır:

- Birincisi, politika tasarımı düzeyinde, yerel yönetimlerin özel gruplara (öğretmen, yaşlı ve engelli yolcular) yönelik hedefli müdahaleler geliştirmesi gerekmektedir. Bu kapsamda, DİĞER_O kategorisindeki kullanıcıların demografik özelliklerini ve seyahat kalıplarını detaylı şekilde yansıtan sosyo-ekonomik değişkenlerin modele entegrasyonu önerilmektedir. Örneğin, yaşlı nüfus oranı yüksek bölgelerdeki istasyonlara yönelik özel erişilebilirlik önlemleri veya öğretmenler için esnek ücretlendirme politikaları, modelin tahmin performansını artırırken aynı zamanda sosyal adaleti de destekleyecektir.
- İkincisi, teknik altyapı düzeyinde, modelin güvenilirliğini artırmak için en az üç yıllık uzun vadeli ve mevsimsel etkileri içeren veri setlerinin kullanılması elzemdir. Bu sayede, DİĞER_O değişkeninin yarattığı aşırı duyarlılık etkileri daha iyi anlaşılabilir ve dengelenebilir. Ayrıca, modelin farklı şehirlerdeki raylı sistemlerde validasyonu yapılarak, bulguların genellenebilirliği test edilmelidir.
- Üçüncüsü, operasyonel yönetim düzeyinde, DİĞER_O değişkenindeki dalgalanmaların yaratacağı etkileri öngörebilmek için gerçek zamanlı izleme sistemlerinin kurulması tavsiye edilmektedir. Bu sistemler sayesinde, özel günlerde veya demografik değişim dönemlerinde ortaya çıkabilecek talep şoklarına karşı proaktif önlemler alınabilir. Özellikle %15-20 aralığındaki kritik eşik değerlerinin sürekli izlenmesi, sistemin sürdürülebilirliği açısından hayati önem taşımaktadır.

Bu çalışmanın geleceğe yönelik çıkarımları bağlamında, modelin tahmin kapasitesini geliştirmek için iki temel araştırma yönü öne çıkmaktadır: Birincisi, raylı sistem verilerinin

otobüs ve feribot gibi diğer toplu taşıma modlarıyla entegre edilerek daha geniş bir analiz çerçevesi oluşturulması; ikincisi ise gelir dağılımı ve göç hareketleri gibi makro sosyo-ekonomik göstergelerin modele dahil edilerek tahminlerin bağlamsal geçerliliğinin artırılmasıdır. Bu kapsamlı yaklaşım, yalnızca modelin tahmin doğruluğunu iyileştirmekle kalmayacak, aynı zamanda ulaşım planlamasının sosyal adalet boyutunu da güçlendirecektir. Nihayetinde, makine öğrenimi algoritmaları ile açıklanabilir yapay zeka (XAI) tekniklerinin stratejik birlikteliği, akıllı şehirler vizyonu çerçevesinde hem operasyonel verimliliği optimize eden hem de sosyal eşitliği gözeten bütüncül bir ulaşım yönetim sisteminin inşasına önemli katkılar sunacaktır.

KAYNAKÇA

- Borucka, A., Mazurkiewicz, D., & Lagowska, E. (2021). Mathematical Modelling As An Element of Planning Rail Transport Strategies. *Transport*, 36(4):354-363.
- Chen, E., Zhirui, Y., Wang, C., & Xu, M. (2020). Subway Passenger Flow Prediction for Special Events Using Smart Card Data. *IEEE Transactions on Intelligent Transportation Systems*, 21(3):1109-1120.
- EUROSAI Project Group. (2023, Mayıs). *Un-Sai Country-Level Collaboration On SDG Audits*. <https://turkiye.un.org/sites/default/files/2023-05/UN-SAI-COUNTRY-LEVEL-SDG-final-final.pdf> adresinden alındı
- Jing, Z., & Yin, X. (2020). Neural Network-Based Prediction Model for Passenger Flow in a Large Passenger Station: An Exploratory Study. *IEEE Access*, 8:36876 - 36884.
- Kusonkhum, W., Srinavin, K., Leungbootnak, N., & Chaitongrat, T. (2022). Using a Machine Learning Approach to Predict the Thailand Underground Train's Passenger. *Hindawi Journal of Advanced Transportation*.
- Li, H., Hao, Y., Xie, C., Han, Y., & Wang, Z.-R. (2023). Emerging technologies and policies for carbon-neutral transportation. *International Journal of Transportation Science and Technology*, 331-332.
- Liu, Y., Liu, Z., & Jia, R. (2019). DeepPF: A deep learning based architecture for metro passenger flow prediction. *Transportation Research Part C: Emerging Technologies*, 101:18-34.
- Madanu, R., Abbod, M., Hsiao, F.-J., & Chen, W.-T. (2022). Explainable AI (XAI) Applied in Machine Learning for Pain Modeling: A Review. *Technologies*, 1-15.
- Nar, M., & Araslankaya, S. (2022). Passenger Demand Forecasting For Railway Systems. *Open Chemistry*, 105-119.
- Proje, B. (2017, Temmuz). *İzmir Ulaşım Ana Planı Sonuç Raporu Yönetici Özeti*. https://www.izmir.bel.tr/CKYuklenen/dokumanlar_2018/upi_sonuc_ozeti.pdf adresinden alındı
- Sariyer, G., Mangla, S., Sozen, M., Li, G., & Kazancoglu, Y. (2024). Leveraging Explainable Artificial Intelligence In Understanding Public Transportation Usage Dates Dor Dustainable Development. *Omega*.
- Utku, A., & Kaya, S. (2023). New Deep Learning-Based Passenger Flow Prediction Mode. *Transportation Research Record: Journal of the Transportation Research Board*, 2677(3): 1-17.
- Zhang, Q., Liu, X., Spurgeon, S., & Yu, D. (2021). A two-layer modelling framework for predicting passenger flow on trains: A case study of London underground trains. *Transportation Research Part A: Policy and Practice*, 151: 119-139.