

MAKİNE ÖĞRENMESİ İLE PAYLAŞIMLI BİSİKLET KULLANIMINA AİT TALEP TAHMİNİ: KONYA ÖRNEĞİ

Gözde CAN ATASAGUN^{1*}, Ece YAĞMUR²

¹ Konya Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Endüstri Mühendisliği Bölümü, Konya, ORCID No: <https://orcid.org/0000-0003-4921-1557>

² Konya Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Endüstri Mühendisliği Bölümü, Konya, ORCID No: <https://orcid.org/0000-0001-5865-3483>

Anahtar Kelimeler	Öz
Paylaşımlı bisiklet sistemleri Makine öğrenmesi Yapay sinir ağları Rassal orman Talep tahmini	<i>Sürdürülebilirlik kavramının her geçen gün öneminin arttığı çağımızda geleneksel ulaşım araçlarının yerine çevreye daha duyarlı alternatif yöntemler araştırılmaktadır. Bu yenilikçi çözümlerden birisi karbon emisjonsuz ulaşımı sağlayan paylaşımlı bisiklet kullanımudur. Paylaşımlı bisiklet kullanımı, ulaşımında karbon ayak izini büyük ölçüde azaltmakta, ayrıca şehir içi trafik yoğunluğunu önleyerek çevresel sürdürülebilirliğe katkı sağlamaktadır. Ekonomik açıdan uygun fiyatlı ve erişilebilir olması nedeniyle de sıklıkla tercih edilen paylaşımlı bisikletler son yıllarda pek çok ülkede yaygın bir şekilde kullanılmaktadır. Paylaşımlı bisiklet sistemlerinden en yüksek faydayı sağlamak adına arz ve talep dengesi kurmak için etkin bir talep tahmin prosedürü geliştirilmelidir. Bu amaçla yapılan çalışmada Türkiye'nin en uzun bisiklet yoluna sahip olan Konya ilinde 2022-2023 yıllarına ait paylaşımlı bisiklet kullanım verileri incelenmiştir. Öncelikle verilere ait saatlik ve günlük bazda tanıttıcı istatistikler verilerek veriler özetlenmiş, ardından tahmin modellerinde kullanılan değişkenler tanımlanmıştır. Paylaşımlı bisiklet kullanımına olan talebin tahmininde Çoklu Doğrusal Regresyon (ÇDR), Ridge Regresyon (RR), Elastik Net (EN), Rassal Orman (RO), Karar Ağaçları (KA), Destek Vektör Makineleri (DVM) ve Yapay Sinir Ağları (YSA) kullanılmıştır. Geliştirilen günlük ve saatlik talep tahmin modelleri ve kullanılan makine öğrenme algoritmaları determinasyon katsayısı (R^2), ortalama mutlak hata (MAE), ortalama hata karesi (MSE) ve kök ortalama hata karesi (RMSE) performans ölçütlerine göre karşılaştırılmıştır. Günlük verilerin değişken olarak kullanıldığı model, incelenen tüm performans ölçütlerine göre saatlik verilerin kullanıldığı modele göre üstünlük sağlamıştır. Algoritmalar karşılaştırıldığında saatlik veriler için en iyi performansa sahip teknik YSA iken, günlük veriler için ise en iyi performansı veren teknik RO olarak belirlenmiştir.</i>

DEMAND FORECASTING FOR BIKE SHARING SYSTEMS BY MACHINE LEARNING: A CASE STUDY IN KONYA

Keywords	Abstract
Bike-sharing systems Machine learning Artificial neural networks Random forest Demand forecasting	<i>As the concept of sustainability increases day by day, eco-friendly techniques are researched instead of traditional transportation vehicles. One of these innovative solutions is the bike-sharing system, which provides carbon-emission-free transportation. The system greatly reduces the carbon footprint and also contributes to environmental sustainability by preventing urban traffic congestion. Bike-sharing systems, which are frequently preferred because they are economically affordable and accessible, have been widely used in many countries recently. An effective demand forecasting procedure should be developed to balance supply and demand. In the study for this purpose, bike sharing data for the years 2022–2023 in Konya, which has the longest bicycle path in Turkey, is examined. First, the data are summarized with descriptive statistics on an hourly and daily basis, and then the variables used in the forecasting models are defined. Multiple Linear Regression (MLR), Ridge Regression (RR), Elastic Net (EN), Random Forest (RF), Decision Trees (DT), Support Vector Machines (SVM), and Artificial Neural Networks (ANN) are used to estimate the actual demand. The hourly and daily demand forecasting models and the machine learning algorithms used are compared according to the coefficient of determination (R^2), mean absolute error (MAE), mean square error (MSE), and root mean square error (RMSE). The model in which daily data is used is superior to the hourly model in all performance measures examined. When the algorithms are compared, the best technique for the hourly model is found to be ANN, while it is determined to be RF for the daily model.</i>

Araştırma Makalesi	Research Article
Başvuru Tarihi : 02.05.2024	Submission Date : 02.05.2024
Kabul Tarihi : 23.08.2024	Accepted Date : 23.08.2024

* Sorumlu yazar: gcatasagun@ktun.edu.tr
<https://doi.org/10.31796/ogummf.1477372>



Bu eser, Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) hükümlerine göre açık erişimli bir makaledir. This is an open access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>).

1. Giriş

Hızlı kentleşme neticesinde tüm dünyada şehirler hava ve su kirliliği, trafik sıkışıklığı, temel hizmetlere erişim sorunları gibi pek çok problemle karşı karşıya kalmaktadır. Artan kentsel dinamikler ve çevresel problemler sonucunda şehirler, yenilenebilir enerji kullanımı, ekolojik yatırımlar, sürdürülebilir şehir planlama gibi çevre yönetimini önceliklendiren politikalar benimsemektedir. Özellikle kentsel ulaşım dünya çapında şehirlerin sürdürülebilirliğinin şekillenmesinde önemli rol oynamaktadır. Şehir merkezlerinde fosil yakıtla çalışan araçların yaygın bir şekilde kullanımı pek çok çevresel problemi beraberinde getirerek sürdürülebilirliği olumsuz bir şekilde etkilemektedir. Bu nedenle ulaşım altyapısı sağlanırken toplu taşıma odaklı planlamaların yapılması, halk sağlığı açısından bisiklet, scooter gibi araçların kullanıldığı aktif ulaşımın teşvik edilmesi, elektrikli araç kullanımının artırılması gibi uygulamalar tüm dünyada hızla hayata geçirilmektedir. Böylece şehir içi ulaşım sistemlerinde etkinliğin ve verimliliğin sağlanarak araçların birbiri ile rekabet etmeden ve birbirini tamamlayarak işletilebilmesi, toplu taşıma, yaya ve bisiklet sistemlerinin bir bütün olarak planlanması ile mümkün olacağı fikri benimsenmeye başlanmıştır (Candan, 2003; Karagöz, 2019). Bu strateji dahilinde öne çıkan uygulamalardan birisi paylaşımlı bisiklet kullanımıdır. Paylaşımlı bisikletler kısa mesafeler için uygun ve esnek bir ulaşım aracı olup kişisel araçlara olan bağımlılığı azaltarak şehir merkezlerindeki trafik yoğunluğunu büyük ölçüde azaltmaktadır. Ayrıca sıfır emisyonla ulaşımı sağlayarak hava kalitesi ve çevresel sürdürülebilirliğe katkıda bulunmakta ve aktif bir ulaşım aracı olduğundan kullanıcıların fiziksel hareketliliğine katkıda bulunarak halk sağlığını olumlu yönde etkilemektedir. Literatürde üçüncü nesil bisiklet paylaşım sistemi olarak da adlandırılan paylaşımlı bisiklet sistemleri için geliştirilen mobilite uygulamaları kullanıcılara geleneksel ulaşım araçlarına kıyasla daha fazla esneklik sunmaktadır. Kullanıcılar yakındaki bisiklet ve scooterlara akıllı telefon uygulamaları ile kolayca erişim sağlayabilmektedir. Özellikle kısa mesafeli kullanımlar için oldukça uygun fiyatlı olması ve bireysel araçları olmayan veya toplu taşıma araçlarına erişimi kısıtlı olan bireyler için erişilebilir olması nedeniyle de paylaşımlı bisikletler sıklıkla tercih edilmektedir.

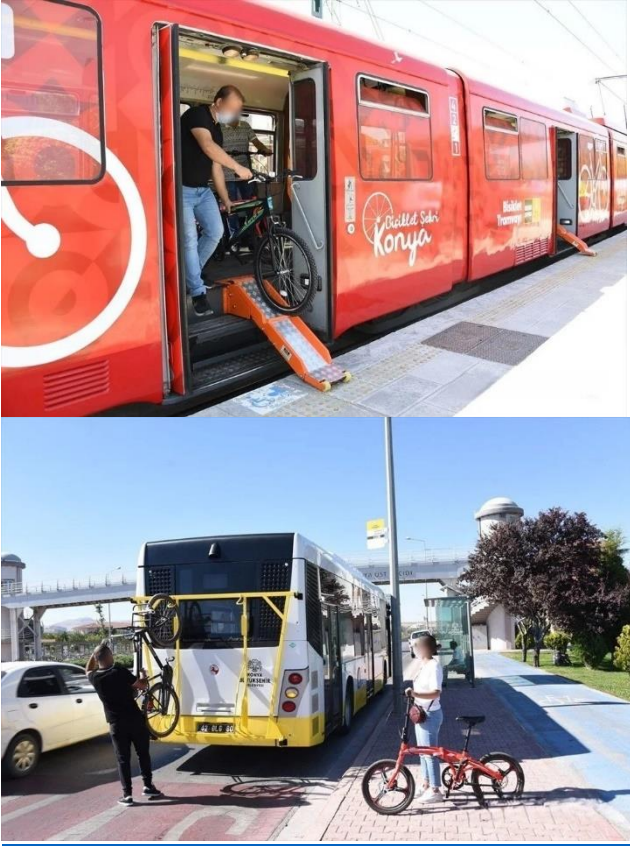
Pek çok ülkede paylaşımlı bisiklet kullanımını teşvik etmek ve bu konuda farkındalık yaratmak için çeşitli çalışmalar yapılmıştır. Paylaşımlı bisiklet sistemi ilk olarak 1998'de Rennes'te kullanılmaya başlanmış, 2000 yılında Münih'te, 2005 yılında Lyon'da ve 2007 yılında Paris'te büyük bir kullanım oranına ulaşmıştır (Midgley ve Affairs, 2011; Shaheen, Guzman ve Zhang, 2010). 2008 yılında Brezilya, Şili, Çin, Yeni Zelanda, Güney Kore, Tayvan ve ABD'de de uygulamaya alınmış,

sonrasında ise tüm dünyaya yayılmıştır (DeMaio, 2009; Madak ve Bardakçı, 2021; Pucher ve Buehler, 2008).

Türkiye'de de pek çok şehirde yerel yöneticiler tarafından desteklenen paylaşımlı bisiklet kullanımı için farklı uygulamalar geliştirilmiştir (Eren, Katanalp, Yıldırım ve Uz, 2018). Türkiye'nin yüzölçümü açısından en büyük şehri olan Konya'da da bisiklet kullanımını teşvik edecek pek çok uygulama mevcuttur. Şehrin topografik yapısı ve iklim koşulları düşünüldüğünde Konya ilinin büyük bir bölümü düze yakın bir arazi üzerinde olduğundan hem bireysel hem de paylaşımlı bisiklet kullanım oranının oldukça yüksek olduğu görülmektedir. 580 kilometre ile Türkiye'nin en uzun bisiklet yoluna sahip olan şehirde yaklaşık 80 adet paylaşımlı bisiklet istasyonu mevcuttur. Bisiklet kullanıcılarının şehir içi ulaşımında sıkıntılar yaşamamasının önüne geçmek için şehir merkezinde kullanılan tramvay ve otobüslerde de bisiklet sürücüleri düşünülerek iyileştirmeler yapılmıştır. Bisiklet tramvayı ve otobüslere eklenen bisiklet taşıma aparatları ile bisiklet sürücülerinin bisikletleri ile seyahat etmeleri mümkün kılınmıştır. Şekil 1'de Konya'da kullanılan bisiklet tramvayı ve bisiklet taşınabilen otobüsler görülmektedir.

Paylaşımlı bisiklet kullanımına hızla artan ilgi sonucunda sisteme yönelik talebin hızlı ve doğru bir şekilde tahmin edilmesi oldukça önemli bir konu haline gelmiştir. İşletmelerin talebi doğru bir şekilde tahmin etmesi bisikletler, bağlantı istasyonları ve bakım ekipleri gibi kaynakların daha verimli bir şekilde kullanılmasına olanak sağlar. İşletmeler sistemin dinamik yapısı gereğince dalgalanan talep karşısında etkin bir talep yönetimi ile bisikletlerin doğru yerde ve zamanda kullanılabilir olmasını sağlayabilir. Böylelikle müşteri ihtiyaçları zamanında karşılanmış olduğundan müşteri memnuniyeti de artırılmış olur. Ayrıca talebin doğru tahmin edilmesi ile işletmeler kendi operasyonel süreçlerini iyileştirebilir ve gelecek dönemler için planlama yapabilirler. Örneğin, bisikletlerin kullanım şekillerini ve hangi bölgede daha çok ihtiyaç duyulduğunu analiz ederek talebin çok olduğu bölgelere bisikletleri taşıyabilir ya da talep dalgalanmalarına karşı fiyat politikalarını güncelleyebilir böylelikle satışlarını ve karlarını artırabilirler. Tüm bu avantajları düşünüldüğünde paylaşımlı bisiklet kullanım sistemleri için etkin bir talep tahmin prosedürünün geliştirilmesi önem arz etmektedir.

Makine öğrenimi, bilgisayarların verilere dayalı öğrenme ile tahminler yapmasını sağlayan yapay zekanın bir alt alanıdır. Bu nedenle literatürde pek çok farklı alanda talep tahmini için makine öğrenme algoritmalarından sıklıkla yararlanılmıştır. Bu modellerin verilerdeki karmaşık kalıpları kolaylıkla keşfederek geleneksel istatistiksel yöntemlere göre daha iyi performans gösterdiği pek çok çalışmada kanıtlanmıştır.



Şekil 1. Bisiklet Tramvayı ve Bisiklet Taşınabilen Otobüs (NTV, 2020)

Yapılan çalışmada Türkiye'deki en uzun bisiklet yoluna sahip olan Konya ilinde 2022-2023 yıllarına ait paylaşımlı bisiklet kullanım verileri incelenerek talep tahmini için hem günlük hem de saatlik bazda tahmin modelleri geliştirilmiştir. Bildiğimiz kadarıyla Türkiye'deki paylaşımlı bisiklet sistemleri için bu zamana kadar yapılmış en kapsamlı çalışma bu çalışmadır. Ayrıca literatürde yer alan çalışmalardan farklı olarak veri setinde bisikletlerin alınıp bırakıldığı istasyon bilgileri de kullanılmıştır. Paylaşımlı bisiklet kullanımına olan talebin tahmininde Çoklu Doğrusal Regresyon (ÇDR), Ridge Regresyon (RR), Elastik Net (EN), Rassal Orman (RO), Karar Ağaçları (KA), Destek Vektör Makineleri (DVM) ve Yapay Sinir Ağları (YSA) kullanılmıştır. Bu yöntemler Türkiye'deki paylaşımlı bisiklet sistemlerinde henüz kullanılmamış yöntemlerdir. Geliştirilen günlük ve saatlik talep tahmin modelleri ve kullanılan makine öğrenme algoritmaları farklı performans ölçütlerine göre değerlendirilmiştir.

Çalışmanın bundan sonraki bölümleri şu şekilde özetlenmiştir: İkinci bölümde paylaşımlı bisiklet sistemleri ile ilgili literatürdeki çalışmalar ele alınmıştır. Üçüncü bölümde verilen materyal ve metot bölümü, veri seti ile talep tahmin algoritmaları olarak iki alt başlıkta incelenmiştir. Dördüncü bölümde algoritmalar sonucunda elde edilen karşılaştırmalı sonuçlar

yorumlanmıştır. Son bölümde ise genel bulgulardan bahsedilerek geleceğe yönelik değerlendirmeler yapılmıştır.

2. Bilimsel Yazın Taraması

Paylaşımlı bisiklet sistemleri ile ilgili literatürde pek çok çalışma mevcuttur. DeMaio (2009) yapmış olduğu çalışmada paylaşımlı bisiklet sistemlerinin kronolojik gelişimini incelemiştir. Çalışmada literatürde yer alan bisiklet paylaşım yaklaşımlarının avantaj ve dezavantajları ortaya konularak sermaye ve işletme maliyetlerinin durumu değerlendirilmiştir. Shaheen ve diğ. (2010)'nın yapmış olduğu çalışmada Avrupa, Amerika ve Asya'daki bisiklet paylaşım sistemlerinin mevcut durumu incelenerek gelecekte yapılması planlanan yenilikler araştırılmıştır. Lorasokkay ve Ağırdır (2011), Konya'da bulunan bisiklet kullanıcıları ile yapmış oldukları anketler vasıtasıyla kullanıcıların bisiklet kullanımlarındaki şikâyet ve önerilerini incelemiş ve şehir içi ulaşımda bisiklet kullanımının artırılarak ulaşımın iyileştirilmesini amaçlamışlardır. Shaheen, Cohen ve Martin (2013) Amerika Birleşik Devletleri'nde ve Kanada'da faaliyet gösteren paylaşımlı bisiklet kullanıcılarını kapsayan ankete dayalı bir araştırma yapmış ve Kuzey Amerika'daki bisiklet paylaşım yaklaşımının büyüme potansiyelini incelemişlerdir. Fishman, Washington ve Haworth (2013)'ün yapmış oldukları çalışmada bisiklet paylaşım sistemlerine ilişkin bir literatür taraması sunulmuş, programların kullanıcı motivasyonlarına ve bisiklet sahibi olma eğilimlerine etkileri araştırılmıştır. Zhao, Deng ve Song (2014) tarafından yapılan çalışmada Çin'deki 69 adet bisiklet paylaşım noktası temel alınmış, kentsel özelliklerin ve sistem karakteristiklerinin bisiklet paylaşım sisteminin günlük kullanımı ve devir hızı üzerindeki etkileri incelenmiştir. Bordagaray, Dell'Olio, Fonzone ve Ibeas (2016) paylaşımlı bisiklet sistemindeki akıllı kart işlemlerinin kullanım potansiyelini incelemişler ve bisiklet kullanım verilerini analiz ederek çevrimdışı bir veri madenciliği prosedürü önermişlerdir. Morton (2018) tarafından Londra'daki paylaşımlı bisiklet sistemi ile ilgili yapılan çalışmada üyelerin hizmet kalitesini nasıl algıladığına, genel memnuniyet düzeylerine, üyeliklerini yenileme eğilimlerine ve programı başkalarına tavsiye etme isteklerine odaklanılmıştır. Ayrıca üyelerin deneyimleri göz önüne alınarak sistemin başarısı için yeni stratejiler önerilmiştir. Zhao, Zhang, Banks ve Xiong (2018) yapmış oldukları çalışmada Çin'deki paylaşımlı bisiklet sistemlerinin tarihsel gelişim aşamalarını ve bu süreçlerdeki mevcut sorunları ele alarak bu sorunlara alternatif çözüm önerileri sunmuşlardır. Jahanshahi, Van Wee ve Kharazmi (2019) İran'ın Mashhad kentinde ülkenin ilk paylaşımlı bisiklet sistemini ve bu sistemin kabul edilebilirliğini etkileyen faktörleri incelemişlerdir. Çalışmada sosyo-kültürel, ekonomik, altyapı ve sistem işlevi gibi faktörler, 134 kullanıcının

katılımıyla analiz edilmiştir. Araştırma, sistemin genellikle gençler, öğrenciler, düşük gelirli bireyler ve araç sahibi olmayan kişiler tarafından kullanıldığını ortaya koymuştur. Si, Shi, Wu, Chen ve Zhao (2019) 2010 ve 2018 yılları arasında bisiklet paylaşım sistemi ile ilgili araştırmaları haritalandırmış ve bu yıllardaki eğilimleri ve katkıları raporlamışlardır. Zhang, Meng, Koh ve Wong (2021) dünya çapında 106 şehirden elde edilen verileri kullanarak paylaşımlı bisiklet sistemlerinin faaliyet süresini etkileyen faktörleri incelemiştir. Çalışmada altyapı, sosyal faktörler ve ekonomik faktörler gibi unsurlar genelleştirilmiş ve bir doğrusal model kullanılarak analiz edilmiştir. Araştırma sonucunda, kapsama alanı, sistem kapasitesi, ödeme türü, finansal destek ve kişi başına satın alma gücü gibi unsurların sistemin sürdürülebilirliğini etkileyen temel faktörler olduğu belirtilmiştir.

Lin, He ve Peeta (2018) yapmış oldukları çalışmada istasyon düzeyinde saatlik talebi tahmin etmek amacıyla bir YSA yaklaşımı önermiş ve modelin farklı mimarilerini incelemiştir. Xu, Ji ve Liu (2018) tarafından yapılan çalışmada istasyonsuz paylaşımlı bisiklet sistemleri incelenerek bu sistemlerdeki bisiklet paylaşımına ait hareketlilik modelleri analiz edilmiş ve dinamik talep tahmini yapmak için derin öğrenme teknikleri kullanılmıştır. Mekânsal ve zamansal analizler yapılarak elde edilen bulgular, Uzun-Kısa Süreli Bellekli (UKSB) sinir ağlarının farklı zaman aralıklarında paylaşımlı bisiklet seyahatlerini tahmin etmede etkili olduğunu göstermektedir. McBain ve Caulfield (2018) İrlanda'da faaliyet gösteren bir paylaşımlı bisiklet sistemini ele almış ve seyahat sürelerindeki değişkenlikleri incelemiştir. Yapılan çalışmada 2015 ve 2016 yılına ait veriler analiz edilerek seyahat sürelerindeki varyasyonu etkileyen mekânsal ve zamansal faktörler belirlenmeye çalışılmıştır. Wang ve Kim (2018) UKSB, Geçitli Tekrarlayan Birimler (GTB) ve RO algoritmalarını kullanarak bisiklet istasyonlarındaki bisiklet sayılarını tahmin etmeyi amaçlamışlardır. Eğitim süresi açısından RO algoritması daha avantajlı olmasına rağmen uzun vadeli tahminler için UKSB'nin daha iyi sonuçlar verdiği görülmüştür. Pan, Zheng, Zhang ve Yao (2019) paylaşımlı bisiklet talebini tahmin etmek için gerçek zamanlı bir yöntem önermişler ve geçmiş veriler, hava durumu verileri ve zaman verilerine dayanarak bisiklet taleplerini belirlemeyi amaçlamışlardır. Seul'deki paylaşımlı bisiklet kullanımını etkileyen faktörler Sathishkumar, Park ve Cho (2020) tarafından analiz edilmiş, sıcaklık, saat, hava koşulları ve kiralanan bisiklet sayısı arasındaki ilişkiler anlamlı bulunmuştur. Tahmin için önemli özellikleri seçmek üzere Boruta algoritması ve RO algoritması kullanılmış, ardından regresyon modelleri eğitilip değerlendirilmiştir. Sathishkumar ve Yongyun (2020) tarafından yapılan bir diğer çalışmada paylaşımlı bisiklet talep tahmini için kurala dayalı bir regresyon tahmin modeli sunulmuştur. İki farklı veri

seti için beş farklı istatistiksel model kullanılarak veri setlerinin eğitimi ve değerlendirilmesi gerçekleştirilmiştir. Kural tabanlı modelin, her iki veri setinde de yüksek tahmin performansı gösterdiği belirtilmiştir. Ayrıca değişken önem analizi sonucunda sıcaklık ve saatin en etkili değişkenler olduğu vurgulanmıştır. Chibwe, Heydari, Imani, Scurtu (2021) Londra'daki bisiklet paylaşım sistemini ele almış ve Ocak 2012'den Haziran 2020'ye kadar olan günlük bisiklet kiralama sayılarındaki eğilimi incelemiştir. Hava durumu, zamansal faktörler ve istasyonların sayısı gibi değişkenler de analize dahil edilmiştir. Jiang (2022) yapmış olduğu çalışmada bisiklet paylaşım sistemlerinde kullanılan derin öğrenme modellerini incelemiş ve bu konuda geniş bir literatür araştırması sunmuştur. Derin öğrenmenin, bisiklet paylaşımı kullanımlarının tahmininde başarılı sonuçlar elde etme konusunda büyük bir avantaj sağladığı belirtilmiştir. Bu çalışma derin öğrenme teknikleriyle paylaşımlı bisiklet kullanımı tahmini için gerçekleştirilen ilk çalışmadır. Ergül Aydın, İçmen Erdem ve Erzurum Çiçek (2023) yapmış oldukları çalışmada paylaşımlı bisiklet kullanımındaki talebi tahmin etmek için gradyan artırma yöntemlerini kullanmışlardır. XGBoost, LightGBM ve CatBoost yöntemlerini Konya ve Washington, D.C. veri setlerine uygulamışlardır. İleri (2024), Güney Kore'de yer alan bir istasyondan 2017 ve 2018 yılları arasında gerçekleşen saatlik bisiklet kiralama sayılarını mevsim bazında incelemiş ve tahmin için Yarasa Algoritması ile optimize edilen Gradyan Artırma Makinesi yöntemini önermiştir. Önerilen yöntemin performansı aynı veri setini inceleyen Sathishkumar ve Yongyun (2024)'un çalışmasında kullandıkları yöntemlere göre daha başarılı bulunmuştur.

Literatürdeki çalışmalar incelendiğinde birçok çalışmada paylaşımlı bisiklet sistemlerinin kullanımı ile ilgili genel bilgiler verildiği ve kullanımı etkileyen faktörlerin incelendiği söylenebilir. Aynı şekilde paylaşımlı bisiklet sürecinin tarihsel gelişimi ve gelecekteki durumu ile ilgili yapılan çalışmalar da mevcuttur. Paylaşımlı bisiklet kullanımına ait talep tahmini ile ilgili yapılan çalışmalarda ise pek çok farklı metod uygulanmış ve etkin sonuçlar elde edilmiştir. Literatürdeki çalışmaların çoğunda şehir nüfusunun oldukça yoğun, talebin de oldukça dinamik olduğu Seul ve New York'ta yer alan paylaşımlı bisiklet sistemi verileri kullanılmış ve analiz edilmiştir. Türkiye'de paylaşımlı bisiklet sistemi uygulamalarına olan talep de sağladıkları avantajlar göz önünde bulundurulduğunda her geçen gün artmaktadır. Uygulamada sıklıkla kullanılmasına rağmen bu alanda literatürde yapılan çalışmalar oldukça sınırlı sayıda olup Türkiye'de paylaşımlı bisiklet talebine yönelik olarak gerçekleştirilen tek çalışmanın yalnızca Ergül Aydın ve diğ. (2023) tarafından yapılmış olan çalışma olduğu görülmüştür. Bu çalışmada Konya ilindeki paylaşımlı

bisiklet sistemi ele alınmış ve 2022 yılının son üç ayına ait bisiklet verileri kullanılmıştır.

Yapmış olduğumuz bu çalışmada Konya'daki paylaşımlı bisiklet sistemi ele alınmış olup 2022 ve 2023 yılına ait Konya Büyükşehir Belediyesi tarafından yayımlanan veriler analiz edilmiştir. Analizler hem saatlik hem de günlük olarak değerlendirilmiş ve farklı performans ölçütlerine göre ayrıntılı olarak sunulmuştur. Ayrıca talebi etkilediği düşünülen meteorolojik veriler ile zamana ait kategorik veriler ve bisiklet istasyonlarından elde edilen lokasyon bilgilerine göre kullanım mesafesi ve kullanım süreleri de tahmin modeline dahil edilmiştir. Tüm bu yönleriyle bildiğimiz kadarıyla yapılan bu çalışma Türkiye'deki paylaşımlı bisiklet sistemini ele alan en kapsamlı çalışmadır.

3. Materyal ve Metot

Bu bölüm veri setinin incelenmesi ve kullanılan talep tahmini algoritmaları olmak üzere iki alt başlıkta incelenmiştir. Çalışmanın her aşamasında araştırma ve yayın etiğine uyulmuştur.

3.1. Veri Seti

Çalışmada kullanılan veri seti için Konya Büyükşehir Belediyesi (2023) tarafından farklı alanlarda verilerin paylaşıldığı açık veri portalından yararlanılmıştır. Belirtilen web sitesinden 01.01.2022-13.12.2023 tarih aralığına ait paylaşımlı bisiklet kullanım verileri (<https://acikveri.konya.bel.tr/dataset/kiralik-bisiklet-kullanim-verileri>) (Konya Açık Veri Portalı, 2024)) alınmıştır. Ayrıca paylaşımlı bisiklet istasyonlarına ait konum bilgileri de <https://aarbike.com.tr/istasyonlar/> web sitesinde yer almaktadır (AAR Bike, 2024). Belirtilen tarih aralığındaki meteorolojik veriler için ise "Python Meteosat" kütüphanesi kullanılmıştır. Belirtilen kütüphane kullanılarak meteorolojik veriler için belirli bir döneme ait saatlik, günlük, haftalık, aylık ve yıllık olmak üzere farklı periyotlarda veriler raporlanabilmektedir.

Paylaşımlı bisiklet verileri ilgili web sitesinde saatlik olarak raporlanmış, ancak yapılan çalışmada bu veriler hem saatlik hem de günlük bazda ele alınmıştır. Bu nedenle meteorolojik veriler de ilgili kütüphaneden hem saatlik hem de günlük bazda çekilmiştir. Bisiklet talebine ait saatlik ve günlük bazda kiralanan bisiklet sayılarına ait önemli tanıtıcı istatistik değerleri Tablo 1'de verilmiştir. Tablo 1'de görüldüğü üzere bisiklet kiralama sayıları saatlik olarak ele alındığı zaman toplam 12767 adet veri mevcut iken günlük olarak 703 adet veri mevcuttur. Ortalama kiralanan bisiklet sayısı saatlik olarak yaklaşık 11 adet iken bu sayı günlük veri setinde yaklaşık olarak 196 dettir. Bir saat diliminde en fazla 119 adet bisiklet kiralılırken günlük olarak en fazla 722 adet bisiklet kiralınmıştır. Bunun yanı sıra her bir veri seti için birinci dördül (%25), medyan (%50) ve

üçüncü dördül (%75) değerleri ile standart sapma değerleri de verilmiştir.

Tablo 1. Bisiklet Talebine Ait Veriler

	Saatlik Kiralanan Bisiklet Sayısı	Günlük Kiralanan Bisiklet Sayısı
Veri	12767	703
Ortalama	10,79	195,93
Standart Sapma	11,95	143,81
En Küçük	1	1
%25	3	79,50
%50	7	191
%75	15	269
En Büyük	119	722

Tablo 2'de Saatlik veri seti (S_Veri) ve Günlük veri seti (G_Veri)'nde kullanılan değişkenler ve tipleri açıklanmıştır. Tablo 2'de görüldüğü üzere veri seti çok sayıda değişken içermektedir. Değişkenlerin tipi (D_Tipi) kategorik (Kat.) ve nümerik (Nüm.) olarak sınıflandırılmıştır. Ayrıca bu değişkenlerin birçoğu kategorize edilerek anlamlı hale getirilmiştir. Yapılan çalışmada literatürde yer alan çalışmalardan farklı olarak istasyon bilgileri de kullanılmıştır. Kullanım mesafesi kategorik veri olarak tanımlanmış olup öncelikle bisikletin alınıp bırakıldığı istasyonlar belirlenerek bu istasyonlar arasındaki mesafe Haversine uzaklık formülü ile hesaplanmıştır. Ardından bu mesafeler çok yakın, yakın, orta, uzak ve çok uzak olmak üzere beş kategoriye ayrılmıştır. Kullanım süresi ise nümerik bir değişken olup bisikletin alındığı zamandan bırakıldığı zamana kadar geçen sürenin dakika cinsinden ölçümü olarak elde edilmiştir. Saatlik ve günlük veriye dönüşüm yapılırken ilgili periyotlarda kiralanan bisiklet sayısına göre ortalama değerler alınmıştır.

Günlük ve saatlik veriler için korelasyon matrisi sırasıyla Şekil 2 ve 3'te gösterilmiştir. Korelasyon matrisleri "Dython" kütüphanesi kullanılarak oluşturulmuştur. Korelasyon matrisi için renkler açıldıkça ilişkilerin arttığı söylenebilir. Buna göre Şekil 2'de gösterilen saatlik veriler için korelasyon matrisi incelendiğinde bisiklet sayısının en çok "saat", "saat dilimi" ve "ay" değişkenleri ile ilişkili olduğu söylenebilir. Şekil 3'te gösterilen günlük veriler için korelasyon matrisinde ise kullanılan bisiklet sayısının en çok "sıcaklık" (en büyük, en küçük ve ortalama sıcaklık), "ay" ve "mevsim" değişkenleri ile ilişkili olduğu görülmektedir.

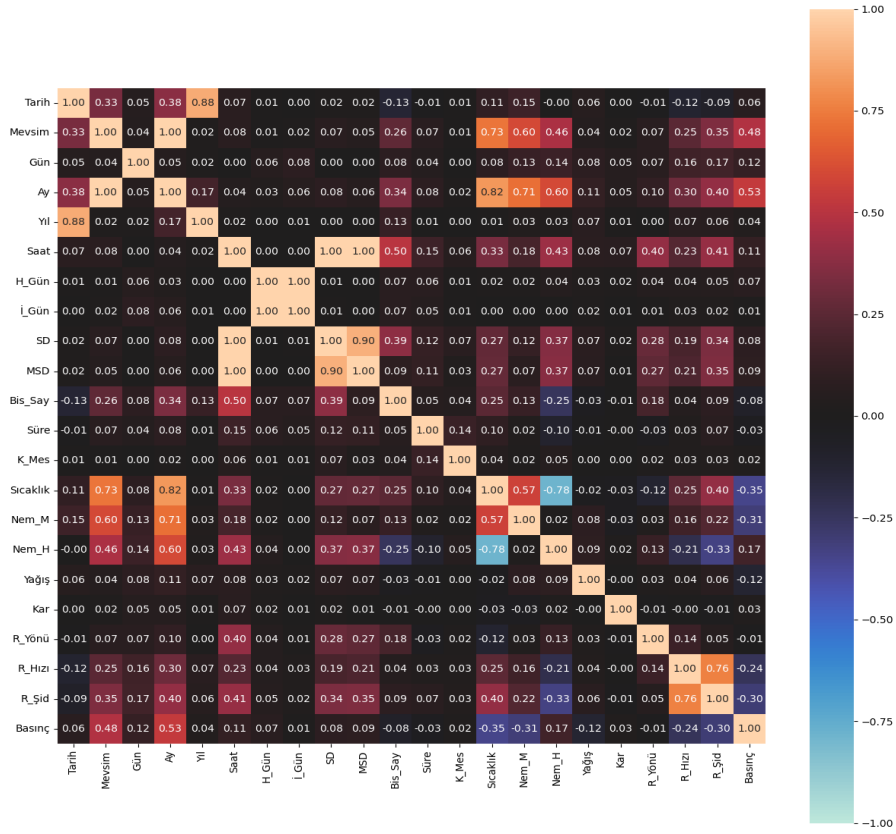
Şekil 4 ve Şekil 5'te sırasıyla saatlik ve günlük veriler için bisiklet talep ortalamasının farklı kategorilere göre değişkenliğini gözlemlemek için kutu diyagramları verilmiştir. Şekil 4'e göre paylaşımlı bisiklet talebi en fazla yaz, en az kış mevsiminde gerçekleşmektedir. Benzer şekilde Mayıs, Haziran ve Temmuz, bisiklet

kullanım talebinin en yoğun olduğu aylar olup Nisan ayında uç değerlerin fazla olduğu görülmektedir. Talebin yoğun olduğu saatler, saat bazında akşam saatleri olup saat dilimi bazında ise 16:00-00:00 arasındadır. Son olarak bisikletlerin genellikle çok yakın mesafeler için tercih edildiği söylenebilir.

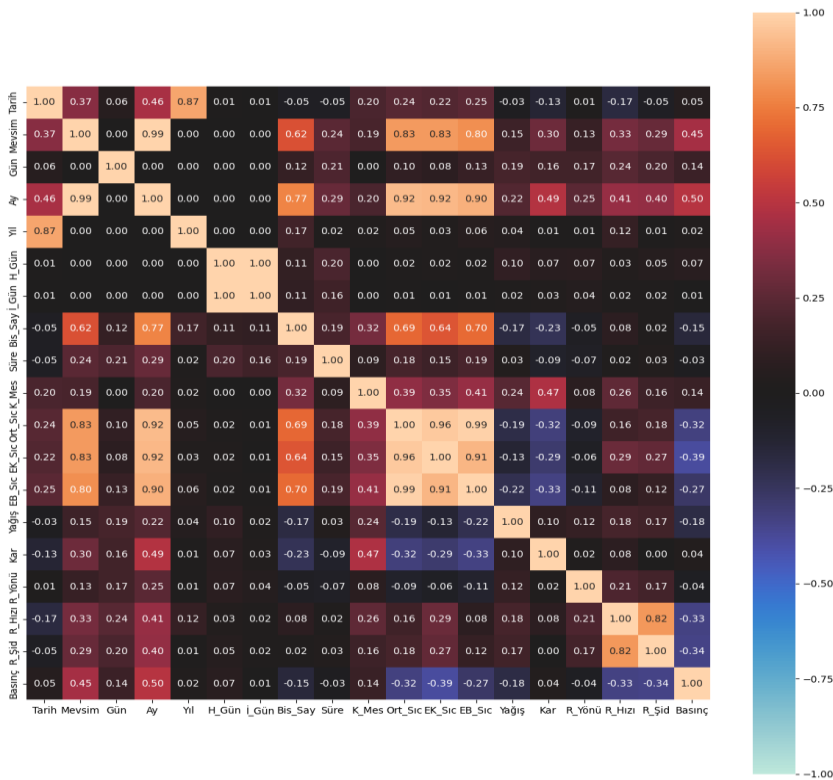
Günlük veriler için hazırlanan kutu diyagramı incelendiğinde ise kategori etkileri daha net bir şekilde görülürken uç değerlerin de saatlik verilere göre azaldığı söylenebilir. Şekil 5'e göre değişkenliğin en fazla olduğu mevsim İlkbahar iken talebin en fazla olduğu mevsim Yaz mevsimidir.

Tablo 2. Tahmin Modelinde Kullanılan Değişkenler

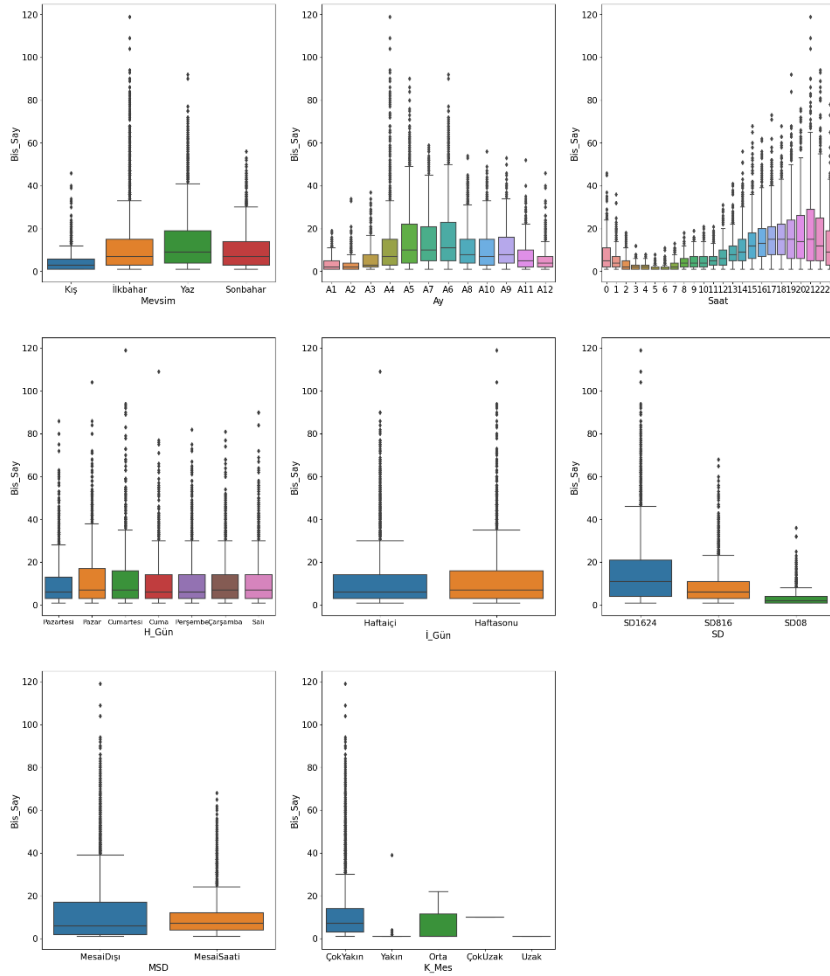
Değişken	D_Tipi	Açıklama	G_Veri	S_Veri
Mevsim	Kat.	1: Kış, 2: İlkbahar, 3: Yaz, 4: Sonbahar	✓	✓
Gün	Kat.	1-31	✓	✓
Ay	Kat.	1-12	✓	✓
Yıl	Kat.	1: 2022, 2: 2023	✓	✓
Haftanın Günleri (H_Gün)	Kat.	1: Pazartesi, 2: Salı, 3: Çarşamba, 4: Perşembe, 5: Cuma, 6: Cumartesi, 7: Pazar	✓	✓
İş Günü (İ_Gün)	Kat.	1: Hafta sonu, 2: Hafta içi	✓	✓
Saat Dilimi (SD)	Kat.	1: 00:00-08:00, 2: 08:00-16:00, 3: 16:00-00:00	x	✓
Mesai Saat Dilimi (MSD)	Kat.	1: Mesai dışı (17:00-08:00), 2: Mesai Saati (08:00-17:00)	x	✓
Kullanım Mesafesi (K_Mes)	Kat.	1: Çok Yakın, 2: Yakın, 3: Orta, 4: Uzak, 5: Çok Uzak	✓	✓
Saat	Kat.	0-23	x	✓
Sıcaklık	Nüm.	Ölçülen sıcaklık değeri	✓	✓
Nem (Humidity) (Nem_H)	Nüm.	Havadaki ölçülen su buharı miktarı	x	✓
Nem (Moisture) (Nem_M)	Nüm.	Belli bir hacim hava içindeki ölçülen buhar	x	✓
Yağış	Nüm.	Ölçülen yağış miktarı	✓	✓
Kar	Nüm.	Ölçülen kar miktarı	✓	✓
Rüzgâr Yönü (R_Yönü)	Nüm.	Belirlenen rüzgâr yönü	✓	✓
Rüzgâr Hızı (R_Hızı)	Nüm.	Ölçülen rüzgâr hızı	✓	✓
Rüzgâr Şiddeti (R_Şid)	Nüm.	Ölçülen rüzgâr şiddeti	✓	✓
Atmosferik Basınç	Nüm.	Ölçülen atmosferik basınç	✓	✓
Kullanma Süresi	Nüm.	Kiralanan bisikletin kullanım süresi	✓	✓
Kullanılan Bisiklet Sayısı (Bis_Say)	Nüm.	Kiralanan bisiklet sayısı	✓	✓



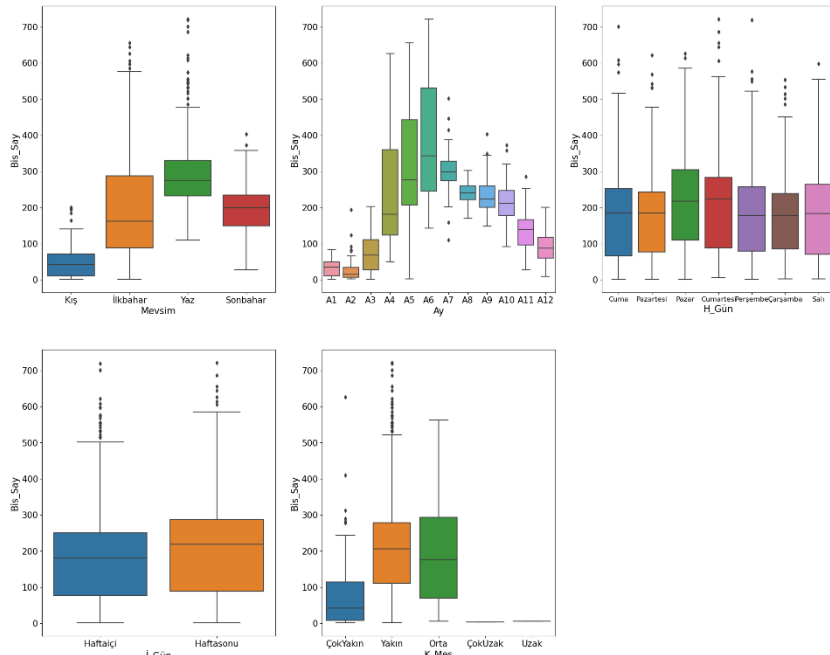
Şekil 2. Saatlik Veriler için Korelasyon Matrisi



Şekil 3. Günlük Veriler için Korelasyon Matrisi



Şekil 4. Saatlik Veriler için Kutu Diyagramı



Şekil 5. Günlük Veriler için Kutu Diyagramı

Talebin en yoğun olduğu aylar ise Nisan, Mayıs ve Haziran ayları iken en az olduğu ay Şubat'tır. İş günü kategorisine göre talebin hafta sonu daha yoğun olduğu söylenebilir. Yakın mesafe bisiklet kullanımının günlük veriler için daha fazla olduğu görülmektedir.

3.2. Talep Tahmini Algoritmaları

Tahmin algoritmaları uygulanmadan önce veriler kullanıma uygun hale getirilmelidir. Literatürde verilerin işlenmesi olarak ele alınan bu aşamada yapılan işlemler şu şekilde özetlenebilir: Makine öğrenmesinde kullanılan algoritmalar değişkenlerin ölççeklerine duyarlı olabilir. Bu nedenle veriler "sklearn" kütüphanesinde yer alan "MinMaxScaler" ile normleştirilmiştir.

Algoritmalarda kullanılacak değişkenler belirlenirken Özyinelemeli Özellik Seçimi (ÖÖS), algoritmasından yararlanılmıştır. Bu teknik makine öğrenimi algoritmalarında kullanılan ve modelin performansını artırmak için en uygun öznelik kümesini belirlemek amacıyla kullanılır. ÖÖS algoritması, başlangıçta tüm özelliklerin bir modelde eğitilmesiyle başlar. Daha sonra, her bir özellik için, modelin performansını etkileyen özelliklerin belirlenmesinde kullanılacak olan bir önem puanı hesaplanır. En düşük önem puanına sahip olan özellikler birer birer çıkarılarak model yeniden eğitilir. Belirli bir kriter veya sınır değeri (örneğin, belirli bir sayıda özellik veya bir performans ölçütü) sağlanana kadar süreç devam eder. Bu süreç sonunda, en uygun özelliklerden oluşan bir özellik kümesi elde edilir. ÖÖS, modelin karmaşıklığını azaltarak aşırı uyum (overfitting) riskini azaltmaya yardımcı olur ve modelin genelleme yeteneğini artırabilir. ÖÖS algoritması, ÇDR için uygulanmış ve algoritma ile saatlik veriler için 93 adet öznelikten 29 adet, günlük veriler içinse 67 adet öznelikten 20 adet öznelik seçilmiştir. Seçilen öznelikler tüm algoritmalar için aynen kullanılmış böylelikle algoritmaların etkinliğinin ölçülmesi amaçlanmıştır. Seçilen özneliklerin hangileri olduğu Ek 1'de verilmiştir. Verilerin işlenmesi adımından sonra, veriler %70'i eğitim ve %30'u test verisi olmak üzere iki gruba ayrılmıştır. Eğitim verileri modelin eğitilmesi aşamasında kullanılırken, test verileri ise geliştirilen modelin performansını değerlendirmek için kullanılır.

Makine öğrenmesinde kullanılan algoritmaların performansı problemin yapısı, veri özellikleri ve veri boyutuna göre farklılık göstermektedir. Çalışma kapsamında kullanılan yöntemler aşağıda ayrıntılı olarak verilmiştir.

Çoklu Doğrusal Regresyon: En basit makine öğrenme tekniğidir ve bağımlı ve bağımsız değişkenler arasındaki neden-sonuç ilişkisini matematiksel olarak ortaya koyar. Doğrusal regresyonda model parametreleri "En Küçük Kareler Yöntemi"ne göre belirlenmektedir. Temel amaç, modelin tahmin ettiği değerler ile gerçek

gözlemler arasındaki farkın karelerinin toplamını (hata kareler toplamını) enküçüklemektir.

Ridge Regresyon: Hoerl ve Kennard (1970) tarafından geliştirilen RR, hata kareler toplamını enküçükleyen eden katsayıları, bu katsayılara bir ceza uygulayarak bulmaktadır. RR ilgisiz değişkenleri çıkarmadan sadece katsayılarını sıfıra yaklaştırarak, tüm değişkenler ile model kurar. Böylece model, katsayıların büyük değerler almasını önleyerek daha dengeli hale gelir.

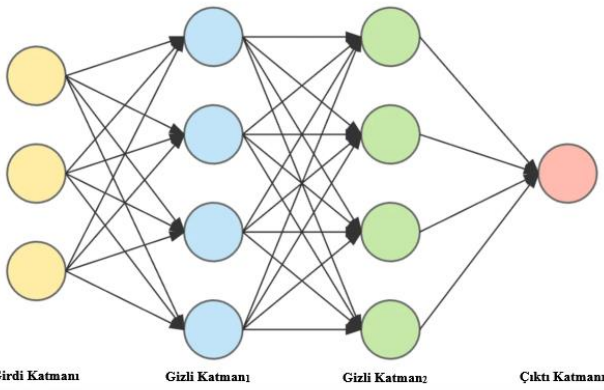
Elastik Net: Zou ve Hastie (2005) tarafından önerilmiştir. RR tarzı cezalandırma ve Lasso regresyon tarzı değişken seçimi yaparak her iki düzenlemenin zayıf yönlerini dengelemekte böylece daha esnek bir model sunmaktadır. Özellikle yüksek boyutlu veri setlerinde ve birçok değişkenin olduğu durumlarda kullanışlıdır.

Karar Ağaçları: Hem sınıflandırma hem de regresyon modelleri için kullanılan ve ağaç yapısı şeklinde ifade edilen bir yöntemdir. Bu yapı düğümler, dallar ve yapraklardan oluşmaktadır. Ağaçta yer alan her bir düğüm bir karar noktasını, her dal bir karar kuralını ve her yaprak ise hedef değişkenin tahmin değerini ifade etmektedir. KA algoritması, veri setindeki değişkenleri ve hedef değişkeni analiz eder ve en uygun karar kuralını belirleyen ağacı oluşturur. Ayrıca uç değerlere karşı duyarlı olmadığından uç değerlerin çok olduğu veri setleri için kullanışlıdır.

Rassal Orman Algoritması: İçeriğinde birden fazla karar ağacı kullanılan ve her karar ağacı için öğrenme sonuçlarının entegre edildiği bir yöntemdir. Her bir karar ağacı, ana veri setinden rassal örneklemeler alınarak oluşturulan farklı alt veri setiyle eğitilmektedir. Yöntemin doğruluğu ormanın büyüklüğü ile bireysel ağaçların gücü ve korelasyonuna bağlıdır. Çok sayıda verinin bulunduğu ve birçok değişkenin kullanıldığı veri setleri için avantajlıdır.

Destek Vektör Makineleri: DVM istatistiksel öğrenme teorisinden gelen yapısal risk minimizasyonu ilkesine dayanan bir yöntemdir. Yapısal risk minimizasyonunun amacı, görünmeyen ve rastgele seçilmiş bir test örneğindeki gerçek hatayı azaltmaktır.

Yapay Sinir Ağları: Biyolojik sinir ağlarının sinir hücrelerinden oluşması gibi, YSA da birbiriyle bağlantılı çok sayıda yapay sinir hücresinden meydana gelmektedir. Bu nedenle biyolojik sinir hücrelerinin sanal bir modeli olarak düşünülebilir. Büyük miktarda veri üzerinde karmaşık desenler bulma ve belirli görevleri gerçekleştirme yeteneğine sahip bir dizi matematiksel modelden oluşan YSA'ya ait bir mimari örneği Şekil 6'da verilmiştir.



Şekil 6. Bir YSA'nın Mimarisi

Bir YSA, birbirinden bağımsız olarak bir girdi katmanı, bir çıktı katmanı ve en az bir tane gizli katmandan oluşmaktadır. Her katman birbirinden farklı sayıda düğüme sahip olabilir. Girdi katmanında bilgiler alınarak gizli katmana aktarılmakta, gizli katmanda ise bu bilgiler işlenerek çıktı katmanına iletilmektedir. YSA'da yer alan gizli katmanlar ve bu katmanlardaki düğüm (nöron) sayıları problem yapısına göre değişiklik gösterebilir.

4. Deneysel Çalışmalar

Çalışmada kullanılan yöntemlere ilişkin hiperparametre değerleri Tablo 3'te verilmiştir. Bu çalışma kapsamına hiperparametreler için bir optimizasyon çalışması dahil edilmemiş ancak pilot koşullarla parametrelerin farklı seviyeleri değerlendirilerek Tablo 3'teki değerler belirlenmiştir.

Geliştirilen modeller determinasyon katsayısı (R^2), ortalama mutlak hata (MAE), ortalama hata karesi (MSE) ve ortalama hata karesi kökü (RMSE) performans ölçütlerine göre karşılaştırılmıştır. Bu performans ölçütleri aşağıda tanımlanmıştır.

- Determinasyon katsayısı: bağımlı değişkendeki (y) değişkenliğin ne kadarının bağımsız değişkenlerdeki (x) değişkenlikle açıklanabildiğini ifade eder. R^2 0 ile 1 arasında bir değer alır ve 1'e yaklaştıkça, modelin veriler üzerindeki uyumu artar.
- Ortalama Mutlak Hata: MAE, her tahminin mutlak değerli hatasının yani model tahmini ile hedef değer arasındaki mutlak farkın ortalamasını verir.
- Ortalama Hata Karesi: MSE model tahmini ile hedef değer arasındaki farkların karelerinin ortalamasını verir. MAE, hatanın mutlak değerlerini alırken, MSE hatanın karelerini alır. Bu nedenle, MSE büyük hatalara daha duyarlıdır.
- Kök Ortalama Hata Karesi: RMSE, MSE'nin karekökü alınarak elde edilir. Hem hatanın büyüklüğünü hem de hataların varyansını dikkate aldığı için modelin performansını değerlendirmede faydalı bir araçtır.

Tablo 3. Algoritmaların Hiperparametre Değerleri

Algoritma	Hiperparametre	Değer
Ridge Regresyon	α solver	1 auto
Elastik Net	α l1_ratio	0,001 0,300
Karar Ağaçları	maksimum derinlik	10
	En küçük örnek sayısı (iç düğüm)	5
	En büyük yaprak düğüm sayısı	40
	splitter	best
Rassal Orman	Ağaç sayısı	150
	Maksimum özellik	auto
	Maksimum derinlik	none
Destek Vektör Makineleri	C (Ceza Parametresi)	1
	ϵ (Sapma için eşik değer)	0,100
	γ (Yayımla parametresi)	0,100
	Kernel fonksiyonu	"rbf" (Radyal tabanlı çekirdek fonksiyonu)
Yapay Sinir Ağları	Küme boyutu	10
	Girdi katmandaki düğüm sayısı	15
	Çıktı katmandaki düğüm sayısı	1
	Aktivasyon fonksiyonu	ReLU (Doğrultulmuş doğrusal birimler)
	Kayıp fonksiyonu	MSE
	Optimizör	adam (adaptif moment tahmini)
	Epochs	100

y_i : i . gözlemin hedef değeri; \hat{y}_i : i . gözlemin tahmin edilen değeri; \bar{y} : hedef değerlerin ortalaması, n : örneklem büyüklüğü ve e_i : i . gözleme ilişkin hata ve $e_i = y_i - \hat{y}_i$ olmak üzere, Eşitlik (1)-(4)'te performans ölçütlerini hesaplamak için kullanılan formülasyonlar verilmiştir.

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

$$MAE = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (2)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n e_i^2 \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (4)$$

Performans ölçütleri kullanılarak yapılan yöntemlerin karşılaştırmaları Tablo 4'te verilmiştir.

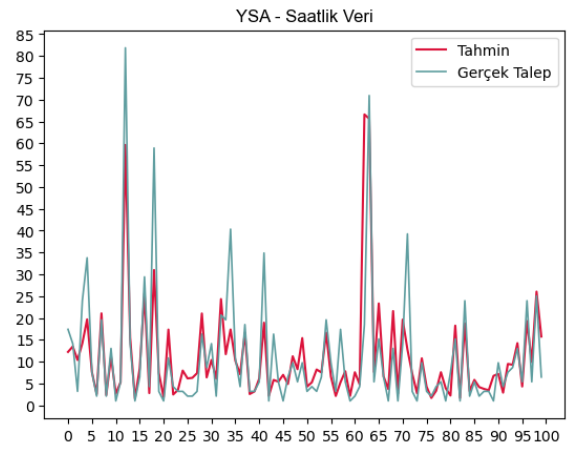
Elde edilen veriler hem saatlik hem de günlük olarak analiz edilmiş ve karşılaştırılmıştır. Buna göre saatlik olarak elde edilen verilerde R^2 değeri en yüksek olan model 0,67 ile YSA yöntemidir. En düşük R^2 değeri ise 0,43 ile EN yöntemi ile elde edilmiştir. MAE, MSE ve RMSE değerleri açısından ise bu değerler tüm modellerde birbirine çok yakın olmakla birlikte yine YSA yönteminde en düşük değerlere ulaşılmıştır.

Tablo 4. Yöntemlerin Karşılaştırılması

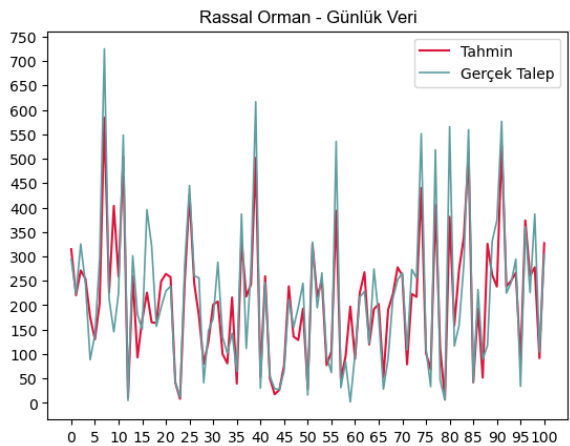
Modeller	Performans Ölçütleri	Saatlik Veri	Günlük Veri
Çoklu Doğrusal Regresyon	R^2	0,450	0,780
	MAE	0,054	0,079
	MSE	0,006	0,012
	RMSE	0,079	0,109
Ridge Regresyon	R^2	0,450	0,680
	MAE	0,054	0,082
	MSE	0,006	0,013
	RMSE	0,079	0,116
Elastik Net	R^2	0,430	0,690
	MAE	0,055	0,080
	MSE	0,006	0,013
	RMSE	0,080	0,114
Karar Ağaçları	R^2	0,540	0,660
	MAE	0,046	0,075
	MSE	0,005	0,014
	RMSE	0,071	0,119
Rassal Orman	R^2	0,640	0,810
	MAE	0,041	0,060
	MSE	0,004	0,008
	RMSE	0,064	0,090
Destek Vektör Makineleri	R^2	0,600	0,770
	MAE	0,050	0,071
	MSE	0,005	0,010
	RMSE	0,067	0,099
Yapay Sinir Ağları	R^2	0,670	0,780
	MAE	0,039	0,067
	MSE	0,004	0,009
	RMSE	0,061	0,096

Günlük veriler için elde edilen sonuçlara göre, R^2 değeri en yüksek olan model RO yöntemidir. ÇDR ve YSA ise 0,78 R^2 değeri ile RO modeline çok yakın sonuçlar elde etmişlerdir. En düşük performans ise KA yöntemine aittir. MAE, MSE ve RMSE sonuçlarına göre de en iyi sonuçlar RO ile elde edildiğinden günlük veriler için en iyi performansın RO yöntemi ile elde edildiği söylenebilir.

Şekil 7 ve 8 'de sırasıyla saatlik ve günlük veriler için en iyi performansın elde edildiği YSA ve RO algoritmaları ile rassal olarak seçilen 100 adet test verisi için elde edilen tahmin değerleri karşılaştırılmıştır. Şekil 7 ve 8'e göre, grafikteki örtüşmeler değerlendirildiğinde her iki yöntem için de gerçekleşen talep ile tahmin edilen talebin birbirlerine çok yakın olduğu söylenebilir.



Şekil 7. Saatlik Veri YSA Algoritması Tahmin Performansı



Şekil 8. Günlük Veri Rassal Orman Algoritması Tahmin Performansı

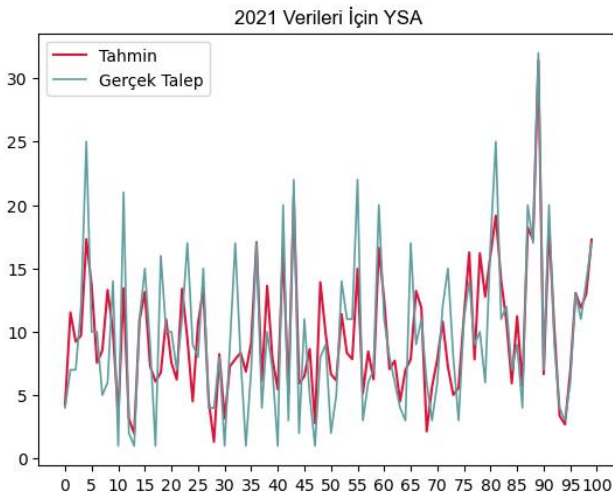
Ergül Aydın ve diğ. (2023) tarafından yapılan çalışmada 2021 yılı Ekim, Kasım ve Aralık aylarında paylaşımlı bisiklet talep verileri kullanılarak XGBoost, LightGBM ve CatBoost yöntemleri ile elde edilen tahmin sonuçları karşılaştırılmıştır. Yapılan analizlerde performansı en

iyi yöntem CatBoost yöntemi olup R^2 değeri 0,66 olarak elde edilmiştir. Yazarlar veri setini saatlik bazda ele aldıklarından bu çalışmada saatlik veriler için en iyi sonucun elde edildiği YSA algoritması aynı veri setine uygulanmıştır. Elde edilen sonuçlar Tablo 5'te raporlanmıştır. ÖÖS yöntemi ile bu veri seti için hangi özneliklerin kullanıldığı ise Ek 2'de verilmiştir.

Tablo 5. 2021 Verileri için Karşılaştırma

Modeller	Performans Ölçütleri	Saatlik Veri
CatBoost (Ergül Aydın ve diğ. (2023))	R^2	0,6564
	MAE	1,3966
	MSE	6,3267
	RMSE	0,2889
YSA Algoritması	R^2	0,6793
	MAE	0,0864
	MSE	0,0128
	RMSE	0,1130

Tablo 5'te görüldüğü üzere YSA algoritması ile CatBoost algoritmasından tüm performans ölçütleri için daha iyi sonuçlar elde edilmiştir. Şekil 9'da 2021 yılı verileri kullanılarak rassal olarak seçilen 100 adet test verisi için YSA ile elde edilen tahmin değerleri ile gerçek talep değerleri karşılaştırılmıştır. Grafikteki örtüşmeler değerlendirildiğinde özellikle test verilerinin genelinde gerçekleşen talep ile tahmin edilen talebin birbirlerine çok yakın olduğu söylenebilir.



Şekil 9. 2021 Verileri için YSA Algoritması Tahmin Performansı

5. Sonuçlar ve Tartışma

Paylaşımlı bisiklet sistemleri tüm Dünya'da yaygın olarak kullanılmakta olup literatürde bu konuda pek çok çalışma yapılmıştır. Türkiye'de ise paylaşımlı bisikletlere ait ilginin her geçen gün arttığı ancak yapılan çalışmaların henüz oldukça sınırlı olduğu görülmüştür.

Paylaşımlı bisiklet sistemlerinde hangi bölgelere bisiklet istasyonu kurulacağı, hangi istasyonda ne kadar bisiklet olması gerektiği gibi konular kaynakların verimli kullanılması ve müşteri memnuniyeti açısından oldukça önemlidir. Bir istasyonda talep edilenden çok daha fazla bisiklet bulundurulması başka bir istasyonda gerekli sayıda bisikletin bulunmamasına sebep olabilir. Ayrıca müşteriler sistemin sürekliliği garanti edilmedikçe paylaşımlı bisiklet kullanımından kaçınılmaktadırlar. Sistem sağlayıcıları için hangi dönemlerde hangi istasyonlardaki bisikletlerin servis ve bakım ihtiyacı olacağı da oldukça önemlidir. Belirli istasyonlardaki talep artışının belirlenmesi bakım planlamasını kolaylaştırmakta ve bisikletlerin sürekli kullanılabilir hale getirilmesinin sağlanmasına yardımcı olmaktadır. Tüm bu süreçten en yüksek düzeyde yararlanabilmek adına sürecin etkin bir şekilde yönetilmesi gerekmektedir. Bu nedenle, talep tahmini konusu paylaşımlı bisiklet sistemlerinin hem kullanıcıları hem de yöneticileri açısından oldukça kritik bir öneme sahiptir. Doğru talep tahmini, sistemin büyümesi veya iyileştirilmesi için yapılan yatırım kararlarını da etkilemektedir. Bu çalışmada Türkiye'nin en uzun bisiklet yoluna sahip olan Konya ilinde kullanılmakta olan paylaşımlı bisiklet sistemleri incelenmiş ve detaylı bir talep tahmin prosedürü gerçekleştirilmiştir. Bildiğimiz kadarıyla Türkiye'deki paylaşımlı bisiklet sistemleri için bu zamana kadar yapılmış en kapsamlı çalışma bu çalışmadır. Çalışmada Konya Büyükşehir Belediyesi tarafından paylaşılan 01.01.2022-13.12.2023 tarih aralığına ait paylaşımlı bisiklet kullanım verileri kullanılmıştır. Bu veri seti hem saatlik hem de günlük olarak analiz edilmiştir. Verilerin işlenmesi aşamasında nümerik değişkenler normleştirilmiş, ayrıca değişken seçimi ile modele katkısı olmayan ve model karmaşıklığını artıran bazı değişkenler modelden çıkarılmıştır. Talep tahmininde ÇDR, RR, EN, RO, KA, DVM ve YSA yöntemleri kullanılmıştır. Geliştirilen yıllık ve saatlik talep tahmin modelleri ve kullanılan makine öğrenme algoritmaları farklı performans ölçütlerine göre karşılaştırılmıştır. Buna göre saatlik veriler için talep tahmininde en iyi sonuçlar YSA ile elde edilirken, günlük veriler için RO'nun daha iyi performans gösterdiği görülmüştür. Geliştirilen tahmin modelleri ile saatlik veriler için R^2 değeri 0,67 ve günlük veriler için de R^2 değeri 0,81 olarak elde edilmiştir. Buna göre günlük modelde bisiklet talebine ait değişkenliğin yaklaşık %80'inin kullanılan bağımsız değişkenlerdeki değişkenlik ile açıklanabildiği söylenebilir.

Büyük verilerin analizinde basit istatistiksel analizlerin performansı kısıtlı olacağından, yapılan çalışmada makine öğrenmesi teknolojilerinden yararlanılarak hem güncel bir konu ele alınmış hem de gerçeğe yakın tahminlerin yapılması sağlanmıştır.

Yapılan çalışmada veri setinde paylaşılan bilgilerin tümü değişken olarak anlamlı bir şekilde kullanılmaya

çalışılmıştır. Bu veri seti Türkiye’de bisiklet paylaşımına ait paylaşılan en kapsamlı veri setlerinden birisidir. Buna karşın çalışmanın bazı sınırlandırmaları bulunmaktadır. Özellikle kullanıcılara ait yaş, cinsiyet gibi kişisel özellikler anonim yapıda verildiğinden analize dahil edilememiştir. Bunun dışında bisiklet istasyonları hastane, üniversite, alışveriş merkezleri gibi halkın yoğun olduğu bölgelere yakınlıklarına göre analiz edilerek istasyonların hangi alanlarda kurulmasının işletmeler açısından daha faydalı olacağı araştırılabilir. Bir başka çalışmada toplu taşıma ve paylaşımlı bisiklet maliyetleri dikkate alınarak farklı ulaşım türlerine olan talep yoğunlukları karşılaştırılabilir. Gelecekte yapılması planlanan çalışmalarda Türkiye’de paylaşımlı bisiklet sistemini kullanan tüm iller için aynı ya da farklı talep tahmin yöntemleri ve performans ölçütleri kullanılarak çalışma genişletilebilir. Ayrıca farklı algoritma parametreleri de ayrıntılı bir şekilde incelenebilir.

Araştırmacıların Katkısı

Bu çalışmada; Gözde CAN ATASAGUN, verilerin analizi, kaynak araştırması, istatistiksel analizler, makalenin ve sonuçlarının oluşturulması; Ece YAĞMUR, verilerin analizi, kavramsal tasarım, istatistiksel analizler, makalenin ve sonuçlarının oluşturulması konularında katkı sağlamışlardır.

Çıkar Çatışması

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir.

Kaynaklar

AAR Bike. Erişim Adresi: <https://aarbike.com.tr/istasyonlar/> (Erişim Tarihi: 09.01.2024)

Bordagaray, M., Dell’Olio, L., Fonzone, A., & Ibeas, Á., (2016). Capturing the conditions that introduce systematic variation in bike-sharing travel behavior using data mining techniques. *Transportation Research Part C: Emerging Technologies*, 71, 231-248. <https://doi.org/10.1016/j.trc.2016.07.009>

Candan, S. (2003). Ulaşım sistemlerinin bütünleştirilmesi açısından Ankara uygulamalarının değerlendirilmesi ve geliştirme önerileri (Yüksek Lisans Tezi). Gazi Üniversitesi Fen Bilimler Enstitüsü, Ankara.

Chibwe, J., Heydari, S., Imani, A. F., & Scurtu, A. (2021). An exploratory analysis of the trend in the demand for the London bike-sharing system: From London Olympics to Covid-19 pandemic. *Sustainable Cities and Society*, 69, 102871. <https://doi.org/10.1016/j.scs.2021.102871>

DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of Public Transportation*, 12(4), 41-56. <http://doi.org/10.5038/2375-0901.12.4.3>

Eren, E., Katanalp, B. Y., Yıldırım, Z. B. ve Uz, V. E. (2018). Türkiye’deki bisiklet paylaşım programları, SETSCI-Conference Proceedings, 1507-1513.

Ergül Aydın, Z., İçmen Erdem, B., & Erzurum Çiçek, Z. I. (2023). Prediction bike-sharing demand with gradient boosting methods. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 29(8), 824-832. <https://doi.org/10.5505/pajes.2023.39959>

Fishman, E., Washington, S., & Haworth, N. (2013). Bike share: a synthesis of the literature. *Transport Reviews*, 33(2), 148-165. <https://doi.org/10.1080/01441647.2013.775612>

Hoerl, A. E. & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1), 55-67. <https://doi.org/10.1080/00401706.1970.10488634>

İleri, K. (2024). Yarasa algoritması ile optimize edilmiş GBM modeli kullanarak mevsim bazlı bisiklet kiralama sayılarının tahmini. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 39(4), 2631-2642. <https://doi.org/10.17341/gazimmfd.1362302>

Jahanshahi, D., Van Wee, B., & Kharazmi, O. A. (2019). Investigating factors affecting bicycle sharing system acceptability in a developing country: The case of Mashhad, Iran. *Case Studies on Transport Policy*, 7(2), 239-249. <https://doi.org/10.1016/j.cstp.2019.03.002>

Jiang, W. (2022). Bike sharing usage prediction with deep learning: A survey. *Neural Computing Applications*, 34(18), 15369-15385. <https://doi.org/10.1007/s00521-022-07380-5>

Karagöz, B. (2019). Yerel yöneticilerin bisiklet ulaşımına bakış açısı: Konya örneği (Yüksek Lisans Tezi). Konya Teknik Üniversitesi Lisansüstü Eğitim Enstitüsü, Konya.

Konya Açık Veri Portalı. Erişim Adresi: <https://acikveri.konya.bel.tr/dataset/kiralik-bisiklet-kullanim-verileri> (Erişim Tarihi: 09.01.2024)

Lin, L., He, Z., & Peeta, S. (2018). Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transportation Research Part C: Emerging*

- Technologies, 97, 258-276.
<https://doi.org/10.1016/j.trc.2018.10.011>
- Lorasokkay, M. A. ve Ağırdir, M. L. (2011). Konya kentçi ulaşımında bisiklet. *Engineering Sciences*, 6(4), 870-881. Erişim adresi: <https://dergipark.org.tr/tr/download/article-file/186142>
- Madak, N. ve Bardakçı, A. (2021). İstasyonsuz bisiklet paylaşım sistemlerinin yükselişi ve düşüşü üzerine. *Pamukkale Üniversitesi İşletme Araştırmaları Dergisi*, 8(2), 668-689. <https://doi.org/10.47097/piar.1026156>
- McBain, C. & Caulfield, B. (2018). An analysis of the factors influencing journey time variation in the cork public bike system. *Sustainable Cities and Society*, 42, 641-649. <https://doi.org/10.1016/j.scs.2017.09.030>
- Midgley, P. & Affairs, S. (2011). Bicycle-sharing schemes: enhancing sustainable mobility in urban areas. *United Nations, Department of Economic*, 8, 1-12.
- Morton, C. (2018). Appraising the market for bicycle sharing schemes: Perceived service quality, satisfaction, and behavioural intention in London. *Case Studies on Transport Policy*, 6(1), 102-111. <https://doi.org/10.1016/j.cstp.2017.11.003>
- NTV. Konya'da sadece bisikletliler için tasarlanan tramvay yolculuğa başladı. (2020). Erişim Adresi: https://www.ntv.com.tr/galeri/seyahat/konyada-sadece-bisikletliler-icin-tasarlanan-tramvay-yolculuga-basladi,o2C70nJLvUm812BHRrCBtA/Ne04gVqc70yTU9R_tseSuA (Erişim Tarihi: 25.12.2023)
- Pan, Y., Zheng, R. C. Zhang, J., & Yao, X., (2019). Predicting bike sharing demand using recurrent neural networks. *Procedia Computer Science*, 147, 562-566. <https://doi.org/10.1016/j.procs.2019.01.217>
- Pucher, J. & Buehler, R. (2008). Cycling for everyone: lessons from Europe. *Transportation Research Record*, 2074(1), 58-65. <https://doi.org/10.3141/2074-08>
- Sathishkumar, V. E., Park, J., & Cho, Y. (2020). Using data mining techniques for bike sharing demand prediction in metropolitan city. *Computer Communications*, 153, 353-366. <https://doi.org/10.1016/j.comcom.2020.02.007>
- Sathishkumar, V. E. & Yongyun, C. (2020). A rule-based model for Seoul bike sharing demand prediction using weather data. *European Journal of Remote Sensing*, 53(sup1), 166-183. <https://doi.org/10.1080/22797254.2020.1725789>
- Sathishkumar, V. E. & Yongyun, C. (2024). Season wise bike sharing demand analysis using random forest algorithm. *Computational Intelligence*, 40(1), e12287. <https://doi.org/10.1111/coin.12287>
- Shaheen, S. A., Cohen, A. P., & Martin, E. W. (2013). Public bikesharing in North America: Early operator understanding and emerging trends. *Transportation Research Record*, 2387(1), 83-92. <https://doi.org/10.3141/2387-10>
- Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, present, and future. *Transportation Research Record*, 2143 (1), 159-167. <https://doi.org/10.3141/2143-20>
- Si, H., Shi, J.-g., Wu, G., Chen, J., & Zhao, X. (2019). Mapping the bike sharing research published from 2010 to 2018: A scientometric review. *Journal of Cleaner Production*, 213, 415-427. <https://doi.org/10.1016/j.jclepro.2018.12.157>
- Wang, B. & Kim, I. (2018). Short-term prediction for bike-sharing service using machine learning. *Transportation Research Procedia*, 34, 171-178. <https://doi.org/10.1016/j.trpro.2018.11.029>
- Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, 95, 47-60. <https://doi.org/10.1016/j.trc.2018.07.013>
- Zhang, J., Meng, M., Koh, P. P., & Wong, Y. D. (2021). Life duration of bike sharing systems. *Case Studies on Transport Policy*, 9(2), 674-680. <https://doi.org/10.1016/j.cstp.2021.03.005>
- Zhao, J., Deng, W., & Song, Y. (2014). Ridership and effectiveness of bikesharing: The effects of urban features and system characteristics on daily use and turnover rate of public bikes in China. *Transport Policy*, 35, 253-264. <https://doi.org/10.1016/j.tranpol.2014.06.008>
- Zhao, N., Zhang, X., Banks, M. S., & Xiong, M. (2018). Bicycle sharing in China: Past, present, and future. *SAIS 2018 Proceedings*, 11.
- Zou, H. & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

Ekler**Ek 1. Seçilen Öznitelikler**

Günlük veriler için öznitelikler "'Ort_Sıc', 'EK_Sıc', 'Kar', 'R_Hızı', 'R_Şid', 'Basınç', 'Sonbahar', 'Yaz', 'İlkbahar', 'G10', 'G11', 'G15', 'A3', 'A4', 'A5', 'A6', 'A8', 'A9', 'Y2023', 'Haftasonu'";

Saatlik veriler için öznitelikler "'Sıcaklık', 'Nem_M', 'Nem_H', 'Kar', 'R_Şid', 'Basınç', 'Yaz', 'İlkbahar', 'A3', 'A5', 'A6', 'A8', 'Y2023', 'S1', 'S13', 'S14', 'S15', 'S16', 'S17', 'S18', 'S19', 'S2', 'S20', 'S21', 'S22', 'S23', 'S8', 'SD1624', 'MesaiSaati'" olarak belirlenmiştir.

Ek 2. 2021 Veriseti için Seçilen Öznitelikler

2021 veriseti için öznitelikler 'Sıcaklık', 'Yağış', 'Kar', 'R_Hızı', 'R_Şid', 'Sonbahar', 'G24', 'A12', 'Cumartesi', 'Pazar', 'S1', 'S10', 'S11', 'S12', 'S13', 'S14', 'S15', 'S16', 'S17', 'S2', 'S3', 'S4', 'S5', 'S6', 'S7', 'S9', 'Haftasonu', 'SD1624', 'SD816', 'MesaiSaati' olarak belirlenmiştir.