



sustainability



Makale

Benzinli ve Dizel Araçlar için GPS Tabanlı Makine Öğrenimi Modelleri Kullanarak Yakıt Tüketimi ve Emisyonların Tahmin Edilmesi

Fahd Alazemi, Asmaa Alazmi, Mubarak Alrumaidhi ve Nick Molden

Konu

Yeşil Operasyon ve Sürdürülebilir Çevrenin Teşvik Edilmesi için Yapay Zeka ve IoT

Düzenleyen

Dr. Chia-Huei Wu ve Prof. Dr. Wei Liu



<https://doi.org/10.3390/su17062395>

Makale

Benzinli ve Dizel Araçlar için GPS Tabanlı Makine Öğrenimi Modelleri Kullanarak Yakıt Tüketimi ve Emisyonların Tahmin Edilmesi

Fahd Alazemi ¹, Asmaa Alazmi ^{2,*} , Mubarak Alrumaidhi ³  ve Nick Molden ⁴ ¹ Güvenlik Departmanı, Kuveyt Belediyesi, Kuveyt Şehri 12027, Kuveyt; falazemi@vt.edu² İnşaat Proje Departmanı, Kamu Çalışmaları Bakanlığı, Güney Surra 12011, Kuveyt³ İnşaat Mühendisliği Bölümü, Teknolojik Çalışmalar Koleji, Uygulamalı Eğitim ve Öğretim Kamu Kurumu, Shuwaikh 70654, Kuveyt; ms.alrumaidhi@paaet.edu.kw⁴ Emissions Analytics, C R Bates Industrial Estate, 2, Stokenchurch, High Wycombe HP14 3PD, Birleşik Krallık; nick@emissionsanalytics.com

* Yazışma: engine85er@vt.edu veya dr_asmaa_alazmi@mpw.gov.kw

Özet: Ulaşım sektörü insanların, malların ve hizmetlerin hareketini sağlamada hayati bir rol oynamaktadır, ancak aynı zamanda enerji tüketimine ve sera gazı emisyonlarına da önemli bir katkıda bulunmaktadır. Yakıt tüketimi ve kirlетici emisyonların doğru modellenmesi, etkin ulaşım yönetimi ve çevresel sürdürülebilirlik için kritik öneme sahiptir. Bu çalışma, yakıt tüketimi ve egzoz emisyonlarını (CO_2 ve NO_x) modellemek için benzinli ve dizel araçlardan alınan gerçek dünya sürüş verilerinin kullanımını araştırmaktadır. Modeller, hem araç hızı hem de GPS hız verilerinden türetilen girdilerle birlikte topluluk torbalı ve karar ağacı algoritmaları kullanılarak geliştirilmiştir. Sonuçlar, tüm veri kümelerinde karar ağacı modelinden sürekli olarak daha iyi performans gösteren toplu torbalı model ile yüksek tahmin doğruluğu göstermektedir. Özellikle, GPS hızına dayalı modeller araç hızına dayalı modellerle karşılaştırılabilir performans göstererek gerçek zamanlı tahminler için GPS verilerini kullanmanın fizibilitesini ortaya koymuştur. Ayrıca, kombine benzinli ve dizel motor veri seti CO_2 emisyon tahminlerinin doğruluğunu artırırken, sadece benzinli veri seti yakıt tüketimi için en yüksek doğruluğu sağlamıştır. Bu bulgular, gerçek zamanlı izleme ve politika oluşturmayı geliştirmek için GPS tabanlı makine öğrenimi modellerini Akıllı Ulaşım Sistemlerine (ITS) entegre etme potansiyelinin altını çizmektedir. Gelecekteki araştırmalar, tahmin yeteneklerini daha da geliştirmek için ağır hizmet araçlarının, ek kirlетicilerin ve gelişmiş modelleme tekniklerinin dahil edilmesini araştırmalıdır.



Akademik Editör: Bin Ji Alındı: 11

Ocak 2025

Revize: 5 Mart 2025

Kabul edildi: 7 Mart 2025

Yayınlanma tarihi: 9 Mart 2025

Alıntı: Alazemi, F.; Alazmi, A.; Alrumaidhi, M.; Molden, N. Benzinli ve Dizel Araçlar için GPS Tabanlı Makine Öğrenimi Modelleri Kullanarak Yakıt Tüketimi ve Emisyonların Tahmin Edilmesi. *Sürdürülebilirlik* **2025**, *17*, 2395. <https://doi.org/10.3390/su17062395>

Telif hakkı: © 2025 yazarlar tarafından. Lisans sahibi MDPI, Basel, İsviçre. Bu makale Creative Commons Attribution (CC BY) lisansı (<https://creativecommons.org/licenses/by/4.0/>) .hüküm ve koşulları altında dağıtılan açık erişimli bir makaledir

Anahtar Kelimeler: yakıt tüketimi; egzoz emisyonları; araç hızı; GPS hızı; benzinli araç; dizel araç

1. Giriş

Ulaşım sistemi, insanların, malların ve hizmetlerin bölgeler arasında hareket etmesini sağlayarak ekonomik büyümenin ve toplumsal gelişmenin temel taşlarından biridir. Bununla birlikte, küresel enerji tüketimine ve sera gazı (GHG) emisyonlarına da önemli ölçüde katkıda bulunarak ciddi çevre ve halk sağlığı sorunlarına yol açmaktadır. Avrupa Birliği'nde ulaştırma sektörü 2017 yılında ulaştırma kaynaklı sera gazı emisyonlarının %72'sini oluştururken, binek araçlar tek başına %44'üne katkıda bulunmuştur [1]. Benzer şekilde, Pekin'de binek otomobiller 2012 yılında 11,3 milyon ton karbondioksit (CO_2) salınımına yol açarak şehrin toplam CO_2 salınımının %75,5'ini oluşturmuştur [2]. Bu rakamlar, politikaları bilgilendirmek ve çevresel etkileri azaltmak için yakıt tüketimi ve araç emisyonlarının doğru bir şekilde tahmin edilmesine yönelik acil ihtiyacın altını çizmektedir.

Araç emisyonları araç tipi, motor özellikleri, operasyonel parametreler (örn. hız, ivme ve sürüş şekilleri) ve çevresel koşullar gibi çeşitli faktörlerden etkilenmektedir. Önceki çalışmalar, araç hızının, özellikle de anlık hızın, sürüş davranışındaki dinamik değişiklikleri yakaladığı için emisyon modellemesinde çok önemli bir rol oynadığını göstermiştir [3]. Yakıt türü ve sınıf ve kimyasal bileşim gibi özellikleri de emisyonları önemli ölçüde etkiler. Motor teknolojisindeki ilerlemeler, daha düşük tekerlekten tekerleğe emisyonu sahip dizel motorlar da dahil olmak üzere daha verimli motorlara yol açmış olsa da [4], ulaşım sektörü kentsel alanlardaki kötü hava kalitesine önemli bir katkıda bulunmaya devam etmektedir.

Geleneksel olarak, emisyonlar ve yakıt tüketimi ampirik, istatistiksel ve analitik yöntemler kullanılarak modellenmiştir. Ancak bu yaklaşımlar karmaşık, çok boyutlu veri kümelerini ele almada genellikle sınırlamalarla karşılaşmaktadır. Makine öğrenimi (ML) teknikleri, doğrusal olmayan ilişkileri modelleme ve büyük veri kümelerini yüksek doğrulukla analiz etme yeteneği sunan güçlü bir alternatif olarak ortaya çıkmıştır. Yapay sinir ağları (ANN), destek vektör makineleri (SVM) ve topluluk yöntemleri gibi makine öğrenimi yaklaşımları emisyon modellemesine başarıyla uygulanmıştır. Son çalışmalar, tahmin edilen ve ölçülen değerler arasında güçlü korelasyonlar ile CO_2 ve NO_x dahil olmak üzere çeşitli emisyonları tahmin etme yeteneklerini göstermiştir [5,6].

Bu gelişmelere rağmen, bazı zorluklar devam etmektedir. İlk olarak, çoğu çalışma tek tek motor tiplerine (benzinli veya dizel) odaklanırken, karma veri setlerinin model genelleştirilebilirliğini geliştirme potansiyeli yeterince araştırılmamıştır. İkincisi, tahmine dayalı modelleme için birincil girdi olarak araç hızına güvenmek, lojistik açıdan karmaşık ve maliyetli olabilen kapsamlı veri toplama altyapısı gerektirir. GPS hız verileri, yaygın olarak erişilebilir olduğu ve minimum donanım değişikliği gerektirdiği için umut verici bir alternatif sunmaktadır. Bununla birlikte, özellikle benzin ve dizel karışık veri kümeleri için tahmine dayalı modellemede güvenilirliği ve doğruluğu tam olarak değerlendirilmemiştir.

Bu araştırma, araç hızı ve GPS hız verilerini kullanarak yakıt tüketimi ve emisyonları tahmin eden makine öğrenimi modelleri geliştirerek bu eksiklikleri gidermeyi amaçlamaktadır. Model performansını değerlendirmek için benzinli ve dizel motorlu araçlardan gerçek dünya sürüş koşullarında toplanan verilerden yararlanılarak ensemble bagged ve karar ağacı algoritmaları kullanılmıştır. Karma ve bireysel veri kümelerini karşılaştıran bu çalışma, veri kümesi kompozisyonunun tahmin doğruluğu üzerindeki etkisini araştırmaktadır. Sonuçlar, çevresel etkilerin gerçek zamanlı izlenmesi ve yönetimi için GPS tabanlı modellerin akıllı ulaşım sistemlerine (ITS) entegre edilmesinin fizibilitesine ilişkin uygulanabilir bilgiler sağlamaktadır.

Bu araştırmanın bulguları, emisyon ve yakıt tüketimi modellerinin doğruluğunu ve uygulanabilirliğini artırarak sürdürülebilir ulaşım çözümleri geliştirmeye yönelik devam eden çabalara katkıda bulunmaktadır. Ayrıca bu çalışma, kentsel ulaşım ve çevresel sürdürülebilirlik alanındaki kritik zorlukları ele almak için makine öğrenimi ve GPS verileri gibi gelişmekte olan teknolojilerden yararlanmanın önemini vurgulamaktadır.

2. Literatür Taraması

Taşıt yakıt tüketimi ve emisyonlarının doğru modellenmesi, ulaştırma sektörünün sera gazı emisyonları ve hava kirliliğine önemli katkısı nedeniyle önemli bir araştırma konusu olmuştur. Çok sayıda çalışma, tahmin doğruluğunu artırmak ve politika oluşturmayı desteklemek için çeşitli modelleme tekniklerini ve etkileyen faktörleri araştırmıştır.

Önceki araştırmalar, yakıt tüketimi ve emisyon tahmininde mevcut araştırmaları şekillendiren temel katkıları ele almaktadır. Araç emisyonları, aracın fiziksel özellikleri, operasyonel özellikleri, çevre koşulları ve yakıt özellikleri gibi çeşitli faktörlerden etkilenmektedir. Sullivan ve diğerleri (2004) dizel motorların genellikle benzinli motorlardan daha düşük emisyon ürettiğini göstermiştir.

temeline dayanmaktadır [4]. Bu arada, Clark ve diğerleri (2011) ve Zhang ve diğerleri (2013) tarafından yapılanlar gibi çalışmalar, emisyon modellemesi için kritik parametreler olarak araç hızını ve ivmesini vurgulayarak bu etkileyen faktörleri kategorize etmiştir [7,8]. Özellikle hız, hem anlık hem de ortalama ölçümlerin doğru tahminler için hayati önem taşıdığı kanıtlanarak kapsamlı bir şekilde araştırılmıştır [3]. İlk çalışmalar, özellikle araç hızı ve sürüş davranışı ile ilgili olarak emisyon tahmininde kullanılan temel ilkelerin, yöntemlerin ve modellerin oluşturulmasında çok önemliydi. Bu çalışmalar, araç çalışma koşulları (hız gibi) ve emisyonlar arasındaki ilişkiyi anlamaya odaklanmış ve daha sonraki araştırmalar için temel bilgiler sağlamıştır.

Emisyon modelleme metodolojilerinin evrimi, çeşitli yaklaşımların geliştirilmesine tanık olmuştur. Faris ve diğerleri (2011) ile Wang ve McGlinchy (2009) emisyon ve yakıt tüketimi modellerini beş kategoride gruplandırmıştır: ölçek girdi modelleri (mikroskobik, makroskobik, mezoskopik), formülasyon modelleri (analitik, ampirik, istatistiksel, grafiksel), boyut modelleri, açıklayıcı değişken modelleri ve durum değişkeni değer modelleri [9,10]. Bunlar arasında mikroskobik modeller, araç davranışındaki gerçek zamanlı değişimleri yakalama yetenekleri nedeniyle tercih edilmiştir, ancak genellikle önemli hesaplama kaynakları gerektirirler. Bu çalışmalar, mevcut metodolojileri etkilemeye devam eden temel modelleri ve çerçeveleri oluşturmuştur. Araç emisyonları, özellikle NO_x ve CO_2 ve yakıt kullanımına ilişkin tahmin doğruluğu, GPS veri uygulaması ile birlikte son makine öğrenimi (ML) uygulamaları ile artırılmıştır. Modern gelişmeler, mevcut emisyon modelleme sorunlarına çözüm geliştirmede tarihsel olarak önemli olduğunu kanıtlayan önceki araştırma bulguları üzerine inşa edilmiştir. Makine öğrenimi yaklaşımları, geleneksel yöntemlere göre önemli avantajlar sunan büyük ve karmaşık veri kümelerini işleme yetenekleri nedeniyle son yıllarda önem kazanmıştır [11-13]. Yapay sinir ağları (ANN), destek vektör makineleri (SVM) ve topluluk yöntemleri gibi teknikler, araç emisyonları ve yakıt tüketimi için yüksek tahmin doğruluğu göstermiştir [14]. Derin öğrenme teknikleri de bu alana girmiştir, ancak uygulamaları sınırlı kalmıştır. Altug ve Kucuk (2019) ve Shin ve diğerleri (2021) tarafından yapılan çalışmalar, sırasıyla egzoz borusu NO_x emisyonlarını ve geçici motor emisyonlarını tahmin etmek için derin öğrenmeyi araştırmıştır [15,16]. Bu yaklaşımlar, emisyon verilerindeki karmaşık ilişkileri yakalama potansiyelini göstermektedir, ancak kapsamlı hesaplama kaynakları ve büyük veri kümeleri.

Ekvador'daki araştırmacılar, yaptıkları incelemeler sonucunda hafif araçlardan kaynaklanan kirletici emisyonlarını değerlendirmek için yeni bir yaklaşım geliştirmiştir [17]. Araştırmacılar, CO_2 , CO, HC ve NO_x emisyonlarını tahmin etmek için araca özgü kütle ve motor kapasitesi verileriyle birlikte GPS verilerinin hız ve eğim bilgilerini kullanan bir yapay sinir ağı geliştirmiştir [17]. İncelenen modeller, CO_2 tahminleri için 0,735 R^2 değeri ve NO_x tahminleri için 0,798 R^2 değeri ile güçlü tahmin gücü göstermiş, bu da emisyon tahmininde GPS verileri ve makine öğrenimi tekniklerinin etkili kombinasyonunu ortaya koymuştur [17]. Liu ve arkadaşları (2023), CO_2 ve NO emisyonlarını tahmin etmek için Motorlu Araç Emisyon Simülatorü (MOVES) modelini ML yöntemleriyle entegre ederek tahmin edilen ve ölçülen değerler arasında güçlü korelasyonlar elde etmiştir [6]. Ahmed ve arkadaşları (2021), metanol-benzin karışımları kullanan kıvılcım ateşlemeli motorların emisyonlarını modellemek için YSA uygulamış ve %99'luk bir doğruluk bildirmiştir [5]. Bu çalışmalar, makine öğrenimi yaklaşımlarının çeşitli bağlamlarda uyarlanabilirliğinin ve güvenilirliğinin altını çizmektedir.

Daha yeni gelişmeler, gerçek dünya sürüş verilerine ve doğru veri toplama için taşınabilir emisyon ölçüm sistemlerinin (PEMS) entegrasyonuna odaklanmıştır [18,19]. Farklı araştırmalar, derin sinir ağları (DNN'ler) aracılığıyla hafif hizmet aracı emisyonları ve yakıt tüketimi için üst düzey modeller geliştirmiştir [20]. Sistem, GPS tanımlama verileriyle birlikte hız parametrelerini ve motor kontrollerini kaydederek gerçek zamanlı veri toplama için (PEMS) kullanmıştır [20]. DNN'lerin uygulanması araştırmacıların

Sürüş modelleri ve emisyon çıktıları arasındaki karmaşık doğrusal olmayan ilişkilerin tahminlerini iyileştirerek doğruluk seviyelerini artırmıştır [20]. Lee ve arkadaşları (2021) gerçek dünya koşullarında dizel araç emisyonlarını tahmin etmek için araca özgü güç, ortam sıcaklığı ve egzoz gazı devridaim oranları gibi çeşitli girdi parametrelerini içeren YSA kullanmıştır [21]. Cha ve arkadaşları (2021) dizel araçlardan kaynaklanan CO₂ emisyonlarını tahmin etmek için regresyon modelleri kullanmış ve anlık hız ve ivmenin en etkili tahminciler arasında olduğunu bulmuştur [22]. Araştırmacılar, makine öğrenimi yöntemlerini kullanarak gerçek sürüş CO₂ ile birlikte NO_x ve yakıt tüketimini tahmin etmek için başka bir çalışma yürütmüştür [23]. Araştırma grubu, emisyonlar ve yakıt kullanımı değerlendirmesi için güvenilir sonuçlar elde eden öngörücü makine öğrenimi modelleri geliştirmek için gerçek ölçülmüş sürüş verilerini kullanmıştır [23]. Analiz, makine öğrenimi yöntemlerinin gerçek sürüş koşullarında emisyon tahmin doğruluğunu artırmak için karmaşık verileri işleme yeteneğini göstermektedir [23]. Wen ve arkadaşları (2021), gradyan artırıcı regresyon (GBR) modellerini kullanarak, farklı koşullar altında çalışan dizel araçlar için yakıt tüketimi modellemesi ile birlikte NO_x ve CO₂ emisyonları tahmini için ML'nin kapasitesini göstermiştir [24]. Wen ve diğerlerine (2021) göre eğitim sürecinde, egzoz akış hızı özellikleriyle birlikte kütsel hava akış hızını da içeren PEMS verileri kullanılmıştır [24]. GBR modelleri, çeşitli sürüş durumlarında 0,99'a kadar ulaşan R² değerleri ile olağanüstü tahmin doğruluğunu teyit etmiştir [24]. Alfaseeh ve arkadaşları (2020) tarafından yapılan bir çalışmada, derin dizi öğrenme modellerinin, özellikle de uzun kısa süreli bellek (LSTM) ağlarının, yol ağlarında meydana gelen sera gazı emisyonlarını nasıl tahmin edebileceği araştırılmıştır [25]. Hız ve trafik yoğunluğu girdisini içeren zaman serisi verileri, bu modellerin yüksek zamansal çözünürlüklerde CO₂ eşdeğeri emisyon tahminleri yapmasını sağlar [25]. Bu yöntemin entegrasyonu, uzmanların ekoyönlendirme stratejileri geliştirmesini ve gerçek zamanlı emisyon izleme sistemlerini çalıştırmasını sağlar. YSA ve DVM yaygın olarak kullanılsa da, torbalanmış ve güçlendirilmiş ağaçlar gibi topluluk yöntemleri emisyonların modellenmesi için güçlü araçlar olarak ortaya çıkmıştır. Kocев ve diğerleri (2007) karmaşık, yüksek boyutlu veri kümelerinin işlenmesinde topluluk öğrenme tekniklerinin etkinliğini göstermiştir [26].

ML algoritmaları, yol üstü uzaktan algılama verileriyle birlikte kullanıldığında araca özgü emisyonların hızlı bir şekilde değerlendirilmesini sağlar. Bilim insanları, hafif hizmet benzinli araçlardan kaynaklanan CO, HC ve NO_x emisyonlarını değerlendirmek için sinir ağlarını ve aşırı gradyan artırmayı rastgele ormanlarla birleştiren topluluk modeli teknolojisini kullanmışlardır [27]. Xia ve diğerlerine (2021) göre, politika yapıcılara ve çevre ajanslarına pratik işlevsellik sağlayan bu metodoloji sayesinde yol üstü araç emisyon denetimi mümkün hale gelmektedir [27].

Gerçek zamanlı hız takibi için GPS verilerinin kullanılması ve araca özgü verilerin entegrasyonu, tahmin modellerine yeni boyutlar eklemiştir. Ayrıca, hesaplama gücündeki artış ve makine öğrenimi tekniklerinin geliştirilmesi, araştırmacıların büyük veri kümelerini daha etkili bir şekilde ele almalarını ve daha doğru, gerçek zamanlı emisyon tahminleri üretmelerini sağlamıştır. Daha yeni çalışmalar, emisyon tahmininin doğruluğunu ve uygulanabilirliğini artırmak için gerçek zamanlı GPS verilerini ve makine öğrenimi gibi gelişmiş hesaplama tekniklerini birleştirmeyi önemli ölçüde genişletmiştir. GPS hız verileri son zamanlarda emisyon modellemesinde araç hız verilerine bir alternatif olarak dikkat çekmektedir. Akkamis ve arkadaşları (2021) tarafından yapılan çalışmalar, özellikle yüksek ivmelenme senaryolarında GPS hız sensörlerinin doğruluğunu vurgulamıştır [28]. Bununla birlikte, karışık araç veri kümelerinde tahmine dayalı modelleme için GPS verilerinin güvenilirliği iyi belirlenmemiştir ve daha fazla keşif için bir fırsat yaratmaktadır.

Modern araştırmalar, son yirmi yılda gerçekleştirilen temel araştırmaların önemli konumunu korumaya devam ettiğini göstermektedir. İlk aşamadaki çalışmalar, emisyon modelleri geliştirilirken hız ve ivme değişkenlerini önemli bileşenler olarak belirlemiştir ve bu değişkenler artık GPS sistemleri ile birlikte gelişmiş makine öğrenimi yöntemleriyle birleştirilmektedir. Geliştirilen sofistike makine öğrenimi tekniklerine dayalı modern tahminler

Geleneksel ampirik modellerden elde edilen sonuçlar, doğru emisyon sonuçlarının tahmin edilmesine yönelik mevcut araştırma çabasında geçerliliğini sürdürmektedir. Önemli ilerlemeler kaydedilmiş olsa da, makine öğrenimi tekniklerinin karışık benzin ve dizel veri kümelerine uygulanmasında ve GPS hız verileri gibi alternatif girdi parametrelerinin kullanımında eksiklikler devam etmektedir. Bu çalışma, yakıt tüketimi ve emisyonlar için GPS tabanlı modellerin tahmin doğruluğunu değerlendirmek için toplu torbalanmış modeller ve karar ağacı algoritmalarından yararlanarak bu boşlukları ele almayı amaçlamaktadır. Bu araştırma, benzinli ve dizel motorlardan elde edilen veri kümelerini birleştirerek, gerçek dünyada uygulanabilirliği artıran ve sürdürülebilir ulaşım yönetimi için eyleme geçirilebilir içgörüler sağlayan genelleştirilmiş modeller geliştirmeyi amaçlamaktadır.

3. Materyaller ve Yöntemler

3.1. Metodoloji

Bu çalışmada, benzinli ve dizel araçlardan elde edilen gerçek dünya sürüş verilerini kullanarak yakıt tüketimi ve egzoz emisyonlarını analiz etmek için öngörücü modeller geliştirmek üzere makine öğrenimi teknikleri, özellikle de karar ağacı ve Torbalı Ağaçlar algoritmaları kullanılmıştır. Modeller, karbondioksit (CO_2) ve azot oksitler (NO_x) dahil olmak üzere yakıt verimliliğini ve kirletici emisyonlarını etkileyen temel faktörleri belirlemek için sürüş veri kümeleri üzerinde eğitilmiştir. Modelin doğruluğunu ve güvenilirliğini sağlamak için ortalama çözünen hata (MAE), ortalama karesel hata (MSE), kök ortalama karesel hata (RMSE) ve belirleme katsayısı ($R^{(2)}$) dahil olmak üzere çeşitli değerlendirme ölçütleri kullanılmıştır. Ayrıca, aşırı uyumu önlemek ve modelin genelleme kabiliyetini geliştirmek için 5 kat çapraz doğrulama uygulanmıştır. Aşağıdaki bölümler, veri ön işleme, model eğitimi, doğrulama stratejileri ve değerlendirme kriterleri dahil olmak üzere metodolojiyi detaylandırmaktadır.

3.1.1. Karar Ağacı

Karar ağaçları (DT), hem sayısal hem de kategorik verilerin işlenmesindeki yorumlanabilirlikleri ve etkinlikleri nedeniyle sınıflandırma ve regresyon görevleri için yaygın olarak kullanılan bir makine öğrenimi tekniğidir. Bir karar ağacı, her bir iç düğümün bir özelliğe dayalı bir karar kuralını temsil ettiği, dalların olası sonuçları gösterdiği ve yaprak düğümlerinin nihai tahmine veya sınıf etiketine karşılık geldiği düğümler, dallar ve yapraklardan oluşan hiyerarşik bir yapıdan oluşur [29]. Bir karar ağacının oluşturulması, veri kümesinin verileri en iyi şekilde ayıran seçilen bir özelliğe göre yinelemeli olarak bölündüğü özyinelemeli bir bölünme yaklaşımını izler. Bölme kriteri, problemin türüne bağlı olarak değişir; sınıflandırma görevleri için Gini safsızlığı veya bilgi kazancı (entropiye dayalı) gibi ölçütler yaygın olarak kullanılırken, regresyon ağaçları tipik olarak ortalama kare hata (MSE) minimizasyonuna dayanır.

Matematiksel olarak, her bir bölünmedeki safsızlık azaltımı aşağıdaki gibi ifade edilebilir. Sınıflandırma için entropi kullanan ağaçlar:

$$I(S) = - \sum_{i=1}^k p_i \log_2 p_i$$

Burada $I(S)$ bir S veri kümesinin entropisini, k sınıf sayısını ve p_i de i sınıfına ait örneklerin oranını gösterir. Optimum bölme, bölmeden önce ve sonra entropi farkı olarak hesaplanan bilgi kazancının maksimize edilmesiyle belirlenir.

Bir karar ağacının oluşturulması üç temel adımdan oluşur. İlk olarak, algoritma veri kümesini seçilen bir kritere göre bölmek için en bilgilendirici özelliği seçer. İkinci olarak, veri kümesi, maksimum ağaç derinliği veya yaprak düğümü başına minimum örnek sayısı gibi bir durdurma koşulu karşılanana kadar özyinelemeli olarak alt kümelere bölünür. Son olarak, ortaya çıkan

ağaç yapısı, gerekirse, tahmin performansına çok az katkıda bulunan dalları kaldırarak aşırı uyumu önlemek için budanır.

3.1.2. Torbalı Ağaçlar

Breiman (1996) tarafından tanıtilen Bagging (bootstrap aggregating) Trees algoritması, model varyansını azaltarak tahmin doğruluğunu artırmak için tasarlanmış yaygın olarak kullanılan bir topluluk öğrenme tekniğidir [30]. Bu yöntem, eğitim veri kümelerinin veri noktalarının değiştirilerek rastgele seçilmesiyle oluşturulduğu bootstrap örneklemesini kullanarak çoklu karar ağacı (DT) modelleri oluşturur. Örneklenen her veri kümesi bağımsız bir DT'yi eğitmek için kullanılır ve modeller arasında çeşitlilik sağlanır [31]. Eğitim sırasında, eğitim örneklerinin toplam sayısı iterasyonlar boyunca aynı kalır, ancak dağılımları yeniden örnekleme süreci nedeniyle değişir ve daha genelleştirilmiş bir modele yol açar [32]. Nihai tahmin, sınıflandırma görevlerinde çoğunluk oylaması veya regresyon görevlerinde ortalama alma yoluyla tüm karar ağaçlarının çıktılarının toplanmasıyla elde edilir. Bu süreç matematiksel olarak şu şekilde ifade edilebilir

$$\hat{z} = \frac{1}{M} \sum_{i=1}^M \phi_i(x),$$

Burada $\hat{z}$ nihai model çıktısını, M DT modellerinin toplam sayısını ve $\phi_i(x)$ ise birinci karar ağacı tarafından yapılan tahmini temsil eder.

Torbalı Ağaçların geliştirilmesi üç ana adımdan oluşur. İlk olarak, temel öğrenciler olarak birden fazla DT modeli oluşturulur. İkinci olarak, verilerin yaklaşık üçte ikisinin eğitim için kullanıldığı çeşitli eğitim setleri oluşturmak için bootstrap örneklemesi uygulanırken, kalan üçte biri model performansını değerlendirmek için torba dışı (OOB) veri olarak ayrılır. Son olarak, bireysel ağaç tahminleri nihai çıktıyı üretmek için toplanır, kararlılığı artırır ve aşırı uyumu azaltır [33,34]. Torbalı Ağaçlar, yüksek tahmin güvenilirliğini korurken ulaşım verilerindeki karmaşık ilişkileri yakalama yetenekleri nedeniyle ulaşım çalışmalarında, özellikle kaza şiddeti tahmini, trafik akışı analizi ve sürücü davranışı modellemesinde yaygın olarak kullanılmaktadır.

3.1.3. Değerlendirme Metrikleri

Tahmine dayalı bir modelin performansını değerlendirmek, doğruluğunu, güvenilirliğini ve genelleme yeteneğini sağlamak için çok önemlidir. Regresyon analizinde, modelin tahminlerinin gerçek gözlemlenen değerlerle ne kadar uyumlu olduğunu değerlendirmek için çeşitli istatistiksel ölçütler kullanılır. En yaygın kullanılan değerlendirme ölçütleri arasında ortalama mutlak hata (MAE), ortalama karesel hata (MSE), kök ortalama karesel hata (RMSE) ve belirleme katsayısı (R^2) bulunmaktadır.

Model doğrulaması, veri kümesi üç alt kümeye bölünerek gerçekleştirilmiştir. Birinci ve ikinci alt kümeler model eğitimi için kullanılırken, geliştirilen model dışa aktarılmış ve performansını simüle etmek için üçüncü alt kümeye uygulanmıştır. Sağlamlığı sağlamak ve aşırı tahmin riskini en aza indirmek için, veri kümesini beş alt kümeye bölen, modeli dört kat üzerinde iteratif olarak eğiten ve kalan bir kat üzerinde doğrulayan yaygın olarak kullanılan bir yeniden örnekleme tekniği olan 5 kat çapraz doğrulama kullandık. Bu işlem beş kez tekrarlanmış ve model performansını değerlendirmek için değerlendirme metriklerinin ortalaması hesaplanmıştır. Seçilen değerlendirme metrikleri, model doğruluğu ve hata dağılımı hakkında bilgi sağlar.

Ortalama Mutlak Hata (MAE)

Ortalama mutlak hata (MAE), yönlerini dikkate almadan tahmin edilen ve gerçek değerler arasındaki hataların ortalama büyüklüğünü ölçer. Şu şekilde hesaplanır

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

Burada n toplam gözlem sayısı, y_i gerçek değer ve $\hat{y}_i$ tahmin edilen değerdir. MAE, hedef değişkenle aynı birimde sezgisel bir hata ölçüsü sağlayarak yorumlanmasını kolaylaştırır. Ancak, büyük hataları karesel hata metrikleri kadar ciddi bir şekilde cezalandırmaz.

Ortalama Karesel Hata (MSE)

Ortalama karesel hata (MSE), gerçek ve tahmin edilen değerler arasındaki karesel farkların ortalamasını hesaplar. Şu şekilde ifade edilir

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2.$$

MSE, hataların karesini alarak daha büyük sapmalara daha yüksek bir ceza uygular ve bu da onu aykırı değerlere karşı daha hassas hale getirir. Bu özellik, önemli tahmin hatalarını en aza indirmesi gereken modellerin değerlendirilmesine yardımcı olur.

Ortalama Karesel Hatanın Kökü (RMSE)

Ortalama karesel hata kökü (RMSE), MSE'den karekök alınarak türetilir ve bu da metriğin orijinal verilerle aynı birimde kalmasını sağlar. Tarafından verilir

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

RMSE yaygın olarak kullanılır çünkü yorumlanabilirliği korurken daha büyük hataları daha ağır bir şekilde cezalandırır. Daha düşük bir RMSE, daha iyi tahmin doğruluğuna sahip bir modeli gösterir.

Belirleme Katsayısı (R^2)

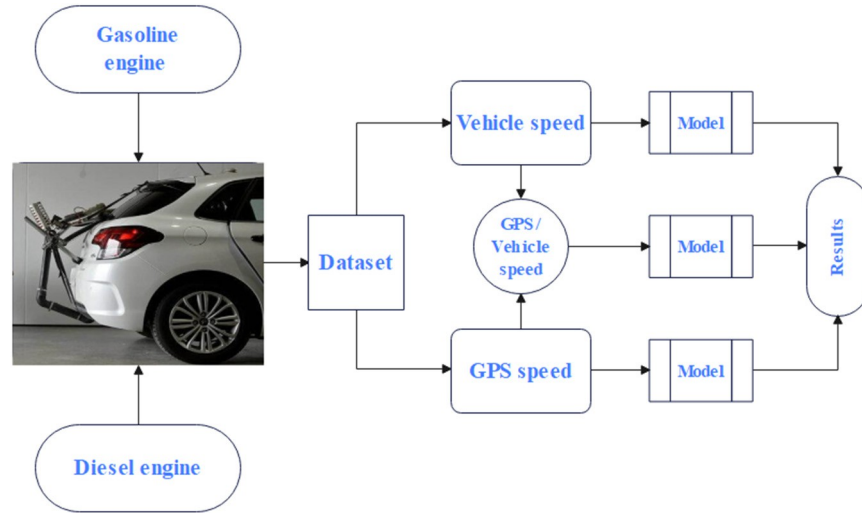
Belirleme katsayısı (R^2), modeldeki bağımsız değişkenler tarafından açıklanan bağımlı değişkendeki varyans oranını ölçer. Şu şekilde tanımlanır

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

Burada $\bar{y}$ gözlenen değerlerin ortalamasıdır. R^2 değeri 0 ile 1 arasında değişir ve 1'e yakın değerler daha iyi bir uyuma işaret eder. Yüksek bir R^2 , modelin hedef değişkendeki değişkenliğin önemli bir kısmını yakaladığını gösterirken, düşük bir R^2 zayıf bir uyuma işaret eder.

3.2. Veri Toplama ve Ön İşleme

Bu çalışmada kullanılan veriler SEMTECH-LDV sensörleri kullanılarak taşınabilir bir emisyon ölçüm sisteminden (PEMS) elde edilmiştir. PEMS kurulumu Şekil 1'de gösterilmektedir. Gerçek dünya testi Emisyon Analitiği Araştırma Merkezi tarafından gerçekleştirilmiştir. Testte ikisi benzinli ve ikisi dizel motorlu olmak üzere dört araç yer almıştır. Araştırmada kullanılan tüm test araçları hafif hizmet tipi, otomatik şanzımanlı, EPA kademe 3 standardındaydı. Benzinli motor verileri Nissan Altima (2WD ve Tier 3 Bin 30) ve Volvo S60 (4-AWD Tier 3 Bin 30) 2020 model sedan araçlardan alınmıştır. Dizel motor verileri 2020 model Jeep Wrangler (4-AWD Tier 3 Bin 160) spor hizmet aracı ve 2021 model Jeep Gladiator (4-AWD Tier 3 Bin 160) pikap aracından alınmıştır. PEMS, GPS, hava durumu istasyonu ve taşınabilir analizör içeren kompakt bir cihazdır. Bu bileşenler PEMS'in egzoz kütle akışını, yakıt oranını, kirleticileri, araç özelliklerini, GPS verilerini, yol özelliklerini ve hava durumu probu verilerini gerçek zamanlı olarak ölçmesini mümkün kılar. PEMS'ten alınan veriler bir CSV dosyasına aktarılmış ve cihazın bekleme moduna karşılık gelen veri noktaları çıkarılmıştır.



Şekil 1. PEMS kurulumu ve araştırma çerçevesi. PEMS kurulumu ve araştırma çerçevesi.

Seçilen girdi parametreleri çevresel faktörler, egzoz parametreleri, motor ve araç hızıdır (Tablo 1). Benzinli motor veri kümesi 21.098 örnekten oluşurken, dizel motorla 20.500 örnek kaydedilmiştir.

Tablo 1. Girdi ve çıktı parametreleri. Giriş ve çıkış parametreleri hakkında ayrıntılar.

Girdi Kategorileri			Çıktılar
Araç Hızı	GPS Hızı Birim		
	GPS Yüksekliği	m	
	GPS Hızı	mph	
	GPS Yer Hızı	km/h	
Hava Durumu Probu Nem Probu	Hava Durumu	%RH	Yakıt Oranı (gal/s)
	Nem		
Ortam Basıncı	Ortam Basıncı	mbar	Anlık Kütle
Hava Probu Sıcaklığı	Hava Probu Sıcaklığı	°	$C \quad CO_2 (g/s)$
Egzoz Kütle Akış Hızı	Egzoz Kütle Akış Hızı	kg/h	Düzeltilmiş Anlık Kütle $NO_{(x)} (g/s)$
Egzoz Sıcaklığı	Egzoz Sıcaklığı	°C	
Manifold Basıncı	Manifold Basıncı	kPa	
Motor Devri	Motor Devri		
RPM Abs Gaz Kelebeği Postn Abs Gaz Kelebeği Direği		%	
Bağıl Gaz Kelebeği Sonrasın Rel. Gaz Kelebeği Direği		%	

İki vaka çalışması gerçekleştirilmiştir. İlk olarak, modelleri eğitmek için dizel ve benzin veri tabanları bir araya getirilmiştir. Bu durumda veriler, karışık dizel/benzin veri tabanından oluşan araç hızı ve GPS hızı veri kümelerine ayrılmıştır. Veri tabanları topluluk torbalı ve ağaç algoritmaları ile eğitilmiştir. İkinci senaryoda, araç ve GPS hızı girdi kategorileri için ayrı dizel ve benzin veri tabanları kullanılarak modelleri eğitmek için yalnızca topluluk torbalı algoritması kullanılmıştır.

Veri kümesi, ağ ağırlığının erken doygunluğunu önlemek için 0 ile 1 arasında normalleştirildi, bu da daha hızlı yakınsamaya yol açacaktı. Model MATLAB kullanılarak eğitilmiştir 9.10 regresyon araç kutusu. Modeli eğitmek için topluluk torbalı ve karar ağacı algoritmaları kullanılmıştır.

3.3. Yakıt Tüketimi Hesaplamaları

Araç emisyonları yakıt tüketimi ile yakından ilişkilidir [35-37]. Referans [38], PEMS test sonuçlarından karbon dengesi yöntemini kullanarak yakıt tüketimini hesaplayan hafif hizmet ve ağır hizmet araç yöntemlerini özetlemiştir. Ölçülen salınan CO_2 , CO ve toplam hidrokarbonlardan (THC), gram cinsinden yakıt tüketimi tahmin edilebilir [38].

4. Sonuçlar

GPS hızları ulaşım araştırmalarında yaygın olarak kullanılmaktadır. GPS hız sensörleri tipik olarak radar hız sensörlerinin yerine kullanılır çünkü hız verilerini darbe sinyalleri halinde sunarlar [28]. Akkamis ve arkadaşlarına (2021) göre, daha yüksek güncelleme frekanslı GPS hız sensörleri, özellikle daha yüksek hızlanma durumlarında daha iyi doğruluk sağlar [28]. Bu araştırmanın ana katkılarından biri, yakıt tüketimi ve araç emisyonlarını tahmin etmek için GPS hız verilerini bir ML modelinde kullanmanın verimliliğini araştırmaktır. GPS hız ölçümünün yüzde hatası, araç hız verileriyle karşılaştırıldığında %1,4'tür. ML'nin yakıt tüketimi ve emisyon faktörlerini tahmin etmek için bu düşük hatalı GPS hız verilerini nasıl doğru bir şekilde kullanabileceğini araştırdık. Derlenen motor tipi veri seti kullanılmış ve gerçek araç hızı ile GPS araç hızı arasında bir karşılaştırma yapılmıştır. Ayrıca bu bölümde, ML modelinin bu iki motor türüne uygulandığında gösterdiği performansı açıklamak için benzinli ve dizel motorlar arasındaki fark tartışılmıştır. Son olarak, gerçek dünya (farklı araç motor tiplerinin derlenmesi) ve ideal (tek motor tipi) koşullar arasındaki farkı ölçmek için derlenmiş dizel ve benzinli veri seti ile ayrı veri setleri arasında bir karşılaştırma yapılmıştır.

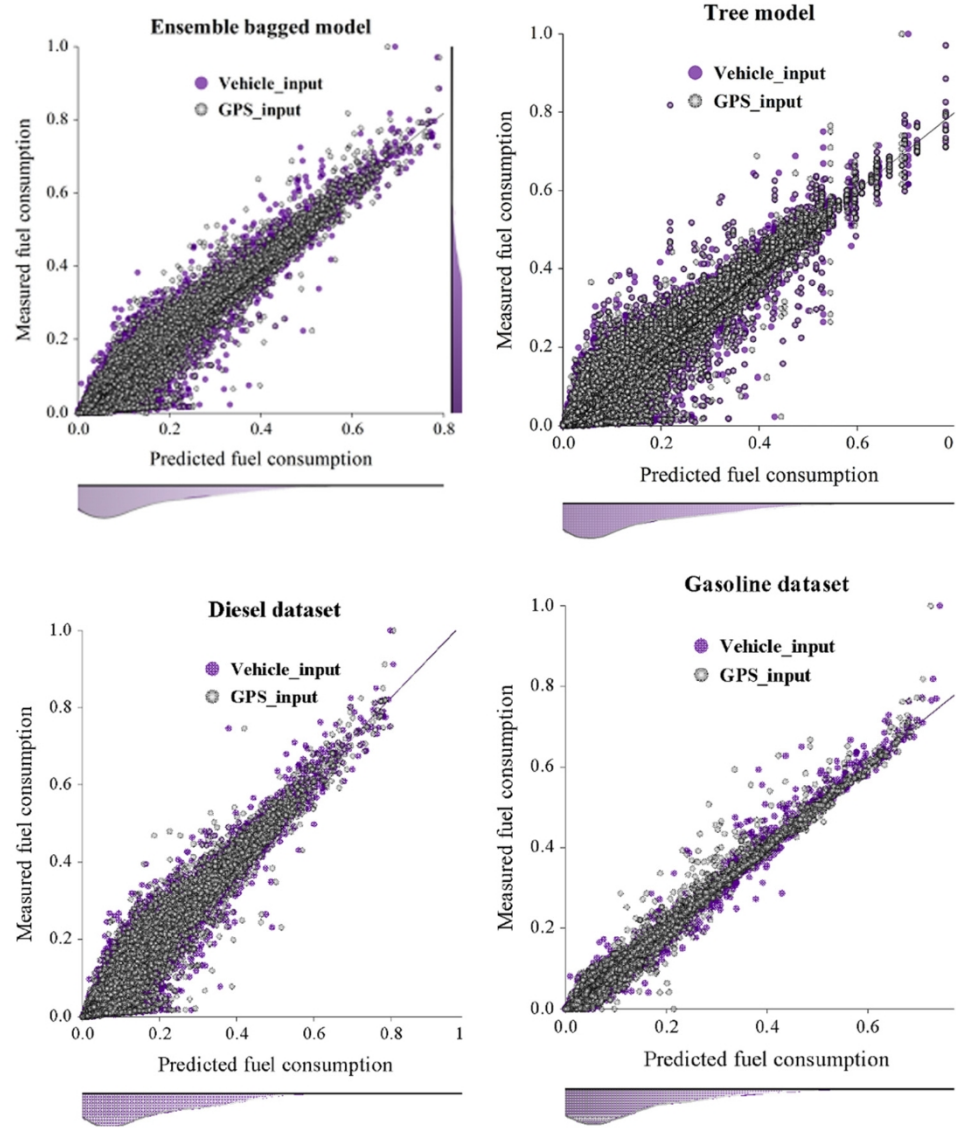
Araştırma veri tabanında (dizel ve benzinli motorlar), ortalama araç hızı 46,5 mph ve ortalama GPS hızı 47,2 mph idi. İzlenen tüm çalışmalar arasında ortalama yakıt oranı ve emisyon faktörleri 0,0003 gal/s yakıt oranı, 2,93 g/s CO_2 ve 0,00029 g/s NO_x idi. İki egzoz kirleticisinin anlık yakıt tüketimini ölçen yakıt oranını tahmin etmek için topluluk ve karar ağacı modelleri kullanılmıştır: CO_2 ve NO_x . Model sonuçları Tablo 2'de, tahmin edilen ve ölçülen değer ilişkisini gösteren regresyon dağılım grafiği ise Şekil 2'de gösterilmektedir.

Sonuçlar iki kategoride sunulmuştur. Birincisi, araç ve GPS hızı girdi sınıfları için karışık dizel ve benzin veri setini kullanarak yakıt tüketimini modellemenin bir değerlendirmesidir.

İkinci kategori, tek dizel ve benzin veri kümesi üzerinde topluluk torbalı algoritmanın performansını değerlendirmektedir. araç ve GPS hız girdi sınıfları için benzin veri kümeleri.

Tablo 2. Yakıt oranı tahmini için model performansı.

			RMSE (mph)	MAE (mph)	MSE	R ²
Karışık	Araç hızı	Topluluk Torbalı Yöntem	0.025	0.014	0.001	0.967
		Ağaç Yöntemi	0.031	0.016	0.001	0.95
	GPS hızı	Topluluk Torbalı Yöntem	0.025	0.014	0.001	0.968
		Ağaç Yöntemi	0.031	0.016	0.001	0.95
Topluluk Torbalı Yöntem						
Tek kişilik	Dizel veri seti	Araç Hızı	0.035	0.023	0.001	0.947
		GPS parametresi	0.034	0.022	0.001	0.949
	Benzin veri seti	Araç Hızı	0.011	0.005	0.000	0.993
		GPS parametresi	0.012	0.006	0.000	0.991



Şekil 2. Tahmin edilen ve ölçülen yakıt tüketimi modelleri için dağılım grafiği.

4.1. Yakıt Tüketiminin Modellenmesi

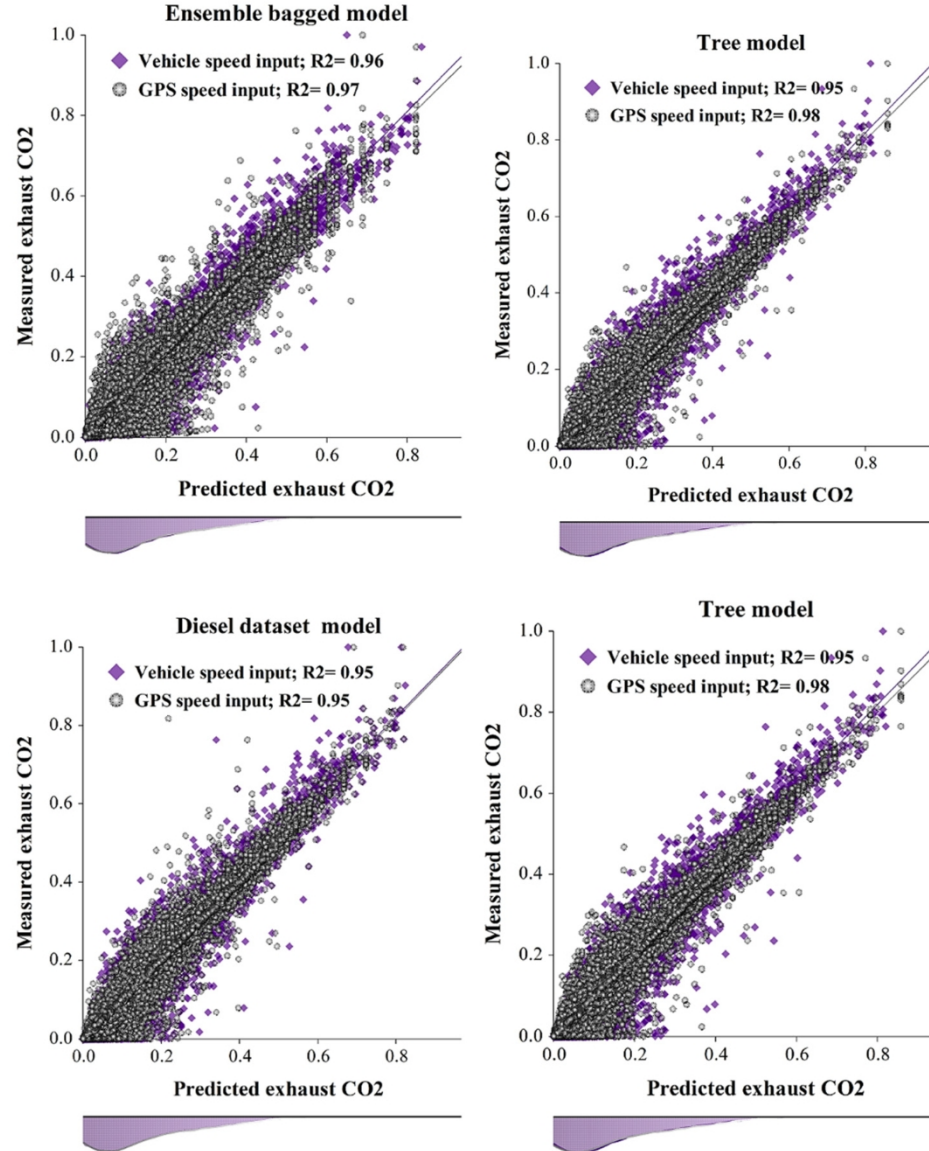
Yakıt tüketiminin modellenmesi, sadece sürdürülebilir bir akıllı ulaşım sisteminin (ITS) uygulanması için değil, aynı zamanda tüketilen enerji miktarının ekonomik olarak gerekçelendirilmesi için de dikkate alınması gereken önemli bir konu haline gelmiştir. Yine Tablo 2'de model performansının sonuçları sunulmaktadır.

Belirleme katsayısı (R^2) değerine göre, topluluk torbalı model performansı hem araç hem de GPS hızı girdi kategorilerinde karar ağacı modelinin performansını aşmıştır.

Şekil 2'de sunulan tahmin edilen ve ölçülen yakıt tüketiminin dağılım grafiği modellerin tahmin doğruluğunu göstermektedir. Sonuçlar, topluluk torbalı model performans doğruluğunun ağaç algoritmasından daha yüksek olduğunu göstermektedir. Dağılım grafiği veri noktaları eşitlik çizgisine ağaç modeli grafiğinden daha yakındır. Bu, daha iyi model performansının bir göstergesidir.

Ayrıca, Tablo 2'de sunulan RMSE, MAE ve MSE değerleri de topluluk torbalı modelin performansının daha iyi doğruluğa sahip olduğunu göstermektedir. Elde edilen her iki modelin normalleştirilmiş MSE değeri 0,001 olmasına rağmen, torbalanmış model için 0,025 ve 0,014 olan daha düşük RMSE ve MAE değeri, araç hızı ve GPS parametreleri kullanılarak yakıt tüketimi tahmini için sırasıyla 0,967 ve 0,968 olan daha yüksek R^2 değerini açıklamaktadır.

Şekil 3 ve 4'te sunulan araç ve GPS hızının üst üste bindirilmiş dağılım grafiği, GPS girdisine dayalı modellerin eşitlik çizgisine çok daha yakın olduğunu göstermektedir. Ayrıca, aykırı değerler için modellerin çizim konumlarındaki benzerlikler, yakıt tüketiminin tahmininde GPS verilerinin kullanılma potansiyelini güçlü bir şekilde ortaya koymaktadır. Sonuçlar ayrıca benzinli motor veri seti ile yakıt tüketiminin modellenmesinin hem araç hızı hem de GPS parametre tabanlı modeller için yüksek performans doğruluğuna sahip olduğunu göstermektedir. Elde edilen belirleme katsayısı (R^2) benzin veri seti için yaklaşık %99'dur. Ayrıca, dizel motor veri kümesi modellerinin düşük performansı, makine hatasından kaynaklanan düşük veri kalitesine veya seçilen girdinin dizel motorlarda tüketimi tahmin etmek için yetersiz olmasına bağlanabilir.



Şekil 3. Tahmin edilen ve ölçülen CO_2 modelleri için dağılım grafiği.

4.2. CO_2 Egzozunun Modellenmesi

Taşımacılık sektörü, özellikle kentsel alanlarda, sadece küresel ısınmayı artırma potansiyelini yükseltmekle kalmayıp aynı zamanda kötü hava kalitesine de katkıda bulunan büyük CO_2 emisyonlarına neden olmaktadır. Bu araştırma, araç ve GPS parametrelerini kullanarak topluluk torbalı ve karar ağacı modellerinin model performansını değerlendirmiştir. Şekil 3 ve Tablo 3'teki verilere dayanarak, GPS parametresi tabanlı modeller aşağıdaki durumlarda daha yüksek performans doğruluğuna sahiptir

hem dizel hem de benzinli karma ve tekli veri tabanlarında araç hızı tabanlı modellerle karşılaştırılmıştır. Normalize edilmiş tahmin edilen ve ölçülen değerler arasında elde edilen belirleme güveni, GPS girdi tabanlı modelde topluluk ve ağaç modelleri için sırasıyla 0,966 ve 0,975'tir. Sonuçlar ayrıca $CO_{(2)}$ 'nin gerçek zamanlı izlenmesinin uygulanabilir olduğunu ve politika yapıcıların ulaşım sisteminde optimizasyonlar yapmasına yardımcı olacak faydalı bir araç olabileceğini göstermektedir.

Tablo 3. CO_2 tahmini için model performansı.

			RMSE (mph)	MAE (mph)	MSE	R ²
Karışık	Araç hızı	Topluluk Torbalı Yöntem	0.028	0.017	0.001	0.960
		Ağaç Yöntemi	0.032	0.019	0.001	0.945
	GPS hızı	Topluluk Torbalı Yöntem	0.026	0.014	0.001	0.966
		Ağaç Yöntemi	0.022	0.011	0.000	0.975
Topluluk Torbalı Yöntem						
Tek kişilik	Dizel veri seti	Araç Hızı	0.036	0.024	0.001	0.947
		GPS parametresi	0.027	0.014	0.001	0.949
	Benzin veri seti	Araç Hızı	0.011	0.005	0.000	0.993
		GPS parametresi	0.012	0.006	0.000	0.991

4.3. NO_x Egzozunun Modellenmesi

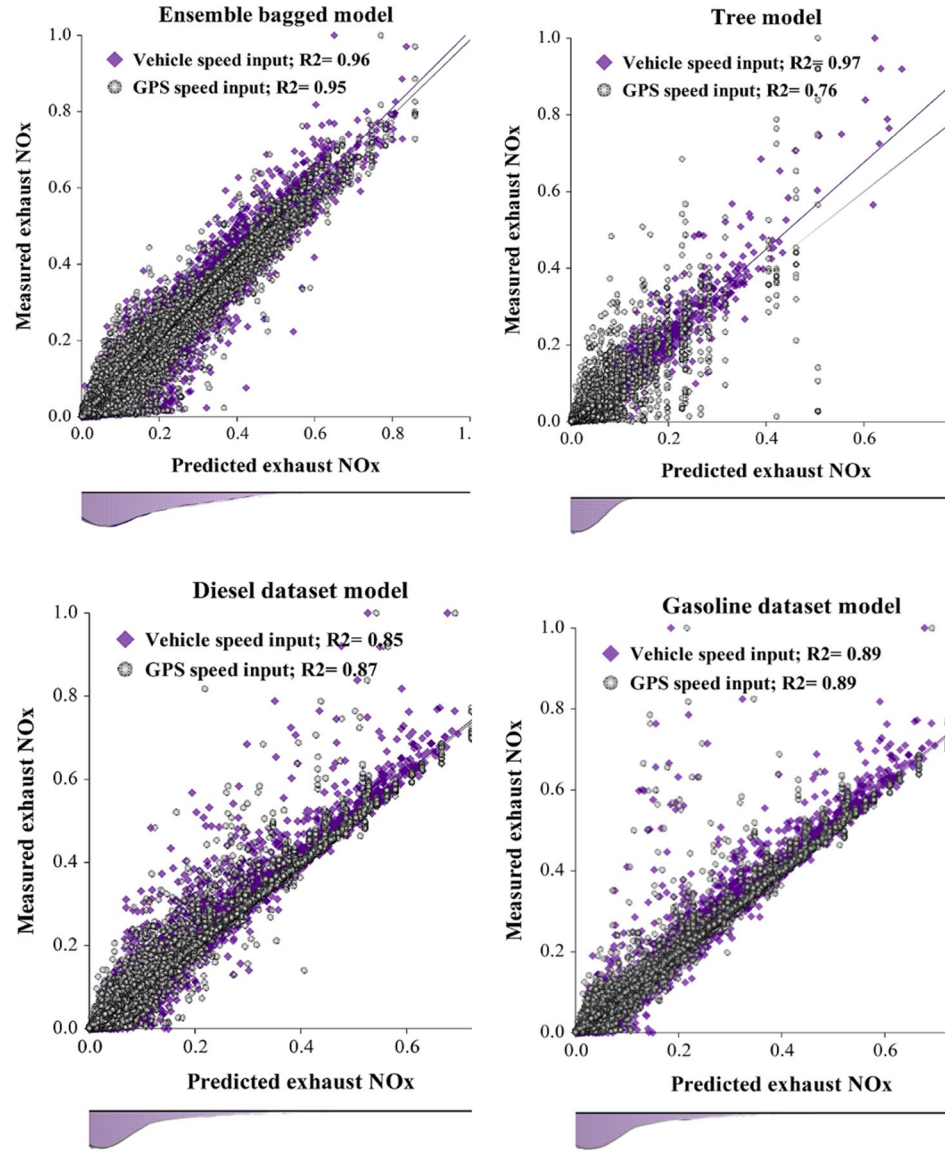
NO_x egzoz emisyonlarının model performansı CO_2 modellerine kıyasla biraz farklı bir eğilim göstermektedir. NO_x modeli GPS girdi parametresini kullanarak düşük performans göstermektedir. Topluluk torbalı modeli ile elde edilen performans doğruluğu %95 iken, ağaç modelinde bu oran %76'ya düşmüştür. Bununla birlikte, topluluk modelinin yüksek performans doğruluğu, GPS parametresinin NO_x egzoz emisyonlarını tahmin etmek için kullanılabileceğini de göstermektedir.

Tablo 4 ayrıca bir modelin performansını doğrulamak için değerlendirme parametrelerinin (RMSE, MAE, MSE ve R²) bir kombinasyonunun gerekli olduğunu göstermektedir.

Tablo 4. NO_x tahmini için model performansı.

			RMSE (mph)	MAE (mph)	MSE	R ²
Karışık	Araç hızı	Topluluk Torbalı Yöntem	0.028	0.017	0.001	0.960
		Ağaç Yöntemi	0.025	0.015	0.001	0.967
	GPS hızı	Topluluk Torbalı Yöntem	0.008	0.002	0.000	0.950
		Ağaç Yöntemi	0.016	0.004	0.000	0.760
Topluluk Torbalı Yöntem						
Tek kişilik	Dizel veri seti	Araç Hızı	0.019	0.006	0.000	0.853
		GPS parametresi	0.016	0.005	0.000	0.869
	Benzin veri seti	Araç Hızı	0.019	0.006	0.000	0.889
		GPS parametresi	0.019	0.006	0.000	0.890

NO_x tahmininde, tek dizel ve benzin veri tabanı modelinin performans doğruluğu Şekil 4'te görüldüğü gibi %89'u geçmemesine rağmen, şaşırtıcı bir şekilde, dizel/benzin karışık veri tabanındaki topluluk torbalı yöntemler %96'ya kadar daha yüksek bir performans doğruluğuna sahipti.



Şekil 4. Tahmin edilen ve ölçülen NO_x modelleri için dağılım grafiği.

5. Tartışma

Artan sera gazı emisyonları ve bunların çevresel etkileri küresel bir endişe haline gelmiştir ve emisyonları ölçmek ve azaltmak için yenilikçi yaklaşımlar gerektirmektedir. Emisyonların ve yakıt tüketiminin gerçek zamanlı tahmini, verimli Akıllı Ulaşım Sistemleri (ITS) geliştirmek ve çevresel ve ekonomik olarak sürdürülebilir ulaşım çözümleri uygulamak için kritik öneme sahiptir. Bu çalışma, GPS hızı ve araç hızı girdilerine dayalı makine öğrenimi modellerini değerlendirerek bu ihtiyaçları ele almaktadır.

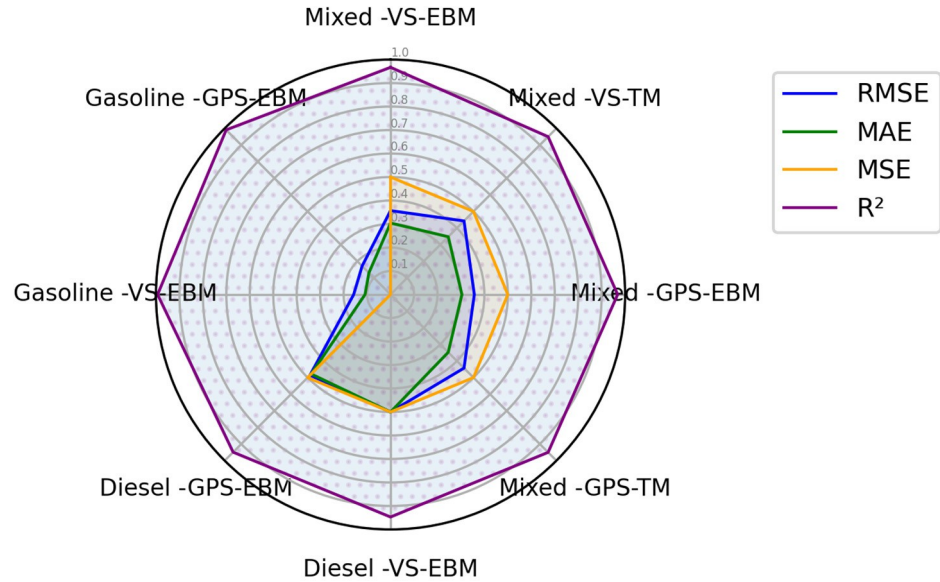
Son zamanlarda yapılan çalışmalar, dinamik, konuma özgü emisyon tahminleri geliştirmek için hız değişimleri ve hızlanma modelleri de dahil olmak üzere gerçek zamanlı araç davranış verilerini entegre etmeye odaklanmıştır. Bu ilerlemelere rağmen, daha fazla araştırma gerektiren önemli zorluklar hala mevcuttur. Örneğin, GPS sinyal doğruluğundaki değişkenlik, farklı araç tiplerinin etkisi ve tahmin modellerinin yeni emisyon standartlarına uyarlanması

ve gerçek dünya sürüş koşulları iyileştirilmesi gereken alanlar olmaya devam etmektedir. Ayrıca, makine öğrenimi ve yapay zekanın gerçek zamanlı GPS verileriyle entegre edilmesi, emisyon tahminlerinin iyileştirilmesi ve elektrikli araçlar ve yeni yakıt teknolojileri gibi yeni ortaya çıkan endişelerin ele alınması için fırsatlar sunmaktadır. Bu çalışma, GPS hızı ve araç hızı girdilerine dayalı makine öğrenimi modellerini değerlendirerek bu ihtiyaçları ele almaktadır.

Bulgular, araç hızına alternatif olarak GPS hızının yakıt tüketimi ve egzoz emisyonlarını tahmin etmek için güvenilir ve etkili bir parametre olduğunu göstermektedir. GPS hızı ve araç hızı ölçümleri arasındaki düşük hata yüzdesi (%1,4), GPS verilerinin gerçek zamanlı izleme için kullanılmasının fizibilitesini göstermektedir. Topluluk torbalı modeli, tüm veri kümeleri ve parametreler arasında karar ağacı modeline kıyasla üstün performans sergilemiş ve karmaşık, yüksek boyutlu veriler için uygulanabilirliğini güçlendirmiştir.

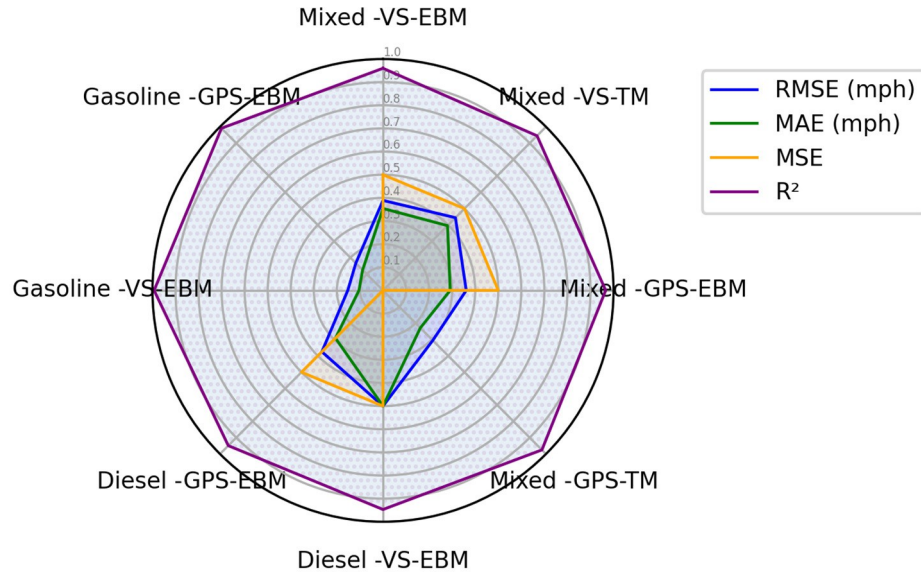
Çalışmada ayrıca benzinli ve dizel motor veri setlerinin birlikte kullanımı da araştırılmıştır. İlginç bir şekilde, sadece benzinli veri seti yakıt tüketimi tahminlerinde en yüksek doğruluğu elde ederken, birleştirilmiş veri seti CO_2 emisyonları için tahmin doğruluğunu artırmıştır. Bu bulgular, veri kümelerinin birleştirilmesinin daha geniş varyasyonları yakalayabileceğini ve özellikle emisyon tahminleri için model sağlamlığını artırabileceğini göstermektedir.

Benzin ve dizel veri kümeleri üzerinde eğitilen modellerin birleştirilmiş veri kümesi ile karşılaştırılması, makine öğrenimi uygulamalarında veri çeşitliliğinin önemini vurgulamaktadır. Benzinli veri kümeleriyle elde edilen yüksek doğruluk, girdi parametrelerinin tutarlılığına bağlanabilirken, dizel motorlar için biraz daha düşük performans, veri kalitesindeki potansiyel değişkenliğe veya motor davranışındaki doğal farklılıklara işaret etmektedir. Şekil 5-7, bir radar grafiği kullanarak modellerin karşılaştırmasının ayrıntılı bir özetini sunmaktadır. Bu Şekillerde, R^2 'nin ayırt edilebilir kalmasını sağlarken görsel temsillerini ölçeklendirmek için performans metrikleri 0 ile 0,5 arasında normalleştirilmiştir.

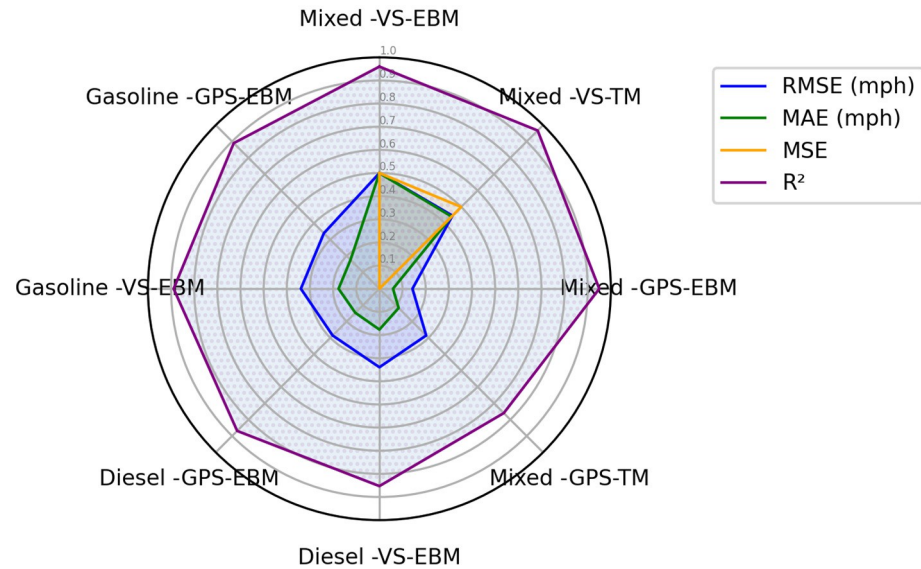


Şekil 5. Yakıt tüketimi için modellerin performansını özetleyen radar grafiği. Burada, VS araç hızını, EBM topluluk torbalı yöntemini ve TM ağaç yöntemini temsil etmektedir.

Bu çalışmanın önemli bir katkısı, GPS verilerinin, özellikle hız sensörlerinin kurulmasının pratik olmayabileceği senaryolarda, araç hız verilerine uygulanabilir bir alternatif olarak gösterilmesidir. Sonuçlar, araca entegre sistemlerle ilişkili lojistik ve yasal karmaşıklıklar olmadan gerçek zamanlı yakıt tüketimini ve emisyonları ölçebilen ITS çözümlerinin önünü açmaktadır.



Şekil 6 CO₂ için modellerin performansını özetleyen radar grafiği. Burada, VS araç hızını, EBM topluluk torbalı yöntemini ve TM ağaç yöntemini temsil etmektedir.



Şekil 7. NO_x için modellerin performansını özetleyen radar grafiği. Burada, VS araç hızını, EBM topluluk torbalı yöntemi ve TM ağaç yöntemini temsil etmektedir.

Ancak, bu çalışmanın kapsamı hafif hizmet araçları ve belirli egzoz kirleticileri (CO₂ ve NO_x) ile sınırlıdır. Bu bulguları ağır hizmet araçlarına, ilave kirleticilere ve derin öğrenme gibi alternatif makine öğrenimi tekniklerine genişletmek için daha fazla araştırma yapılması gerekmektedir. Gelecekteki araştırmalar, daha genelleştirilmiş ve ölçeklenebilir modeller geliştirmek için değişen yol koşullarının, araç türlerinin ve çevresel faktörlerin etkisini de incelemelidir.

Bu araştırma, enerji kullanımını optimize etmeyi ve çevresel etkileri en aza indirmeyi amaçlayan politika yapıcılar ve ulaşım planlamacıları için pratik bilgiler sunarak sürdürülebilir ulaşım için makine öğreniminden yararlanmada ileriye doğru atılmış bir adımı temsil etmektedir.

6. Sonuçlar

Bu çalışmada, benzinli araçlarda yakıt tüketimi ve egzoz emisyonlarının (CO₂ ve NO_x) tahmin edilmesi için topluluk torbalı ve karar ağacı algoritmalarının uygulanması araştırılmıştır.

ve dizel motorlu araçlar için araç hızı ve GPS tabanlı hız girdi parametreleri kullanılarak geliştirilmiştir. Araştırma ayrıca model doğruluğunu artırmak için benzinli ve dizel motor veri kümelerini birleştirmenin etkisini de araştırmıştır.

Sonuçlar, hem topluluk torbalı hem de karar ağacı modellerinin, hem araç hızı hem de GPS hız girdileri için yakıt tüketimi tahminlerinde %95'i aşan performansla yüksek tahmin doğruluğu elde ettiğini göstermiştir. Topluluk torbalı modeli, karar ağacı modelinden sürekli olarak daha iyi performans göstererek karmaşık veri kümelerini işlemdeki üstün kabiliyetini ortaya koymuştur. Özellikle, GPS tabanlı girdi modelleri araç hızı tabanlı modellerle karşılaştırılabilir bir performans göstererek akıllı ulaşım sistemlerinde (ITS) gerçek zamanlı modelleme uygulamaları için GPS verilerinin uygulanabilirliğini vurgulamıştır.

Ayrıca analiz, benzin ve dizel veri kümelerinin kombinasyonunun CO₂ emisyonları için tahmin doğruluğunu artırdığını ve karma veri kümesinin tek tek motor veri kümelerinden daha iyi performans gösterdiğini ortaya koymuştur. Bununla birlikte, benzinli motor veri kümesi, her iki girdi değişkeni kategorisinde de %99'a ulaşarak en yüksek tahmin doğruluğunu sergilemiştir.

Bu araştırma, yakıt tüketimi ve egzoz emisyonlarının gerçek zamanlı izlenmesi için makine öğrenimi modellerinde GPS hız verilerinden yararlanma potansiyelini vurgulamaktadır. Bu tür gelişmeler, politika yapıcılara ulaşım sistemlerini optimize etmede ve çevresel etkileri azaltmada önemli ölçüde yardımcı olabilir. Gelecekteki çalışmalar, bu araştırmanın kapsamını ek egzoz kirleticileri, ağır hizmet araçları ve tahmin yeteneklerini daha da geliştirmek ve diğer makine öğrenimi yöntemleriyle karşılaştırmalı analiz yapmak için derin öğrenme tekniklerinin entegrasyonunu içerecek şekilde genişletmeye odaklanmalıdır.

Yazar Katkıları: Çalışma konsepti ve tasarımı: F.A.; veri toplama: F.A., A.A., M.A. ve N.M.; sonuçların analizi ve yorumlanması: F.A., A.A. ve M.A.; taslak makale hazırlama: F.A., A.A. ve M.A. Tüm yazarlar makalenin yayınlanan versiyonunu okumuş ve kabul etmiştir.

Finansman: Bu araştırma herhangi bir dış fon almamıştır.

Veri Kullanılabilirlik Beyanı: Bu çalışmada sunulan veriler ilgili yazarın talebi üzerine temin edilebilir.

Teşekkür: Yazarlar, makalede kullanılan verileri sağladığı için Emissions Analytics'e teşekkür eder.

Çıkar Çatışmaları: Yazarlar herhangi bir çıkar çatışması beyan etmemişlerdir.

Referanslar

1. Avrupa'da Ulaşımdan Kaynaklanan Sera Gazı Emisyonları . Çevrimiçi olarak mevcuttur: <https://www.eea.europa.eu/data-and-maps/indicators/transport-emissions-of-greenhouse-gases/transport-emissions-of-greenhouse-gases-12> (14 Ocak 2025 tarihinde erişilmiştir).
2. Wang, Z.; Chen, F.; Fujiyama, T. Pekin'de Kentsel Yolcu Taşımacılığından Kaynaklanan Karbon Emisyonu. *Transp. Res. Part D Transp. Environ.* **2015**, *41*, 217-227. [CrossRef]
3. Int Panis, L.; Broekx, S.; Liu, R. Anlık Trafik Emisyonlarının Modellenmesi ve Trafik Hız Sınırlarının Etkisi. *Sci. Total Environ.* **2006**, *371*, 270-285. [CrossRef]
4. Sullivan, J.L.; Baker, R.E.; Boyer, B.A.; Hammerle, R.H.; Kenney, T.E.; Muniz, L.; Wallington, T.J. Dizel (Benzine karşı) Motorlu Araçların CO₂ Emisyon Faydası. *Environ. Sci. Technol.* **2004**, *38*, 3217-3223. [CrossRef]
5. Ahmed, E.; Usman, M.; Anwar, S.; Ahmad, H.M.; Nasir, M.W.; Malik, M.A.I. Benzin-Metanol Karışımları Kullanan SI Motorun Performansını ve Emisyonlarını Tahmin Etmek için YSA Uygulaması. *Sci. Prog.* **2021**, *104*, 00368504211002345. [CrossRef]
6. Liu, R.; He, H.; Zhang, Z.; Wu, C.; Yang, J.; Zhu, X.; Peng, Z. Hafif Hizmet Benzinli Araçlardan CO₂ ve NO Tahmini için Entegre MOVES Modeli ve Makine Öğrenme Yöntemi. *J. Clean. Prod.* **2023**, *422*, 138612. [CrossRef]
7. Clark, N.N.; Kern, J.M.; Atkinson, C.M.; Nine, R.D. Ağır Hizmet Dizel Araç Emisyonlarını Etkileyen Faktörler. *J. Air Waste Manag. Doç.* **2002**, *52*, 84-94. [CrossRef]
8. Zhang, K.; Yao, L.; Li, G. Araç Emisyonlarını Etkileyen Faktörler ve Emisyon Modelleri. *ICTE 2013 içinde: Safety, Speediness, Intelligence, Low-Carbon, Innovation*; American Society of Civil Engineers: Reston, VA, ABD, 2013; s. 2848-2853. [CrossRef]
9. Faris, W.F.; Rakha, H.A.; Kafafy, R.I.; Idres, M.; Elmoselhy, S. Taşıt Yakıt Tüketimi ve Emisyon Modellemesi: Derinlemesine Bir Literatür Taraması. *Int. J. Veh. Syst. Model. Test.* **2011**, *6*, 318. [CrossRef]

10. Araç Emisyon Modellemesi ve Yeni Zelanda için Sorunların İncelenmesi. Çevrimiçi olarak mevcuttur: https://www.researchgate.net/publication/289082445_Review_of_vehicle_emission_modelling_and_the_issues_for_New_Zeland (14 Ocak tarihinde erişilmiştir 2025).
11. Alrumaidhi, M.; Farag, M.M.G.; Rakha, H.A. Sentetik Yeniden Örnekleme Teknikleri Kullanarak Yaşlı Sürücüler Arasında Yol Kazası Şiddetini Tahmin Etmek İçin Parametrik ve Parametrik Olmayan Veriye Dayalı Modellerin Karşılaştırmalı Analizi. *Sürdürülebilirlik* **2023**, *15*, 9878. [CrossRef]
12. Alrumaidhi, M.; Rakha, H.A. Çarpışma Şiddetini Etkileyen Faktörlerin Zamansal Değişkenliğini Keşfetmek için Ekonometrik Bir Analiz COVID-19 Nedeniyle. *Sürdürülebilirlik* **2024**, *16*, 1233. [CrossRef]
13. Alrumaidhi, M.; Rakha, H.A. Yaşlı Sürücüler Arasında Çarpışma Şiddetini Etkileyen Faktörler: Çok Düzeyli Ordinal Lojistik Regresyon Yaklaşımı. *Sürdürülebilirlik* **2022**, *14*, 11543. [CrossRef]
14. Alazmi, A.; Rakha, H. Makine Öğreniminin Rafine Edilmemiş Partikül Hava Kirliliği Mobil İzleme Verilerini Rastgele, Mekansal ve Mekansal-Zamansal Olarak İşleme Yeteneğinin Değerlendirilmesi ve Doğrulanması. *Int. J. Environ. Res. Halk Sağlığı* **2022**, *19*, 10098. [CrossRef] [PubMed]
15. Altug, K.B.; Küçük, S.E. Denetimli Öğrenme Algoritmaları Kullanarak Egzoz Borusu NO_x Emisyonunun Tahmin Edilmesi. In Proceedings of the 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 11-13 October 2019; pp. 1-5.
16. Shin, S.; Lee, Y.; Park, J.; Kim, M.; Lee, S.; Min, K. Zaman Serisi Verileri Kullanarak Geçici Dizel Motor NO_x Emisyonlarının Tahmin Edilmesi Derin Öğrenme Modelleri ile Ön İşleme. *Proc. Inst. Mech. Eng. Bölüm D J. Automob. Eng.* **2021**, *235*, 3170-3184. [CrossRef]
17. Rivera-Campoverde, N.D.; Arenas-Ramírez, B.; Muñoz Sanz, J.L.; Jiménez, E. GPS Verileri ve Makine Öğrenimi Araçları, Hafif Araç Emisyonlarını Tahmin Etmek için Pratik ve Uygun Maliyetli Bir Kombinasyon. *Sensörler* **2024**, *24*, 2304. [CrossRef]
18. Ma, dziel, M. Sivilleştirilmiş Petrol Gazı ile Çalışan Araç CO₂ Emisyonunun Taşınabilir Emisyon Ölçüm Sistemi, Araç İçi Teşhis Verileri ve Gradient-Boosting Makine Öğrenmesine Dayalı Modellenmesi. *Energies* **2023**, *16*, 2754. Çevrimiçi olarak mevcuttur: <https://www.mdpi.com/1996-1073/16/6/2754> (14 Ocak 2025 tarihinde erişilmiştir). [CrossRef]
19. Yao, Y.; Li, J.; He, C.; Chen, Y.; Yu, H.; Wang, J.; Yang, N.; Zhao, L. Hafif Hizmet Hibrit Elektrikli Araçların Gerçek Sürüşte Partikül Emisyonları Üzerine Araştırma. *Atmosfer. Pollut. Res.* **2025**, *16*, 102332. [CrossRef]
20. Motallebiaraghi, F.; Rabinowitz, A.; Jathar, S.; Fong, A.; Asher, Z.; Bradley, T.; Motallebiaraghi, F.; Rabinowitz, A.; Jathar, S.; Fong, A.; vd. *Derin Sinir Ağları Kullanılarak Hafif Hizmet Araç Emisyonu ve Yakıt Ekonomisinin Yüksek Doğruluklu Modellemesi*; SAE International: Warrendale, PA, ABD, 2021.
21. Lee, J.; Kwon, S.; Kim, H.; Keel, J.; Yoon, T.; Lee, J. Gerçek Sürüş Döngüsü Altında Dizel Aracın NO_x Tahminine Uygulanan Makine Öğrenimi. *Appl. Sci.* **2021**, *11*, 3758. [CrossRef]
22. Cha, J.; Park, J.; Lee, H.; Chon, M.S. Hafif Hizmet Dizel Araçları ile Gerçek Dünya CO₂ Emisyonları için Regresyon Analizine Dayalı Tahmin Çalışması. *Int. J. Automot. Technol.* **2021**, *22*, 569-577. [CrossRef]
23. Shahariar, G.M.H.; Bodisco, T.A.; Surawski, N.; Komol, M.M.R.; Sajjad, M.; Chu-Van, T.; Ristovski, Z.; Brown, R.J. Gerçek Sürüş CO₂, NO_x ve Yakıt Tüketiminin Makine Öğrenimi Yaklaşımlarıyla Tahmini. *Next Energy* **2023**, *1*, 100060. [CrossRef]
24. Wen, H.-T.; Lu, J.-H.; Jhang, D.-S. Dizel Araçların Gerçek NO_x ve CO₂ Emisyon Tahminlerinin Özellik Önem Analizi Gradient Boosting Regresyon Modeline Dayalı Yol Sürüşü. *Int. J. Environ. Res. Public Health* **2021**, *18*, 13044. [CrossRef]
25. Alfaseeh, L.; Tu, R.; Farooq, B.; Hatzopoulou, M. Derin Dizi Öğrenme Kullanarak Yol Ağı Üzerinde Sera Gazı Emisyonu Tahmini. *Transp. Res. Part D Transp. Environ.* **2020**, *88*, 102593. [CrossRef]
26. Kocev, D.; Vens, C.; Struyf, J.; Džeroski, S. Çok Amaçlı Karar Ağaçları Toplulukları. *Makine Öğrenimi* içinde: *ECML 2007*; Kok, J.N., Koronacki, J., de Mantaras, R.L., Matwin, S., Mladenić, D., Skowron, A., Eds.; Springer: Berlin/Heidelberg, Almanya, 2007; pp. 624-631.
27. Xia, Y.; Jiang, L.; Wang, L.; Chen, X.; Ye, J.; Hou, T.; Wang, L.; Zhang, Y.; Li, M.; Li, Z.; et al. Yol Üstü Uzaktan Algılama ve Makine Öğrenimi Kullanılarak Hafif Hizmet Benzinli Araç Emisyonlarının Hızlı Değerlendirilmesi. *Bilim Toplamı. Environ.* **2022**, *815*, 152771. [CrossRef] [PubMed]
28. Akkamis, M.; Keskin, M.; Şekerli, Y.E. Farklı Veri Güncelleme Frekanslarına Sahip Üç Düşük Maliyetli GPS Hız Sensörünün Karşılaştırmalı Değerlendirmesi. *AgriEngineering* **2021**, *3*, 423-437. [CrossRef]
29. Quinlan, J.R. Karar Ağaçlarının İndüksiyonu. *Mach. Learn.* **1986**, *1*, 81-106. [CrossRef]
30. Breiman, L. Tahmin Edicileri Torbalama. *Mach. Learn.* **1996**, *24*, 123-140. [CrossRef]
31. Peng, L.; Zheng, S.; Zhong, Q.; Chai, X.; Lin, J. Ray Denetim Verilerini Kullanarak Tren Gövdesi Titreşimlerini Tahmin Etmek İçin Çoklu Korelasyon Katsayılarına Sahip Yeni Bir Torbalı Ağaç Ensemble Regresyon Yöntemi. *Mech. Sist. Sinyal Süreci.* **2023**, *182*, 109543. [CrossRef]
32. Sun, J.; Lang, J.; Fujita, H.; Li, H. DTE-SBD ile Dengesiz Kurumsal Kredi Değerlendirmesi: SMOTE ve Farklılaştırılmış Örnekleme Oranları ile Torbalama Tabanlı Karar Ağacı Topluluğu. *Inf. Sci.* **2018**, *425*, 76-91. [CrossRef]
33. Chen, K.; Peng, Y.; Lu, S.; Lin, B.; Li, X. Belediye Katı Atık Yakma Fırınlarından PCDD/Fs Emisyonunun Modellenmesi için Torbalama Tabanlı Topluluk Öğrenme Yaklaşımları. *Chemosphere* **2021**, *274*, 129802. [CrossRef]

34. Elbeltagi, A.; Heddami, S.; Katipoglu, O.M.; Alsumaiei, A.A.; Al-Mukhtar, M. RReliefF Algoritması ile Geliştirilmiş Makine Öğrenimi Modellerine Dayalı 1958-2021 Yılları İçin Nemli İklimlerde Gelişmiş Uzun Vadeli Gerçek Evapotranspirasyon Tahmini. *J. Hydrol. Reg. Stud.* **2024**, *56*, 102043. [[CrossRef](#)]
35. Gis, W.; Bielaczyc, P.; Gis, W.; Bielaczyc, P. *Emission of CO₂ and Fuel Consumption for Automotive Vehicles*; SAE International: Warrendale, PA, ABD, 1999.
36. Harrington, W. Yakıt Ekonomisi ve Motorlu Araç Emisyonları. *J. Environ. Econ. Manag.* **1997**, *33*, 240-252. [[CrossRef](#)]
37. Pinto, G.; Oliver-Hoyo, M.T. Araç Yakıt Tüketimi ve CO₂ Emisyonları Arasındaki İlişkinin Kimyasal İlkelerini Açıklamak İçin Kullanılması. *J. Chem. Eğitim.* **2008**, *85*, 218. [[CrossRef](#)]
38. Pavlovic, J.; Fontaras, G.; Broekaert, S.; Ciuffo, B.; Ktistakis, M.A.; Grigoratos, T. Gerçek Dünya Operasyonunda Araç Yakıt Tüketimini Ne Kadar Doğru Ölçebiliriz? *Transp. Res. Part D Transp. Environ.* **2021**, *90*, 102666. [[CrossRef](#)]

Sorumluluk Reddi/Yayıncının Notu: Tüm yayınlarda yer alan ifadeler, görüşler ve veriler yalnızca yazar(lar)a ve katkıda bulunan(lar)a aittir, MDPI ve/veya editör(ler)e ait değildir. MDPI ve/veya editör(ler), içerikte atıfta bulunulan herhangi bir fikir, yöntem, talimat veya üründen kaynaklanan insanlara veya mallara gelebilecek herhangi bir zarar için sorumluluk kabul etmez.