

## Araçlarda CO<sub>2</sub> Emisyonlarının Farklı Yapay Sinir Ağı Modelleri Kullanılarak Tahminlerinin Karşılaştırılması

Emrah ASLAN\*<sup>1</sup> ORCID 0000-0002-0181-3658

<sup>1</sup>Dicle Üniversitesi, Silvan Meslek Yüksekokulu, Bilgisayar Programcılığı Bölümü, Diyarbakır

Geliş tarihi: 14.03.2024 Kabul tarihi: 27.06.2024

Atıf şekli/ How to cite: ASLAN, E., (2024). Araçlarda CO<sub>2</sub> Emisyonlarının Farklı Yapay Sinir Ağı Modelleri Kullanılarak Tahminlerinin Karşılaştırılması. Çukurova Üniversitesi, Mühendislik Fakültesi Dergisi, 39(2), 309-324.

### Öz

İklim değişikliği, insanlık için en büyük çevresel tehlikelerden biridir. İklim değişikliğinde karbondioksit (CO<sub>2</sub>), sera etkisinin başlıca sebeplerindedir. Ulaşım sektörü, büyük CO<sub>2</sub> emisyon kaynaklarından birini oluşturmaktadır. Bu makale, araçlarının anlık CO<sub>2</sub> emisyonlarını tahmin etmek için bir yapay sinir ağı (YSA) modeli sunmaktadır. Araçlarda CO<sub>2</sub> emisyonlarını tahmin etmek için Linear Regresyon, XGBoost Regresör ve K-Nearest Neighbours Regresörü olmak üzere üç regresyon modeli kullanılarak kapsamlı bir yaklaşım kullanılmıştır. Araştırma, araçlardaki CO<sub>2</sub> emisyonlarını tahmin etmek ve analiz etmek için bu yapay sinir ağlarının yeteneklerinden yararlanmaya odaklanmaktadır. Farklı modellerin kullanılması, doğruluk ve verimlilik açısından performanslarının karşılaştırmalı olarak değerlendirilmesine olanak sağlamaktadır. Yüksek doğruluk ve uygulanabilirlik sağlayan bu yöntem, motor hacmi, silindiri, şehir içi ve şehir dışı yakıt tüketimi gibi parametreler ile egzoz emisyonlarının öngörücüleri olarak kullanılmaktadır. Her parametrenin emisyon tahminlerine olan önemi, test ve eğitim doğruluğu, kök ortalama kare hatası, ortalama mutlak hata, R<sup>2</sup> skor gibi sonuçlar karşılaştırılarak kapsamlı bir şekilde analiz edilmiştir. Bu çalışma, özellikle araç emisyonları bağlamında CO<sub>2</sub> emisyon tahmin metodolojilerinin ilerlemesine katkıda bulunmayı amaçlamaktadır. Bu araştırmanın bulguları, ulaştırma sektöründe karbon ayak izlerini azaltmak için sürdürülebilir çözümler arayan politika yapıcılar, çevreciler ve otomotiv mühendisleri için önem taşımaktadır.

**Anahtar Kelimeler:** CO<sub>2</sub> emisyonu, Yapay sinir ağı, Yakıt tüketimi

### Comparison of CO<sub>2</sub> Emissions Prediction in Vehicles Using Different Artificial Neural Network Models

#### Abstract

Climate change is one of the greatest environmental threats to humanity. Carbon dioxide (CO<sub>2</sub>) is one of the main causes of the greenhouse effect in climate change. The transportation sector is one of the major sources of CO<sub>2</sub> emissions. This paper presents an artificial neural network (ANN) model for estimating the instantaneous CO<sub>2</sub> emissions of vehicles. A comprehensive approach using three regression models, namely

\*Sorumlu yazar (Corresponding Author): Emrah ASLAN, [emrah.aslan@dicle.edu.tr](mailto:emrah.aslan@dicle.edu.tr)

Linear Regression, XGBoost Regressor and K-Nearest Neighbors Regressor, is used to predict CO<sub>2</sub> emissions from vehicles. The research focuses on leveraging the capabilities of these artificial neural networks to predict and analyze CO<sub>2</sub> emissions from vehicles. The use of different models allows for a comparative evaluation of their performance in terms of accuracy and efficiency. This method, which provides high accuracy and applicability, uses parameters such as engine displacement, cylinder, urban and non-urban fuel consumption as predictors of exhaust emissions. The importance of each parameter to emission predictions is comprehensively analyzed by comparing results such as test and training accuracy, root mean square error, mean absolute error, R<sup>2</sup> score. This study aims to contribute to the advancement of CO<sub>2</sub> emission estimation methodologies, especially in the context of vehicle emissions. The findings of this research are important for policy makers, environmentalists and automotive engineers seeking sustainable solutions to reduce carbon footprints in the transportation sector.

**Keywords:** CO<sub>2</sub> emissions, ANN, Fuel Consumption

## 1. GİRİŞ

Çoğunlukla insan faaliyetlerinden kaynaklanan CO<sub>2</sub> emisyonları, küresel iklim değişikliğinin giderek artan zorluklarına önemli bir katkıda bulunmaktadır. Bu emisyonların azaltılması zorunluluğu, sera etkisini artırmadaki önemli rollerinden kaynaklanmakta ve ekosistemler, hava durumu ve deniz seviyeleri üzerinde olumsuz etkilere yol açmaktadır. CO<sub>2</sub> emisyonlarının farklı kökenlerinin tanınması, çevrenin korunmasına yönelik bütüncül stratejilerin formüle edilmesi için temel teşkil etmektedir [1].

Çevresel sürdürülebilirlik konusundaki çağdaş söylemde, karbondioksit (CO<sub>2</sub>) emisyonlarının doğru tahmin edilmesi ve azaltılması zorunlu hedefler olarak öne çıkmaktadır. Toplumlar iklim değişikliğinin sonuçlarıyla boğuşurken, önemli bir karbon ayak izine sahip olan ulaştırma sektörü, araştırma ve müdahale için bir odak noktası haline gelmiştir [2]. Bu çalışma, etkinliklerini ve nüanslarını ayırt etmek için çeşitli Yapay Sinir Ağı (YSA) modellerinin karşılaştırmalı bir analizini kullanarak araçlarda CO<sub>2</sub> emisyonu tahminini araştırmaktadır.

Karbondioksit (CO<sub>2</sub>) emisyonlarının daha geniş bağlamında, ulaşım sektörü önemli bir rol üstlenmekte ve çevresel zorluklara önemli bir katkıda bulunmaktadır [3]. Geleneksel araçlarda yaygın olan ve yakıtın yanması sırasında önemli miktarlarda CO<sub>2</sub> açığa çıkaran içten yanmalı motorlar özellikle dikkat çekmektedir. Bu emisyon

kaynağı, kentleşmenin hızlanması ve buna bağlı olarak küresel ulaşım talebindeki artış göz önüne alındığında daha da önem kazanmaktadır. Bu eğilimlerin kesişmesi, genişleyen kentsel manzara ve artan ulaşım ihtiyaçları, geleneksel araçlardan kaynaklanan CO<sub>2</sub> emisyonlarının etkisini toplu olarak artırdığından, araç emisyonlarını ele alma aciliyeti vurgulamaktadır [4,5].

Araçlardan kaynaklanan CO<sub>2</sub> emisyonlarının kesin bir şekilde anlaşılması çok önemlidir. Böyle bir kavrayış, çevresel etkileri azaltmayı amaçlayan bilinçli karar alma ve stratejik planlama için gereklidir [6]. Ulaşım sektörünün çok önemli rolünü kabul eden bu çalışma, kapsamlı bir karşılaştırma için yapay sinir ağı modelleri kullanarak CO<sub>2</sub> emisyonlarını tahmin etme ve öngörme gibi incelikli bir görevi ele almaktadır. Bu sayede, sürdürülebilir uygulamaların ilerletilmesi ve araç faaliyetleriyle ilişkili çevresel ayak izinin azaltılmasına yönelik devam eden arayışta politika değerlendirmelerinin bilgilendirilmesi için çok önemli iç görüleri katkıda bulunmayı amaçlamaktadır [7].

Yapay zeka (YZ) alanındaki son adımlar, çevresel zorluklara yenilikçi çözümler için yollar açmıştır. Araç emisyonları bağlamında, Yapay Sinir Ağları (YSA'lar) CO<sub>2</sub> emisyonlarını tahmin etmek ve öngörmek için sofistike bir araç sunmaktadır. İnsan beyninin karmaşık mimarisinden esinlenen bu modeller, verilerdeki karmaşık örüntüleri ayırt etmede mükemmeldir. Araştırmacılar, YSA'nın yeteneklerinden yararlanarak, araçlardaki CO<sub>2</sub>

emisyonlarını etkileyen çok yönlü faktörleri inceleyebilir, daha doğru tahminler ve bilinçli karar vermeyi kolaylaştırabilir [8].

Bu araştırma, çeşitli YSA modellerinin kapsamlı bir karşılaştırmasını yaparak araçlardaki CO<sub>2</sub> emisyonları hakkındaki söylemlere katkıda bulunmayı amaçlamaktadır. Bu çalışma sayesinde, farklı sinir ağı mimarilerinin güçlü yönlerine ve sınırlamalarına ışık tutmayı ve sürdürülebilir ve çevreye duyarlı bir geleceğe giden yolda ilerlemeye çalışan politika yapımcılar, araştırmacılar ve endüstri paydaşları için değerli bilgiler sağlamayı amaçlamaktadır.

Bu çalışmanın Bölüm 1 kısmında problem hakkında genel bilgilere yer verilmiştir. Çalışmanın ilerleyen kısımları şu şekildedir. Bölüm 2 CO<sub>2</sub> emisyon tahmini için yapılmış literatür çalışmalarını içermektedir. Bölüm 3 uygulamalar için gerekli olan materyal ve metodu, Bölüm 4 yapılan çalışmalar ve tartışmayı kapsamaktadır. Bölüm 5'te ise, sonuçlar ve gelecekteki çalışmalar açıklanmıştır.

## 2. LİTERATÜR TARAMASI

Zeng ve arkadaşları, çalışmalarında CO<sub>2</sub> tahmini için Yapay Sinir Ağı, Destek Vektör Makinesi (DVM) ve VT-Micro tekniklerini kullanmışlardır. Ayrıca, araçlar için çevresel olarak en uygun rotayı bulmak için, On-Board Diagnostics (OBD) portu aracılığıyla 30 saniye aralıklarla hız, ivme ve gaz verisi kullanmışlardır. Yazarlar tarafından belirtilen sorun, sadece 2 özellikle CO<sub>2</sub> emisyonunun kötü tahmin edilmesidir [9]. Oduro ve ark., çalışmalarında aracın hızı ve ivmesini kullanarak CO<sub>2</sub> tahmininin regresif analizini yapmışlardır. Elde edilen sonuçlar, araç hızı ve ivmesinin CO<sub>2</sub>, hız ve ivme arasında lineer bir ilişki gösterdiğini göstermektedir. Hız, ivmeye kıyasla CO<sub>2</sub> emisyon faktörü ile daha güçlü bir ilişkiye sahiptir [10].

Razak ve arkadaşları, çeşitli yüklerde bir dizel motorunun GHG emisyonlarını azaltmak için farklı yakıtların optimal karışımını bulmak için lineer programlama yaklaşımını kullanmıştır. Çalışmaları, %70 dizel, %20 biyodizel ve %10 butanol

karışımının kullanılmasının sülfür ve CO<sub>2</sub> emisyonlarında sırasıyla %35 ve %36 azalmaya yol açtığını göstermiştir [11].

Shim ve arkadaşları, çalışmalarında bir çift yakıtlı dizel motorunun toplu söndürme olayını aşmak için eşdeğerlik oranı ve birincil yüklem sıcaklığının etkisini değerlendirdi. Araştırmaları, şarj sıcaklığının ve hava miktarının kontrolünün düşük yüklerde toplu söndürme olayını hafifletmek için kullanışlı bir teknik olabileceğini ortaya koydu [12].

Prabhu ve arkadaşları, CO<sub>2</sub> ve NO<sub>x</sub> gibi kirletici emisyonları ölçmek için bir SCD motorunu ele almıştır. Yakıt ön ısıtılmış biyogaz içeriyordu ve ön ısıtmanın ısı yükünü sağlamak için motordan çıkan ısının bir ısı değiştiricisinde geri kazanılması sağlanmıştır. Sonuçlar, ön ısıtma işleminin %40,5'lik bir fren biyogaz tüketimine yol açtığını kanıtlamıştır [13].

Roy ve arkadaşları, dört zamanlı bir SCD'nin motor performansını ve emisyonlarını tahmin ettiler. Bu çalışmada, ESFC,  $\eta_{th}$ , CO<sub>2</sub>, NO<sub>x</sub> ve PM gibi parametreler değerlendirildi. Sarace ve ark., AIN kullanarak cerium oksit nanopartiküllerinin dizel motor performansı ve emisyonları üzerindeki etkisini incelediler. Sonuçları, cerium oksit nanopartiküllerinin NO<sub>x</sub> ve HC'yi önemli ölçüde azalttığını, CO emisyon oranlarını hafifçe arttırdığını ve yakıt tüketimini azalttığını gösterdi [14].

Alfaseeh ve arkadaşları, uzun-kısa dönem belleği kullanarak kentsel bir alandaki sera gazı emisyonlarını link düzeyinde tahmin etmek için çeşitli sayıda sinir ağı katmanına sahip çoklu modeller geliştirdiler ve daha derin sinir ağı modelinin gelişmiş tahmin performansı sergilediğini buldular [15].

Claudio ve arkadaşları, Dinamik Programlama (DP) algoritması tarafından tahmin edilen tanktan tekerleğe karbondioksit emisyonlarının tahmini için yenilikçi, derin sinir ağlarına dayalı bir model sunulmuştur. Kapsamlı bir simülasyon seti sırasında %91'in üzerinde sınıflandırma

performansları ve %1'in altında ortalama regresyon hataları elde etmişlerdir [16].

Jigu ve Sungwook yüksek derecede doğruluk ve pratiklik sunan YSA kullanarak dizel araçların karbondioksit nitrojen oksit ve karbon monoksit emisyonlarını tahmin etmek için yeni bir yaklaşım sunmuşlardır. Çalışmalarında dikkate alınan altı parametreye ek olarak gaz pedalı konumu, gaz kelebeği konumu, VSP, egzoz gazı sıcaklığı, yol eğimi ve vites konumu gibi diğer parametreler de YSA girişi olarak kullanılmıştır [17].

Natarajan ve arkadaşları, emisyonları tahmin etmek için araç özelliklerinden yararlanan toplu öğrenme modelleri oluşturmak amacıyla, hafif ticari araç tasarımları için karbondioksit emisyonlarını tahmin etmek amacıyla güçlendirme ve diğer regresyon modellerinin geliştirilmesini önermişlerdir. Önerilen güçlendirme modeli Catboost, girdi olarak yalnızca bir araba özelliği olsa bile CO<sub>2</sub> emisyonlarını doğru bir şekilde tahmin etme kapasitesine sahiptir. Catboost, literatürde önerilen diğer algoritmalara göre veri işlemeyi daha kısa sürede ve daha az bellekle gerçekleştirir. Çalışma % 95 bir doğruluk ile çalışmaktadır [18].

Al-Nefae ve arkadaşları, çalışmalarında araçlardaki CO<sub>2</sub> emisyonlarını tahmin etmek için derin öğrenme tabanlı LSTM ve BiLSTM modelleri geliştirerek, bu modellerin performansını Kaggle'dan elde edilen veri setleri üzerinde değerlendirmiştir. BiLSTM modeli, en yüksek doğruluk ve tahmin değerleriyle R<sup>2</sup> %93.78 en iyi performansı göstermiştir [19].

Dong ve arkadaşları, çalışmalarında hibrid elektrikli araçların soğuk başlangıç emisyonlarını tahmin etmek için genetik algoritma ile optimize edilmiş bir sinir ağı modeli geliştirmiştir. Düşük maliyetli ve yüksek tahmin doğruluğuna sahip olan bu model, farklı soğuk başlangıç sıcaklıkları ve girdi parametreleri altında CO, CO<sub>2</sub>, hidrokarbon (THC), NO<sub>x</sub> ve partikül madde sayısı (PN) emisyonlarını başarıyla tahmin edebilmişlerdir [20].

Paul ve arkadaşları, yapay zeka modellerinin enerji tüketimine ve CO<sub>2</sub> emisyonlarına olan etkilerini değerlendirerek, bu modellerin olumlu ve olumsuz etkilerini nicelendiren bir yöntem önermişlerdir. Üç farklı YZ modeli üzerinde yapılan değerlendirmeler, modellerin CO<sub>2</sub> emisyonlarında azalmaya katkıda bulunabileceğini, ancak bu azalmanın bağlama bağlı olarak değişkenlik gösterdiğini göstermiştir [21].

Wang ve arkadaşları, çalışmalarında ağır ticari dizel araçların NO<sub>x</sub> emisyonlarını tahmin etmek için geliştirilmiş bir çift-gizli-katmanlı BP sinir ağı modeli sunmuştur. Yeni bir özellik mühendisliği işleme yöntemi kullanılarak model, gerçek yol emisyon testlerinde yüksek doğrulukla çalışmaktadır. Çalışma sonucunda test doğruluğunu %87.024 olarak bulunmuştur [22].

Mądziel çalışmasında, start-stop teknolojisiyle donatılmış bir araç için CO<sub>2</sub> emisyonlarını tahmin etmek için yapay zeka teknikleri kullanarak bir hesaplama modeli oluşturma metodolojisini sunmuştur. Üç farklı makine öğrenimi tekniği analiz edilmiş ve en iyi tahmin sonuçları gradyan artırma yöntemiyle elde edilmiştir. Geliştirilen modeller, CO<sub>2</sub> emisyonlarını mikro ölçekte çevresel analiz için yeni bir yöntem sunmaktadır [23].

Literatür taraması, mevcut CO<sub>2</sub> tahmini ve modelleme yöntemlerinin temel olarak hız ve sıkışıklık verileri ile birlikte özel sensörlerden oluştuğunu göstermektedir. Hız ve ivme, CO<sub>2</sub> emisyonu ile yüksek derecede ilişkili olmasına rağmen, aracın emisyon karakteristikleri hakkında sınırlı bilgi sağlayabilir. CO<sub>2</sub> modellemesi için özel sensörler, doğru olabilir ancak geniş ölçekli uygulamalar için çok uygun değildir. Dahası, literatürdeki CO<sub>2</sub> tahmini yöntemleri bireysel araç emisyonlarını etkili bir şekilde izlemekten acizdir. Mevcut literatürün incelenmesi, motorların performansını ve kirliliğini araştıran farklı çalışmalar olduğunu göstermektedir. Ancak, yazarların bilgisine göre, motor performans değerlendirmesi ve kirlilik üzerinde giriş hava ve soğutma suyunun etkisini inceleyen deneysel

çalışmalarda bir eksiklik bulunmaktadır. Başka bir deyişle, literatür hala farklı tasarım değişkenlerinin dikkate alınarak motor performansı ve kirletici emisyonların eş zamanlı olarak değerlendirilmesini ve tahmin edilmesini içeren kapsamlı araştırmalara ihtiyaç duymaktadır. Bu eksikliğin yapay zeka yöntemleri kullanılarak giderilebileceği önerilmektedir.

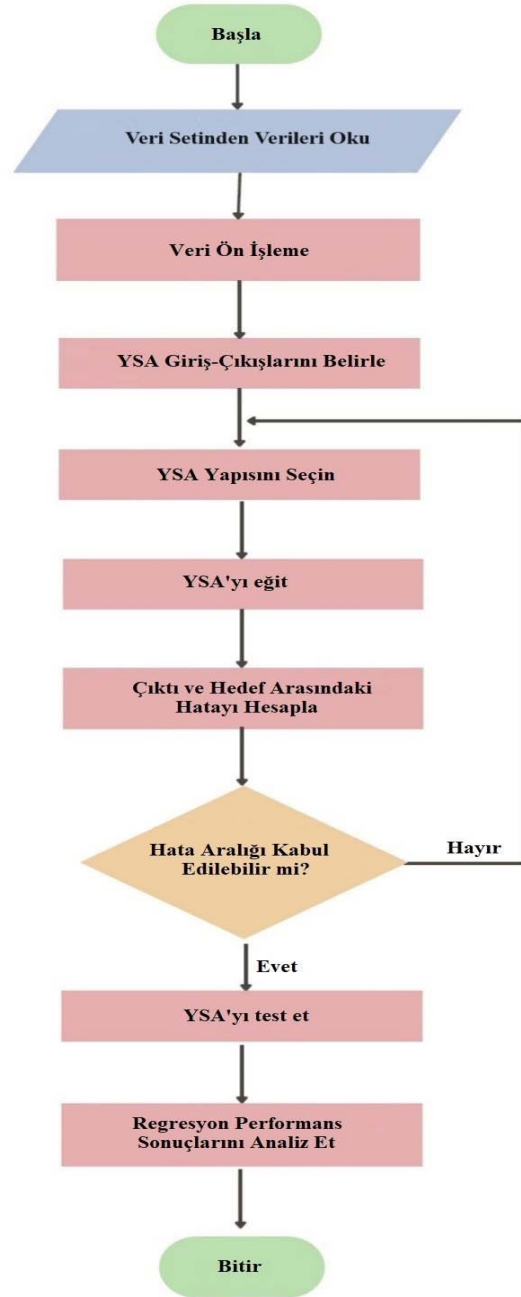
### 3. YÖNTEM

Araçların oluşturduğu emisyonu tespit etmek için Linear, K-Nearest Neighbours, XGBoost yapay sinir ağları kullanılmıştır. Model geliştirme sürecinin akış diyagramı gösterimi Şekil 1'de sunulmuştur. Çalışmada, gerçek araç verilerinden elde edilen verilerden oluşan bir veri seti kullanılmıştır. Veri setinde yer alan verilerin %80 oranında eğitim ve %20 oranında test kümelerine ayrılmıştır. Eğitim verileri, model girdisi için önemli özellikleri belirlemek üzere korelasyon testi de dahil olmak üzere ön işleme tabi tutulmuştur. Ön işlemenin ardından veriler, CO<sub>2</sub> emisyonlarını tahmin etmek üzere regresyon modellerini eğitmek için kullanılır. Modellerin performansı daha sonra tahmin kabiliyetini doğrulamak için test verileri kullanılarak değerlendirilir. Verilen regresyon algoritmalarıyla karşılaştırmalı analiz yapılmıştır.

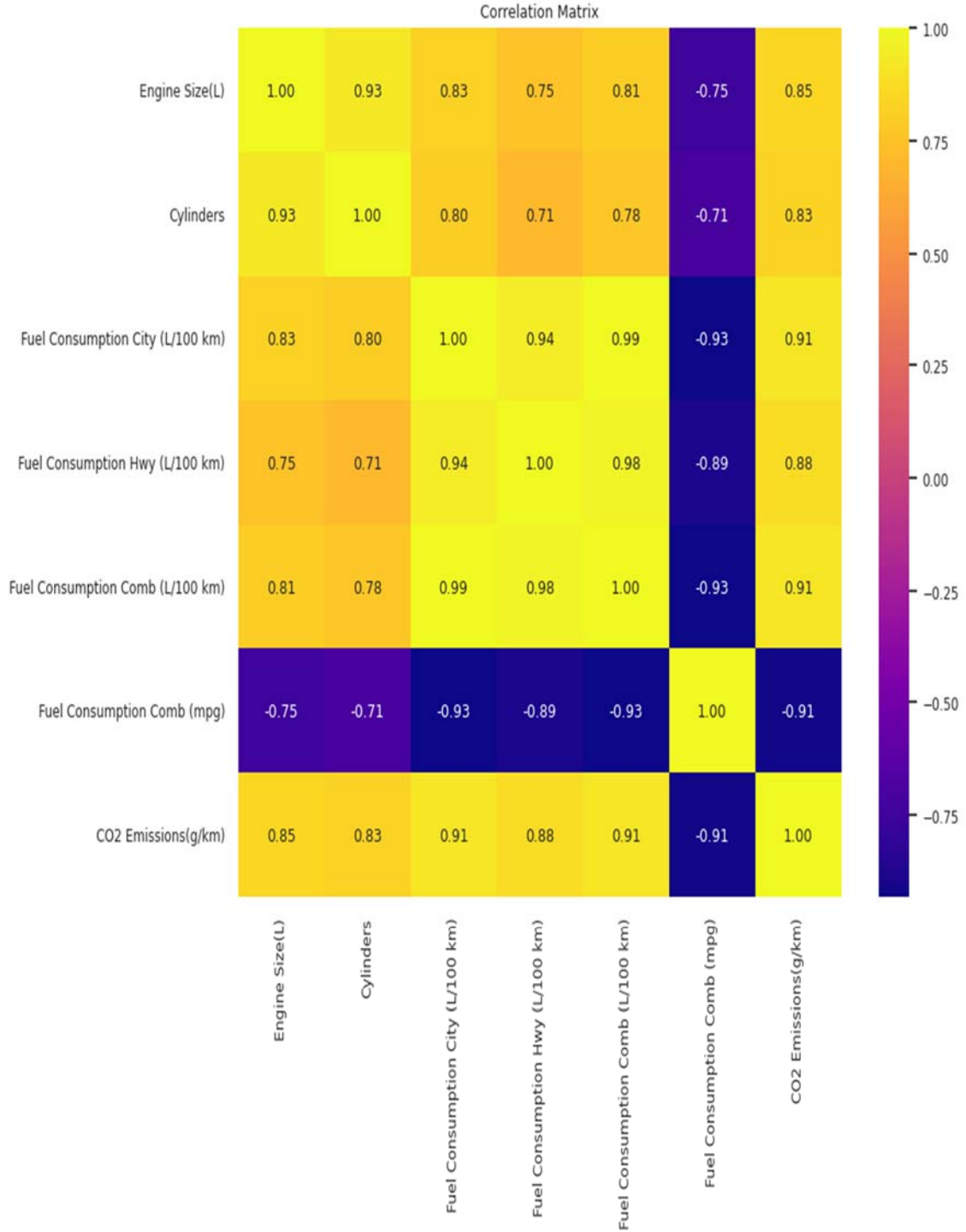
#### 3.1. Veri Seti

Veri seti Kanada'da hafif ticari araçların yakıt tüketimlerini ve CO<sub>2</sub> emisyon verilerini içermektedir. CO<sub>2</sub> emisyonu veri seti farklı marka ve modelde araçlar için toplam 7385 satır, 12 sütundan oluşan veri bulundurmaktadır [24]. Veri seti içerisinde vites türü, vites sayısı, yakıt türü, şehir içi kullanım, şehirlerarası kullanım, CO<sub>2</sub> emisyonu, motor hacmi, silindir sayısı gibi teknik bilgilerin yanında marka model, üretim yılı, araç sınıfı gibi araç kimlik bilgileri de yer almaktadır. Veri setinin parametreleri arasındaki ilişkiyi anlamak için, korelasyon testi sonucu Şekil 2'de görselleştirilmiştir. Ardından, her bir giriş parametresinin korelasyonu ve giriş ile hedef parametreler arasındaki ikili korelasyona dayalı olarak ilişki analiz edilmektedir. Şekil 3'de korelasyon matrisi veriler arasındaki ilişki

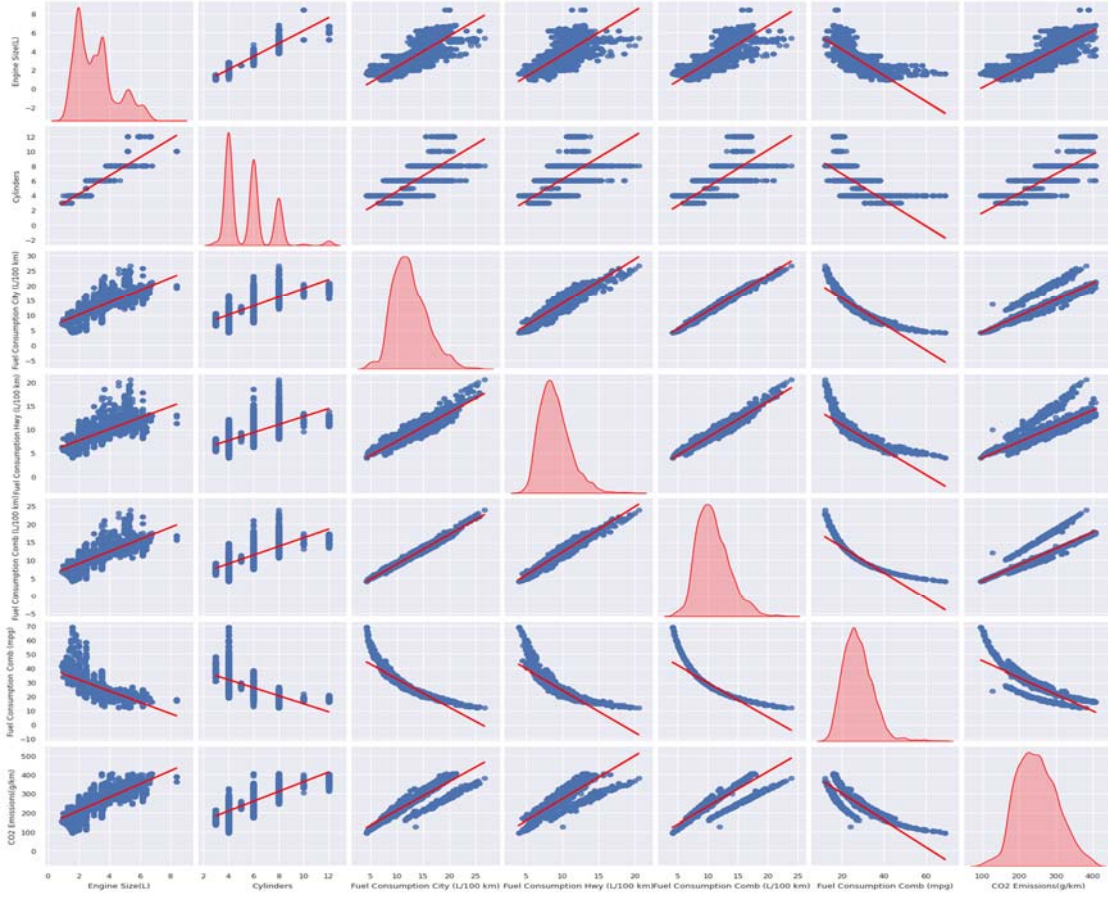
histogram olarak verilmiştir. Şekil 4'te yakıt tipi ve CO<sub>2</sub> emisyon verileri arasındaki ilişki gösterilmiştir.



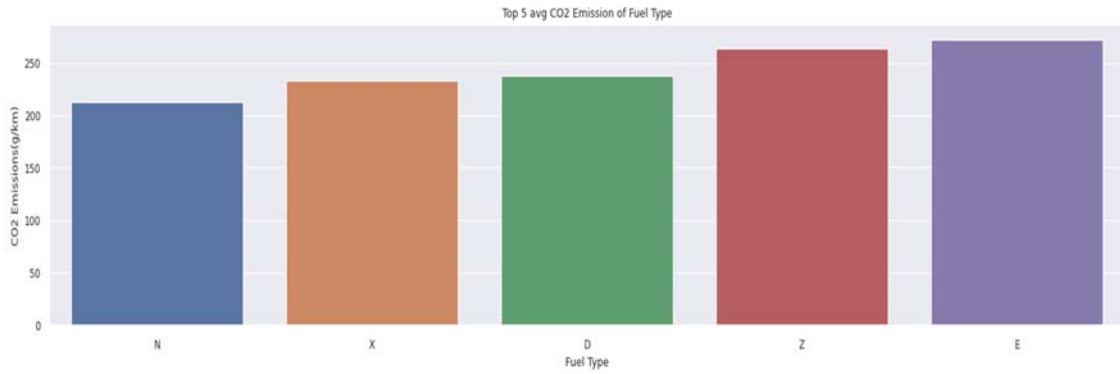
Şekil 1. Akış diyagramı



Şekil 2. Korelasyon matrisi

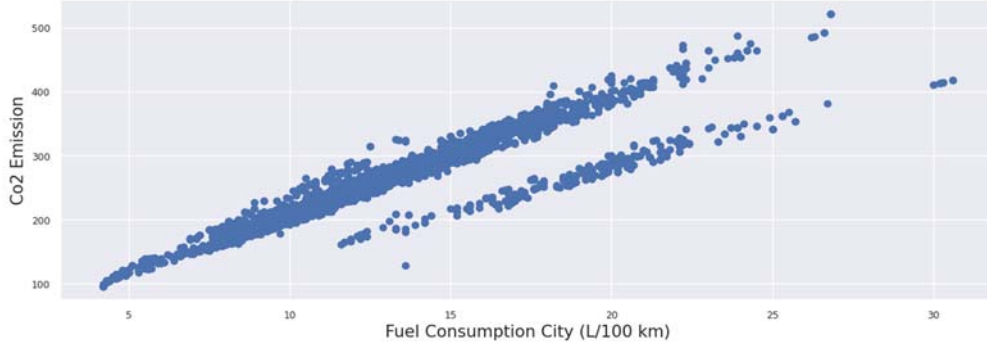


Şekil 3. Korelasyon matrisi histogram grafiği

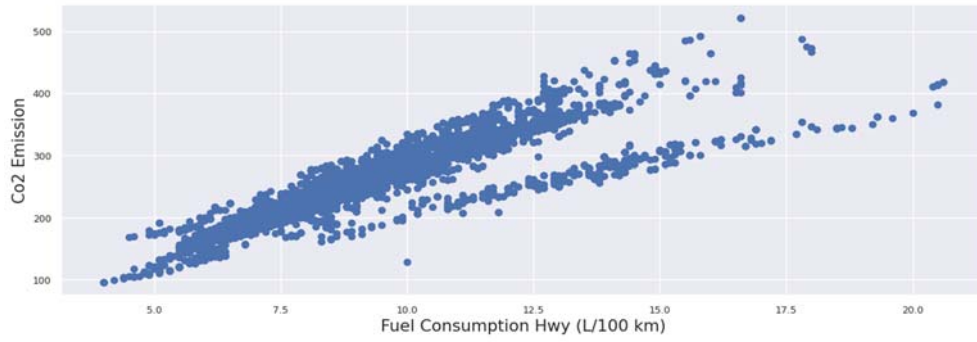
Şekil 4. Yakıt tipi ve CO<sub>2</sub> emisyon grafiği

Şehir içi yakıt tüketiminin CO<sub>2</sub> emisyonu ile arasındaki ilişki Şekil 5'de şehirler arası yakıt tüketiminin CO<sub>2</sub> emisyonu ile ilişkisi Şekil'6 da

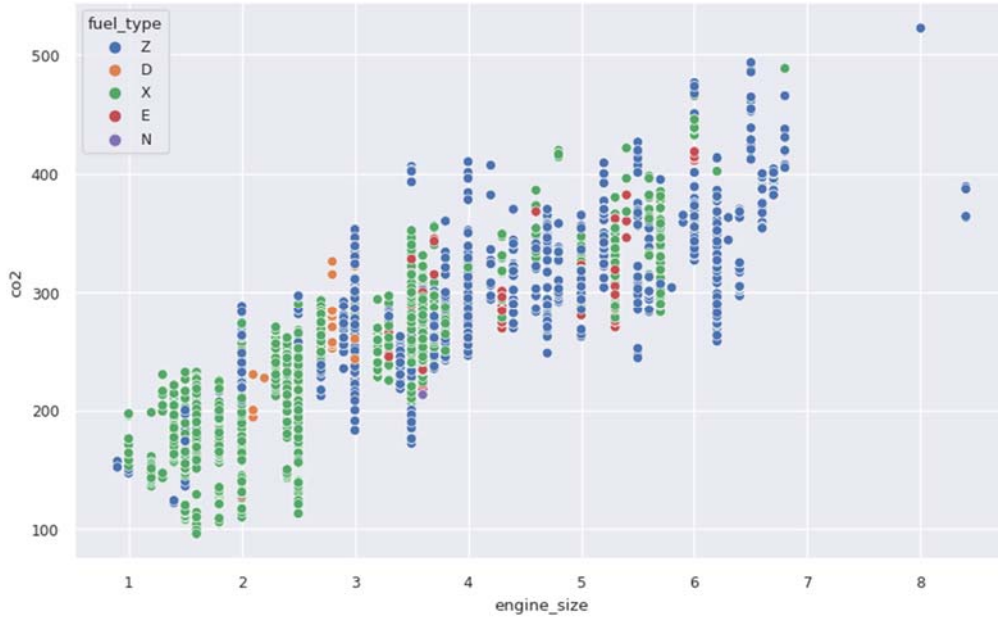
verilmiştir. Motor hacmine göre farklı yakıt türlerine sahip araçların oluşturduğu CO<sub>2</sub> emisyonu Şekil 7'de gösterilmiştir.



Şekil 5. Şehir içi yakıt tüketimi ve CO<sub>2</sub> emisyon grafiği



Şekil 6. Şehirlerarası yakıt tüketimi ve CO<sub>2</sub> emisyon grafiği



Şekil 7. Motor hacmine göre farklı yakıt türlerine sahip araçların oluşturduğu CO<sub>2</sub> emisyonu



### 3.2. Lineer Regresyon

Lineer regresyon, istatistik ve makine öğrenmesinde sıkça kullanılan temel bir modeldir. Bu model, bir bağımlı değişkenin bir veya daha fazla bağımsız değişkenle olan ilişkisini doğrusal bir denklemle modellemeyi amaçlamaktadır. Lineer regresyon, bu ilişkiyi doğrusal bir denklemle temsil etmektedir. Temel lineer regresyon modeli genellikle şu formda ifade edilmektedir:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \epsilon \quad (1)$$

Burada, Y bağımlı değişken veya sonuç, X bağımsız değişken,  $\beta$  modelin katsayısı ve  $\epsilon$  hata terimidir. Lineer regresyonun temel amacı, bağımlı değişkenle bağımsız değişkenler arasındaki ilişkiyi açıklamak ve bu ilişkiyi temsil eden bir doğru veya düzlem oluşturmaktır. Bu doğru veya düzlem, veri noktalarının dağılımını en iyi şekilde yakalamaya çalışmaktadır.

Lineer regresyon, özellikle bağımsız değişkenlerle bağımlı değişken arasındaki ilişkiyi anlamak ve tahminler yapmak için sıkça kullanılmaktadır. Lineer regresyonun geniş bir uygulama yelpazesi bulunmaktadır. Örneğin, tahmin yapma, korelasyon analizi ve optimizasyon gibi alanlarda sıklıkla kullanılmaktadır. Bu model, bir evin satış fiyatını tahmin etmekten hasta tedavi sonuçlarını analiz etmeye kadar birçok farklı alanda değerli bilgiler sağlayabilmektedir.

Bu modelin avantajları arasında basitlik ve anlaşılabilirlik gelmektedir. Ayrıca, hesaplama açısından verimli olduğu için büyük veri setleri üzerinde hızlı sonuçlar elde etmek mümkündür. Ancak, lineer regresyonun sınırlamaları da bulunmaktadır. Lineer ilişkileri modellemeye uygun olmayan veri setleri için bu modelin kullanılması doğru sonuçlar vermeyebilir. Eksik veya yanıltıcı veriler, modelin performansını olumsuz etkileyebilmektedir.

Sonuç olarak, lineer regresyon istatistik ve makine öğrenmesinde temel bir araç olarak karşımıza çıkmaktadır. Doğru bir şekilde uygulandığında, bu model veri analizi, tahminleme ve karar verme süreçlerinde büyük avantajlar sağlayabilmektedir.

Ancak, her model gibi, lineer regresyonun da sınırlamaları ve dikkat edilmesi gereken noktaları bulunmaktadır.

### 3.3. K-Nearest Neighbors Regresyon (K-NN)

K-Nearest Neighbors Regresyon (K-NN) algoritması, hem sınıflandırma hem de regresyon problemlerinde sıkça kullanılan temel bir makine öğrenmesi algoritmasıdır. Algoritmanın temel amacı, bir veri noktasını çevresindeki en yakın k komşunun etiketi veya değeri ile tahmin etmektir. KNN algoritmasının çalışma prensibi oldukça basit ve anlaşılırdır. Veri noktası, benzerlik ölçütüne göre diğer tüm veri noktalarına olan mesafeleri hesaplayarak, en yakın k veri noktasını seçmektedir.

Sınıflandırma durumunda, en sık görülen sınıf veya etiket tahmin olarak verilmektedir. Örneğin, bir e-postanın spam olup olmadığını tahmin etmek için kullanılabilir. Regresyon durumunda ise, en yakın k veri noktasının ortalama değeri tahmin olarak kullanılmaktadır.

K-NN algoritmasının avantajları arasında basitlik, kolay uygulanabilirlik ve anlaşılabilirlik bulunmaktadır. Başlangıç seviyesi kullanıcılar için ideal bir tercih olmasını sağlamaktadır. Ancak, büyük veri setleri ve yüksek boyutlu özellik uzaylarında hesaplama maliyeti artabilmektedir. Bu durum, algoritmanın pratikte kullanımını sınırlayabilmektedir.

Ayrıca, K-NN algoritmasının önemli hiperparametreleri vardır. Özellikle k değerinin seçimi ve benzerlik ölçütünün belirlenmesi, modelin performansını önemli ölçüde etkileyebilmektedir. Bu parametrelerin doğru bir şekilde ayarlanması, algoritmanın başarısını artırmak için kritik öneme sahiptir.

Sonuç olarak, K-NN algoritması, veri analizi, tahmin etme ve karar verme süreçlerinde etkili bir araç olarak kullanılabilir. Ancak, doğru bir şekilde uygulanması için hiperparametrelerin dikkatlice seçilmesi ve algoritmanın sınırlamalarının bilinmesi gerekmektedir.

### 3.4. XGBoost Regresyon

XGBoost (Extreme Gradient Boosting) son zamanlarda makine öğrenimi yarışmalarında ve endüstriyel uygulamalarda büyük başarı elde etmiş bir öğrenme algoritmasıdır. Gradient Boosting yöntemlerine dayanır, ancak özellikle ölçeklenebilirlik, hız ve doğruluk açısından geliştirilmiştir. XGBoost'un temel çalışma prensibi, birçok zayıf tahmin ediciyi bir araya getirerek güçlü bir tahminci oluşturmaktır. Model, önceki tahminlerin hatalarını düzeltmeye odaklanarak ardışık olarak bu tahmin edicileri eğitmektedir. Bu, bir sonraki tahmin ediciyi eğitmek için kullanılan veri setindeki hataları minimize etmek için bir gradyan inişi yaklaşımıdır. XGBoost'un bazı önemli özellikleri şunlardır:

**Ölçeklenebilirlik:** Büyük veri setleriyle başa çıkabilir ve çoklu işlemcileri veya dağıtılmış sistemleri kullanarak hızlı bir şekilde eğitim yapabilmektedir.

**Hız:** Optimizasyonları sayesinde diğer gradient boosting kütüphanelerine göre daha hızlıdır.

**Regülerizasyon:** Aşırı öğrenmeyi önlemek için çeşitli regülerizasyon tekniklerini desteklemektedir.

**Esneklik:** Sınıflandırma, regresyon, sıralama ve diğer birçok problem türünü desteklemektedir.

XGBoost, endüstriyel uygulamalarda ve makine öğrenimi yarışmalarında liderlik eden modellerin oluşturulmasında yaygın olarak kullanılmaktadır. Özellikle regresyon analizlerinde, XGBoost'un hızlı ve ölçeklenebilir yapısı sayesinde etkileyici sonuçlar alınabilir.

## 4. ARAŞTIRMA BULGULARI

Bu çalışmada, araçlardaki CO<sub>2</sub> emisyon tahmininde kullanılan regresyon algoritmalarının performansı değerlendirilmiştir. Linear, K-Nearest Neighbours, XGBoost Regressor algoritmaları üzerinde eğitim gerçekleştirilmiştir. Modelin potansiyelini detaylı

bir şekilde incelemek amacıyla regresyon algoritmalarını tanımlayan değişkenler ve hiperparametreler üzerinde kapsamlı bir araştırma gerçekleştirilmiştir. Çalışmada kullanılan Linear Regresyon ile eğitimde kullanılan parametreler Çizelge 1'de verilmiştir.

**Çizelge 1.** Linear Regresyon Parametreleri

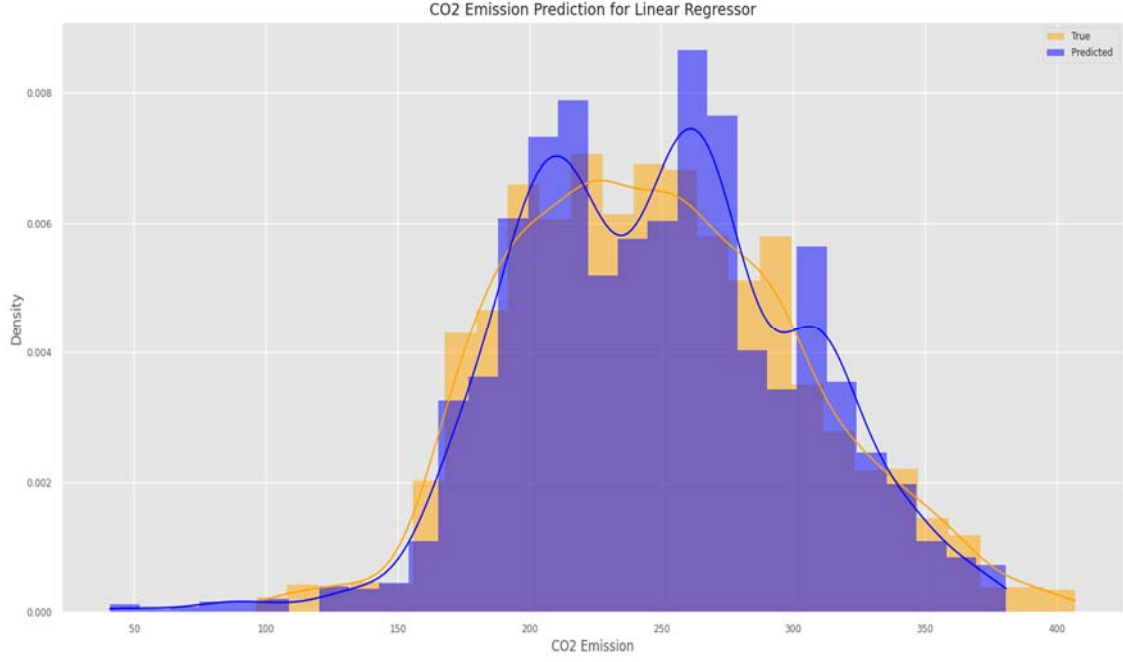
| Parametre     | Değer |
|---------------|-------|
| Fit_intercept | true  |
| copy_X        | true  |
| n_job         | none  |
| positive      | true  |
| Max_iter      | 10    |
| Random_state  | 10    |

Linear Regresyonun eğitim sonucunda araçlarda CO<sub>2</sub> emisyonunu tahmin etmede %90.69 eğitim performansı ve %90.76 test performansı elde edilmiştir. RMSE 16.65, MAE 11.06, R<sup>2</sup> Skoru 90.76'dir. Linear regresyon algoritması ile tahmin edilen araçlardaki CO<sub>2</sub> emisyonu ve gerçek emisyon değerlerinin tahminini gösteren grafik Şekil 8'de verilmiştir.

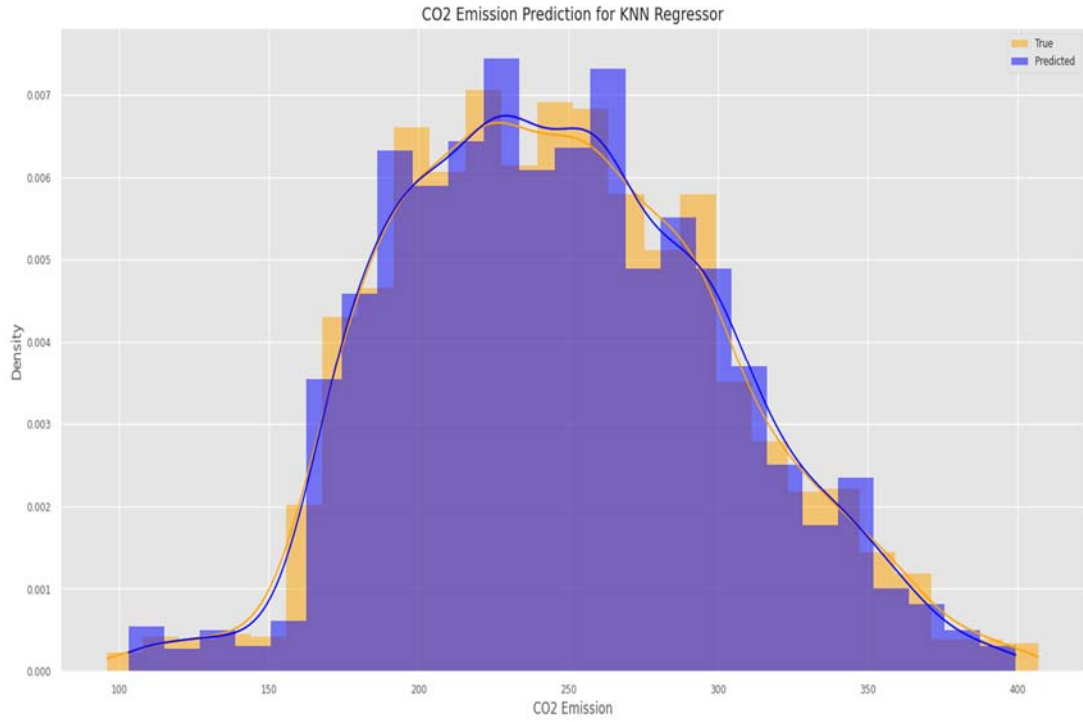
**Çizelge 2.** K-Nearest Neighbour Regressor Parametreleri

| Parametre    | Değer |
|--------------|-------|
| Probability  | True  |
| Epsilon      | 0.1   |
| Degree       | 3     |
| Max_iter     | 10    |
| Random_state | 10    |

K-Nearest Neighbour Regressor ile eğitimde kullanılan parametreler Çizelge 2'de verilmiştir. K-Nearest Neighbour Regresyonun eğitim sonucunda araçlarda CO<sub>2</sub> emisyonunu tahmin etmede %97.76 eğitim performansı ve %97.39 test performansı elde edilmiştir. RMSE 8.84, MAE 3.75, R<sup>2</sup> Skoru 97.39'dir. K-Nearest Neighbour Regresyon algoritması ile tahmin edilen araçlardaki CO<sub>2</sub> emisyonu ve gerçek emisyon değerlerinin tahminini gösteren grafik Şekil 9'da verilmiştir.



Şekil 8. Linear Regresyon ile gerçek değer ve tahmin edilen değer grafiği

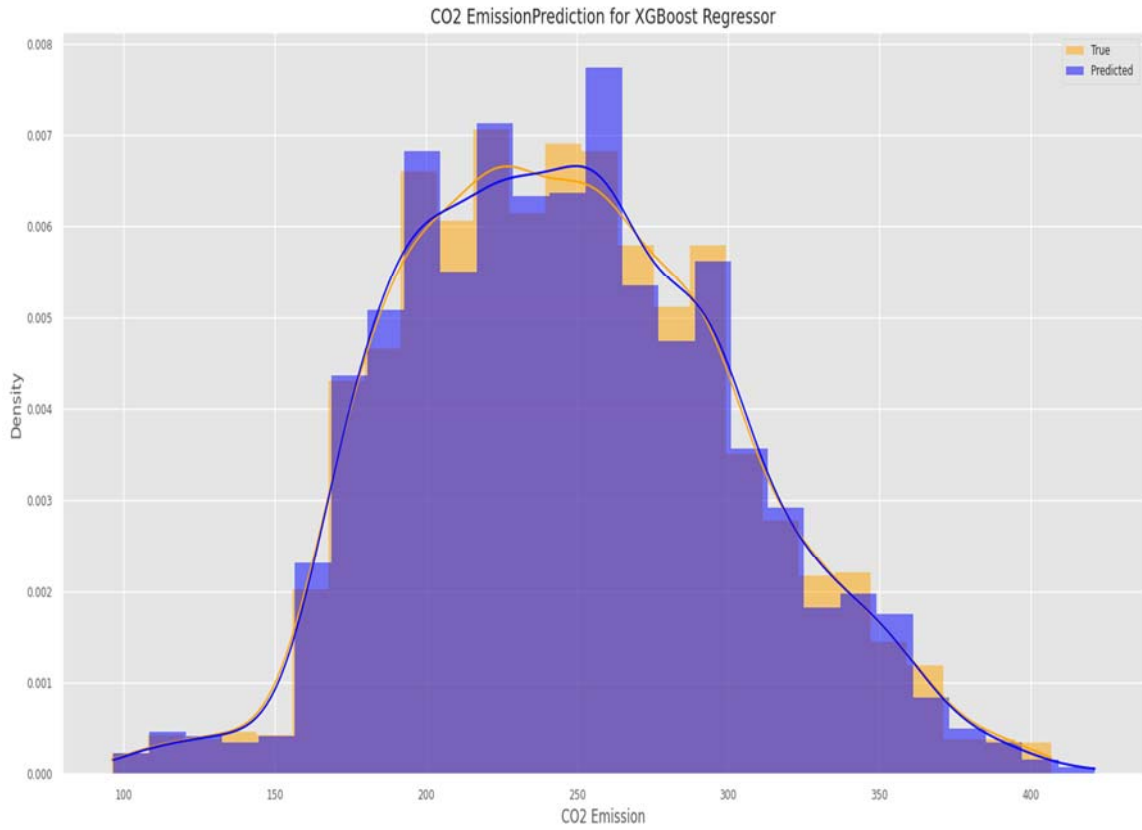


Şekil 9. K-Nearest Neighbour Regresyon ile gerçek değer ve tahmin edilen değer grafiği

**Çizelge 3.** XGBoost Regresyon Parametreleri

| Parametre     | Değer  |
|---------------|--------|
| booster       | gbtree |
| verbosity     | 1      |
| eta           | 0.3    |
| learning_rate | 0.99   |
| max_dept      | 6      |
| n_estimators  | 50     |

XGBoost Regressor ile eğitimde kullanılan parametreler Çizelge 3'te verilmiştir. XGBoost Regresyonun eğitim sonucunda araçlarda CO<sub>2</sub> emisyonunu tahmin etmede %99.64 eğitim performansı ve %98.75 test performansı elde edilmiştir. RMSE 6.45, MAE 2.81, R<sup>2</sup> Skoru 98.75'tür. XGBoost Regresyon algoritması ile tahmin edilen araçlardaki CO<sub>2</sub> emisyonu ve gerçek emisyon değerlerinin tahminini gösteren grafik Şekil 10'da verilmiştir.



**Şekil 10.** XGBoost Regresyon ile gerçek değer ve tahmin edilen değer grafiği

**Çizelge 4.** Regresyon algoritmalarının karşılaştırılması

| Regresyon algoritması | Eğitim doğruluğu | Test doğruluğu | RMSE skor | MAE skor | R <sup>2</sup> skor |
|-----------------------|------------------|----------------|-----------|----------|---------------------|
| Lineer                | 90.69            | 90.76          | 16.65     | 11.06    | 90.76               |
| K-Nearest neighbours  | 97.76            | 97.39          | 8.84      | 3.75     | 97.39               |
| XGBoost               | 99.64            | 98.75          | 6.45      | 2.81     | 98.75               |

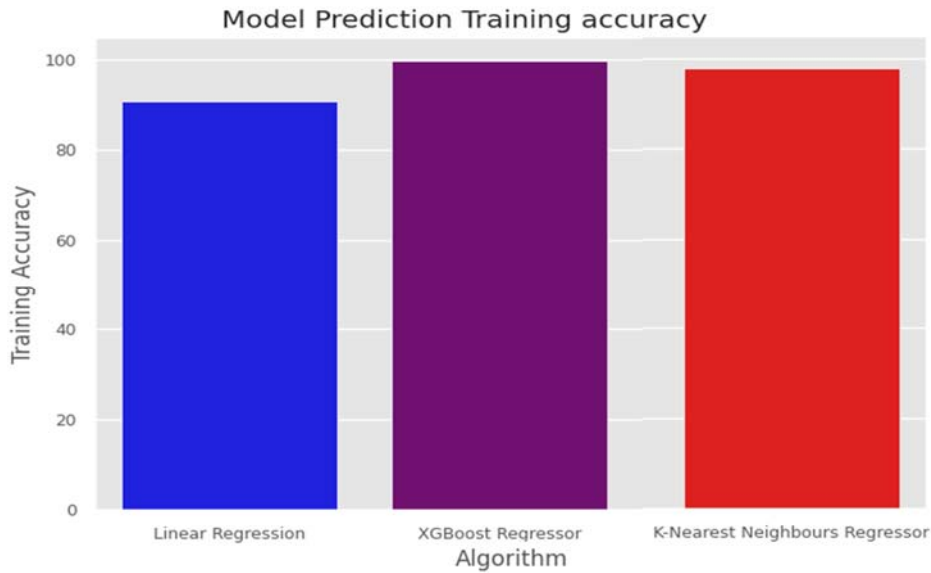
Çalışmada uygulanan 3 farklı yöntem ile araçlardaki CO<sub>2</sub> emisyon tahmini gerçekleştirilmiştir. Regresyon algoritmalarının performanslarının değerlendirme metrikleri ve algoritmaların başarı ölçeklendirilmesine ait veriler Çizelge 4’da sunulmuştur. Şekil 11 ve Şekil 12’de sırasıyla eğitilen modellerin eğitim ve test doğrulukları her bir regresyon modeli için grafiksel olarak gösterilmiştir. Şekil 13’te ise regresyon modelleri için R<sup>2</sup> skoru grafik ile gösterilmiştir.

Çizelge 4 incelendiğinde XGBoost Regresyon’un, eğitim doğruluğunda %99.64 ve test doğruluğunda %98.75 ile en yüksek performansı gösterdiğini ortaya koymaktadır. Bu değerler, XGBoost’un yüksek tahmin doğruluğu, düşük hata oranı ve hızlı işlem süresi ile öne çıktığını göstermektedir.

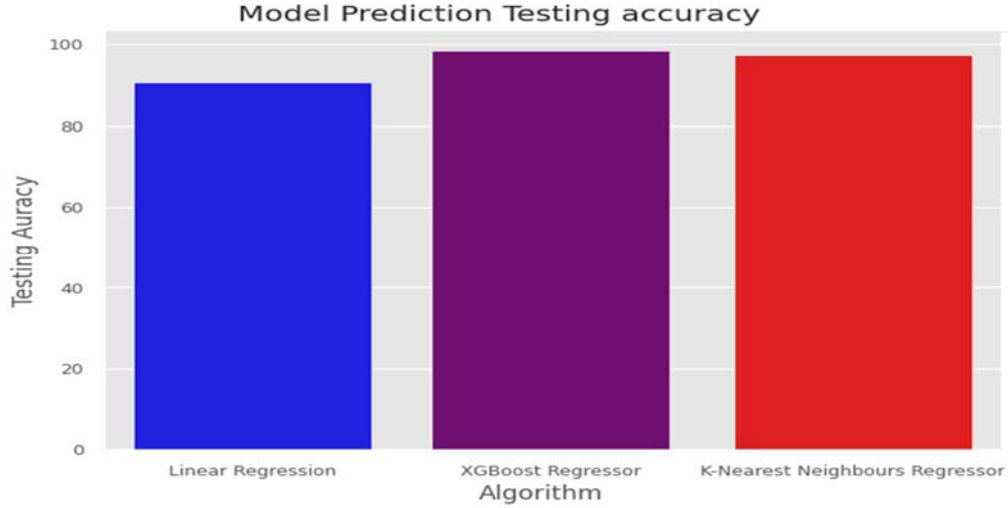
Çalışmanın sonuçlarına göre, XGBoost Regresyon en yüksek tahmin doğruluğu ve en düşük hata oranlarıyla dikkat çekerken, Linear Regresyon daha sınırlı bir performans göstermektedir. K-NN ise esnekliğiyle ve doğrusal olmayan ilişkileri yakalama kapasitesiyle önemli bir alternatif olarak

karşımıza çıkmaktadır. Bu nedenle, tahmin problemi ve veri setinin özelliklerine bağlı olarak en uygun algoritmanın seçilmesi önem arz etmektedir.

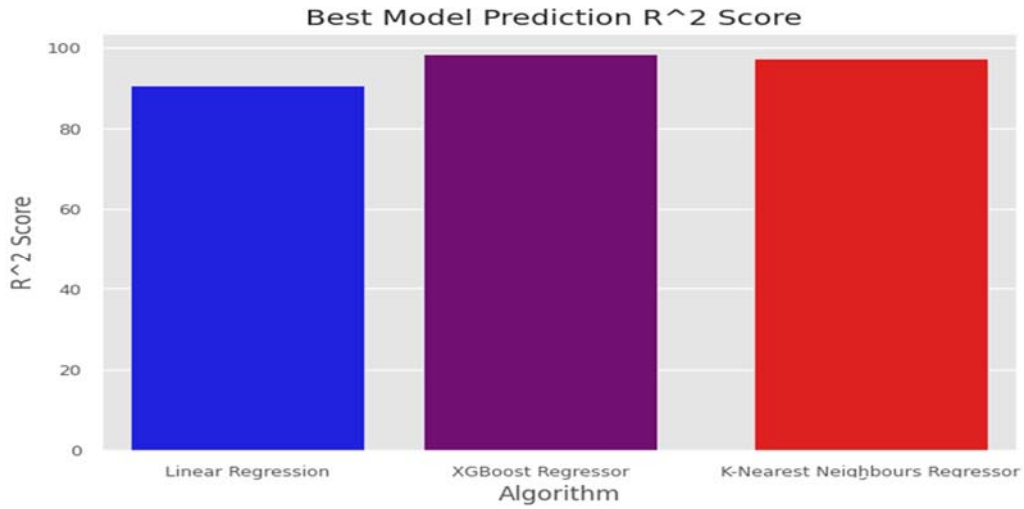
Bu çalışma, araçlardaki CO<sub>2</sub> emisyon tahmininde kullanılan regresyon algoritmalarının performansını değerlendiren kapsamlı bir analizi sunmaktadır. Özellikle, Lineer, K-Nearest Neighbours ve XGBoost Regresyon algoritmaları üzerinde gerçekleştirilen eğitim ve test süreçleri, detaylı bir karşılaştırmalı analiz sağlamaktadır. Bu tür bir performans değerlendirmesi, literatürde sıkça karşılaşılan uygulamalardan biri olup, farklı algoritma türlerinin avantajları ve sınırlamaları hakkında önemli bilgiler sağlamaktadır. Ancak, bu çalışmanın özgünlüğü, bu algoritmalara özgü hiperparametrelerin ve değişkenlerin kapsamlı bir şekilde incelenmesi ve bu parametrelerin tahmin performansı üzerindeki etkilerinin derinlemesine değerlendirilmesindedir. Bu yaklaşım, araçlardaki CO<sub>2</sub> emisyon tahmininde kullanılan regresyon algoritmalarının daha etkili bir şekilde seçilmesine ve uygulanmasına katkıda bulunabilecektir.



Şekil 11. Modellerin eğitim doğruluğu



Şekil 12. Modellerin test doğruluğu



Şekil 13. Modellerin R<sup>2</sup> skoru

## 5. SONUÇLAR

Araçlardan kaynaklanan CO<sub>2</sub> emisyonlarının eksiksiz bir biçimde anlaşılması son derece önemlidir. Bu anlayış, çevresel etkileri azaltmaya yönelik bilinçli karar alımı ve stratejik planlama için temel bir gerekliliktir. Linear Regresyon, XGBoost Regresör ve K-Nearest Neighbours algoritmaları kullanılarak arabalar için CO<sub>2</sub> emisyonu tahmini sonuçları, bu modellerin tahmin yetenekleri hakkında değerli bilgiler sunmaktadır.

Birincisi, Linear Regresyon modeli %90,69 eğitim doğruluğu ve %90,76 test doğruluğu ile iyi bir performans sergilemektedir. Orta düzeydeki 16,65'lik Ortalama Karesel Hata (RMSE) makul bir hassasiyet düzeyine işaret ederken, 11,06'lık Ortalama Mutlak Hata (MAE) CO<sub>2</sub> emisyonlarının tahmininde makul bir doğruluğa işaret etmektedir. 90,76'lık R<sup>2</sup> skoru, tahmin edilen ve gerçek değerler arasında güçlü bir korelasyon olduğunu göstermektedir. İkincisi, K-Nearest Neighbours modeli, %97,76'lık dikkate değer bir eğitim

doğruluğu ve %97,39'luk bir test doğruluğu elde ederek üstün performans göstermektedir. Doğrusal Regresyona kıyasla 8,84'lük daha düşük RMSE ve 3,75'lik MAE, gelişmiş doğruluk ve hassasiyete işaret etmektedir. Son olarak, XGBoost algoritması %99,64 gibi yüksek bir eğitim doğruluğu ve %98,75 gibi etkileyici bir test doğruluğu ile etkili algoritma olarak ortaya çıkmaktadır. En düşük 6,45 RMSE, 2,81 MAE ve %98,75 en yüksek R<sup>2</sup> skoru ile XGBoost modeli, araçlar için CO<sub>2</sub> emisyonlarını tahmin etmede iyi bir doğruluk, hassasiyet ve tahmin yetenekleri sergilemektedir. Sonuç olarak, bu regresyon modellerinin karşılaştırılması, araçlardan kaynaklanan CO<sub>2</sub> emisyonlarının tahmin edilmesinde makine öğrenimi algoritmalarının etkinliğinin altını çizmektedir. Sonuçlar, XGBoost modelinin diğer modellere göre üstünlüğünü vurgulayarak, çevresel araştırma ve politika oluşturma alanlarındaki uygulamalar için potansiyelini ortaya koymaktadır. Bu bulgular, CO<sub>2</sub> emisyonu tahmininin, sürdürülebilir ve çevreye duyarlı ulaşım çözümleri arayışında daha doğru ve bilinçli karar vermenin yolunu açacaktır.

## 6. KAYNAKLAR

- Özupak, Y., 2024. Evrişimli Sinir Ağı (ESA) Mimarileri ile Hücre Görüntülerinden Sıtmanın Tespit Edilmesi. Çukurova Üniversitesi Mühendislik Fakültesi Dergisi, 39(1), 197-210.
- Zacharof, N., Fontaras, G., Ciuffo, B., Tansini, A., Prado-Rujas, I., 2021. An Estimation of Heavy-duty Vehicle Fleet CO<sub>2</sub> Emissions Based on Sampled Data. Transport. Res. Transport Environ., 94, 102784.
- Ganesan, P., Rajakarunakaran, S., Thirugnanasambandam, M., Devaraj, D., 2015. Artificial Neural Network Model to Predict the Diesel Electric Generator Performance and Exhaust Emissions. Energy, 83, 115-124.
- Çay, Y., 2013. Prediction of a Gasoline Engine Performance with Artificial Neural Network. Fuel, 111, 324-331.
- Hawkes, A.D., 2010. Estimating Marginal CO<sub>2</sub> Emissions Rates for National Electricity Systems. Energy Policy, 38, 5977-5987.
- Labecki, L., Cairns, A., Xia, J., Megaritis, A., Zhao, H., Ganippa, L.C., 2012. Combustion and Emission of Rapeseed Oil Blends in Diesel Engine. Applied Energy, 95, 139-146.
- Tasdemir, S., Saritas, I., Ciniviz, M., Allahverdi, N., 2011. Artificial Neural Network and Fuzzy Expert System Comparison for Prediction of Performance and Emission Parameters on a Gasoline Engine. Expert Systems with Applications, 38, 13912-23.
- Anderson, T.R., Hawkins, E., Jones, P.D., 2016. CO<sub>2</sub>, the Greenhouse Effect and Global Warming: From the Pioneering Work of Arrhenius and Callendar to Today's Earth System Models. Endeavour, 40(3), 178-187.
- Zeng, W., Miwa, T., Morikawa, T., 2016. Prediction of Vehicle CO<sub>2</sub> Emission and Its Application to Eco-routing Navigation. Transportation Research Part C: Emerging Technologies, 68, 194-214.
- Odoro, S., Metia, S., Duc, H., Ha, Q., 2013. CO<sub>2</sub> Vehicular Emission Statistical Analysis with Instantaneous Speed and Acceleration as Predictor Variables. In Proceedings of the International Conference on Control, Automation and Information Sciences, 158-163.
- Razak, N.H., Hashim, H., Yunus, N.A., Klemes, J.J., 2022. Integrated Linear Programming and Analytical Hierarchy Process Method for Diesel/Biodiesel/Butanol in Reducing Diesel Emissions. Journal of Cleaner Production, 337, 130297.
- Shim, E., Park, H., Bae, C., 2018. Intake Air Strategy for Low HC and CO Emissions in Dual-fuel (CNG-Diesel) Premixed Charge Compression Ignition Engine. Applied Energy, 225, 1068-77.
- Prabhu, A.V., Avinash, A., Brindhadevi, K., Pugazhendhi, A., 2021. Performance and Emission Evaluation of Dual Fuel CI Engine Using Preheated Biogas-air Mixture. Science of the Total Environment, 754, 142389.
- Soukht, S.H., Taghavifar, H., Jafarmadar, S., 2017. Experimental and Numerical Consideration of the Effect of CeO<sub>2</sub> Nanoparticles on Diesel Engine Performance and Exhaust Emission with the Aid of Artificial Neural Network. Applied Thermal Engineering, 113, 663-72.
- Alfaseeh, L., Tu, R., Farooq, B., Hatzopoulou, M., 2020. Greenhouse Gas Emission Prediction

- on Road Network Using Deep Sequence Learning. *Transport. Res. Transport Environ*, 88, 102593.
16. Claudio, M., Daniela, M., Alessandro, D.M., Ezio S., 2021. A Deep Neural Network Based Model for the Prediction of Hybrid Electric Vehicles Carbon Dioxide Emissions. *Energy and AI*, 5, 100073, 2666-5468.
  17. Jigu, S., Sungwook, P., 2023. Optimizing Model Parameters of Artificial Neural Networks to Predict Vehicle Emissions. *Atmospheric Environment*, 294, 119508, 1352-2310.
  18. Natarajan, Y., Wadhwa, G., Sri, K.R., Paul, A., 2023. Forecasting Carbon Dioxide Emissions of Light-Duty Vehicles with Different Machine Learning Algorithms. *Electronics*, 12, 2288.
  19. Al-Nefaie, A.H., Aldhyani, T.H.H., 2023. Predicting CO<sub>2</sub> Emissions from Traffic Vehicles for Sustainable and Smart Environment Using a Deep Learning Model. *Sustainability*, 15, 7615.
  20. Dong, T., Zhen, Z., Lun, H., Jinchong, P., Yang, X., 2023. Prediction of Cold Start Emissions for Hybrid Electric Vehicles Based on Genetic Algorithms and Neural Networks. *Journal of Cleaner Production*, 420, 138403.
  21. Paul, D., Dieudonné, T., Guillaume, C., 2023. Method and Evaluations of the Effective Gain of Artificial Intelligence Models for Reducing CO<sub>2</sub> Emissions. *Journal of Environmental Management*, 331, 117261, 0301-4797.
  22. Wang, Z., Feng, K., 2024. NO<sub>x</sub> Emission Prediction for Heavy-Duty Diesel Vehicles Based on Improved GWO-BP Neural Network. *Energies*, 17, 336.
  23. Maźziel, M., 2024. Instantaneous CO<sub>2</sub> Emission Modelling for a Euro 6 Start-Stop Vehicle Based on Portable Emission Measurement System Data and Artificial Intelligence Methods. *Environ Sci Pollut Res*, 31, 6944-6959.
  24. CO<sub>2</sub> Emission by Vehicles Dataset, <https://open.canada.ca/data/en/dataset/98f1a129-f628-4ce4-b24d-6f16bf24dd64#wb-auto-6>, Erişim tarihi: 02.02.2024.