

GAZİ

JOURNAL OF ENGINEERING SCIENCES

Forecasting Light Commercial Vehicle Sales in Turkey by Machine Learning Methods

Mehmet Kayakuş^{a*}, Mustafa Terzioğlu^b, Ayten Yağmur^c, Dilşad Erdoğan^d

Submitted: 19.11.2023 Revised: 07.12.2023 Accepted: 08.12.2023 doi:10.30855/gmbd.0705S11

ABSTRACT

Light commercial vehicles such as vans, pick-up trucks, panel vans, and minibuses are a class of vehicles used in large quantities, especially in the trade and service sectors. Changes in demand for this class of vehicles are also perceived as an indicator of the economic vitality of countries. In this study, it is thought that forecasting the sales and imports of light commercial vehicles, which is accepted as a macroeconomic indicator, will contribute to the evaluation of general economic indicators and will be useful for automotive companies operating in this market in terms of effective corporate resource planning and efficient use of resources from a micro perspective. The sales forecasting model designed in this study was created by analysing previous studies in the literature and including macroeconomic variables that are thought to affect light commercial vehicle sales in the model. The sales forecasting model designed in this study is constructed by analysing previous studies in the literature and including macroeconomic variables that are thought to affect light commercial vehicle sales in the model. Three machine learning methods, namely artificial neural network (ANN), multiple linear regression (MLR) and decision tree (DT) regression, were used to measure the forecasting success of the model. As a result of the study, the R2 value was found to be 94.6% for ANN, 64.1% for MLR, and 82.2% for DT. According to the results obtained, it is concluded that the model designed for the prediction of light commercial vehicle sales in Turkey performs very successful predictions with ANN method.

Keywords: Light commercial vehicle, vehicle demand forecast, Turkey economy, machine learning

^{a,*} Akdeniz Üniversitesi,
Manavgat Sosyal ve Beşerî Bilimler
Fakültesi,
Yönetim Bilişim Sistemleri Bölümü
07600 - Antalya, Türkiye
Orcid: 0000-0003-0394-5862
e mail: mehmetkayakus@akdeniz.edu.tr

^b Akdeniz Üniversitesi,
Korkuteli Meslek Yüksekokulu,
Muhasebe ve Vergi Bölümü
07800 - Antalya, Türkiye
Orcid: 0000-0002-4614-7185

^c Akdeniz Üniversitesi,
İktisadi ve İdari Bilimler Fakültesi,
Çalışma Ekonomisi ve Endüstri İlişkileri
Bölümü
07070 - Antalya, Türkiye
Orcid: 0000-0003-2138-240X

^d Akdeniz Üniversitesi,
Korkuteli Meslek Yüksekokulu,
Finans, Bankacılık ve Sigortacılık Bölümü
07800 - Antalya, Türkiye
Orcid: 0000-0002-9117-5994

*Corresponding author:
mehmetkayakus@akdeniz.edu.tr

Anahtar Kelimeler: Hafif ticari araç, taşıt talep tahmini, Türkiye ekonomisi, makine öğrenmesi

Türkiye'deki Hafif Ticari Araç Satışlarının Makine Öğrenmesi Yöntemleriyle Tahmin Edilmesi

ÖZ

Kamyonet, pikap, panelvan, minibüs gibi hafif ticari araçlar, özellikle ticaret ve hizmet sektörlerinde büyük miktarlarda kullanılan bir araç sınıfıdır. Bu sınıftaki araçlara yönelik talepteki değişimler, ülkelerin ekonomik canlılığının da bir göstergesi olarak algılanmaktadır. Bu çalışmada, makroekonomik bir gösterge olarak kabul edilen hafif ticari araç satış ve ithalatının tahmin edilmesinin, genel ekonomik göstergelerin değerlendirilmesine katkı sağlayacağı ve bu pazarda faaliyet gösteren otomotiv firmaları için mikro bakış açısıyla etkin kurumsal kaynak planlaması ve kaynakların verimli kullanılması açısından faydalı olacağı düşünülmektedir. Bu çalışmada tasarlanan satış tahmin modeli, literatürdeki önceki çalışmaların analiz edilmesi ve hafif ticari araç satışlarını etkileyebileceği düşünülen makroekonomik değişkenlerin modele dahil edilmesi ile oluşturulmuştur. Bu çalışmada tasarlanan satış tahmin modeli, literatürdeki önceki çalışmalar analiz edilerek ve hafif ticari araç satışlarını etkileyebileceği düşünülen makroekonomik değişkenler modele dahil edilerek oluşturulmuştur. Modelin tahmin başarısını ölçmek için yapay sinir ağı (YSA), çoklu doğrusal regresyon (MLR) ve karar ağacı regresyonu (DTR) olmak üzere üç makine öğrenmesi yöntemi kullanıldı. Çalışma sonucunda R2 değeri YSA için %94,6, ÇDR için %64,1 ve DTR için %82,2 olarak bulunmuştur. Elde edilen sonuçlara göre, Türkiye'deki hafif ticari araç satışlarının tahmini için tasarlanan modelin YSA yöntemi ile oldukça başarılı tahminler gerçekleştirdiği sonucuna varılmıştır.

1. Giriş (Introduction)

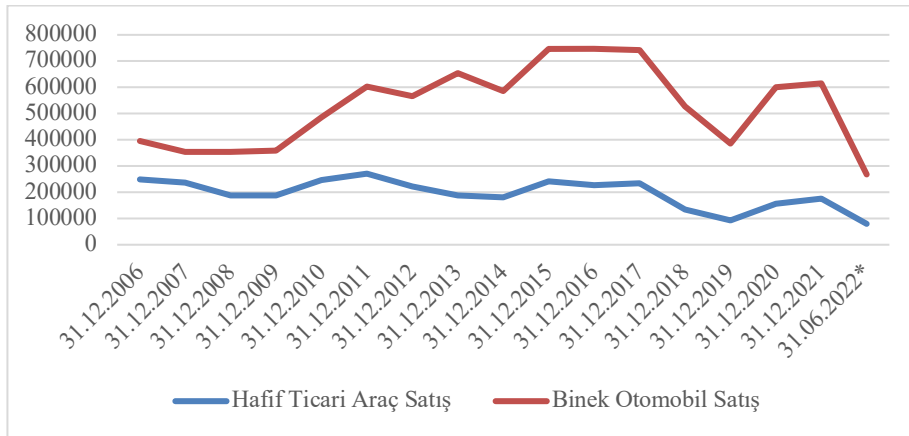
Otomotiv sektörü; teknolojik gelişmelerle birlikte ülke ekonomisi için yarattığı katma değer, istihdam alanı yaratması ve diğer sektörlerle olan yakın ilişkisi sebebiyle önem taşımaktadır [1]. Diğer sektörlerle bağlantılı olan otomotiv sektörü teknolojiye gerçekleşen hızlı gelişmeler ve küreselleşmenin de etkisiyle yaşanan uluslararası rekabet ortamında ülkelerin ekonomik görünümü açısından önemi daha da artmıştır [2, 3]. Otomotiv sektörü aynı zamanda sahip olduğu ürün çeşitliliği ile teknolojik gelişmelerin en hızlı uygulandığı, üretim tekniklerinin hızla geliştiği ve rekabetin en fazla yaşandığı sektörlerdendir. Bu açıdan hem gelişmiş hem de gelişmekte olan ülkeler için önem taşımaktadır. Aynı zamanda otomotiv sektörünün hizmet sektörüyle olan bağlantısı sebebiyle satış, bakım-onarım sigortacılık, kiralama hizmetlerine de katkı sağlamaktadır [2].

Otomotiv sektöründe uluslararası piyasada fiyat rekabeti söz konusuysen günümüzde kalite, ürün çeşitliliği ve geleceğe yatırım rekabet için belirleyici unsurlar olmuştur. Satışlar tüketici tercihlerine göre belirlendiği için ürün geliştirme, marka ve model geliştirme gibi faktörler daha önemli hale gelmiştir [4].

Otomotiv sektöründe üretim planlarının yapılabilmesi için uzun dönemli talep tahminlerinin yapılması önemlidir. Talep ise belli bir dönemde piyasadaki tüketicilerin bir ürüne yönelik olarak ortaya çıkan satın alma isteğidir. Tüketicilerin seçimleri birçok faktöre bağlı olmakla birlikte en önemli faktör malın kendi fiyatıdır. Bu bağlamda rekabetin fazla olduğu otomobil sektöründe, otomobil talebi sadece kendi fiyatı değil beraberinde diğer otomobil fiyat ve kalitesinden de etkilenmektedir. Ayrıca ülkenin genel ekonomik görünümü, taşıt finansmanı olanakları, hane halklarının tasarruf kabiliyeti gibi faktörler de otomotiv satışlarını etkileyen diğer değişkenlerdir [5, 6].

Tahmin, geçmiş yıllar verilerinden hareketle çeşitli yöntemler kullanarak gelecek hakkında farklı yöntemler uygulayarak geleceği tahmin etmektir. Talep tahmini yapmak ise; sektördeki firmaların satış gelirlerini, verimliliği artırma, gibi amaçları için önem taşımaktadır [7]. Farklı yöntemlerle yapılan tahminler firmaların piyasa performansına katkı sağlayarak, kar elde etmelerine etmelerini sağlar [8]. Aynı zamanda satış tahmini firma yöneticilerin belirsiz bir ortamda uygun kararlar almasına yardımcı olmaktadır.

Türkiye’de hafif ticari araçların yıllara göre satışı bakıldığında yıllar içinde hafif dalgalanmalar seyretse de yatay bir seyir izlediği görülmektedir. Özellikle binek otomobil satışlarındaki daha büyük dalgalanmaların hafif ticari araçlardaki satışlarda aynı büyüklükte oluşmadığı Şekil 1’de gösterilmektedir.



Şekil 1. Yıllara Göre Türkiye’de Hafif Ticari Araç ve Binek Otomobil Satışı
(Light commercial vehicle and passenger car sales in turkey by years)

Not: Zaman aralığı araştırmanın yapıldığı dönemi içermektedir. * İlk altı aylık satışlar dikkate alınmıştır.

Şekil 1’de görüldüğü gibi hafif ticari araç satışları araştırmanın yapıldığı 2006- 2022 döneminde 100.000 adet ile 300.000 adet aralığında değişmektedir. Özellikle taşıt piyasasının temel pazarı binek otomobil piyasasına göre 2011-2015 arasında satışların seyri ayrılmaktadır. Bunun nedeni 2008 yılında başlayarak 2012 yılının sonlarına kadar devam eden küresel ekonomik krizin Türkiye’deki özellikle ticaret ve üretim alanındaki etkileri olarak düşünülebilir. Bu nedenle hafif ticari araçların satışını ileri tahmin metodlarından olan makine

öğrenmesi yöntemleri ile tahmin etmek bu alanda Türkiye gibi bu alanda dışa bağımlı olan ülkelerin kaynak kullanımını için büyük önem taşımaktadır.

Bu çalışmada makine öğrenmesi yöntemleri kullanılarak Türkiye’de satılan hafif ticari araç satış miktarı tahmini gerçekleştirilmiştir. Otomotiv sektörü ve buna bağlı hizmet sektörleri olan satış, sigortacılık, bakım ve kiralama hizmetleri bu çalışma sonuçlarına gelecek bütçe ve planlarını şekillendirebileceklerdir. Otomotiv sektörünün faaliyet alanlarını büyütme ve yeni yatırımlar yapma kararlarına katkı sağlayacaktır.

Türkiye’de otomobil satış tahmini [9-11] yapılmış olmakla birlikte; bu çalışmanın hem kapsam olarak Türkiye üzerinde daha önce çalışılmamış hafif ticari araçlarının satışlarını ele alması hem de kullanılan veri setindeki bağımsız değişken seçimi ve kullanılan yöntemler çalışmayı özgünlük kazandırmakta ve benzer çalışmalardan ayırmaktadır.

2. Literatür (Literature)

Alper ve Serdar çalışmalarında Türkiye’de 1996-1999 yılları arasında otomobil piyasasının bir görünümünü ortaya koyarak otomobil fiyat fonksiyonu tahmin etmişler. 2005 yılı için talep tahmini yapmışlardır. 120 ayrı marka ve modele ait aylık fiyat, satış, kalite ve teknik özellik verileriyle birlikte talep eğrisinin tahmininde, bahsi geçen verilerin yansıma, reel faiz, reel tüketici kredileri, enflasyon ve enflasyonun değişkenliği verileri de kullanılmıştır. Çalışmalarında panel veri analizini kullanmışlardır. Çalışmanın sonucunda yapılan tahminler sonucunda, yerli, Avrupa Birliği’nden (AB) ve diğer ülkelerden ithal edilen otomobil taleplerinin yapısal değişiklikler gösterdiği belirlenmiştir [11].

Wang ve arkadaşları çalışmalarında otomobil satış miktarı, tesadüfi gösterge, öncü gösterge, toptan eşya fiyat endeksi ve gelir olmak üzere yedi değişkenden oluşan bir satış tahmin yöntemi önermiştir. Tayvan’da otomobil satışı üzerine tahminleme yapmışlardır. Veri seti aylık tahminlerden oluşmaktadır. Yapay sinir ağları (ANN), ANFIS (Adaptive-Network Based Fuzzy Inference Systems), ARIMA (autoregressive integrated moving average) yöntemleri kullanılmıştır. Çalışmanın ampirik sonuçlarına göre ANFIS yönteminin diğer iki yönteme göre daha başarılı sonuç verdiğini paylaşmışlardır [12].

Karaatlı ve arkadaşları çalışmalarında Türkiye’deki Ocak 2007 ve Haziran 2011 yılları arasındaki aylık otomobil satışları verileriyle yapay sinir ağları kullanılarak satış tahmini yapılmıştır. Çalışmalarında Türkiye’de 2007- 2011 yıllarındaki verilerle yeni otomobil satışı tahmin edilmiştir. Çalışmada otomobil satışını etkileyen değişkenler gayri safi yurtiçi hasıla, reel kesim güven endeksi, tüketici güven endeksi, yatırım harcamaları, tüketim harcamaları, tüketici fiyat endeksi, araç alım satımı, dolar ve zaman olarak belirlenmiştir. Çalışmada MAPE (Mean Absolute Percentage Error) değeri %16,82 çıkarak modelin doğru tahminler sınıfında gerçekleştiği araştırmacılar tarafından paylaşılmıştır. Gerçek değerlerle tahminler karşılaştırıldığında tahmin edilen ve gerçekleşen değerlerin birbirlerine yakın olduğu görülmektedir [9].

Akyurt çalışmasında Ocak 2011 – Eylül 2015 dönemi için Türkiye’deki yerli otomobil satışına ilişkin aylık verilerle yerli otomobil satış büyüklükleri tahmin edilmiştir. Geri yayılım algoritmasıyla eğitilmiş ileri beslemeli yapay sinir ağları modelini kullanılmıştır. Çalışma sonucunda, bulgular mevsimsel etkilerden arındırılmış yerli otomobil satış serinin kullandığı, iki gizli katmanın olduğu ve her gizli katmanda yedi nöronun olduğu modelin en iyi tahmin sonucunu verdiği saptanmıştır [10].

Kuvvetli ve arkadaşları çalışmalarında farklı segment ve markaları kullanarak, ekonomik ve çevresel parametrelerle aylık otomobil satışlarını tahmini yapmışlardır. Girdi olarak taşıt kredilerinde faizler, geçmiş zamanda gerçekleşmiş satış adetleri, gayrisafi milli hasıla, yakıt tüketimi, CO2 emisyonları gibi farklı parametreler seçilmiştir. 2008 – 2012 yılları arasındaki satış miktarlarını tahmin edebilecek bir ANN modeli önerilmiştir. İleri beslemeli yapay sinir ağları kullanılmış ve modelin eğitimi için ise Levenberg-Marquadt algoritması seçilmiştir. Yapay sinir ağı sonuçları lineer regresyon modeli ile kıyaslanmıştır. Yapay sinir ağı modeli ile lineer regresyon modeline göre daha doğru tahminler elde edilmiştir [13].

Pai ve Liu çalışmalarında ABD’de Şubat 2008’den Ağustos 2017’ye kadar aylık araç satış talebini tahmin etmek için çok değişkenli bir regresyon modeli ve zaman serisi analizi önermiştir. Twitter’den “araba satın al”, “yol satın al” ve “araç satın al” olmak üzere üç anahtar kelime ile veri toplayarak bu verileri otomatik duygu çözümlü ile analiz etmişlerdir. Çalışmaları, her iki tweetin duygu puanlarını içererek geliştirilen hibrit modelin araç satış talebinde daha üstün olduğunu ortaya çıkarmışlardır [14].

Aslankaya ve Öz çalışmalarında talep tahmin yöntemlerini kullanarak otomobil satışlarını tahmin etmeyi amaçlamışlardır. Zaman serisi analizi ve yapay sinir ağları kullanılmıştır. Kayıtlı araç, GSYİH (Gayri safi yurt içi hasıla), tüketici fiyat endeksi, dolar kuru, reel sektör güven endeksi, tüketici güven endeksi, aylık çalışma saatleri, üretilen model sayısı bağımsız değişken olarak kullanılmıştır. Satılan toplam araç sayısı da bağımlı değişken olarak alınmıştır. Karşılaştırma yapıldığında yapay sinir ağlarını yöntemindeki değerlerin birbirine daha yakın olduğu görülmüştür. Yapay sinir ağlarıyla elde edilen MAPE değeri %7,44 dür. Bu sonuç çoklu regresyon, hareketli ortalama ve basit üstel düzeltme yöntemlerine göre daha iyi bir tahmin gücünü işaret etmektedir [15].

3. Materyal ve Yöntem (Material and Method)

Bu çalışmada makine öğrenmesi yöntemleri kullanılarak Türkiye’de satılan hafif ticari araç satış miktarı tahmini gerçekleştirilmiştir. Çalışmada sekiz bağımsız değişken ve bir bağımlı değişkenden oluşan 1782 verinin yer aldığı özgün bir veri seti kullanılmıştır. Çalışmada tahmin için makine öğrenmesi yöntemlerinden yapay sinir ağları, çoklu doğrusal regresyon ve karar ağaçları regresyon yöntemleri kullanılmıştır. Bu bölümde çalışmanın materyal ve yöntemi hakkında bilgi verilmiştir.

3.1. Veri seti (Data set)

Çalışmada Ocak 2006- Haziran 2022 dönemini kapsayan aylık veriler kullanılmıştır. Veriler, Türkiye İstatistik Kurumu (TÜİK), Otomotiv Distribütörleri Derneği (ODD) ve Türkiye Cumhuriyet Merkez Bankası Elektronik Veri dağıtım Sistemi (TCMB EVDS) üzerinden elde edilen açık kaynak verilerden düzenlenmiştir. Çalışmada talep tahmini modeli oluştururken bağımlı değişken olarak ODD’nin aylık raporlarından derlenen aylık hafif ticari satış adedi kullanılmıştır. Modelin tüm değişkenleri Tablo 1’de gösterilmektedir.

Tablo 1. Bağımlı ve bağımsız değişkenler (Dependent and independent variables)

Bağımlı Değişken	Bağımsız Değişkenler
Hafif Ticari Satış Adedi	Tüketici Fiyat Endeksi (TÜFE) (2003=100)
	Sanayi Güven Endeksi
	İmalat Sanayi Kapasite Kullanım Oranı
	Taahhüt Kredi Faiz Oranı
	ABD dolar TL Kuru
	Euro TL Kuru
	Sanayi Üretim Endeksi (2015=100)
Trafikten Silinen Hafif Ticari Araç Sayısı	

Çalışmada hafif ticari araç satış adedi tahminini gerçekleştirmek için modele Türkiye’nin genel ekonomik durumunu yansıtabilmek için tüketici fiyat endeksi (TÜFE), imalat sanayi kapasite kullanım oranı, sanayi üretim endeksi, ABD doları kuru, Euro kuru ve bunun yanında sanayi sektöründeki ekonomik beklenti ve eğilimleri ortaya koyan sanayi güven endeksi de bağımsız değişkenler olarak seçilmiştir. Ayrıca modele taşıt satışı ile ilgili taşıt kredi faiz oranı ve trafikten silinen hafif ticari araç sayısı bağımsız değişken olarak dahil edilmiştir. Taahhüt kredi faiz oranındaki değişikliklerin satın almada belirleyici bir finansman kaynağı olması düşünüldükçe modele dahil edilmiştir. Trafikten silinen hafif ticari araç sayısı bağımsız değişkeni ise yeni araç talebi oluşturabileceği için modele ilave edilmiştir.

3.2. Yöntem (Method)

3.2.1. Yapay Sinir Ağları (Artificial Neural Networks)

Yapay sinir ağları, insanların gerçek bir biyolojik sinir sisteminin çalışma yapısını bilgisayar ve makine sistemlerine uyarlamak ve insanlar gibi öğrenme, tahmin ve sınıflandırma işlevlerini gerçekleştirmek için kullanılan yapay zeka disiplini [16]. Özellikle doğrusal olmayan problemlerin çözümünde sağladıkları başarılı sonuçlar nedeniyle yapay sinir ağlarının kullanım alanı ve yaygınlığı artmıştır. ANN belirli bir konuda, mevcut verileri kullanarak sonucu bilinmeyen örneklerle ilgili tahmin yapabilme yeteneğine sahiptir [15, 17].

Yapay sinir ağları, mimari yapısına göre ileri ve geri beslemeli olarak ikiye ayrılmaktadır. İleri beslemeli ağlar, bir giriş, bir gizli ve bir çıkış katmanından oluşmaktadır. Gizli ve çıkış katmanlarında bulunan işlem

birimlerine nöron adı verilir ve her nöron kendi aktivasyon fonksiyonu ve toplama fonksiyonundan oluşmaktadır [18]. Aktivasyon fonksiyonu, ağın verilerdeki karmaşık kalıpları öğrenmesine yardımcı olmak için bir yapay sinir ağına eklenen bir fonksiyondur ve bir nöronun ağda aktive edilip edilmeyeceğine karar vermektedir. Nöronun ağa girişinde matematiksel işlemler kullanarak tahmin sürecinde önemli olup olmadığına karar vermesini sağlamaktadır [19]. Gizli katmandaki aktivasyon fonksiyonu seçimi, ağ modelinin eğitim veri kümesini ne kadar iyi öğrendiğini kontrol etmesini sağlamaktadır [20]. Çıktı katmanındaki aktivasyon fonksiyonu seçimi, modelin yapabileceği tahminlerin türünü tanımlamaktadır. Aktivasyon fonksiyonu olarak adimsal, doğrusal ve sigmoid gibi fonksiyonlarından biri seçilebilir. Nöronlar daha sonra diğer katmanlardan nöronlara kısmen veya tamamen bağlanmaktadır. Her bağlantının bir ağırlığı bulunmaktadır. Ağırlıklar, sinir ağının istenen hedefe ulaşması için ayarlanan birer parametredir. Her katman girdilere karşılık gelen ağırlıklarıyla çarpılmaktadır. Daha sonra nöronlardaki aktivasyon fonksiyonuna aktarılır [21]. Geri beslemeli sinir ağının temel amacı, sinir ağının çıktısının istenen hedefe ulaşması için ağ ağırlıklarının her birini güncellemesidir. Geri beslemeli sinir ağı en çok kullanılan yapay sinir ağı türüdür [22]. Sinir ağı çıktısı ile istenen hedef arasındaki hata, bir maliyet fonksiyonu olarak ifade edilmektedir. Amaç, ağırlıkları güncelleyerek bu maliyet fonksiyonunu en aza indirmektir [23].

Yapay sinir ağı ile beyin arasındaki bir diğer benzerlik ise öğrenme yaklaşımıdır. İnsan beyni gibi yapay sinir ağının da eğitilmesi gerekir. Yapay sinir ağının eğitimi, nöronların girdileri üzerinde en iyi ağırlıkları ayarlama işlemidir. Yapay sinir ağlarının en önemli özelliklerinden biri, daha önce hiç görmedikleri soruların (girdileri) cevaplarını önceki bilgilerinden çıkarabilmeleridir. Bu, yapay sinir ağlarının genelleme yeteneği olarak adlandırılmaktadır [24].

3.2.2. Çoklu Doğrusal Regresyon (Multiple Linear Regression)

Regresyon modeli, bir bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkinin gücünü ve karakterini belirlemeye çalışan finans, yatırım ve diğer disiplinlerde kullanılan istatistiksel bir yöntemdir [25, 26]. Değişkenler arasındaki bu ilişki doğrusal veya doğrusal olmayan düzlemler ile ifade edilebilmektedir. [27]

Doğrusal regresyon, regresyon analizinde kullanılan bir modelleme tekniğidir. Doğrusal regresyon, iki değişken arasındaki ilişkiyi düz bir çizgi boyunca kurmaya çalışmaktadır. Doğrusal regresyon bağımsız ve bağımlı değişken olmak üzere iki sürekli değişkene sahip olduğunda kullanılabilir. Bağımsız değişken, bağımlı değişkeni veya sonucu hesaplamak için kullanılan değişkendir [28].

Doğrusal regresyon, değişken sayısına göre basit ve çoklu doğrusal regresyon olmak üzere ikiye ayrılmaktadır. Basit doğrusal regresyon, bağımlı değişkenin değerini bir bağımsız değişken kullanarak tahmin etmeyi sağlamaktadır [27, 29].

Tek bir bağımsız değişkenle açıklanamayan bağımlı değişkenler için çoklu regresyon kullanılmaktadır. Çoklu doğrusal regresyon, bağımlı bir değişkenin sonucunu tahmin etmek için iki veya daha fazla bağımsız değişken kullanan istatistiksel bir tekniktir [30].

m adet bağımsız değişken olan ve değişkenler arasında doğrusal bir ilişki olan bir modelde:

$$y = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots + \beta_k X_{ik} + \varepsilon_{ij} \quad (1)$$

Burada X bağımsız değişkenin değerini, β regresyon katsayısını, e hata terimini, k bağımsız değişken sayısını göstermektedir. β katsayısı, bağımsız değişkenin bir birim değişmesine karşılık bağımlı değişkende meydana gelecek değişme miktarını göstermektedir.

3.2.3. Karar Ağaçları Regresyonu (Decision Trees Regression)

Makine öğrenmesi teorisine dayanan karar ağacı, sınıflandırma ve regresyon problemlerinin çözümü için kullanılan bir yöntemdir. Karar ağacı, her düğümde rasgele sayıda düğüm ve daldan oluşan bir ağaç veri yapısıdır [31, 32]. Tekrarlayan bölme işlemi yoluyla tahmin için kullanılacak kurallar kümesi oluşturmaktadır. Karar ağacının diğer modelleme tekniklerine göre en büyük avantajı, yorumlanabilir kuralları veya mantık ifadelerini temsil edebilecek bir model üretmesidir [33, 34]. Sınıflandırma karmaşık hesaplamalar yapılmadan gerçekleştirilebilmektedir. Ayrıca hem sürekli hem de kategorik değişkenler için

kullanılabilmektedir [35].

Sınıflandırmayı tek bir karar adımında gerçekleştirmek için bir dizi özellik kullanan diğer sınıflandırma yaklaşımlarından farklı olarak, karar ağacı çok aşamalı ve hiyerarşik bir karar şemasına dayanan ağaca benzeyen bir yapı oluşturmaktadır. Ağaç, bir kök düğümden (tüm verileri içeren), bir dizi iç düğümden (bölünmeler) bir dizi terminal düğümden (yapraklar) oluşmaktadır. Karar ağacı yapısının her düğümü, bir sınıfı diğer sınıflardan ayıran ikili bir karar vermektedir. Ağaç yapısı yaprak düğüme ulaşılan kadar aşağı doğru hareket ettirilerek gerçekleştirilmektedir [36].

Bir karar ağacının temel amacı, karmaşık bir kararı, yorumlanması daha kolay bir çözüme götürecektir birkaç basit karara bölmektir. Bir karar ağacı yaklaşımında, verilerin özellikleri tahmin değişkenleri iken, eşleştirilecek sınıf hedef değişken olarak adlandırılmaktadır [37].

Hedef değişken ayrık olduğunda karar ağacı sınıflandırması, hedef değişkenler sürekli olduğunda karar ağacı regresyonu olarak adlandırılmaktadır. Karar ağacı regresyon yaklaşımı, özellikler ve hedef nesnel arasındaki ilişkilerin doğrusal ya da doğrusal olmadığı varsayım üzerine kurulmaktadır. Bir karar ağacında, sınıflandırma için maksimum bilgi taşıyan öznitelikler otomatik olarak seçilmekte ve kalan öznitelikler reddedilerek hesaplama verimliliği artmaktadır [38].

Karar ağacı, verilen örneklerden bir karar ağacı oluşturan bir algoritma oluşturmaktadır. Gerçekleştirilen algoritma, uygunluk fonksiyonunu minimize ederek optimal karar ağacını bulmayı amaçlamaktadır. Başlangıçta, ağacın yapısını belirlemek için tüm eğitim örnek için kullanılmaktadır [39]. Veri kümesi, her bağımsız değişken için birkaç bölme noktasına bölünmektedir. Her bölme noktasında, gerçekleştirilen algoritma, önceden tanımlanmış uygunluk fonksiyonuyla ilgili olarak tahmin edilen değer ile gerçek değerler arasındaki hatayı hesaplamaktadır. Değişkenler arasındaki bölünme noktası hataları karşılaştırılmakta ve en düşük uygunluk fonksiyonu değerini veren değişken bölme noktası olarak seçilmektedir [40, 41]. Bu süreç yinelemeli olarak her bir düğüm, kullanıcı tarafından belirlenen minimum düğüm boyutuna (düğümdeki eğitim örneklerinin sayısı) ulaşılan ve bir uç düğüm haline gelene kadar devam etmektedir [42, 43].

En yaygın ağaç yöntemleri arasında CHAID, CART, C4.5 ve C5 bulunur [44].

3.3. Model performans değerlendirilmesi

Çalışma sonuçlarını analiz etmek ve yorumlamak için R^2 , MSE (Mean Squared Error) ve MAPE (Mean Absolute Percentage Error) teknikleri kullanılmıştır.

R^2 , belirli bir olayın sonucunu tahmin ederken, bağımlı değişkendeki değişikliklerin bağımsız bir değişkendeki farkla nasıl açıklanabileceğini inceleyen istatistiksel bir ölçümdür [45]. Başka bir deyişle, daha yaygın olarak R^2 olarak bilinen bu katsayı, doğrusal ilişkinin iki değişken arasında ne kadar güçlü olduğunu değerlendirir. R^2 değeri 0 ile 1 arasında bir sayıdır [46]. 0 ile 1 arasında olması modelin sonucu kısmen tahmin ettiğini gösterir. 1 olması mükemmel bir uyumu gösterdiğini; 0 ise modelin verileri doğru bir şekilde modelleyemediğini gösterir [47, 48]. R^2 formülü Denklem 2'de gösterilmiştir:

$$R^2 = 1 - \frac{\text{Açıklanamayan Varyasyon}}{\text{Toplam Varyasyon}} \quad (2)$$

MSE (Mean Squared Error), regresyon modelinin verilere ne kadar iyi veya ne kadar kötü uyduğunu ölçmek için kullanılan istatistiksel bir yöntemdir [49]. Bir regresyon çizgisinin bir dizi noktaya ne kadar yakın olduğunu söyler. Bunu, noktalardan regresyon çizgisine olan mesafelerin (bu mesafeler hata olarak adlandırılır) karesini alarak yapar [50]. Makine öğrenmesinde gözlenen ve tahmin edilen değerler arasındaki ortalama kare farkını değerlendirir. Bir modelde hata olmadığında, MSE sıfıra eşittir. Bir modelin MSE değeri düşüğe, yani 0'a yakınsa, veri seti için iyi bir uyumdur. Tersine, 0'dan uzak yüksek MSE değerine sahip bir model, veri seti için uygun değildir. Model hatası arttıkça değeri de artar [51, 52]. MSE formülü Denklem 3'te gösterilmiştir:

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

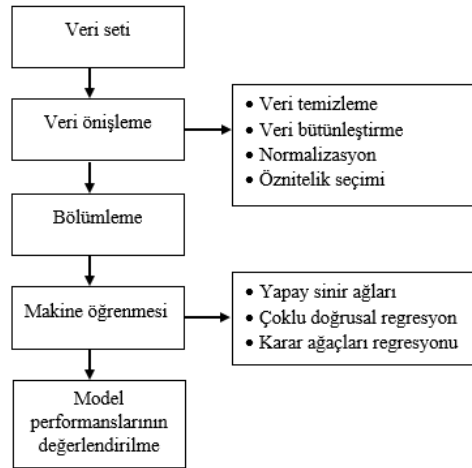
Burada n veri sayısını, e hata değerini göstermektedir.

MAPE (Mean Absolute Percentage Error) bir tahmin sisteminin doğruluğunu yüzdesel olarak ölçen istatistiksel bir yöntemdir. MAPE, bir makine öğrenmesi yönteminin iyi performans gösterip göstermediğini belirlemek için yaygın olarak kullanılan bir yöntemdir [53]. Bir veri kümesindeki her girişin mutlak yüzde hatalarının ortalamasını temsil eder ve ortalama olarak, tahmin edilen miktarların gerçek miktarlarla karşılaştırıldığında ne kadar doğru olduğunu gösterir. MAPE genellikle daha büyük veri kümelerini analiz etmek için etkilidir [54]. Veri kümelerinin MAPE'sini sıfır değerlerle hesaplamak mümkün değildir. Bunun nedeni, hesaplamanın sıfıra bölmeyi gerektirmesidir. MAPE ne kadar küçük olursa, tahmin o kadar iyi olur. MAPE, yorumlanması ve açıklanması kolay önemli bir ölçümdür [55]. MAPE formülü Denklem 4'te gösterilmiştir:

$$MAPE = \frac{100}{n} \sum_j^n \frac{|e_j|}{|A_j|} \quad (4)$$

4. Sonuçlar ve Tartışma (Results and Discussion)

Çalışmada Türkiye'deki ticari araç satış adetini tahmin etmek için 198 veri kullanılmıştır. Çalışmada oluşturulan modelde bağımlı değişken olarak Türkiye'deki hafif ticari araç satış adeti; bu bağımlı değişkeni tahmin etmek için de sekiz bağımsız değişken kullanılmıştır. Çalışmada çoklu doğrusal regresyon, karar ağacı regresyonu ve yapay sinir ağları olmak üzere 3 farklı makine öğrenmesi tekniği kullanılmıştır. Modellerin başarı ve hata analizlerini yapmak için ise R2, MSE ve MAPE istatistiksel yöntemleri kullanılmıştır. Çalışmada açık kaynak program olan Knime programı kullanılmıştır. Knime programı veri analizi, raporlama, entegrasyon işlemlerinde kullanılan grafiksel bir arayüze sahip olan bir platformdur. Çalışmanın temel çalışma yapısı Şekil 2'de görülmektedir.



Şekil 2. Çalışmanın yapısı (Structure of the study)

Şekil 2'de görüldüğü gibi veri seti oluşturulduktan sonra yapılacak ilk işlem veri seti üzerinde ön işleme aşamasıdır. Bu aşamanın da ilk adımı veri temizlemedir.

Veri toplama aşamasında hatalı veri elde edilebilir veya veri toplama araçlarından hatalı bilgiler gelebilmektedir. Bu yüzden veri kümeleri makine öğrenmesi yöntemlerinin doğru olarak anlayamadığı ve yorumlayamadığı veriler içerebilir [56]. Veri kümesindeki bu hatalara gürültü veri adı verilmektedir. Gürültülü veriler, depolama alanı miktarını gereksiz yere artırır ve herhangi bir veri madenciliği analiz sonucunu olumsuz etkileyebilir [57, 58]. Gürültülü verileri veri kümesinden tespit etmek ve ortadan kaldırmak için yapılabilecekler:

Normalleştirme, genellikle makine öğrenmesi için veri hazırlamanın bir parçası olarak uygulanan bir tekniktir [59]. Normalleştirme amacını, değer aralıklarındaki farklılıkları bozmadan veya bilgi kaybetmeden veri

kümesindeki sayısal sütunların değerlerini ortak bir ölçek kullanacak şekilde değiştirmektedir [60, 61]. Bazı algoritmaların verileri doğru modellemesi için normalleştirme de gereklidir. Normalleştirme, kaynak verilerdeki genel dağılımı ve oranları koruyan yeni değerler oluştururken, değerleri modelde kullanılan tüm sayısal sütunlara uygulanan bir ölçek içinde tutarak bu sorunları önler [62]. Bu çalışmada veriler Min-Max yöntemi kullanılarak orijinal veriler üzerinde doğrusal bir dönüşüm gerçekleştirilmiştir. Her özellik için o özelliğin minimum değeri 0'a, maksimum değeri 1'e ve diğer tüm değerler 0 ile 1 arasında ondalık sayıya dönüştürülür. Normalleştirme de orijinal veri değerleri arasındaki ilişkileri korunur [63, 64]. Min-Max normalizasyon formülü:

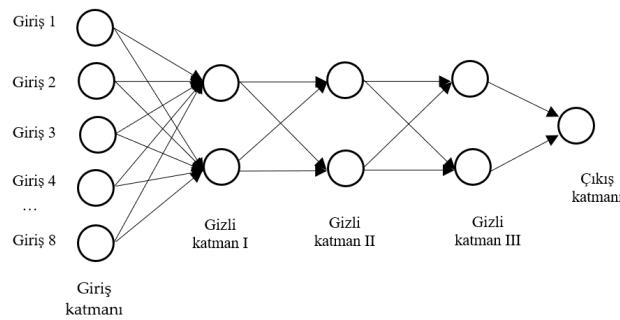
$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (5)$$

Burada x' normalize edilmiş veriyi, x_i girdi değerini, x_{min} veri kümesindeki en küçük veriyi, x_{max} veri kümesindeki en büyük veriyi göstermektedir.

Çalışmanın bir diğer aşaması verilerin eğitim ve test olmak üzere ikiye bölünmesidir [65]. Eğitim aşamasında seçilen makine öğrenmesi için ideal parametrelerin belirlendiği ve hatanın minimum düzeye indirildiği aşamadır. Test aşamasında ise eğitim aşamasında belirlenen parametrelerin veri setinde daha önce kullanılmamış veriler üzerinde test edildiği ve istatistiksel olarak değerlendirildiği aşamadır [66]. Eğitim ve test verilerinin ayrılması için kesin bir kural olmamakla birlikte çalışmanın niteliğine ve veri sayısına göre veri kümesinin %60'ı ile %80'i arasında bir oran eğitim verisi olarak ayrılmaktadır. Bu çalışmada veri kümesinin %70 eğitim ve %30 test verisi olarak yani 139 veri eğitim, 59 veri test için ayrılmıştır. Verilerin eğitim ve test olmak üzere ikiye ayrılmasında farklı yöntemler kullanılmaktadır. Veri kümesinin en üstten alma, doğrusal örnekleme ve rastgele veri alma kullanılabilir veri seçim yöntemlerinden bazılarıdır. Çalışmada iki modelin sonuçlarını karşılaştırabilmek için veri seçiminde doğrusal örnekleme yöntemi tercih edilmiştir.

Çalışmada üç farklı makine öğrenmesi tekniği kullanılmıştır. Bunlar yapay sinir ağları, çoklu doğrusal regresyon ve karar ağaçları regresyonudur.

Yapay sinir ağları yöntemi için sekiz giriş, bir çıkış nöronundan oluşan geri beslemeli bir model geliştirilmiştir. İdeal gizli katman ve her gizli katmandaki nöron sayısına modelde yapılan çok sayıda test sonucu karar verilmiştir. Çalışmada en başarılı sonucu verdiği için üç gizli katman ve her gizli katmanda üç nöron olan yapının daha başarılı sonuçlar ürettiği görülmüştür. Geliştirilen modelin yapısı Şekil 3'te görülmektedir.



Şekil 3. Tasarlanan ANN modeli (ANN model)

Aktivasyon fonksiyonu için çeşitli fonksiyonların denenmesinin ardından sigmoid fonksiyonun geliştirilen model ve veri kümesi için en iyi sonucu verdiği görülmüştür. Çalışmada hata değerlerini minimize edilmesi için geri yayılım sinir ağı tercih edilmiştir. Modelde en iyi sonucu almak için 1000 iterasyon gerçekleştirilmiştir.

Çoklu doğrusal regresyonda değişkenlerin sisteme olan etkilerini ölçebilmek için öncelikle bir önem değeri (significance value) belirlendi. En yüksek p-value değerine (probability value) sahip değişken belirlenip ve şayet $P > SL$ ise değişken sistemden kaldırıldı. Model yeniden oluşturulup ardından bu adım tekrarlandı. Tüm değişkenler için $P < SL$ olduğunda ise eleme sonlandırıldı. Tasarlanan modelde p değerleri için 0,05'in altında kalan bağımsız değerler olmadığı için modelin anlamlı olduğu görülmüştür.

Çalışmada kullanılan bir diğer makine öğrenmesi yöntemi olan karar ağacı regresyonu yönteminde bölmenin hesaplandığı kalite ölçüsü olarak "Gini Index" kullanılmıştır. Her düğümde en az üç minimum kayıt sayısını bulunmaktadır. Kayıt sayısı bu sayıdan küçük veya ona eşitse, ağaç daha fazla yetiştirilmez. Ağaçta görünüm için 10.000 kayıt saklanmaktadır. Çalışmada 8 tane iş parçacığından ve dolayısıyla işlemciden veya çekirdekten yararlanılmıştır. Bu da çalışmanın performansının artmasını sağlamıştır.

ANN, MLR ve DT yöntemlerinin karşılaştırılması Tablo 2'de gösterilmiştir.

Tablo 2. Modellerin başarı ve hata karşılaştırılması (Success and error comparison of the models)

	Yapay sinir ağları	Çoklu doğrusal regresyon	Karar ağaçları regresyonu
R ²	0,946	0,641	0,822
MSE	0,001	0,055	0,041
MAPE	0,130	0,654	0,295

Verilerin doğrusal bir eğriye ne kadar iyi uyduğunu gösteren R² değerinin 1 olması, test verilerinin doğrusal bir eğri sağlandığını göstermektedir. Çalışmanın sonucunda R² değeri ANN için %94,6, MLR için %64,1, DT için %82,2 olmuştur ve ideal değere çok yakın olduğu görülmüştür. Makine öğrenmesi modellerinin performansını hata değerine göre ölçen MSE değerinin sıfıra yakın olması modellerin daha iyi ve daha az hata gösterdiğini ifade etmektedir. Bu yüzden çalışmada kullanılan modeller için MSE değerinin sıfır değerine yakın olması istenmektedir. Çalışmada MSE değerinin ANN için 0.001, MLR için 0.055 ve DT için 0.041 olduğu ve ideal değere yakın olduğu görülmüştür. MAPE değerinin yüzde 10 değerinin altında olan modeller çok iyi kabul edilmektedir [67]. Çalışmada ANN için 1,30, MLR için 6.54 ve DT için 2.95 olduğu görülmüştür. Her üç model içinde MAPE değerinin çok iyi olduğu kabul edilmektedir. Hata ve başarı değerlerine bakıldığında en başarılı ve en az hataya sahip modellerin sırasıyla ANN, DT ve MLR olduğu görülmüştür.

5. Sonuçlar (Conclusion)

Bu çalışmada Ocak 2006 Haziran 2022 tarihleri arasında veriler ile oluşturulan modelde makine öğrenmesi yöntemleri kullanılarak Türkiye'de satılan hafif ticari araç satış miktarı tahmini gerçekleştirilmiştir. Çalışma, otomotiv sektörünün bir alt kolu olan hafif ticari araçların satış adedi tahmini üzerine yapılması literatürdeki özellikle Türkiye'deki çalışmalardan farklılaşmaktadır. Hafif ticari araçların odaklanmasındaki temel motivasyon bu grup araçların özellikle binek otomobillerden farklı olarak sadece bireysel kullanım değil ülkenin ticari, üretim ve hizmet olmak üzere üç temel sektörde ticari bir taşıt vasfı ile kullanılmasıdır. Bu üç temel sektördeki dinamizm ülkelerin özellikle ekonomi üzerinde olumlu yönde bir devinim sağlarken daralmada ekonomilerdeki yavaşlama hakkında önemli göstergeler vermektedir. Bu nedenle çalışmada özellikle Türkiye gibi gelişmekte olan ülkelerde yapılacak araştırmalar açısından bir katkı sağlayacağı düşünülmektedir. Bunun yanında hafif ticari araç sektörü başta bu araçların satışı başta olmak üzere bankacılık, sigortacılık, bakım-onarım ve kiralama hizmetleri gibi birçok sektör ile etkileşim içindedir. Bu nedenle hafif ticari araçların satışını tahmin etmek ilgili sektörlerdeki hedef, kaynak planlaması ve bütçeleme faaliyetleri için önem arz etmektedir. Çalışmada en büyük sınırlama veri aralığı ile ilgilidir. Her ne kadar Türkiye'deki benzer otomotiv sektörü satış tahmini çalışmalara göre daha uzun bir veri aralığı kullanılsa da özellikle 1990'lardan başlayarak o yıllarda Türkiye'de yaşanan ekonomik krizlerinde etkilerini makine öğrenmesi tekniklerinde kullanmak modelin başarısının ölçülmesi noktasında daha yararlı olacağı muhakkaktır. Ancak sunulan verilerin çoğu ortak başlangıç noktası Ocak 2006 olduğu için zaman serisi buradan başlatılmıştır.

Bu araçlarının satışını gerçekleştiren otomotiv satış bayileri bu modeli kullanılarak otomotiv fabrikalarından yıllık alım stoklarını belirleyebileceklerdir. Böylelikle bu taşıtları alırken oluşacak sermaye maliyetinden kaçınarak aynı zamanda stok maliyetlerini düşürebilirler. İkinci olarak bankacılık, sigortacılık, bakım onarım, filo hizmetleri veren işletmeler de oluşacak hafif satış miktarı tahminlerine göre satış bütçelerini hazırlayabileceklerdir. Ayrıca özellikle bu tip araçlara bakım onarım hizmetleri veren şirketler yedek parça ihtiyaçlarını bu tahminlere göre şekil verebileceklerdir. Son olarak çalışmada değinildiği gibi hafif ticari araçların satışı ekonominin genel görünümü hakkında bilgiler de verebilir. Ekonominin konjonktürel olarak yükselme ve zirve dönemlerinde ekonominin büyümesi ile bu araçların satışı artacaktır. Bu yüzden bu satışların tahmini uzun dönemli ekonomik kararların alınmasında kullanılabilirliği düşünülmektedir.

Çalışmanın modeli tasarlanırken hem ekonominin genel görünümü hem de hafif ticari araç satışlarını

etkileyeceği düşünülen tüketici fiyat endeksi (TÜFE), sanayi güven endeksi, imalat sanayi kapasite kullanım oranı, sanayi güven endeksi, taşıt kredi faiz oranı, trafikten silinen hafif ticari araç sayısı, ABD dolar TL ve Euro TL kuru bağımsız değişken olarak seçilmiştir. Çalışmada makine öğrenmesi tekniklerinden ANN, MLR ve DT yöntemleri kullanılmıştır. Elde edilen sonuçlar R^2 değeri ANN için %94,6, MLR için %64,1, DT için %82,2 olmuştur ve ideal değere çok yakın olduğu görülmüştür. Bu sonuçlar ANN'nin diğer iki makine öğrenmesi yöntemine göre daha başarılı tahmin gücüne sahip olacağını göstermektedir. Çalışmanın sonuçları Karatlı vd. (2012), Akyurt (2015), Kuvvetli vd. (2015)'nin Türkiye'de otomobil satış tahmini üzerine yaptıkları benzer çalışmalar ile örtüşmektedir. Bu çalışmalarda da ANN başarılı bir makine öğrenmesi tekniği olarak tespit edilmiştir. Ayrıca Vahabi vd. (2016)'da İran'da bir otomobil markası üzerine yapmış olduğu çalışma da benzer şekilde ANN'yı başarılı bir makine öğrenmesi olarak göstermiştir.

Çalışmadaki tasarlanan modelin değişkenlerine ait verilerin ulaşılabilirliği ve uygulanan makine öğrenmesi teknikleri bundan sonraki akademik araştırmalar için karşılaştırılabilir bir önerinin yanında hafif ticari araç satışı yapan bayiler ve sektörün paydaş işletmeleri için de geleceğin planlaması noktasında örnek teşkil etmektedir. Problem Sandgren tarafından araştırıldı. Deb ve Goyal, problemi genetik algoritmaya dayalı akıllı bir yöntem olan genetik uyarlamalı aramaya (GeneAS) uyguladı. Lampinen ve Zelinka, problemi diferansiyel evrim (DE) yöntemiyle çözmeye çalışmışlardır. Lampinen ve arkadaşları, problemi parçacık sürüsü algoritmasıyla araştırdı.

Çıkar Çatışması Beyanı (Conflict of Interest Statement)

Yazarlar tarafından herhangi bir çıkar çatışması bildirilmemiştir.

Kaynaklar (References)

- [1] A. G. Ö. Görener, "Türk otomotiv sektörünün ülke ekonomisine katkıları ve geleceğe yönelik sektörel beklentiler," *Yaşar Üniversitesi E-Dergisi*, vol. 3, no. 10, pp. 1213-1232, 2008.
- [2] S. Yılmaz, K. Taştan, N. Ecek, ve E. Çınar, "Otomotiv sektörünün dünyadaki ve Türkiye'deki değişimi," *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi*, vol. 7, no. 3, pp. 685-695, 2017.
- [3] O. Başkol, "Türk otomotiv sektörünün uluslararası rekabet gücü (1996-2010)," *Paradoks Ekonomi Sosyoloji ve Politika Dergisi*, vol. 7, no. 2, pp. 63-78, 2011.
- [4] T.C. Sanayi ve Teknoloji Bakanlığı, "Otomotiv sektör raporu," [Online]. Available: <https://www.sanayi.gov.tr/assets/pdf/plan-program/OtomotivSektorRaporu2020.pdf> [Accessed: December. 12, 2023].
- [5] M. H. Eken ve M. Çiçek, "Türkiye'de otomotiv sektöründeki ürünlerin kredilerle finansmanının satışlara etkisi," *Maliye ve Finans Yazıları*, vol. 1, no. 84, pp. 61-77, 2009.
- [6] S. Kaynak ve Y. O. Ari, "Türk otomotiv sektöründe yoğunlaşma: binek ve hafif ticari araçlar üzerine bir uygulama," *Ekonomik Yaklaşım*, vol. 22, no. 80, pp. 39-58, 2011.
- [7] T. Efendigil ve Ö. E. Eminler, "Havacılık sektöründe talep tahminin önemi: Yolcu talebi üzerine bir tahmin modeli," *Yaşar Üniversitesi E-Dergisi*, vol. 12, pp. 14-30, 2017.
- [8] M. Soysal ve M. Ömürganülşen, "Türk turizm sektöründe talep tahmini üzerine bir uygulama," *Anatolia: Turizm Araştırmaları Dergisi*, vol. 21, no. 1, pp. 128-136, 2010.
- [9] M. Karaatlı, Ö. C. Helvacıoğlu, N. Ömürbek ve G. Tokgöz, "Yapay sinir ağları yöntemi ile otomobil satış tahmini," *Uluslararası Yönetim İktisat ve İşletme Dergisi*, vol. 8, no. 17, pp. 87-100, 2012. Doi: 10.11122/ijmeh.2012.8.17.290
- [10] İ. Z. Akyurt, "Talep tahmininin yapay sinir ağlarıyla modellenmesi: Yerli otomobil örneği," *Ekonometri ve İstatistik Dergisi*, no. 23, pp. 147-157, 2015.
- [11] C. E. Alper ve A. Mumcu-Serdar, "Türkiye'de otomobil talebinin tahmini," *Ekonomi ve Ekonometri Merkezi, Boğaziçi Üniversitesi*, 2000.
- [12] F. K. Wang, K. K. Chang, and C. W. Tzeng, "Using adaptive network-based fuzzy inference system to forecast automobile sales," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10587-10593, 2011. Doi: 10.1016/j.eswa.2011.02.100
- [13] Y. Kuvvetli, C. Dağsuyu, ve M. Oturakci, "Türkiye'deki araç satışları için ekonomik ve çevresel faktörleri göz önüne alan yapay sinir ağı tabanlı bir tahmin yaklaşımı," *Endüstri Mühendisliği*, vol. 26, no. 3, pp. 23-31, 2015.
- [14] P. F. Pai and C. H. Liu, "Predicting vehicle sales by sentiment analysis of Twitter data and stock market values," *IEEE Access*, vol. 6, pp. 57655-57662, 2018. Doi: 10.1109/ACCESS.2018.2873730

- [15] S. Arslankaya and V. Öz, "Time series analysis of sales quantity in an automotive company and estimation by artificial neural networks," *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, vol. 22, no. 5, pp. 1482-1492, 2018. Doi: 10.16984/saufenbilder.456518
- [16] S. Haykin, "Neural networks and learning machines", 3/E. Pearson Education India, 2009.
- [17] M. Kayakuş and K. K. Çevik, "Estimation the number of visitors of e-commerce website by artificial neural networks during covid19 in turkey," *Electronic Turkish Studies*, vol. 15, no. 4, pp. 615-631, 2020.
- [18] F. Recknagel, M. French, P. Harkonen, and K.-I. Yabunaka, "Artificial neural network approach for modelling and prediction of algal blooms," *Ecological Modelling*, vol. 96, no. 1-3, pp. 11-28, 1997. Doi: 10.1016/S0304-3800(96)00049-X
- [19] B. Karlik and A. V. Olgac, "Performance analysis of various activation functions in generalized MLP architectures of neural networks," *International Journal of Artificial Intelligence and Expert Systems*, vol. 1, no. 4, pp. 111-122, 2011.
- [20] G. Bandyopadhyay and S. Chattopadhyay, "Single hidden layer artificial neural network models versus multiple linear regression model in forecasting the time series of total ozone," *International Journal of Environmental Science & Technology*, vol. 4, pp. 141-149, 2007. Doi: 10.1007/BF03325972
- [21] S. B. Maind and P. Wankar, "Research paper on basic of artificial neural network," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 2, no. 1, pp. 96-100, 2014.
- [22] S. Herzog, C. Tetzlaff, and F. Wörgötter, "Evolving artificial neural networks with feedback," *Neural Networks*, vol. 123, pp. 153-162, 2020. Doi: 10.1016/j.neunet.2019.12.004
- [23] C. Vidal et al., "Hybrid energy storage system state-of-charge estimation using artificial neural network for micro-hybrid applications," *IEEE Transportation Electrification Conference and Expo (ITEC)*, 2018: IEEE, pp. 1075-1081. Doi: 10.1109/ITEC.2018.8450251
- [24] S. Cavaliere, P. Maccarrone, and R. Pinto, "Parametric vs. neural network models for the estimation of production costs: A case study in the automotive industry," *International Journal of Production Economics*, vol. 91, no. 2, pp. 165-177, 2004. Doi: 10.1016/j.ijpe.2003.08.005
- [25] L. E. Eberly, "Multiple linear regression," *Topics in Biostatistics*, pp. 165-187, 2007. Doi: 10.1007/978-1-59745-530-5_9
- [26] M. Kayakuş, D. Erdoğan, and M. Terzioğlu, "Predicting the share of tourism revenues in total exports," *Alphanumeric Journal*, vol. 11, no. 1, pp. 17-30, 2023. Doi: 10.17093/alphanumeric.1212189
- [27] M. Tranmer and M. Elliot, "Multiple linear regression," *The Cathie Marsh Centre for Census and Survey Research (CCSR)*, vol. 5, no. 5, pp. 1-5, 2008.
- [28] S. Kılıç, "Doğrusal regresyon analizi," *Journal of Mood Disorders*, vol. 3, no. 2, pp. 90-92, 2013.
- [29] M. Kayakuş, B. Tutcu, M. Terzioğlu, H. Talaş, and G. F. Ünal Uyar, "ROA and ROE forecasting in iron and steel industry using machine learning techniques for sustainable profitability," *Sustainability*, vol. 15, no. 9, p. 7389, 2023. Doi: 10.3390/su15097389
- [30] X. Yan and X. Su, "Linear regression analysis: theory and computing". *World scientific*, 2009. Doi: 10.1142/6986
- [31] W. Y. Loh, "Classification and regression trees," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 1, no. 1, pp. 14-23, 2011. Doi: 10.1002/widm.8
- [32] M. Kayakuş and F. Y. Açıkgöz, "Classification of News Texts by Categories Using Machine Learning Methods," *Alphanumeric Journal*, vol. 10, no. 2, pp. 155-166, 2022. Doi: 10.17093/alphanumeric.1149753
- [33] P. Perner, U. Zscherpel, and C. Jacobsen, "A comparison between neural networks and decision trees based on data from industrial radiographic testing," *Pattern Recognition Letters*, vol. 22, no. 1, pp. 47-54, 2001. Doi: 10.1016/S0167-8655(00)00098-2
- [34] M. Kayakuş and F. Yiğit Açıkgöz, "Twitter'da makine öğrenmesi yöntemleriyle sahte haber tespiti," *Abant Sosyal Bilimler Dergisi*, vol. 23, no. 2, pp. 1017-1027, 2023. Doi: 10.11616/asbi.1266179
- [35] G. K. Tso and K. K. Yau, "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks," *Energy*, vol. 32, no. 9, pp. 1761-1768, 2007. Doi: 10.1016/j.energy.2006.11.010
- [36] M. Xu, P. Watanachaturaporn, P. K. Varshney, and M. K. Arora, "Decision tree regression for soft classification of remote sensing data," *Remote Sensing of Environment*, vol. 97, no. 3, pp. 322-336, 2005. Doi: 10.1016/j.rse.2005.05.008
- [37] L. Breiman, J. H. Friedman, R. A. Olshen, and C. J. Stone, "Classification and regression trees", Routledge, 2017. Doi: 10.1201/9781315139470
- [38] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55-63, 1968. Doi: 10.1109/TIT.1968.1054102
- [39] C. Kingsford and S. L. Salzberg, "What are decision trees?," *Nature Biotechnology*, vol. 26, no. 9, pp. 1011-1013, 2008. Doi: 10.1038/nbt0908-1011

- [40] J. Elith, J. R. Leathwick, and T. Hastie, "A working guide to boosted regression trees," *Journal of Animal Ecology*, vol. 77, no. 4, pp. 802-813, 2008. Doi: 10.1111/j.1365-2656.2008.01390.x
- [41] W.-Y. Loh and Y.-S. Shih, "Split selection methods for classification trees," *Statistica Sinica*, pp. 815-840, 1997.
- [42] E. Pekel, "Estimation of soil moisture using decision tree regression," *Theoretical and Applied Climatology*, vol. 139, no. 3, pp. 1111-1119, 2020. Doi: 10.1007/s00704-019-03048-8
- [43] B. M. Moret, "Decision trees and diagrams," *ACM Computing Surveys (CSUR)*, vol. 14, no. 4, pp. 593-623, 1982. Doi: 10.1145/356893.356898
- [44] J. R. Quinlan, *C4. 5: Programs for machine learning*. Elsevier, 2014.
- [45] A. Günel, "Regresyon denkleminin başarısını ölçmede kullanılan belirleme katsayısı ve kritiği," *Doğuş Üniversitesi Dergisi*, vol. 4, no. 2, pp. 133-140, 2003. Doi: 10.31671/dogus.2019.313
- [46] D. J. Ozer, "Correlation and the coefficient of determination," *Psychological Nulletin*, vol. 97, no. 2, p. 307, 1985. Doi: 10.1037/0033-2909.97.2.307
- [47] N. J. Nagelkerke, "A note on a general definition of the coefficient of determination," *Biometrika*, vol. 78, no. 3, pp. 691-692, 1991. Doi: 10.1093/biomet/78.3.691
- [48] O. Renaud and M.-P. Victoria-Feser, "A robust coefficient of determination for regression," *Journal of Statistical Planning and Inference*, vol. 140, no. 7, pp. 1852-1862, 2010. Doi: 10.1016/j.jspi.2010.01.008
- [49] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79-82, 2005. Doi: 10.3354/cr030079
- [50] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)," *Geoscientific Model Development Discussions*, vol. 7, no. 1, pp. 1525-1534, 2014. Doi: 10.5194/gmdd-7-1525-2014
- [51] M. Şahan, "Yapay sinir ağları ve Angström-Prescott denklemleri kullanılarak Gaziantep, Antakya ve Kahramanmaraş için global güneş radyasyonu tahmini," *Süleyman Demirel Üniversitesi Fen Edebiyat Fakültesi Fen Dergisi*, vol. 16, no. 2, pp. 368-384, 2021. Doi: 10.29233/sdufeff.953182
- [52] M. Kayakuş, "Yazılım çaba tahmininde yapay sinir ağları için optimum yapının belirlenmesi," *Avrupa Bilim ve Teknoloji Dergisi*, no. 22, pp. 43-48, 2021. Doi: 10.31590/ejosat.847712
- [53] A. De Myttenaere, B. Golden, B. Le Grand, and F. Rossi, "Mean absolute percentage error for regression models," *Neurocomputing*, vol. 192, pp. 38-48, 2016. Doi: 10.1016/j.neucom.2015.12.114
- [54] F. Liantoni and A. Agusti, "Forecasting bitcoin using double exponential smoothing method based on mean absolute percentage error," *JOIV: International Journal on Informatics Visualization*, vol. 4, no. 2, pp. 91-95, 2020. Doi: 10.30630/joiv.4.2.335
- [55] J. McKenzie, "Mean absolute percentage error and bias in economic forecasting," *Economics Letters*, vol. 113, no. 3, pp. 259-262, 2011. Doi: 10.1016/j.econlet.2011.08.010
- [56] E. Karadağ, "Eğitim bilimleri doktora tezlerinde kullanılan araştırma modelleri: Nitelik düzeyleri ve analitik hata tipleri," *Kuram ve Uygulamada Eğitim Yönetimi*, vol. 1, no. 1, pp. 49-71, 2010.
- [57] A. Oğuzlar, "Veri ön işleme," *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, no. 21, pp. 67-76, 2003.
- [58] F. Arslan ve H. T. Kahraman, "Yapay zekâ tabanlı büyük veri yönetim aracı," *Journal of Investigations on Engineering and Technology*, vol. 2, no. 1, pp. 8-21, 2019.
- [59] P. J. M. Ali, R. H. Faraj, E. Koya, P. J. M. Ali, and R. H. Faraj, "Data normalization and standardization: a technical report," *Mach Learn Tech Rep*, vol. 1, no. 1, pp. 1-6, 2014.
- [60] Y. Gültepe, "Makine öğrenmesi algoritmaları ile hava kirliliği tahmini üzerine karşılaştırmalı bir değerlendirme," *Avrupa Bilim ve Teknoloji Dergisi*, no. 16, pp. 8-15, 2019. Doi: 10.31590/ejosat.530347
- [61] S. Yavuz ve M. Deveci, "İstatistiksel normalizasyon tekniklerinin yapay sinir ağı performansına etkisi," *Erciyes Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, no. 40, pp. 167-187, 2012.
- [62] D. Singh and B. Singh, "Investigating the impact of data normalization on classification performance," *Applied Soft Computing*, vol. 97, p. 105524, 2020. Doi: 10.1016/j.asoc.2019.105524
- [63] L. Munkhdalai, T. Munkhdalai, K. H. Park, H. G. Lee, M. Li, and K. H. Ryu, "Mixture of activation functions with extended min-max normalization for forex market prediction," *IEEE Access*, vol. 7, pp. 183680-183691, 2019. Doi: 10.1109/ACCESS.2019.2959789
- [64] D. Borkin, A. Némethová, G. Michalčonok, and K. Maiorov, "Impact of data normalization on classification model accuracy," *Research Papers Faculty of Materials Science and Technology Slovak University of Technology*, vol. 27, no. 45, pp. 79-84, 2019. Doi:

10.2478/rput-2019-0029

[65] M. F. Keskenler, D. Deniz, ve T. Aydın, "Yapay zekâ destekli ÇOKS yöntemi ile kredi kartı sahtekarlığının tespiti," *El-Cezeri*, vol. 8, no. 2, pp. 1007-1023, 2021. Doi: 10.31202/ecjse.908260

[66] M. K. Uçar, M. Nour, H. Sindi, ve K. Polat, "The effect of training and testing process on machine learning in biomedical datasets," *Mathematical Problems in Engineering*, vol. 2020, 2020. Doi: 10.1155/2020/2836236

[67] F. Aslay ve Ö. Üstün, "Meteorolojik parametreler kullanılarak yapay sinir ağları ile toprak sıcaklığının tahmini," *Politeknik Dergisi*, vol. 16, no. 4, pp. 139-145, 2013.

* This paper was presented at the 5th International Conference on Artificial Intelligence and Applied Mathematics in Engineering (ICAIAE 2023) and the abstract was published as an e-book.

This is an open access article under the CC-BY license

