



Open Access Journal
e-ISSN: 2619 – 8991

Araştırma Makalesi (Research Article)

Cilt 8 - Sayı 4: 963-969 / Temmuz 2025
(Volume 8 - Issue 4: 963-969 / July 2025)

ÇEVRESEL TUTUMLARIN SINIFLANDIRILMASINDA ÖZELLİK SEÇİMİ VE MODEL PERFORMANSININ İNCELENMESİ

Rukiye UZUN ARSLAN^{1*}, İrem ŞENYER YAPICI², Fuat ALKAN¹

¹Zonguldak Bulent Ecevit University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 67100, Zonguldak, Türkiye

²Zonguldak Bulent Ecevit University, Faculty of Engineering, Department of Computer Engineering, 67100, Zonguldak, Türkiye

Özet: Çevresel tutumlar, bireylerin sürdürülebilirlik ve ekolojik bilinç düzeylerini yansıtan önemli göstergelerden biridir ve bu tutumların doğru bir şekilde tahmin edilmesi, çevresel politikaların ve farkındalık çalışmalarının etkinliğini artırmada kritik bir rol oynamaktadır. Bu doğrultuda, bu çalışmada bireylerin çevresel tutumlarını tahmin etmek amacıyla çeşitli makine öğrenmesi (MÖ) algoritmaları ve farklı özellik seçimi yöntemleri karşılaştırılmıştır. Analizlerde, Türkiye’de yaşayan 384 katılımcıdan elde edilen açık erişimli bir veri seti kullanılmıştır. Çalışma kapsamında, yedi farklı MÖ algoritması (Rasgele Orman, Lojistik Regresyon, Destek Vektör Makineleri, K-En Yakın Komşu, Çok Katmanlı Algılayıcı, Naive Bayes ve Oylama) ile dört özellik seçimi yöntemi (Ekstra Ağaçlar, Karşılıklı Bilgi, Temel Bileşen Analizi ve Rasgele Orman) değerlendirilmiştir. Elde edilen bulgular, Temel Bileşen Analizi yönteminin Çok Katmanlı Algılayıcı (ÇKA) ve Lojistik Regresyon algoritmalarının performansını anlamlı ölçüde artırdığını göstermiştir. En yüksek doğruluk oranı (%98,7) ve duyarlılık değeri (%100) ÇKA algoritması ile elde edilmiştir. Ayrıca, Karşılıklı Bilgi ve Rasgele Orman yöntemlerinin, Destek Vektör Makineleri ve Oylama algoritmalarının performansını iyileştirdiği belirlenmiştir. Sonuç olarak, bu çalışma, çevresel tutumların tahmininde kullanılan özellik seçimi yöntemleri ile MÖ algoritmalarının kombinasyonlarının etkisini ortaya koymaktadır. Açık erişimli bir veri setinin kullanılması, çalışmanın şeffaflığını ve tekrarlanabilirliğini artırmıştır. Gelecekteki çalışmalarda, daha geniş ve çeşitli veri setleri üzerinde gerçekleştirilecek analizlerle elde edilen bulguların genellenebilirliği değerlendirilebilir.

Anahtar kelimeler: Çevresel tutum tahmini, Özellik seçimi, Model performansı, Sınıflandırma algoritmaları

Feature Selection and Evaluation of Model Performance in Environmental Attitude Classification

Abstract: The environmental attitudes are one of the prominent indicators reflecting the sustainability and ecological awareness levels of individuals, and accurate prediction of these attitudes plays a critical role in increasing the effectiveness of environmental policies and awareness efforts. This study compares various machine learning (ML) algorithms and different feature selection methods to predict the environmental attitudes of individuals in this context. In the analyses, an open-access dataset obtained from 384 participants living in Türkiye has been used. Seven different ML algorithms (Random Forest, Logistic Regression, Support Vector Machines, K-Nearest Neighbours, Multi-Layer Perceptron, Naive Bayes, and Voting) and four feature selection methods (Extra Trees, Mutual Information, Principal Component Analysis, and Random Forest) have been analysed. The obtained results have indicated that the Principal Component Analysis method significantly improved the performance of the Multilayer Perceptron (MLP) and Logistic Regression algorithms. The highest accuracy rate (98.7%) and sensitivity value (100%) were obtained with the MLP algorithm. Furthermore, mutual information and random forest methods were found to improve the performance of support vector machines and voting algorithms. In conclusion, this study demonstrates the effectiveness of combinations of feature selection methods and ML algorithms in predicting environmental attitudes. The use of an open access dataset increased the transparency and reproducibility of the study. In future studies, the generalizability of the findings can be evaluated by analyses on larger and more diverse data sets.

Keywords: Environmental attitude prediction, Feature selection, Model performance, Classification algorithms

*Corresponding author: Zonguldak Bulent Ecevit University, Faculty of Engineering, Department of Electrical and Electronics Engineering, 67100, Zonguldak, Türkiye

E mail: rukiye.uzun@beun.edu.tr (R. UZUN ARSLAN)

Rukiye UZUN ARSLAN <https://orcid.org/0000-0002-2082-8695>

İrem ŞENYER YAPICI <https://orcid.org/0000-0003-0655-340X>

Fuat ALKAN <https://orcid.org/0000-0001-7914-2383>

Gönderi: 18 Mart 2025

Kabul: 29 Nisan 2025

Yayınlanma: 15 Temmuz 2025

Received: March 18, 2025

Accepted: April 29, 2025

Published: July 15, 2025

Cite as: Arslan UR, Yapıcı Şİ, Alkan F. 2025. Feature selection and evaluation of model performance in environmental attitude classification. BSJ Eng Sci, 8(4): 963-969.

1. Giriş

Çevresel tutumlar, bireylerin doğal çevreye yönelik algılarını, inançlarını ve davranışsal eğilimlerini kapsayan çok boyutlu bir kavramdır. Bu tutumlar, sürdürülebilir kalkınma hedeflerine ulaşma sürecinde ve çevre politikalarının etkinliğinin artırılmasında kritik bir rol oynamaktadır (Milfont ve Duckitt, 2010). Özellikle

iklim değişikliği, biyoçeşitlilik kaybı ve çevresel kirlilik gibi küresel ölçekteki ekolojik sorunlar, çevresel tutumların önemini giderek artırmakta ve bu alandaki araştırmalara verilen önceliğin artmasına neden olmaktadır (Calculli vd., 2021). Literatürde, çevresel tutumlar ile sürdürülebilir davranışlar arasında anlamlı bir ilişki olduğu gösterilmiştir. Olumlu çevresel tutumlar, geri dönüşüm, enerji verimliliği ve çevre dostu



politikaları destekleme gibi sürdürülebilir davranışlarla pozitif yönde ilişkilendirilirken, olumsuz çevresel tutumlar ise çevresel sorunlara karşı ilgisizlik, değişime direnç ve çevreye zarar verici davranışlarla bağlantılıdır. Bu durum, ekolojik dengenin bozulmasını hızlandırarak çevresel sürdürülebilirliği tehdit etmektedir. Dolayısıyla, çevresel tutumların analizi ve bu tutumların olumlu yönde dönüştürülmesi, sürdürülebilir kalkınma açısından büyük önem taşımaktadır (Kollmuss ve Agyeman, 2002). Çevresel tutumları değerlendirmek amacıyla genellikle bireylerin kendi beyanlarına dayanan anketler ve istatistiksel analizler kullanılmaktadır (Gifford ve Sussman, 2012). Ancak, bu yöntemler çevresel tutumları belirleyen faktörler arasındaki çok yönlü etkileşimlerini açıklamada yetersiz kalarak, analizlerin kapsamını ve derinliğini sınırlayabilmektedir. Bu nedenle, çevresel tutumların çok boyutlu ve dinamik yapısını daha iyi anlayabilmek için makine öğrenmesi (MÖ) gibi yeni analitik yaklaşımların kullanımı giderek artmaktadır. Örneğin, 2018 yılında gerçekleştirilen bir çalışmada, bireylerin iklim değişikliği ve çevresel bozulmaya yönelik tutumlarını tahmin etmede psikolojik ve demografik faktörlerin göreceli önemi MÖ temelli bir yaklaşımla incelenmiştir. Bu çalışmada, Çin, İsviçre ve Amerika Birleşik Devletleri'nde gerçekleştirilen üç farklı anket çalışmasından elde edilen veriler, rasgele orman (RO) algoritması kullanılarak analiz edilmiştir. Analizlerde bireylerin gelecekteki sonuçları dikkate alma eğiliminin, geleneksel demografik değişkenlere kıyasla iklim ve çevre tutumlarını tahmin etmede daha baskın olduğu tespit edilmiştir (Beiser-McGrath ve Huber, 2018). 2022 yılında yapılan bir çalışmada ise, Çin'deki halka açık firmaların çevre koruma ve karbon emisyonlarına yönelik tutumlarını incelemek amacıyla yenilikçi bir yaklaşım benimsenmiştir. Çalışmada, firma-yatırımcı soru-cevap kayıtları gibi metin tabanlı veriler kullanılarak, doğal dil işleme tabanlı duygu analizi ve RO modelleri gibi MÖ algoritmaları uygulanmıştır. Bu yöntemler, firmaların çevresel tutumlarını niceliksel olarak ölçmek ve zaman içindeki değişimlerini analiz etmek için kullanılmıştır (Li vd., 2022). Aynı yıl yapılan başka bir çalışmada, bireylerin ekonomik büyüme ve çevre koruma arasındaki öncelik tercihlerini tahmin etmek amacıyla makroekonomik ve bireysel düzeydeki değişkenler makine öğrenmesi yöntemleri kullanılarak analiz edilmiştir. Çalışmada 51348 katılımcıya ait veriler kullanılarak, bir dizi değişken seçim algoritması uygulanmış ve bireylerin çevresel öncelik düzeylerini tahmin etmek için RO modelleri eğitilmiştir. Yapılan analizlerde, en optimal modelin tahmin hata oranı %34,15 olarak hesaplanmış ve bireylerin çevre koruma önceliğini belirlemede makroekonomik ve psikolojik değişkenlerin önemli olduğu tespit edilmiştir (Lou vd., 2022). 2022 yılında yapılan başka bir çalışmada ise, üniversite öğrencilerinin çevreci davranışlarını öngörmek amacıyla karar ağacı (KA) temelli bir model geliştirilmiştir. Çalışmada, 334 üniversite öğrencisinden elde edilen verilerin analizi sonucunda, çevresel

sorumluluk bilinci, yenilikçi davranış ve algılanan davranışsal kontrolün çevreci davranışların en önemli belirleyicileri olarak tespit edilmiştir. Önerilen modellerle yapılan analizlerde %72,89'luk doğruluk oranlarına ulaşılmıştır (Wang vd., 2022). 2024 yılında yapılan bir çalışmada, bireylerin çevresel tutumlarının farklı MÖ algoritmalarıyla tahmin edilebilirliği irdelenmiştir. Bu kapsamda, k-en yakın komşu (KNN), KA ve destek vektör makineleri (DVM) olmak üzere üç farklı MÖ algoritmasının performansları karşılaştırılmıştır. En yüksek başarımla %95,76 doğruluk oranı ile DVM tarafından elde edilmiştir (Değirmenci, 2024). Aynı yıl yapılan bir diğer çalışmada, bireylerin 84 katılımcıdan elde edilen verilerle çevreye yönelik tutumları analiz etmek amacıyla lojistik regresyon (LR), DVM ve KA algoritmaları kullanılmıştır. Çalışmada, LR modeli %94,53 doğruluk oranı ile en yüksek başarımla sergilemiştir (Köklü ve Sulak, 2024). Yakın zamanda yapılan bir çalışmada, bireylerin çevresel farkındalık düzeylerini ve tutumlarını analiz etmek amacıyla MÖ tabanlı bir model geliştirilmiştir. Bu çalışmada, DVM, gradyan artırma (GA), çok katmanlı algılayıcı (ÇKA), kuadratik diskriminant analizi (KDA) ve torbalama (Bagging) algoritmalarının performansları karşılaştırılmıştır. Modelin doğruluğunu artırmak ve daha az sayıda özellikle yüksek performans elde etmek için varyans analizi tabanlı bir özellik seçimi yöntemi uygulanmıştır. Bu yöntem, modelin karmaşıklığını azaltarak en anlamlı özelliklerin belirlenmesini sağlamıştır. Yapılan analizler sonucunda, GA algoritması %84,4 doğruluk oranı ile en yüksek performansı göstermiştir (Yapıcı vd., 2025). Ancak, mevcut literatürde özellik seçimi süreçlerinin MÖ temelli modellerin performansları üzerindeki etkilerini kapsamlı bir şekilde inceleyen çalışmalar sınırlıdır. Çevresel tutumların tahmin edilmesinde en belirleyici değişkenlerin (özelliklerin) belirlenmesi, model performansının artırılmasında kritik bir rol oynamaktadır. Bu kapsamda, gerçekleştirilen bu çalışmada çevresel tutumların tahmininde en etkili özellikleri belirlemek amacıyla Karşılıklı Bilgi (Mutual Information, MI), Temel Bileşen Analizi (TBA), Ekstra Ağaçlar (Extra Trees, EA) ve RO yöntemleri kullanılmıştır. Daha sonra, seçilen özellikler kullanılarak yedi farklı MÖ algoritması eğitilmiş ve başarımları karşılaştırılmıştır. MÖ algoritması olarak RO, LR, DVM, KNN, çok katmanlı algılayıcı (ÇKA), Naive Bayes (NB) ve oylama (Voting) algoritmaları kullanılmıştır. Bu sayede, çevresel tutumların tahmin edilmesine yönelik daha etkili bir metodolojik çerçeve sunulması ve literatürdeki boşluğun doldurulması hedeflenmektedir

2. Materyal ve Yöntem

Bu çalışmada, bireylerin çevresel tutumlarını analiz etmek amacıyla oluşturulmuş bir veri seti kullanılmıştır. Türkiye'de yaşayan ve yaşları 18 ile 45 arasında değişen toplam 384 katılımcıdan çevrimiçi anket yöntemiyle elde edilen veri seti, bireylerin çevresel sorumluluk davranışlarını ölçmeye yönelik 37 değişken içermektedir.

Katılımcıların cinsiyet dağılımı incelendiğinde, örneklemin 177 erkek ve 207 kadın bireyden oluştuğu tespit edilmiştir. Veri setinde demografik ve sosyoekonomik faktörler (cinsiyet, yaş, eğitim düzeyi, gelir durumu) ile birlikte çevre eğitimi geçmişi (çevre eğitimi alıp almadıkları, küresel ısınma ve iklim değişikliği dersleri, çevresel okuryazarlık kursları, okul öncesi eğitim) gibi değişkenler yer almaktadır. Ayrıca, çevresel konulara yönelik tutumları ölçmeye yönelik 27 soru içeren anket kapsamında, kirlilik, kaynak tüketimi, geri dönüşüm, enerji kullanımı, çevresel eğitim ve aktivizm gibi konulara dair görüşler toplanmıştır. Çalışmada hedef (çıktı) değişken olarak bireylerin çevresel tutumları analiz edilerek, katılımcılar pozitif veya negatif tutum sergileyen bireyler olarak iki ayrı

sınıfa ayrılmıştır. Veri setine

https://nigmetkoklu.com/dataset/environmental_attitude_dataset.zip adresi üzerinden erişim sağlanabilmektedir (Köklü ve Sulak, 2024).

İlgili veri seti, bireylerin çevresel tutumlarını sınıflandırmaya yönelik MÖ modellerinin geliştirilmesinde kullanılmıştır. Bu kapsamda, veri ön işleme aşamasında eksik değerlerin kontrolü sağlanmış, kategorik değişkenler uygun şekilde dönüştürülmüş ve özellik seçimi gerçekleştirilmiştir. Ardından, farklı MÖ algoritmaları eğitilerek performansları karşılaştırılmış ve en başarılı model belirlenmiştir. Şekil 1'de, veri ön işleme, model eğitimi ve değerlendirme süreçlerini özetleyen sözde kod sunulmaktadır.

```

Adım 1: Veri Kümesini Yükleme ve Ön İşleme
Girdi: Veri Kümesi
Çıktı: Ön işlenmiş veri kümesi (X_train, X_test, y_train, y_test)
    1.1 Veri kümesini yükle
    1.2 Özellik seçimi uygula (TBA, MI, EA, RO vb.)
    1.3 Veri kümesini %80 eğitim ve %20 test olarak ayır
    1.4 Eksik verileri işle (gerekliyse)
    1.5 Sayısal özellikleri normalize et veya standartlaştır (gerekliyse)
    1.6 Kategorik değişkenleri kodla (varsa)

Adım 2: MÖ Modellerinin Eğitimi ve Değerlendirilmesi
Girdi: Ön işlenmiş veri kümesi (X_train, y_train)
Çıktı: Eğitilmiş modeller ve performans metrikleri (modeller, metrikler)
    2.1 Model listesi tanımla = {DVM, MLP, KA, LR, NB, RO, Voting}
    2.2 Model başarı metriklerini saklamak için boş bir sözlük oluştur
    (model_metrikleri)
    Her model için:
    2.3 Modeli başlat
    2.4 k-kat çapraz doğrulama (k=5) kullanarak modeli eğit
    2.5 Test verisi (X_test) üzerinde tahmin yap
    2.6 Performans metriklerini hesapla (doğruluk, kesinlik, duyarlılık, F1-skoru)
    2.7 Modelin adını ve performans metriklerini model_metrikleri sözlüğüne kaydet

Adım 3: En İyi Modelin Seçilmesi
Girdi: model_metrikleri sözlüğü
Çıktı: En iyi model (en_ iyi_model)
    3.1 En yüksek performans metriğine sahip modeli belirle
    3.2 en_ iyi_model değişkenini en yüksek metrik değerine sahip model olarak ata

Adım 4: Sonuçların Raporlanması
Girdi: en_ iyi_model ve model_metrikleri
Çıktı: Sonuç tablosu ve en iyi model bilgileri
    4.1 Tüm modellerin başarı metriklerini tablo halinde göster
    4.2 En iyi modelin adını ve metriklerini raporla
    
```

Şekil 1. Çevresel tutumların tahminine yönelik makine öğrenmesi sürecini özetleyen sözde kod.

Veri ön işleme aşamasında, eksik verilerin olmadığı tespit edilmiş olmasına rağmen, diğer ön işleme adımları titizlikle uygulanmıştır. Sayısal değişkenler arasındaki ölçek farklılıklarını gidermek için StandardScaler kullanılarak normalizasyon işlemi gerçekleştirilmiştir. Bu adım, değişkenlerin aynı ölçekte olmasını sağlayarak

modelin daha dengeli ve etkili bir şekilde öğrenmesine katkıda bulunmuştur. Ayrıca, özellik seçimi sürecinde değişkenlerin sınıflandırma başarısına etkisini değerlendirmek amacıyla EA, MI, TBA ve RO gibi yöntemler kullanılmıştır. Bu yöntemler sayesinde en yüksek bilgi değerine sahip değişkenler belirlenmiş ve

analiz sürecine dahil edilmiştir. Bu adım, hem modelin hesaplama maliyetini azaltmış hem de sınıflandırma performansını artırmayı hedeflemiştir.

Çalışmada, çevresel tutumların tahmin edilmesi amacıyla yedi farklı MÖ algoritması (RO, LR, DVM, MLP, NB, KNN ve Oylama) kullanılmıştır. Bunlardan RO, çok sayıda karar ağacının bir araya gelmesiyle oluşan bir topluluk öğrenme yöntemi olup, yüksek doğruluk oranları ve aşırı öğrenmeye karşı dirençli yapısıyla dikkat çekmektedir. LR, özellikle ikili sınıflandırma problemlerinde kullanılan, doğrusal bir model olup, sonuçların yorumlanabilirliği açısından avantaj sağlamaktadır (Uzun et al. 2019). DVM, veriyi yüksek boyutlu bir özellik uzayında ayıran optimal hiperdüzlemi bulmaya dayalı bir yöntem olup, özellikle küçük ve orta ölçekli veri setlerinde etkilidir (Arslan ve Yapici, 2024). MLP, yapay sinir ağlarının bir türü olup, doğrusal olmayan veri yapılarını modellemek için kullanılmaktadır (Arslan vd., 2020). NB, Bayes teoremine dayalı olasılıksal bir sınıflandırma yöntemi olup, özellikle yüksek boyutlu veri setlerinde hızlı ve etkili sonuçlar vermektedir. KNN, bir veri noktasının sınıfını veya

değerini, etrafındaki en yakın k komşusunun özelliklerine dayanarak tahmin eden, öğrenme aşaması olmayan (lazy learning) ve mesafe temelli çalışan bir MÖ yöntemidir (Chethana, 2021). Oylama algoritması, birden fazla temel modelin tahminlerini birleştirerek daha güçlü bir topluluk modeli oluşturan bir yöntemdir (Kumari vd., 2021).

Model eğitimi sürecinde, veri seti %80 eğitim ve %20 test olarak ayrılmıştır. Ayrıca, modelin genelleme yeteneğini artırmak ve aşırı öğrenmeyi önlemek amacıyla 5 katlı çapraz doğrulama yöntemi uygulanmıştır. Bu yöntem, modellerin farklı veri alt kümeleri üzerinde test edilmesini sağlayarak daha güvenilir sonuçlar elde edilmesine olanak tanımıştır.

Model başarımlarını değerlendirmek için doğruluk, hassasiyet, duyarlılık ve F1-Skoru gibi performans metrikleri kullanılmıştır. Bu metrikler, modellerin çevresel tutumları ne kadar doğru ve etkili bir şekilde sınıflandırdığını ölçmek amacıyla tercih edilmiştir. İlgili metriklere ait denklemler Tablo 1'de verilmiştir.

Tablo 1. Sınıflandırma metrikleri ve formülleri

Metrik	Tanımı	Matematiksel Formülü
Doğruluk	Modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranı.	$\frac{TP + TN}{TP + TN + FP + FN}$
Duyarlılık	Pozitif olarak tahmin edilen örneklerin gerçekten pozitif olma oranı.	$\frac{TP}{TP + FN}$
Hassasiyet	Gerçekte pozitif olan örneklerin doğru tahmin edilme oranı.	$\frac{TP}{TP + FP}$
F1-skor	Duyarlılık ve hassasiyet metriklerinin harmonik ortalaması	$2 * \frac{\text{Duyarlılık} * \text{Hassasiyet}}{\text{Duyarlılık} + \text{Hassasiyet}}$

3. Bulgular

Bireylerin çevresel tutumlarının MÖ algoritmaları ile tahmin edilmesine yönelik bu çalışmada, model başarımını artırmak ve en etkili değişkenleri belirlemek amacıyla dört farklı özellik seçimi yöntemi uygulanmıştır. Bu kapsamda, EA, MI, TBA ve RO yöntemleriyle seçilen özellikler (değişkenler) yedi farklı MÖ algoritması ile değerlendirilmiş ve algoritmaların performansları karşılaştırılmıştır. Seçilen değişkenlerin modeller üzerindeki etkisini daha güvenilir bir şekilde analiz edebilmek için veri seti %80 eğitim ve %20 test olarak bölünmüş ve 5 katlı çapraz doğrulama yöntemi uygulanmıştır. Ardından, belirlenen değişkenler kullanılarak eğitilen modellerin performansları doğruluk, hassasiyet, duyarlılık ve F1-skoru metrikleri üzerinden analiz edilmiştir.

Bu doğrultuda Tablo 2, EA yöntemiyle belirlenen özelliklerin kullanılması durumunda MÖ algoritmalarının performans metriklerini göstermektedir. Sonuçlara göre, DVM algoritması doğruluk (%85,7), hassasiyet (%86,5), duyarlılık (%84,2) ve F1-skoru (%85,3) metriklerinde en yüksek performansı göstererek diğer algoritmalar

arasında öne çıkmıştır. KNN, NB ve Voting algoritmaları, hassasiyet metriklerinde yüksek performans (%88,2) göstermelerine rağmen, duyarlılık (%78,9) ve F1-skoru (%83,3) metriklerinde daha düşük performans sergilemişlerdir. RO algoritması ise tüm metriklerde en düşük performansı göstermiştir (doğruluk: %80,5, hassasiyet: %79,5, duyarlılık: %81,6, F1-skoru: %80,5). LR ve ÇKA algoritmaları ise tüm metriklerde benzer performans sergilemişlerdir (doğruluk: %81,8, hassasiyet: %83,3, duyarlılık: %78,9, F1-skoru: %81,1). Bu sonuçlar, EA yöntemiyle seçilen özelliklerin kullanılması durumunda DVM algoritmasının diğer algoritmalara göre daha etkili olduğunu göstermektedir. Tablo 3, MI yöntemiyle seçilen özelliklerin kullanılması durumunda elde edilen performans değerlerini göstermektedir. DVM ve Voting algoritmalarının doğruluk (%92,2), hassasiyet (%92,1), duyarlılık (%92,1) ve F1-skoru (%92,1) metriklerinde en yüksek performansı sergileyerek diğer algoritmalar arasında öne çıkmıştır. LR algoritması, hassasiyet metriğinde en yüksek değeri (%96,9) elde etmesine karşın, duyarlılık (%81,6) ve F1-skoru (%88,6) açısından daha düşük

performans göstermiştir. RO algoritması ise doğruluk (%88,3), hassasiyet (%89,2), duyarlılık (%86,8) ve F1-skoru (%88,0) değerleriyle dengeli bir performans sergilemiştir. KNN algoritması, duyarlılık metriğinde (%86,8) görece yüksek bir başarı göstermesine rağmen, hassasiyet (%82,5) ve F1-skoru (%84,6) bakımından nispeten daha düşük sonuçlar elde etmiştir. NB algoritması, hassasiyet metriğinde (%91,2) yüksek bir değer sunmasına rağmen, duyarlılık (%81,6) ve F1-skoru

(%86,1) açısından daha düşük performans sergilemiştir. ÇKA algoritması ise doğruluk (%88,3), hassasiyet (%87,2), duyarlılık (%89,5) ve F1-skoru (%88,3) değerleriyle dengeli bir performans ortaya koymuştur. Bu sonuçlar, MI yöntemiyle seçilen özelliklerin kullanımıyla DVM ve Voting algoritmalarının diğer algoritmalarla kıyasla daha etkili olduğunu ortaya koymaktadır.

Tablo 2. EA yöntemiyle seçilen özelliklerin kullanılması durumunda MÖ algoritmalarına ait performans metrikleri

Sınıflandırıcı	Doğruluk	Hassasiyet	Duyarlılık	F1- skor
LR	81,8%	83,3%	78,9%	81,1%
RO	80,5%	79,5%	81,6%	80,5%
DVM	85,7%	86,5%	84,2%	85,3%
K-NN	84,4%	88,2%	78,9%	83,3%
NB	84,4%	88,2%	78,9%	83,3%
ÇKA	81,8%	83,3%	78,9%	81,1%
Voting	84,4%	88,2%	78,9%	83,3%

Tablo 3. MI yöntemiyle seçilen özelliklerin kullanılması durumunda MÖ algoritmalarına ait performans metrikleri

Sınıflandırıcı	Doğruluk	Hassasiyet	Duyarlılık	F1- skor
LR	89,6%	96,9%	81,6%	88,6%
RO	88,3%	89,2%	86,8%	88,0%
DVM	92,2%	92,1%	92,1%	92,1%
K-NN	84,4%	82,5%	86,8%	84,6%
NB	87,0%	91,2%	81,6%	86,1%
ÇKA	88,3%	87,2%	89,5%	88,3%
Voting	92,2%	92,1%	92,1%	92,1%

Tablo 4, TBA yöntemiyle seçilen özelliklerin kullanılması durumunda elde edilen performans metriklerini göstermektedir. ÇKA algoritması, %98,7 doğruluk, %97,4 kesinlik ve %100 duyarlılık ile en yüksek performansı gösterirken, LR algoritması da %97,4 doğruluk ve F-skör ile benzer bir başarı elde etmiştir. RO ve DVM algoritmaları sırasıyla %93,5 ve %92,2 doğruluk oranlarıyla tatmin edici sonuçlar verirken, KNN ve NB

algoritmaları da %90'ın üzerinde doğruluk oranlarıyla performanslarını sergilemiştir. Voting algoritması ise %94,8 doğruluk ve %94,9 F-skör ile dengeli bir performans göstermiştir. Bu sonuçlar, TBA yöntemiyle seçilen özelliklerin, özellikle ÇKA ve LR gibi algoritmalar için etkili olduğunu ve özellik seçiminin model performansını önemli ölçüde artırabileceğini ortaya koymaktadır.

Tablo 4. TBA yöntemiyle seçilen özelliklerin kullanılması durumunda MÖ algoritmalarına ait performans metrikleri

Sınıflandırıcı	Doğruluk	Hassasiyet	Duyarlılık	F1- skor
LR	97,4%	95,0%	100,0%	97,4%
RO	93,5%	90,2%	97,4%	93,7%
DVM	92,2%	88,1%	97,4%	92,5%
K-NN	92,2%	92,1%	92,1%	92,1%
NB	90,9%	86,0%	97,4%	91,4%
ÇKA	98,7%	97,4%	100,0%	98,7%
Voting	94,8%	92,5%	97,4%	94,9%

Son olarak, Tablo 5 RO yöntemiyle belirlenen özelliklerin kullanılması durumunda elde edilen sonuçları göstermektedir. ÇKA algoritması, %98,7 doğruluk, %97,4 kesinlik ve %100 duyarlılık ile en yüksek performansı gösterirken, LR algoritması da %96,1 doğruluk ve F-skör

ile yakın bir başarı sergilemiştir. DVM algoritması %92,2 doğruluk ve %92,5 F-skör ile tatmin edici sonuçlar verirken, RO ve KNN algoritmaları sırasıyla %89,6 ve %89,6 doğruluk oranlarıyla daha düşük performans göstermiştir. NB algoritması ise %85,7 doğruluk ve

%84,9 F-skor ile en düşük performansa sahiptir. Voting algoritması, %96,1 doğruluk ve %96,2 F-skor ile LR ve ÇKA'ya yakın bir performans sergilemiştir. Bu sonuçlar, RO yöntemiyle belirlenen özelliklerin özellikle ÇKA ve LR algoritmaları için avantaj sağladığını ortaya koymaktadır. Genel olarak, uygulanan dört farklı özellik seçimi yönteminin model performansı üzerindeki etkisi

değişkenlik göstermiştir. MI ve RO yöntemleri, DVM ve Voting algoritmalarının performansını artırırken, TBA yöntemi özellikle ÇKA algoritmasının başarımını yükseltmiştir. EA yöntemi ile ise DVM algoritması en iyi sonucu vermiştir. Bu sonuçlar, model performansının kullanılan özellik seçimi yöntemine ve algoritmaya bağlı olarak değiştiğini göstermektedir.

Tablo 5. RO yöntemiyle seçilen özelliklerin kullanılması durumunda MÖ algoritmalarına ait performans metrikleri

Sınıflandırıcı	Doğruluk	Hassasiyet	Duyarlılık	F1- skor
LR	96,1%	94,9%	97,4%	96,1%
RO	89,6%	87,5%	92,1%	89,7%
DVM	92,2%	88,1%	97,4%	92,5%
K-NN	89,6%	94,1%	84,2%	88,9%
NB	85,7%	88,6%	81,6%	84,9%
ÇKA	98,7%	97,4%	100,0%	98,7%
Voting	96,1%	92,7%	100,0%	96,2%

4. Tartışma

Bu çalışmada, bireylerin çevresel tutumlarının tahmin edilmesi amacıyla farklı MÖ algoritmaları ve özellik seçimi yöntemlerinin performansları kapsamlı bir şekilde değerlendirilmiştir. Türkiye'de yaşayan 384 katılımcıdan oluşan açık erişimli bir veri seti kullanılarak yürütülen analizler, çevresel tutum tahmininde özellik seçimi yöntemlerinin ve MÖ algoritmalarının etkinliğini karşılaştırmayı hedeflemiştir. Çalışma kapsamında dört farklı özellik seçimi yöntemi (MI, TBA, EA ve RO) uygulanmış ve seçilen öznitelikler yedi farklı MÖ algoritmasının (RO, LR, DVM, KNN, ÇKA, NB ve Oylama) performansları değerlendirilmiştir.

Elde edilen bulgular, hem özellik seçimi yöntemlerinin hem de MÖ algoritmalarının performansında belirgin farklılıklar olduğunu ortaya koymuştur. EA yöntemiyle seçilen özniteliklerin kullanıldığı durumda DVM algoritması %85,7 doğruluk ve %85,3 F1-skoru ile en yüksek performansı sergilemiştir. Buna karşılık, MI yöntemiyle seçilen öznitelikler kullanıldığında DVM ve Oylama algoritmaları %92,2 doğruluk ve F1-skoru ile daha üstün bir başarı göstermiştir. TBA yönteminin ise özellikle ÇKA ve LR algoritmalarının performansını önemli ölçüde iyileştirdiği gözlemlenmiştir. Bu bağlamda, ÇKA algoritması %98,7 doğruluk ve %100 duyarlılık oranı ile dikkat çekici bir performans sergilemiştir. Benzer şekilde, RO yöntemiyle seçilen öznitelikler de ÇKA ve LR algoritmalarının başarısını artırmıştır.

Bu sonuçlar, çevresel tutum tahmini gibi karmaşık bir problemi çözmek için özellik seçimi yöntemlerinin ve MÖ algoritmalarının kombinasyonunun kritik bir rol oynadığını göstermektedir. Özellikle TBA yöntemi, ÇKA ve LR gibi algoritmalar için en etkili sonuçları sağlarken, MI ve RO yöntemleri DVM ve Oylama algoritmalarının performansını artırmıştır. Bu bulgular, çevresel tutumların analizinde kullanılan metodolojik yaklaşımların, model performansını önemli ölçüde etkileyebileceğini göstermektedir.

Katkı Oranı Beyanı

Yazarların katkı yüzdeleri aşağıda verilmiştir. Yazarlar makaleyi incelemiş ve onaylamıştır.

	R.U.A.	İ.S.Y.	F.A.
K	40	40	20
T	30	40	30
Y	60	20	20
VTI	20	50	30
VAY	40	40	20
KT	20	40	40
YZ	30	30	40
KI	40	30	30
GR	60	20	20

K= kavram, T= tasarım, Y= yönetim, VTI= veri toplama ve/veya işleme, VAY= veri analizi ve/veya yorumlama, KT= kaynak tarama, YZ= Yazım, KI= kritik inceleme, GR= gönderim ve revizyon, PY= proje yönetimi, FA= fon alımı.

Çalışma Beyanı

Yazarlar bu çalışmada hiçbir çıkar ilişkisi olmadığını beyan etmektedirler.

Etik Onay Beyanı

Bu çalışmada hayvanlar ve insanlar üzerinde herhangi bir çalışma yapılmadığı için etik kurul onayı alınmamıştır.

Kaynaklar

- Arslan RU, İşler Y, Toksan M. 2020. Prediction of the success of wart treatment methods. Karaelmas Sci Eng J, 10(1): pp: 44-52.
- Arslan RU, Yapıcı İŞ. 2024. Farklı örnekleme tekniklerine ve farklı sınıflandırıcılara dayanarak kalp yetmezliği tanılı hastaların sağkalımlarının incelenmesi. EMO Bilimsel Dergi, 14(2): pp: 35-47.

- Beiser-McGrath LF, Huber RA. 2018. Assessing the relative importance of psychological and demographic factors for predicting climate and environmental attitudes. *Clim Change*, 149: pp: 335-347.
- Calculi C, D'Uggento AM, Labarile A, Ribecco N. 2021. Evaluating people's awareness about climate changes and environmental issues: A case study. *J Clean Prod*, 324: 129244.
- Chethana C. 2021. Prediction of heart disease using different KNN classifier. In: *Proceedings of the 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, May 2021, IEEE, pp: 1186-1194.
- Değirmenci A. 2024. Comparison of K-Nearest Neighbors, Decision Tree and Support Vector Machines Methods in Predicting Environmental Attitudes. *Emerg Trends Electr Electron Eng*, 4: 69-78.
- Gifford R, Sussman R. 2012. Environmental attitudes. In: Clayton SD, ed. *The Oxford Handbook of Environmental and Conservation Psychology*. Oxford University Press, Oxford, UK, 1st ed., pp: 65-80. <https://doi.org/10.1093/oxfordhb/9780199733026.013.0004>.
- Koklu N, Sulak SA. 2024. Classification of environmental attitudes with artificial intelligence algorithms. *Intell Methods Eng Sci*, 3(2): pp: 54-62.
- Kollmuss A, Agyeman J. 2002. Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior?. *Environ Educ Res*, 8(3): pp: 239-260.
- Kumari S, Kumar D, Mittal M. 2021. An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *Int J Cogn Comput Eng*, 2: pp: 40-46.
- Li C, Li L, Zheng J, Wang J, Yuan Y, Lv Z, Liu W. 2022. China's public firms' attitudes towards environmental protection based on sentiment analysis and random forest models. *Sustainability*, 14(9): pp: 5046.
- Lou X, Lin Y, Li LMW. 2022. Predicting priority of environmental protection over economic growth using macroeconomic and individual-level predictors: Evidence from machine learning. *J Environ Psychol*, 82: 101843.
- Milfont TL, Duckitt J. 2010. The environmental attitudes inventory: A valid and reliable measure to assess the structure of environmental attitudes. *J Environ Psychol*, 30(1): pp: 80-94. <https://doi.org/10.1016/j.jenvp.2009.09.001>.
- Uzun R, İşler Y, Toksan M. 2019. WEKA yazılım paketinin siğil tedavi yöntemlerinin başarısının tahmininde kullanımı. *Düzce Univ Bilim Teknol Derg*, 7(1): pp: 699-708.
- Wang Q, Kou Z, Sun X, Wang S, Wang X, Jing H, Lin P. 2022. Predictive analysis of the pro-environmental behaviour of college students using a decision-tree model. *Int J Environ Res Public Health*, 19(15): pp: 9407.
- Yapıcı İŞ, Arslan RU, Alkan F. 2023. Makine öğrenmesi tabanlı çevresel tutum sınıflandırmasında varyans analizi yaklaşımı. In: *Proceedings of the 2nd International Colosseum Scientific Researches and Innovation Congress*, February 22-23, Italy, pp: 6-16.