

Makine Öğrenmesi ile Okyanus Asidifikasyonu Tahmini

Yunus Emre ÜNGÖR^{1*} 

¹ Savunma Teknolojileri, Lisansüstü Eğitim Enstitüsü, Sivas Bilim ve Teknoloji Üniversitesi,
Türkiye.yunuseu@gmail.com

ÖZ

Günümüzün önemli çevresel sorunlarından biri olan okyanus asidifikasyonu, özellikle deniz ekosistemleri ve deniz ürünlerine dayalı ekonomiler üzerinde ciddi tehditler oluşturmaktadır. Bu çalışma, okyanus asidifikasyonunu ele alarak makine öğrenmesi yöntemleriyle bir tahmin modeli geliştirmeyi amaçlamaktadır. Geliştirilen modelin, karar destek sistemlerinde kullanılarak asidifikasyonun etkilerini önceden belirlemeye ve ilgili önlemleri planlamaya katkı sağlaması hedeflenmektedir. Çalışmada küresel okyanus asidifikasyonu eğilimleri veri seti kullanılmış ve 13 öznitelik içeren bu veri setinde okyanus asidifikasyonu (pH) ile en yüksek korelasyona sahip özniteliklerin sırasıyla fosil yakıtlardan kaynaklanan küresel ısınma katkısı, gaz ve yıl olduğu tespit edilmiştir. Eksik veriler, merkezi eğilim ölçüleri ile doldurularak ön işleme süreci tamamlanmıştır. Model geliştirme aşamasında 11 farklı makine öğrenmesi algoritması test edilmiş ve sonuçlar karşılaştırılmıştır. Ekstra Ağaçlar algoritması, en düşük MSE (3.273×10^{-6}) ve MAE (1.055×10^{-3}) değerlerine sahip olup, R^2 skoru (0.99696) ile en yüksek doğruluk oranını göstermiştir. Karar Ağacı, Gradyan Artırma ve Rastgele Ormanlar gibi diğer topluluk yöntemleri de yüksek performans sergilemiştir. Çalışmanın sonuçları, literatürdeki benzer araştırmalarla uyumludur ve makine öğrenmesi yöntemlerinin okyanus asidifikasyonunu tahmin etmede etkili bir araç olabileceğini göstermektedir. Gelecekteki çalışmalar için, veri setinin örnek sayısının artırılmasıyla modelin genellenebilirliği güçlendirilebilir ve derin öğrenme teknikleri uygulanarak daha kapsamlı tahmin modelleri geliştirilebilir. Ayrıca, okyanus asidifikasyonuna yönelik farkındalık çalışmalarına ağırlık verilmesi gerektiği önerilmektedir.

Anahtar Kelimeler: Okyanus asidifikasyonu, makine öğrenmesi, çevre kirliliği

* Sorumlu yazarın e-postası: yunuseu@gmail.com

Cite as: Üngör, Y.E. (2025). Makine Öğrenmesi ile Okyanus Asidifikasyonu Tahmini, *Journal of Marine and Engineering Technology*, 5(2), 110-126. <https://doi.org/10.58771/joinmet.1759683>

Prediction of Ocean Acidification Using Machine Learning

ABSTRACT

One of today's significant environmental issues, ocean acidification poses a serious threat, particularly to marine ecosystems and economies that rely on seafood. This study aims to address ocean acidification by developing a prediction model using machine learning methods. The developed model is intended to be utilized in decision support systems to predict the impacts of acidification in advance and assist in planning relevant preventive measures. The study uses the Global Ocean Acidification Trends and Impacts dataset, which contains 13 features. Among these, the features with the highest correlation with ocean acidification (pH) were identified as Global Warming Contribution from Fossil Fuels, Gas, and Year, respectively. Missing data were handled through central tendency measures, completing the preprocessing phase. During the model development stage, 11 different machine learning algorithms were tested and compared. The ExtraTrees algorithm achieved the best performance, with the lowest MSE (3.273×10^{-6}) and MAE (1.055×10^{-3}), and the highest R^2 score (0.99696). Other ensemble methods, such as Decision Tree, Gradient Boosting, and Random Forest, also demonstrated high performance. The results of the study are consistent with similar research in the literature and indicate that machine learning methods can be effective tools for predicting ocean acidification. For future studies, increasing the sample size of the dataset could enhance the model's generalizability, and deep learning techniques could be applied to develop more comprehensive prediction models. Furthermore, it is recommended that efforts to raise awareness about ocean acidification be emphasized.

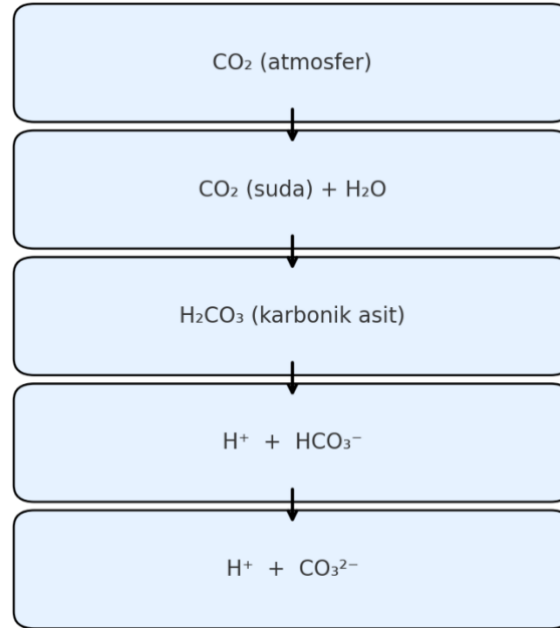
Keywords: Ocean acidification, machine learning, environmental pollution

1 Giriş

Okyanus kıtalar arası kalan ana deniz, umman olarak tanımlanmaktadır (Türk Dil Kurumu [TDK], 2025). Dünyada pasifik okyanusu, Atlantik okyanusu, Hint okyanusu, Güney okyanusu ve Arktik okyanusu olmak üzere beş tane okyanus bulunmakta ve dünyanın %70'ini başta okyanuslar olmak üzere su kütleleri oluşturmaktadır (Taner, 2015). Okyanuslarda kirlenme, ısınma ve asitlenme gibi sorunlar giderek hızla artmaktadır (Taner, 2015). Okyanus asidifikasyonu “küresel ısınmanın kötü ikizi” olarak tabir edilmektedir (Tayyar, 2023). Sanayileşmenin başlamasıyla beraber beşerî faaliyetlere bağlı olarak karbondioksit salınımı artmış bu artış okyanus asidifikasyonunu tetiklemiştir (Tayyar, 2023). Okyanus asidifikasyonu insan kaynaklı oluşan karbondioksit emisyonlarının bir sonucu olarak karşımıza çıkmaktadır (Krasting vd., 2022). Havada bulunan karbondioksit gazı okyanus gibi su kütleleriyle etkileşimde bulunmakta, daha çok reaktif hale gelmekte ve bunun nihayetinde de okyanus asitlenmesi oluşmaktadır (Tayyar, 2023). Havadaki fazla karbondioksiti okyanuslar absorbe ederek iklim değişikliğinin etkilerini azaltırlar ancak bu durum okyanus suyunun kimyasının değişmesine neden olur (Gattuso vd., 2014). Karbondioksit yoğunluk seviyesinin sürekli artması okyanus ve denizlerdeki asit oranını artırmaktadır (Tayyar, 2023). Okyanuslarda karbondioksit konsantrasyonunun artması okyanusları asidik sular haline getirmektedir (Taner, 2015). Okyanuslarda asitleşme oranının yükselmesi okyanuslardaki karbonat oranını azaltmakta ve bu azalış pek çok deniz canlısını negatif olarak etkilemektedir (Tayyar, 2023). Kalsifiye canlılar iskelet ve kabuklarında karbonat iyonlarına gereksinim duyduklarından okyanus asitlenmesinden en fazla etkilenen canlı gruplarından (Özhan vd., 2022). Küresel ısınma dünyadaki okyanus, deniz vb. su alanlarında fizyokimyasal değişimlere sebep olmakta bu değişim sucul canlıların yaşam şartlarını olumsuz etkilemektedir (Atar & Kızılgök, 2018). Okyanus asidifikasyonu doğal türlerin üreme süreçlerini ve larva gelişimlerini olumsuz etkileyerek popülasyonlarının azalmasına yol açarken, sel, kuraklık ve kasırga gibi iklim değişikliği kaynaklı aşırı hava olayları istilacı türlerin yeni yaşam alanları aramasına neden olmakta ve zayıflayan doğal türlerin savunmasız kalması istilacı türlerin bu alanları istila etmesine zemin hazırlamaktadır (Karahana, 2021).

Okyanusların en önemli problemlerinden birisi okyanus asitlenmesidir (Özhan vd., 2022). Okyanus asitlenmesi özellikle pasifik okyanusundaki ekosistemi ciddi olarak tehdit etmektedir (Vasanth vd., 2024). Okyanus asitlenmesinin bütün ekosistemi etkiler olması toplumsal ve iktisadi sorunların oluşmasına sebep olmaktadır (Tayyar, 2023). Okyanus asitlenmesi mavi ekonomileri tehdit edici hale gelmiştir (Tayyar, 2023). Deniz ürünleri, kıyı turizmi, ulaşım ve lojistiği kapsayan mavi ekonomi kavramı okyanus asitlenmesinden nasibini almaktadır (Tayyar, 2023). Okyanuslar önemli deniz ürünlerini ve su kaynaklarını bünyesinde barındırmaktadırlar (Taner, 2015). Bu durumdan balıkçılık sektörü olumsuz etkilemektedir (Atar & Kızılgök, 2018).

Hem okyanus asitlenmesi hem de iklim değişikliği Karbondioksit salınımına sıkı sıkıya bağlı durumlardır (Doney vd., 2009). Atmosferdeki karbondioksit konsantrasyonunun artmasına büyük ölçüde fosil yakıtların yoğun olarak kullanılması sebep olmaktadır (Özhan vd., 2022). Sera gazı çeşitlerinden olan karbondioksit çok tehlikeli bir gazdır (Tayyar, 2023). İklim değişikliğinin en önemli nedenlerinden biri, atmosferdeki karbondioksit (CO_2) konsantrasyonunun artmasıdır (Özhan vd., 2022). Açık okyanus yüzeyi sularındaki CO_2 sisteminde bu denli büyük değişikliklerin, Dünya'nın son 20 milyon yıllık geçmişinde muhtemelen hiç meydana gelmediği düşünülmektedir (Feely, 2004). Atmosferde günden güne yoğunlaşan sera gazları okyanusları değişime uğratmaktadır (Taner, 2015). Atmosferik karbondioksit gazı okyanusların pH seviyesini düşürmektedir (Doney vd., 2009). Okyanus suyu pH değeri, okyanus asitlenmesinin doğrudan bir göstergesi olmasının yanı sıra uzun vadede iklim değişikliğinin çevresel etkilerini anlamada kritik bir öneme sahiptir. Okyanus asitlenmesinin kimyasal oluşum süreci Şekil 1'de verilmiştir.



Şekil 1: Okyanusların Asitlenme Süreci

Şekil 1'den anlaşılacağı üzere Atmosferdeki karbondioksit, okyanus yüzeyine temas ettiğinde suda çözünerek deniz suyuyla karışır. Çözünen karbondioksit, suyla reaksiyona girerek karbonik asit H_2CO_3 oluşturur. Bu bileşik, okyanus asitlenmesinin temel bileşenlerindendir. Karbonik asit, hidrojen iyonu (H^+) ve bikarbonat (HCO_3^-) iyonuna ayrışır. Serbest hale gelen hidrojen iyonları, okyanus suyunun asidik hale gelmesine neden olarak pH seviyesini düşürür. Bikarbonat iyonları daha fazla ayrılarak karbonat iyonları (CO_3^{2-}) ve ek olarak hidrojen iyonları (H^+) oluşturur. Hidrojen iyonlarının artışı, okyanus suyunun daha da asidik hale gelmesine yol açar (Tayyar, 2023).

Makine öğreniminin okyanus asitlenmesine karşı önemli bir öngörü sağladığı, pH tahminlerinde etkili olduğu ve gelecekteki çalışmalar için bir yol haritası sunduğu ifade edilmektedir (Ramdas vd., 2024). Okyanus asidifikasyonu değerlendirmek için makine öğrenmesi vazgeçilemez bir araç olmuştur (Vasanth vd., 2024). Okyanusların geleceği için oldukça bilimsel veri toplanması gerekmektedir (Taner, 2015). Literatürde yer alan okyanus asidifikasyonu ile ilgili makine öğrenmesi çalışmaları şöyledir;

Krasting vd., (2022) çalışmalarında, deniz suyu pH değerinin uzaktan algılanmasını ve makine öğrenmesi yöntemleriyle modellenmesini ele almıştır. Bu çalışmada, deniz yüzeyi karbon dioksit kısmi basıncı (pCO_2) verileri kullanılarak geniş ölçekli bir pH veri seti oluşturulmuş ve rastgele orman (Random Forest) modeli ile yüksek doğrulukta bir pH tahmin modeli geliştirilmiştir. Model, eğitim verisinde korelasyon katsayısı ($R^2 = 0.96$) ve kök ortalama kare hata (RMSE = 0.008), test verisinde ise $R^2 = 0.83$ ve RMSE = 0.017 değerleri ile yüksek performans göstermiştir. Ayrıca, mutlak hata ortalaması (MAE) gibi hata ölçütleri de değerlendirilmiş ve modelin genel olarak düşük hata payına sahip olduğu belirtilmiştir. Modelin, 2004-2019 yılları arasındaki küresel pH değişimlerini başarıyla yansıttığı ve iklim değişikliğine bağlı olarak okyanus yüzey pH seviyelerinde düşüş trendi tespit ettiği ifade edilmektedir. Bu yaklaşım, okyanus karbonat sistemindeki değişimleri anlamak için yeni bir yöntem sunarak, uzaktan algılama tabanlı pH tahminlerinin doğruluğunu artırmayı amaçlamaktadır.

Pasifik okyanusu asitlenmesi için çok değişkenli hibrit modelleme: gelecekteki pH eğilimlerini tahmin etme ve temel biyojeokimyasal sürücülerini analiz etme başlıklı çalışmada Pasifik okyanusundaki gelecekteki pH eğilimini tahmin etmek ve söz konusu eğilimdeki temel biyojeokimyasal etmenlerin etkisini analiz etmek amacıyla çok değişkenli hibrit bir makine öğrenmesi modeli kullanılmıştır. Bu çalışmada, farklı algoritmaların güçlü yönlerini stratejik olarak birleştirerek çeşitli okyanus asitlenmesi parametrelerini tahmin etmeyi hedeflemiş ve elde edilen sonuçlar hibrit modellerin yüksek doğruluk sağladığını ortaya koymuştur. Öngörülen pH eğilimlerinin pasifik okyanusu için endişe verici olduğu ifade edilmektedir. Pasifik Okyanusu'nda önümüzdeki 200 yıl boyunca okyanus asidifikasyonunun etkisini öngörmek amacıyla, pH değerinin tahmini için %25 Karar Ağacı ve %75 Gradyan Artırma (XGBoost) modellerinden oluşan hibrit bir model kullanılmıştır. Bu model, 22.720'lik en düşük ortalama kare hata (MSE) değeriyle yüksek doğruluk ve 0.998 R^2 değeriyle güçlü açıklayıcı olduğu söylenebilir (Vasanth vd., 2024).

Bu bağlamda Mevcut bu araştırmada okyanus asidifikasyonunu etkileyen çevresel veriler üzerinden en yüksek doğrulukla tahmin sağlayan makine öğrenmesi algoritmalarını belirlemek ve bu algoritmaların performansını karşılaştırmaktır. Çalışmanın hedefi ise, elde edilen en iyi modelleri kullanarak okyanus asidifikasyonuna karşı alınacak önlemlerin planlanmasında ve karar destek sistemlerinde uygulanabilir bilgiler sağlamaktır. Bu çalışmada, okyanus asidifikasyonunu incelemek amacıyla, Jayasurya (2024) tarafından sunulan Küresel Okyanus Asitlenmesi Eğilimleri ve Etkileri (Global Ocean Acidification Trends and Impacts) veri seti kullanılmıştır. Veri seti üzerinden, farklı makine öğrenmesi algoritmaları uygulanarak okyanus asidifikasyonunun tahmini gerçekleştirilmiş ve elde edilen modellerin performansları karşılaştırılmıştır. Bu yaklaşım hem veri odaklı analizler yapmayı hem de tahmin sonuçlarının karar destek sistemlerinde kullanım potansiyelini değerlendirmeyi mümkün kılmaktadır.

2 Metodoloji

2.1 Veri Seti

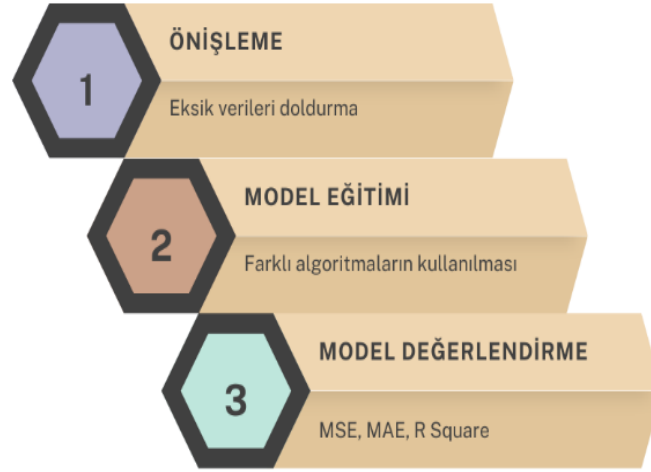
Bu çalışmada kullanılan Küresel Okyanus Asitlenmesi Eğilimleri ve Etkileri veri seti Kaggle ortamından (*Global Ocean Acidification Trends and Impacts*, t.y.) alınmıştır. 13 öznitelik ve 75 kayıttan oluşmaktadır. Veri seti 1950 yılından başlayarak 2023 yılına kadar gözlemsel verileri içermektedir. Söz

konusu veri setinin öznitelikleri şu şekildedir: year (Yıl), Gas (TWh, direct energy) [Gaz (TWh, doğrudan enerji)], Oil (TWh, direct energy) [Petrol (TWh, doğrudan enerji)], Coal (TWh, direct energy) [Kömür (TWh, doğrudan enerji)], Nuclear (% direct primary energy) [Nükleer (% doğrudan birincil enerji)], Renewables (% direct primary energy) [Yenilenebilir enerji (% doğrudan birincil enerji)], Fossilfuels (% directprimaryenergy) [Fosil yakıtlar (% doğrudan birincil enerji)], Annual GHG emission (in billiontons) [Yıllık sera gazı emisyonu (milyar ton cinsinden)], Global average temperature [Küresel ortalama sıcaklık], Global warming contribution from fossil fuels [Fosil yakıtlardan kaynaklanan küresel ısınma katkısı], CO₂ emissions from fossil fuels (in billiontons) [Fosil yakıtlardan kaynaklanan CO₂ emisyonu (milyar ton cinsinden)], Number of deaths from natural disasters [Doğal afetlerden kaynaklanan ölüm sayısı], Ocean acidification (in pH) [Okyanus asitlenmesi (pH cinsinden)]. Tablo 1’de veri setinin öz nitelikleri ve ilk beş kaydı verilmiştir.

Tablo 1: Küresel Okyanus Asitlenmesi Eğilimleri ve Etkileri veri seti

Öznitelikler	0	1	2	3	4
Yıl	1950.0	1951.0	1952.0	1953.0	1954.0
Doğalgaz (TWh, doğrudan enerji)	2.092.000	2.200.000	2.400.000	2.600.000	2.700.000
Petrol (TWh, doğrudan enerji)	5.444.000	6.000.000	6.500.000	7.000.000	7.500.000
Kömür (TWh, doğrudan enerji)	12.603.000	12.800.000	13.000.000	13.200.000	13.500.000
Nükleer (% doğrudan birincil enerji)	0.500000	0.500000	0.500000	0.500000	0.500000
Yenilenebilir (% doğrudan birincil enerji)	1.000.000	1.000.000	1.000.000	1.000.000	1.000.000
Fosil yakıtlar (% doğrudan birincil enerji)	98.500.000	98.500.000	98.500.000	98.500.000	98.500.000
Yıllık sera gazı emisyonu (milyar ton)	16.52	17.18	17.51	17.92	18.35
Küresel ortalama sıcaklık	-0.226622	-0.061154	0.015355	0.077631	-0.116750
Fosil yakıtlardan kaynaklı küresel ısınma katkısı	0.151223	0.155539	0.160021	0.164643	0.169367
Fosil yakıtlardan kaynaklı CO ₂ emisyonu (milyar ton)	5,93E+15	6,38E+15	6,47E+15	6,65E+15	6,79E+15
Doğal afetlerden kaynaklı ölüm sayısı	6845	18227	6886	12516	41847
Okyanus asitlenmesi (pH)	8.150.000	8.150.000	8.150.000	8.150.000	8.150.000

Bu çalışmada eksik veriler; ortalama, medyan, mod ve bir sonraki değer gibi uygun stratejilerle doldurularak ön işleme yapılmış ardından veri setinin %20’si test ve %80’i eğitim olarak ayrılmış, 11 farklı algorithmadan Ekstra Ağaçlar (ExtraTrees), Karar Ağacı (Decision Tree), Gradyan Artırma (Gradient Boosting), Rastgele Ormanlar (Random Forest), Doğrusal Regresyon (Linear Regression), Ridge Regresyonu (Ridge Regression), Uyarlamalı Artırma (AdaBoost), K En Yakın Komşular (K-Nearest Neighbors), Destek Vektör Makinesi (Support Vector Machine), Gauss Süreci (Gaussian Process) ve Sinir Ağı (Neural Network) oluşan bir model tasarlanmış ve bu algoritmaların MSE, MAE ve R^2 değerleri hesaplanarak en iyi sonuç veren algoritmalar tespit edilmiştir. Algoritmalar Python programlama dili kullanılarak analiz edilmiştir. Bu çalışmadaki makine öğrenme süreci Şekil 2’de verilmiştir.



Şekil 2: Makine Öğrenmesi Süreci

2.2 Makine öğrenmesi

Makine öğrenmesi yapay zekanın bir alt kümesidir (Akbulut & Adem, 2023). Makine öğrenmesi dijital veya gerçek ortamlardan elde edilmiş verilerden oluşan veri setleri üzerinde insan zekâsının süreçlerini esas alarak işlem yapmak için algoritma geliştirme alanıdır (El Naqa & Murphy, 2015). Makine öğrenmesi öğrenme sürecinde bilgisayar algoritmaları kullanılarak modeller oluşturulur (Zhou, 2021). Bilgisayarların doğrudan programlama komutları yerine, istatistiksel modeller ve algoritmalar kullanarak veriler üzerinde öğrenme yapması ve çeşitli işlemleri yerine getirmesini ifade eder (Mahesh, 2020). Makine öğrenmesinin amacı istatistiki ve matematiksel işlemlerle tahmin yapmak ve çıkarımda bulunmaktır (Akbulut & Adem, 2023). Makine öğrenme modelleri, çok sayıda katmandan oluşan yapılarıyla büyük veri analizlerinde hayati bir rol oynamaktadır. Bu modeller, verileri giderek daha soyut düzeylerde temsil ederek, büyük miktarda verinin işlenmesini mümkün kılar. (Sonmez vd., 2024). Algoritmaların veri bilimindeki başarıları, değişken sayısı, problem türü ve modele bağlı olarak farklılık gösterebildiğinden en yüksek başarıyı elde etmek için farklı algoritmalar denenmeli ve elde edilen sonuçlar karşılaştırılarak değerlendirilmelidir (Üngör & Adem, 2024).

2.3 Karar Ağacı

Karar Ağaçları sınıflama ve tahmin problemlerinde sıkça kullanılmaktadır (Çalış vd., 2014). Denetimli makine öğrenmesi yöntemidir (Pathan, 2021). Karar ağaçlarının oluşturulmasında en yaygın kullanılan algoritmalar arasında ID3 ve C4.5 yer almakta olup, bu algoritmalar veri kümesi üzerinde en uygun bölünmeleri belirleyerek modelin yapılandırılmasını sağlamaktadır (Çalış vd., 2014). Ağacın kök düğümü tüm veri setini temsil eder ve algoritma her düğümde, en iyi bölünmeyi sağlayacak özelliği ve eşik değerini seçerek veriyi özyinelemeli olarak çocuk düğümlere ayırır; bu süreç, düğümler saf hâle geldiğinde veya önceden belirlenmiş ağaç derinliğine ulaşıldığında sona erer ve ağacın son düğümleri olan yaprak düğümler, sonuçları veya sınıf etiketlerini gösterirken, her bölünme kararı bilgi kazancı, Gini saflığı veya varyans azaltımı gibi matematiksel kriterler kullanılarak alınır (Mienye & Jere, 2024). Karar ağacı algoritması, eğitim verisini tekrar tekrar alt kümeler ayırır; işlem, her alt kümedeki örnekler tek bir sınıfa ait olana veya ağacın maksimum derinliğine ulaşılan kadar sürer (Pathan, 2021). Karar ağaçlarının anlaşılması ve yorumlanmasının görece kolay olması, karar vericilerin süreci daha etkin bir şekilde yönetmelerine ve doğru stratejik adımlar atmalarına katkı sağlamaktadır (Çalış vd., 2014).

2.4 Rastgele Ormanlar

2000’li yıllarda Leo Breiman tarafından önerilen Rastgele Ormanlar, her bir ağacın rastgele seçilmiş veri alt uzaylarında büyütüldüğü ve bu ağaçlardan oluşan bir tahmin topluluğunu ifade eder (Biau & Scornet, 2016). Her bir ağacın diğerlerinden bağımsız olarak ve aynı olasılık dağılımından rastgele örneklenen veriler üzerinden karar verdiği bir ağaç topluluğudur (Breiman, 2001). Rastgele Ormanlar algoritmasının adı, birden fazla karar ağacı kullanılmasından dolayı ‘orman’, her bir ağacın eğitim verisindeki değişkenlerden rastgele seçim yapmasından dolayı ise ‘rastgele’ kavramlarının birleşiminden oluşmaktadır (Canaz Sevgen vd., 2020). Rastgele Ormanlar karar ağaçlarına dayalı makine öğrenmesi algoritmasıdır (Canaz Sevgen vd., 2020). Rastgele orman modeli, özyinelemeli örneklerinden oluşturulan çok sayıda budanmamış karar ağacının, her düğümde rastgele seçilen değişkenler arasından en iyi bölünmeyle büyütülmesi ve sonuçta tahminlerin ortalama (regresyon) veya oy çokluğu (sınıflama) ile birleştirilmesi esasına dayanır (Parlak & Kayri, 2022). Sınıflandırma ve regresyon problemlerinde iyi performans gösteren bir topluluk yöntemidir (Nadkarni vd., 2023). Hızlı sonuç üretebilen, kolay uygulanabilen, yüksek doğrulukta tahminler sağlayabilen ve aşırı öğrenme riski olmadan çalışabilen bir makine öğrenmesi yöntemidir.

2.5 Gradyan Artırımı

Gradyan Artırımı makineleri, farklı kayıp fonksiyonlarına göre öğrenme gibi uygulamaya özgü ihtiyaçlara uyarlanabilen ve geniş bir yelpazede başarılı sonuçlar göstermiş güçlü bir makine öğrenmesi yöntemleri ailesidir (Natekin & Knoll, 2013). Geri uyumlama ve parametrik olmayan regresyonlar kullanarak tahmin modelleri geliştirmeyi hedefler ve tek bir model oluşturmak yerine, önce bir başlangıç modeli kurar ve kayıp fonksiyonunu minimize ederek sürekli yeni modeller ekleyerek en doğru tahmini elde etmeye çalışır (He vd., 2019). Rastgele ormanların basit ortalama yaklaşımının aksine, artırma yöntemleri topluluğa ardışık olarak yeni zayıf öğrenici modeller ekleyerek her iterasyonda önceki topluluğun hatalarını düzeltmeye dayanan yapıcı bir strateji izler (Natekin & Knoll, 2013).

2.6 Ekstra Ağaçlar

Ekstra Ağaçlar bir diğer adıyla Aşırı Rastgeleleştirilmiş Ağaçlar (Extremely randomised trees), makine öğrenmesi alanında karar ağaçları metodunu esas alan topluluk öğrenme algoritmalarındandır (Elma & İlhan, 2024). Random Forest algoritması üzerinden geliştirilmiş olup yeni sayılabilecek bir algoritmadır (Ahmad vd., 2018). Ekstra Ağaçlar algoritması Rastgele Ormanlar algoritması olarak da kabul edilebilmektedir (Berrouachedi vd., 2019). Bu algoritma Rastgele Ormanlar algoritmasına benzer olmakla birlikte karar ağaçlarını daha çok rastgele oluşturur (Elma & İlhan, 2024). Ekstra Ağaçlar Algoritması Rastgele Ormanlardan farklı olarak eğitim için torbalama (bagging) yöntemiyle veri setinden alt küme veri setleri oluşturmaz tüm veri setini kullanır ayrıca düğüm bölme işleminde en iyi özelliği ve bu özelliğin değerini rastgele belirler (Jo vd., 2016). Bu nedenden dolayı aşırı uyum problemlerine daha az eğilimlidir (Ahmad vd., 2018). *Ekstra Ağaçlar* algoritması, budanmamış karar veya regresyon ağaçlarından oluşan bir topluluk oluşturur ve bu işlemi klasik Yukarıdan aşağıya yaklaşımı (top-down) yöntemiyle gerçekleştirir (Geurts vd., 2006). Ekstra Ağaçlar algoritması veri setinin tümünden karar ağacı üretir ve bu karar ağaçlarındaki düğümleri yaprak düğüme ulaşıncaya kadar rastgele iki çocuk düğüme böler sonrasında tüm ağaçların tahminlerini çoğunluk oylamayla birleştirir (Alfian vd., 2022). Ekstra Ağaçlar yöntemi rastgele bölme noktaları ve özellik seçimi gibi süreçlerde rastgele algoritmalar kullanması sebebiyle belleği ve zamanı tasarruflu kullanır (Mastelini vd., 2022). Hesaplama açısından görece daha verimli olup, büyük veya yüksek boyutlu veri setlerinde uygulanabilirliği açısından avantaj sağlamaktadır (Nadkarni vd., 2023).

Bu çalışmada Ekstra Ağaçlar yöntemi tercih edilmiştir, çünkü her düğümde rastgele bölünmeler kullanarak model çeşitliliğini artırmakta ve aşırı uyum riskini azaltmaktadır. Veri setimiz 75 kayıt ve 13 nitelikten oluştuğundan, küçük ve sınırlı sayıda gözlem içeren bir veri seti söz konusudur; bu bağlamda Ekstra Ağaçlar yöntemi hem doğruluk hem de hesaplama verimliliği açısından avantaj sağlamaktadır. Ayrıca, bu yöntem aracılığıyla elde edilen değişken önem skorları, veri setindeki kritik değişkenlerin belirlenmesine ve modelin yorumlanabilirliğinin artırılmasına olanak tanımaktadır.

3 Bulgular

Tablo 2’de öznelilikler arasındaki ilişki gösterilmiştir

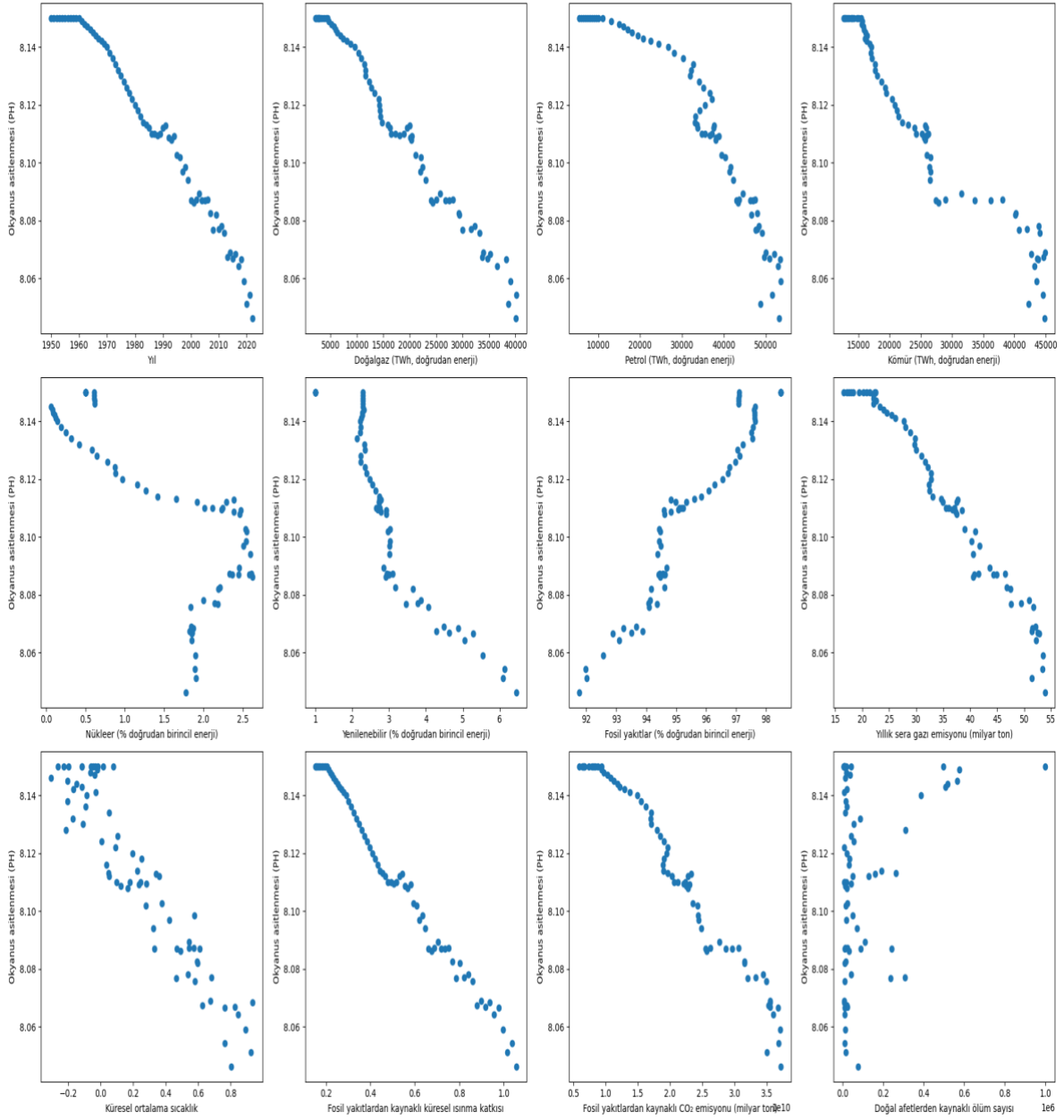
Tablo 2’de görüldüğü üzere Okyanus asidifikasyonu ile en yüksek ilişkili olan üç öz nitelik sırasıyla Fosil Yakıtların Küresel Isınmaya Katkısı (-0.9945), Gaz (-0.9925) ve Yıl (-0.987129) iken en düşük ilişkili olan öz nitelikler sırasıyla Doğal Afetlerden Ölen Kişi Sayısı (0.286849), Nükleer (-0.779401) ve Yenilenebilir Enerji Kaynakları’dır (-0.909861). Şekil 3’te Okyanus asidifikasyonu ile diğer öznelilikler arasındaki ilişki grafiksel olarak gösterilmiştir.

Tablo 2: Öznelilikler arası ilişki

Kod	Öznelilik	Kod	Öznelilik
Ö-1	Yıl	Ö-8	Yıllık sera gazı emisyonu
Ö-2	Doğalgaz	Ö-9	Küresel ortalama sıcaklık
Ö-3	Petrol	Ö-10	Fosil yakıtlardan kaynaklı küresel ısınma katkısı
Ö-4	Kömür	Ö-11	Fosil yakıtlardan kaynaklı CO ₂ emisyonu
Ö-5	Nükleer	Ö-12	Doğal afetlerden kaynaklı ölüm sayısı
Ö-6	Yenilenebilir	Ö-13	Okyanus asitlenmesi
Ö-7	Fosil yakıtlar		

	Ö-1	Ö-2	Ö-3	Ö-4	Ö-5	Ö-6	Ö-7	Ö-8	Ö-9	Ö-10	Ö-11	Ö-12	Ö-13
Ö-1	1.00	0,99	0,97	0,96	0,79	0,90	-0,97	1,00	0,93	0,99	0,99	-0,26	-0,99
Ö-2	0,99	1.00	0,95	0,98	0,74	0,92	-0,97	0,99	0,94	1,00	0,99	-0,27	-0,99
Ö-3	0,97	0,95	1.00	0,89	0,76	0,84	-0,92	0,97	0,85	0,93	0,97	-0,28	-0,93
Ö-4	0,96	0,98	0,89	1.00	0,69	0,89	-0,92	0,97	0,93	0,98	0,97	-0,23	-0,97
Ö-5	0,79	0,74	0,76	0,69	1.00	0,54	-0,82	0,78	0,73	0,76	0,75	-0,30	-0,78
Ö-6	0,90	0,92	0,84	0,89	0,54	1.00	-0,92	0,90	0,86	0,92	0,90	-0,15	-0,91
Ö-7	-0,97	-0,97	-0,92	-0,92	-0,82	-0,92	1.00	-0,96	-0,92	-0,97	-0,95	0,23	0,97
Ö-8	1,00	0,99	0,97	0,97	0,78	0,90	-0,96	1.00	0,93	0,99	1,00	-0,26	-0,98
Ö-9	0,93	0,94	0,85	0,93	0,73	0,86	-0,92	0,93	1.00	0,95	0,92	-0,27	-0,94
Ö-10	0,99	1,00	0,93	0,98	0,76	0,92	-0,97	0,99	0,95	1.00	0,99	-0,27	-0,99
Ö-11	0,99	0,99	0,97	0,97	0,75	0,90	-0,95	1,00	0,92	0,99	1,00	-0,26	-0,98
Ö-12	-0,26	-0,27	-0,28	-0,23	-0,30	-0,15	0,23	-0,26	-0,27	-0,27	-0,26	1,00	0,29
Ö-13	-0,99	-0,99	-0,93	-0,97	-0,78	-0,91	0,97	-0,98	-0,94	-0,99	-0,98	0,29	1,00

Şekil 3'te özniteliklerin eksik değerleri bir grafikte gösterilmiştir.



Şekil 3: Okyanus asidifikasyonu ilişkisel grafiği

Şekil 3'te gösterilen değerler doğrultusunda Şekil 4'te gösterildiği gibi; Yıllık sera gazı emisyonu, Fosil yakıtlardan kaynaklanan küresel ısınma katkısı, Fosil yakıtlardan kaynaklanan CO₂ emisyonları ve Okyanus asidifikasyonu özniteliklerinden 2 adet, Yıl, Gaz, Petrol, Kömür, Nükleer, Yenilenebilir, Fosil yakıtlar, Küresel ortalama ve Sıcaklık özniteliklerinden 1 adet veri eksikliği olup doğal afetlerden kaynaklanan ölüm sayısı özniteliğinden veri eksikliği yoktur. Yıl özniteliğindeki eksiklikler ileri doldurma ile Yıllık sera gazı emisyonu, Gaz, Kömür, Küresel ortalama sıcaklık, Fosil yakıtlardan kaynaklanan küresel ısınma katkısı ve Fosil yakıtlardan kaynaklanan CO₂ emisyonlarındaki eksiklikler ortalamayla, Nükleer, Yenilenebilir, Fosil yakıtlar ve Doğal afetlerden kaynaklanan ölüm sayısındaki eksiklikler medyanla ve Okyanus asidifikasyonundaki öznitelikler modla doldurulmuştur.



Şekil 4: Öznitelik eksik değerler sayısı

MSE: Ortalama Kare Hatası (Mean Squared Error) Tahmin edilen değerler ile gerçek değerler arasındaki farkların karelerinin ortalamasını olarak hesaplanan bir hata ölçütüdür (Jannani vd., 2021).

MAE: Ortalama Mutlak Hata (Mean Absolute Error) tahmin edilen değerler ile gerçek değerler arasındaki farkların mutlak değerlerinin ortalamasını olarak hesaplanan bir hata ölçütüdür (Jannani vd., 2021).

R²: Belirleme Katsayısı (R-squared): modelin bağımlı değişkenin değişkenliğini ne kadar açıkladığını ölçen bir istatistiksel ölçüt olup elde edilen değer +1'e yaklaştıkça modelin veriye uyumu artar. Bu metriklerin matematiksel ifadeleri şöyledir:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - y'_i| \quad (1)$$

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - y'_i)^2 \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_i - y'_i)^2}{\sum_{i=1}^m (y_i - \bar{y})^2} \quad (3)$$

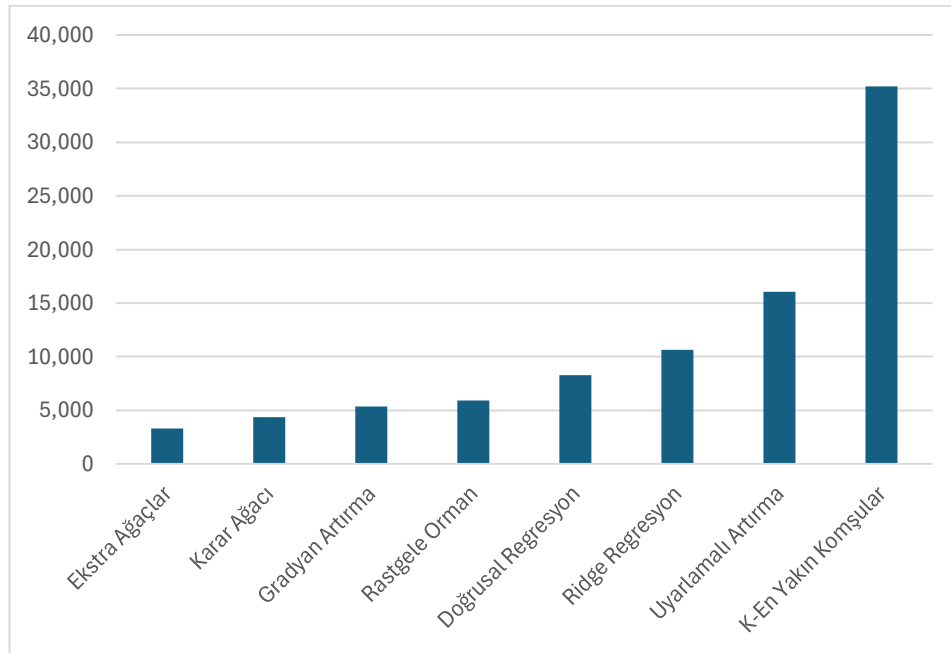
Bu formülde kullanılan m değişkeni test verilerindeki toplam eleman sayısını, y_i i'inci örneğe karşılık gelen gerçek değeri, y'_i i'inci örneğe karşılık gelen tahmin değerini, $\bar{y}$ gerçek değerlerin ortalamasını

ifade eder (Jannani, vd., 2021). Bu çalışmada kullanılan modelin MAE (1), MSE (2) ve R^2 (3) değerleri Tablo 3'te verilmiştir.

Tablo 3: Makine Öğrenmesi Modellerinin Performans Karşılaştırması

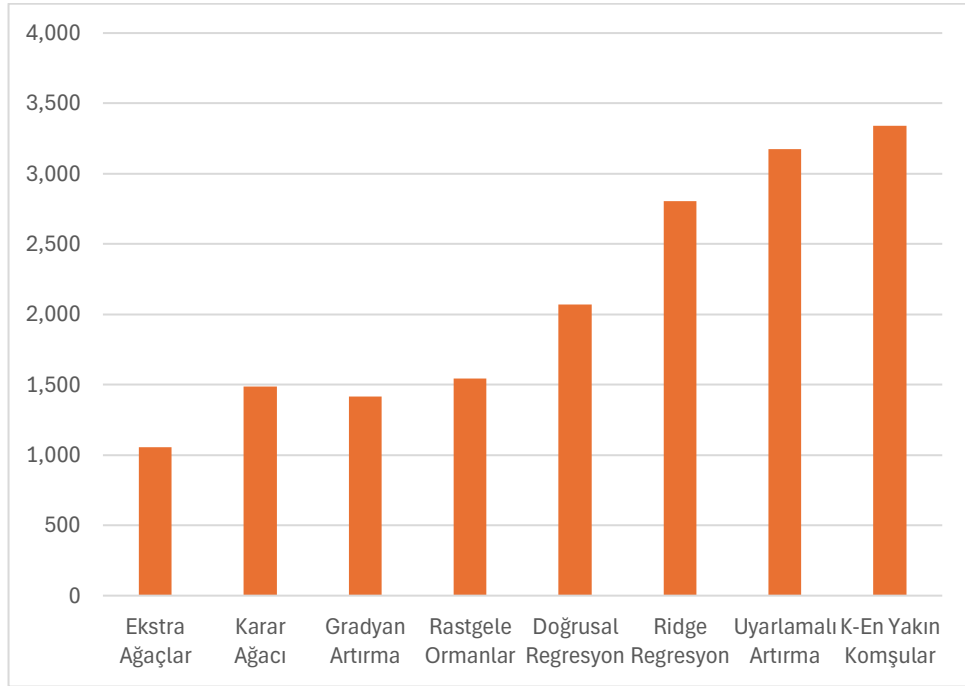
Model	MSE (10^{-6})	MAE (10^{-3})	R^2 Skoru
Ekstra Ağaçlar	3.273	1.055	0.99696
Karar Ağacı	4.384	1.486	0.99593
Gradyan Artırma	5.387	1.416	0.99499
Rastgele Orman	5.909	1.543	0.99451
Doğrusal regresyon	8.307	2.069	0.99229
Ridge Regresyon	10.662	2.803	0.99010
Uyarlamalı Artırma	16.062	3.176	0.98509
K-En Yakın Komşular	35.229	3.342	0.96729
Destek Vektör Makinesi	1339.979	32.622	-0.24424
Gaus Süreci	65843.650	8114.344	-61138.15
Sinir Ağı	5.27×10^{15}	6.39×10^7	-4.89×10^{18}

Tablo 3'te verilen model performans değerleri incelendiğinde, Ekstra Ağaçlar modeli, en düşük MSE (3.273×10^{-6}) ve MAE (1.055×10^{-3}) değerlerine sahip olup, R^2 skoru (0.99696) ile veriye en iyi uyumu gösteren modeldir. Bu, modelin tahminlerinin hem kare hatalar hem de mutlak hatalar açısından düşük olduğunu ve veriyi yüksek doğrulukla modellediğini göstermektedir. Karar Ağacı ($R^2 = 0.99593$) ve Gradyan Artırma ($R^2 = 0.99499$) modelleri de sırasıyla yüksek R^2 skorlarına sahip olup, MSE ve MAE değerleri de birbirine yakın seviyelerde kalmaktadır. Bu, bu modellerin de tahminlerinde düşük hata paylarıyla yüksek doğruluk sağladığını, ancak Ekstra Ağaçlar modelinin performans açısından bir adım önde olduğunu göstermektedir. Rastgele Orman ($R^2 = 0.99451$) modeli ise, özellikle Topluluk öğrenme yöntemleri arasında dikkat çekici bir performans sergileyerek veriye iyi uyum sağlamaktadır. Ancak, Doğrusal Regresyon ve Ridge Regresyon gibi daha basit modellerin R^2 değerleri, bu modellerin performansının gerisinde kalmaktadır. Şekil 5'te modellerin MSE değerleri verilmiştir.



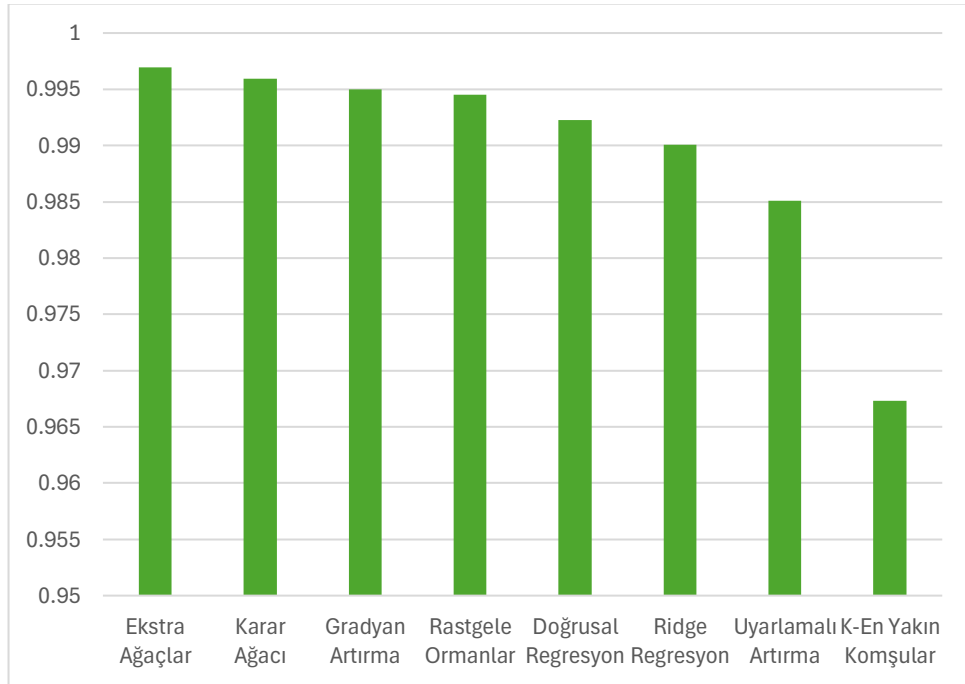
Şekil 5: Modellerin MSE değerleri

En düşük MSE değerleri sırasıyla Ekstra Ağaçlar, Karar Ağaçları, Gradyan Artırma modellerinde bulunmuştur. Şekil 6’da modellerin MAE değerleri verilmiştir.



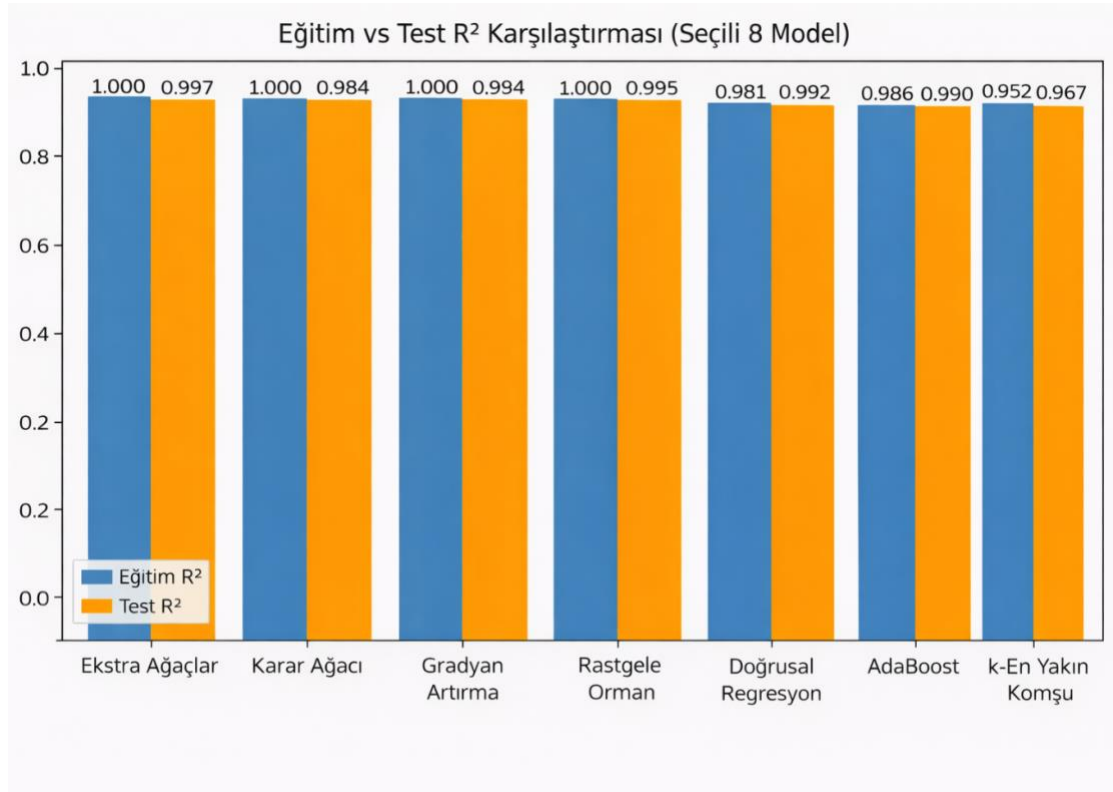
Şekil 6: Modellerin MAE değerleri

En düşük MAE değerleri sırasıyla Ekstra Ağaçlar, Gradyan Artırma, Karar Ağacı, modellerinde bulunmuştur. Şekil 7’de modellerin R^2 değerleri verilmiştir. En yüksek R^2 değerleri Ekstra Ağaçlar, Karar Ağacı, Gradyan modellerinde bulunmuştur.



Şekil 7: Modellerin R^2 Skorları

Şekil 8’de eğitim ve test verileri üzerinde elde edilen R^2 değerleri verilmiştir. Eğitim ve test verileri arasında anlamlı bir performans farklılığı görülmemiştir.



Şekil 8: Eğitim ve Test verisi R^2 Skorları karşılaştırması

Modellerin eğitim setindeki başarıları ile test setindeki başarıları birbirine oldukça yakın olduğu görülmektedir. Bu durum, modellerin eğitim verisine aşırı uyum (overfitting) göstermediğini, dolayısıyla genelleme yeteneklerinin yüksek olduğunu ortaya koymaktadır.

4 Tartışma

Bu çalışmada elde edilen bulgular, topluluk yöntemlerinin doğrusal modellere göre daha iyi bir performans sunduğunu ve özellikle Ekstra Ağaçlar, Karar Ağacı, Gradyan Artırma ve Rastgele Orman gibi modellerin güçlü tahmin yetenekleri sergilediğini ortaya koymaktadır. Genel olarak, topluluk tabanlı modeller (Ekstra Ağaçlar, Gradyan Artırma, Rastgele Orman) veri setinde daha yüksek doğruluk ve düşük hata oranları sağlayarak diğer basit regresyon modellerine göre üstün performans göstermektedir. Bu üstünlük, çok sayıda karar ağacı ve rastgele özellik seçim stratejisi sayesinde aşırı uyumun azaltılması ve veriyi daha iyi modelleme yeteneklerinden kaynaklanmaktadır (Nadkarni vd., 2023). Dolayısıyla, bu çalışma bağlamında topluluk yöntemlerin regresyon problemlerinde tercih edilmesi daha uygun görülmektedir. Ekstra Ağaçlar modeli, diğer modellere göre en düşük hata ve en yüksek uyum gösteren model olmuştur.

Ekstra Ağaçlar algoritması en yüksek R^2 ve en düşük MSE/MAE değerleri ile öne çıkmaktadır; bu, her düğümde rastgele bölünmeler kullanarak model çeşitliliğini artırması ve aşırı uyum riskini azaltması ile açıklanabilir. Karar Ağacı ve Gradyan Artırma algoritmaları da yüksek doğruluk sağlamaktadır; Karar Ağacı veri setindeki değişken ilişkilerini etkili şekilde yakalarken, Gradyan Artırma ardışık ağaç optimizasyonu sayesinde hataları minimize etmektedir. Küçük ve orta boyutlu veri seti yapısı, basit ağaç tabanlı modellerin performansını artırmış, karmaşık modeller ise düşük performans göstermiştir. Benzer

çalışmaları olan Krasting vd. (2022) ve Vasanth vd. (2024) araştırmalarında elde edilen hata değerleri, mevcut araştırmadan daha yüksek ve uyumları daha düşüktür. Ayrıca, bu benzer çalışmalarda da en iyi modeller, mevcut çalışmayla benzer olan topluluk yöntemlerinden Karar Ağaçları ve Rastgele Ormanlardır.

Mevcut çalışmanın aksine, Alver vd., (2025) Kızıldeniz'deki pH değişimlerini tahmin etmek için doğrusal regresyon, karar ağaçları ve yapay sinir ağları gibi makine öğrenmesi algoritmalarını kullanmış ve özellikle doğrusal modellerin yüksek tahmin doğruluğu sağladığı bulunmuştur. Benzer şekilde, Joshi vd., (2022) tarafından merkezi Bengal Körfezi'nde yapılan çalışmada deniz yüzeyi karbondioksit kısmi basıncı tahminleri incelenmiş ve Aşırı Gradyan Artırma (XGBoost)'nın doğrusal ve yapay sinir ağı modellerine kıyasla üstün performans gösterdiği görülmüştür. Bu bulgular, farklı deniz ekosistemlerinde topluluk yöntemler gibi makine öğrenmesi algoritmalarının, sınırlı ve çok boyutlu veri setlerinde dahi güvenilir tahminler sağlayabildiğini ortaya koymaktadır. Mevcut çalışma kapsamında yapılan analizler, küçük veri setlerinde Ekstra Ağaçlar algoritması ile yüksek doğruluk elde edilebileceğini desteklemektedir.

Quayesam vd. (2025) Büyük Set Resifi bölgesinde deniz yüzeyi sıcaklığını tahmin etmek amacıyla çeşitli makine öğrenmesi algoritmalarının karşılaştırmalı bir analizini gerçekleştirmiştir. Çalışmada Ridge Regresyon, LASSO, Rastgele Orman ve Aşırı Gradyan Artırma algoritmalarının performansı değerlendirilmiş; özellikle topluluk yöntemler olan Rastgele Ormanlar ve Aşırı Gradyan Artırma doğrusal modellere göre daha yüksek tahmin doğruluğu sağladığı görülmüştür. Bu bulgular, karmaşık ve çok değişkenli çevresel veriler üzerinde topluluk yöntemlerin doğruluğu artırmada etkili olduğunu göstermektedir. Bizim çalışmamızda da Ekstra Ağaçlar gibi topluluk yöntemleri kullanılarak sınırlı veri setinde yüksek doğruluk elde edilmesi, benzer şekilde topluluk tabanlı algoritmaların küçük ve çok boyutlu veri setlerinde bile üstün performans gösterebileceğini desteklemektedir.

Çalışmada kullanılan veri seti yalnızca 75 kayıttan oluştuğundan, elde edilen modellerin genellenebilirliği sınırlı olabilir. Küçük örneklem boyutu, modelin farklı veri dağılımlarında ve bölgesel koşullarda performansını tam olarak yansıtamayabilir. Gelecekte, daha geniş örneklemli veri setleri ve derin öğrenme tekniklerinin entegrasyonu ile yapılacak çalışmalar, modellerin genellenebilirliğini artıracak ve farklı bölgesel koşullara uyarlanmasını kolaylaştıracaktır. Bununla birlikte, mevcut sonuçlar, sınırlı veri ortamında kullanılan algoritmaların uygulanabilirliğini ve veri setindeki değişkenler arasındaki ilişkileri yakalama yeteneklerini ortaya koymaktadır.

5 Sonuçlar

Bu çalışma, okyanus asidifikasyonunu etkileyen çevresel faktörleri analiz ederek, makine öğrenmesi algoritmalarını kullanmak suretiyle pH seviyesini tahmin etmeye yönelik başarılı bir model geliştirmeyi amaçlamıştır. Kullanılan Küresel Okyanus Asitlenmesi Eğilimleri ve Etkileri veri seti üzerinde gerçekleştirilen analizlerde, okyanus asidifikasyonu en yüksek korelasyona sahip özneliliklerin sırasıyla fosil yakıt kaynaklı küresel ısınma katkısı, gaz tüketimi ve yıl değişkeni olduğu tespit edilmiştir. Eksik verilerin uygun yöntemlerle tamamlanmasının ardından, çeşitli makine öğrenmesi algoritmaları karşılaştırılmış ve Ekstra Ağaçlar algoritması, düşük hata oranları ve yüksek belirleme katsayısı ile en iyi performansı göstermiştir. Bu durum, söz konusu algoritmanın okyanus asidifikasyonunu öngörmede yüksek düzeyde etkili olduğunu ortaya koymaktadır. Çalışmanın sonuçları, makine öğrenmesinin çevresel tahmin modelleri kapsamında güçlü ve uygulanabilir bir araç olduğunu ortaya koyarken, özellikle çevresel karar destek sistemlerine entegre edilebilecek potansiyele sahip olduğunu göstermektedir. Bu tür modeller, çevresel riskleri öngörerek sürdürülebilir politika ve stratejilerin geliştirilmesine katkı sağlayabilir. Daha geniş örneklemeye sahip veri setleri ve derin

öğrenme teknikleriyle desteklenen çalışmaların yapılması, modellerin genellenebilirliğini artıracak ve farklı bölgesel koşullara uyarlanmasını kolaylaştıracaktır.

6 Beyanname

6.1 Çalışmanın Sınırlılıkları

Bu çalışmada elde edilen bulgular, kullanılan veri seti ile sınırlıdır.

6.2 Teşekkür

Bu araştırma için, yazar dışında katkı yapan kişi veya kurum bulunmamaktadır.

6.3 Finansman Kaynağı

Bu araştırma için herhangi bir finansal destek alınmamıştır.

6.4 Çıkar Çatışması

Bu çalışmada herhangi bir çıkar çatışması yoktur.

Kaynakça

- Ahmad, M. W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production*, 203, 810–821. <https://doi.org/10.1016/j.jclepro.2018.08.207>
- Akbulut, S., & Adem, K. (2023). Derin öğrenme ve makine öğrenmesi yöntemleri kullanılarak gelişmekte olan ülkelerin finansal enstrümanlarının etkileşimi ile BIST 100 tahmini. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 12(1), 52–63.
- Alfian, G., Syafrudin, M., Fahrurrozi, I., Fitriyani, N. L., Atmaji, F. T. D., Widodo, T., & Rhee, J. (2022). Predicting breast cancer from risk factors using SVM and extra-trees-based feature selection. *Computers*, 11(9), 136. <https://doi.org/10.3390/computers11090136>
- Alver, D. O., Isik, H., Palabiyik, S., Akkan, B. E., & Akkan, T. (2025). pH acidification in the Red Sea: A machine learning-based validation study. *Journal of Sea Research*, Article 102613.
- Atar, H. H., & Kızılgök, A. B. (2018). Küresel ısınmanın balıkçılığa etkileri. *Third Sector Social Economic Review*, 53(3), 1102–1115.
- Berrouachedi, A., Jaziri, R., & Bernard, G. (2019). Deep cascade of extra trees. In *Trends and applications in knowledge discovery and data mining* (pp. 117–129). Springer.
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, 25(2), 197–227.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Canaz Sevgen, S., & Aliefendioğlu, Y. (2020). Mass appraisal with a machine learning algorithm: Random forest regression. *Bilişim Teknolojileri Dergisi*, 13(3), 301–311.
- Çalış, A., Kayapınar, S., & Çetinyokuş, T. (2014). Veri madenciliğinde karar ağacı algoritmaları ile bilgisayar ve internet güvenliği üzerine bir uygulama. *Endüstri Mühendisliği*, 25(3), 2–19.
- Doney, S. C., Fabry, V. J., Feely, R. A., & Kleypas, J. A. (2009). Ocean acidification: The other problem. *Annual Review of Marine Science*, 1, 169–192.
- Elma, E., & İlhan, N. (2024). Çevrimiçi sosyal ağlarda sahte hesap tespiti. *Harran Üniversitesi Mühendislik Dergisi*, 9(2), 71–89.

- Feely, R. A. (2004). Impact of anthropogenic CO₂ on the CaCO₃ system in the oceans. *Science*, 305(5682), 362–366.
- Gattuso, J. P., Brewer, P. G., Hoegh-Guldberg, O., Kleypas, J. A., Pörtner, H. O., & Schmidt, D. N. (2014). Cross-chapter box on ocean acidification. In C. B. Field et al. (Eds.), *Climate change 2014: Impacts, adaptation, and vulnerability* (pp. 129–131). Cambridge University Press.
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63, 3–42.
- Global Ocean Acidification Trends and Impacts*. (t.y.), <https://www.kaggle.com/datasets/jayasurya666/global-ocean-acidification-trends-and-impacts>
- He, Z., Lin, D., Lau, T., & Wu, M. (2019). Gradient boosting machine: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.1908.06951>
- Jannani, A., Sael, N., & Benabbou, F. (2021). Predicting quality of life using machine learning: Case of world happiness index. 2021 4th International Symposium on Advanced Electrical and Communication Technologies (ISAECT), 1–6.
- Jayasurya, S. (2024). Global ocean acidification trends and impacts [Data set]. *Kaggle*. <https://www.kaggle.com/datasets/jayasurya666/global-ocean-acidification-trends-and-impacts>
- John, V., Liu, Z., Guo, C., Mita, S., & Kidono, K. (2016). Real-time lane estimation using deep features and extra trees regression. In *Image and Video Technology* (pp. 721–733). Springer.
- Joshi, A. P., Kumar, V., & Warrior, H. V. (2022). Modeling the sea-surface pCO₂ of the central Bay of Bengal region using machine learning algorithms. *Ocean Modelling*, 180, 102094.
- Karahan, A. (2021). İklim değişikliğinin istilacı yabancı türlerin başarısına etkisi. *Türkiye Deniz Bilimleri Vakfı Yayınları*, 2(1), 12–25.
- Krasting, J. P., De Palma, M., Sonnewald, M., Dunne, J. P., & John, J. G. (2022). Regional sensitivity patterns of Arctic Ocean acidification revealed with machine learning. *Communications Earth & Environment*, 3(1), 91.
- Mastelini, S. M., Nakano, F. K., Vens, C., & de Leon Ferreira, A. C. P. (2022). Online extra trees regressor. *IEEE Transactions on Neural Networks and Learning Systems*, 34(10), 6755–6767.
- Mienye, I. D., & Jere, N. (2024). A survey of decision trees: Concepts, algorithms, and applications. *IEEE Access*, 12, 86716–86727.
- Nadkarni, S. B., Vijay, G. S., & Kamath, R. C. (2023). Comparative study of random forest and gradient boosting algorithms to predict airfoil self-noise. *Engineering Proceedings*, 59(1), 24.
- Natekin, A., & Knoll, A. (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurobotics*, 7, 21.
- Özhan, K., Kocaman, E., & Ibello, V. (2022). Doğu Akdeniz’de okyanus asitlenmesi ve antropojenik karbonun ekosisteme etkileri. *Deniz Bilimleri Araştırma Dergisi*, 10(2), 45–60.
- Parlak, M. S., & Kayri, M. (2022). Öğretmenlerin e-öğrenme hazırbulunuşluk düzeylerini etkileyen faktörlerin rastgele orman algoritması yöntemi ile incelenmesi. *Van Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 19(3), 670–696.
- Pathan, S. S. (2021). An approach to decision tree induction for classification. *Turkish Journal of Computer and Mathematics Education*, 12(12), 919–928.
- Quayesam, D. L., Akubire, J. A., & Darkwah, O. A. G. (2025). *A comparison of machine learning algorithms for predicting sea surface temperature*. *arXiv*. <https://arxiv.org/abs/2411.15202>
- Ramdas, R., Patnaik, B. S., Parmar, A., Tadepalli, S. K., & Vasanth, K. (2024). Beyond pH levels: A comprehensive survey on ocean acidification. 2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE), 1–9.
- Sonmez, M. E., Gumus, N. E., Eczacioglu, N., Develi, E. E., Yücel, K., & Yildiz, H. B. (2024). Enhancing

- microalgae classification accuracy in marine ecosystems through convolutional neural networks and support vector machines. *Marine Pollution Bulletin*, 205, 116616.
- Taner, A. C., & Odaş, F. M. (2015). Küresel okyanus sularının ısınması ve asitlenmesi sonucu deniz ürünleri çeşitliliği azalmasının önlenmesi ve 2015 BM İklim Anlaşması açmazı (Teknik rapor). Denizcilik Odası Yayınları. <https://www.fmo.org.tr/wp-content/uploads/2011/07/Kuresel-Okyanus-Sularinin-Isinmasi-Asitlenmesi-Sonucu-Deniz-Urunleri-Cesitliliği-Azalmasinin-Onlenmesi-ve-2015-BM-Iklim-Anlasmasi-Acmazi.pdf>
- Tayyar, A. (2023). Küresel ısınmanın kötü ikizi – mavi ekonomi ve endüstri-içi ticaret: Türkiye için ARDL analizi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (57), 21–36.
- Türk Dil Kurumu. (2023). Okyanus. Güncel Türkçe Sözlük. <https://sozluk.gov.tr>
- Üngör, Y. E., & Adem, K. (2025). Makine öğrenmesi yöntemlerini optimize ederek saç dökülmesinin nedenlerini belirlemek. 4. BİLSEL International World Science and Research Congress, İstanbul.
- Vasanth, K., Kishore, R. K., Sugumaran, V., Krishnamoorthy, R., Ramdas, R., & Tadepalli, S. K. (2024). Multi-variate hybrid modeling for Pacific Ocean acidification: Predicting future pH trends and analyzing key biogeochemical drivers. *CSI Transactions on ICT*, 1–18.



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).