



Formlu Keçelerin Akustik Performansı için Makine Öğrenimi Tabanlı Tahmin Modeli

Sena ARSLAN ATMACA¹, Aslı AKSOY²

¹Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Bursa, TURKEY, ORCID ID 0009-0001-6452-4087

²Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Endüstri Mühendisliği Anabilim Dalı, Bursa, TURKEY, ORCID ID 0000-0002-2971-2701

Corresponding Author: Sena ARSLAN ATMACA, sena-arslan-98@hotmail.com

Özet

Otomotiv endüstrisinde artan konfor beklentileriyle birlikte akustik performans kritik bir odak noktası haline gelmiştir. Bu doğrultuda akustik beklentilere uygun malzemeler sunan, hızlı ve doğru bir şekilde ölçüm yapabilen tedarikçiler sektöründe ön plana çıkmaktadır. Bu çalışmada, ses yalıtımı amacıyla kullanılan keçe malzemelerinin ses yutum katsayısını tahmin etmek için üç farklı makine öğrenmesi algoritmasının (doğrusal regresyon, rastgele orman, k-en yakın komşu) performansı karşılaştırılmıştır. Çalışma kapsamında hazır bir veri seti kullanılmamış olup, tüm veriler epoksi keçe malzeme ile hazırlanan numuneler üzerinden gerçekleştirilen akustik performans ölçümlerinden elde edilmiştir. Model performansları ise eğitim ve test veri setlerinden elde edilen çeşitli performans metrikleri ile değerlendirilmiştir. Sonuçlar, rastgele orman algoritmasının diğer algoritmalara kıyasla daha iyi bir performans sergilediğini göstermektedir. Sonuç olarak, bu çalışma, makine öğrenmesi tekniklerinin akustik performansı hızlı ve doğru bir şekilde tahmin etme potansiyelini vurgulamaktadır. Empedans tüpü testi, alfa kabin testi gibi ampirik yöntemler ve analitik yöntemlerle karşılaştırıldığında, makine öğrenmesi temelli tahmin modellerinin zaman, iş gücü ve maliyet açısından sağladığı avantajlar, bu teknolojinin endüstriyel süreçlerdeki önemini daha belirgin hale getirmektedir.

Article Info

Research Article
Received: 10/04/2025
Accepted: 05/01/2026

Anahtar Kelimeler

Makine öğrenmesi, ses yutum katsayısı, lineer regresyon, k-en yakın komşu, rastgele orman

Öne Çıkanlar

Keçe malzemelerinin ses yutum katsayısını tahmin etmek için doğrusal regresyon, rastgele orman ve k-en yakın komşu algoritmaları karşılaştırılmıştır. Rastgele orman algoritmasının diğer yöntemlere kıyasla daha iyi performans gösterdiği belirlenmiştir. Makine öğrenmesi, geleneksel testlere kıyasla zaman ve maliyet açısından önemli avantajlar sunmaktadır.

Machine Learning Based Prediction Model for Acoustic Performance of Formed Felts

Abstract

The automotive industry has an increasing focus on comfort expectations, with acoustic performance becoming a critical focus point. In this regard, suppliers offering materials that meet acoustic expectations and can facilitate rapid and precise measurements are becoming increasingly prominent in the sector. The present study aims to contribute to this emerging field by comparing the performance of three distinct machine learning algorithms – linear regression, random forest, k-nearest neighbour– in predicting the sound absorption coefficient of felt materials employed for sound insulation. It is important to note that the study did not utilise any pre-existing data sets; all data were obtained from acoustic performance measurements performed on samples prepared with epoxy felt material. The performance of each model was evaluated using a range of performance metrics derived from both the training and test data sets. The findings of this study demonstrate that the random forest algorithm exhibits superior performance compared to the other algorithms. The study thus demonstrates the potential of machine learning techniques to predict acoustic performance quickly and accurately. Compared to conventional methods such as impedance tube testing, alpha cabinet testing, and analytical methods, the machine learning-based prediction models demonstrate significant advantages in terms of reduced time, labour, and cost, thereby underscoring the significance of this technology in industrial processes.

Keywords

Machine learning, sound absorption coefficient, linear regression, k-nearest neighbour, random forest

Highlights

A comparative analysis was conducted of linear regression, random forest and k-nearest neighbour algorithms with the objective of predicting the sound absorption coefficient of felt materials. The findings of this study indicated that the random forest algorithm exhibited superior performance in comparison to the other methodologies examined. The employment of machine learning has been shown to offer significant advantages in terms of reduced time and cost when compared to traditional testing methods.

1. Giriş

Son yıllarda teknolojide yaşanan gelişmelerle birlikte, otomobil kullanıcılarının performans ve konfor beklentileri gittikçe artmaktadır. Kabin içerisindeki gürültünün azaltılması sürüş konforunun sağlanması açısından en önemli beklentilerin başında gelir. Bu doğrultuda, kabin içi akustiği iyileştirmek amacıyla çeşitli yenilikçi malzemeler ve teknolojiler geliştirilmektedir. Bu hedef doğrultusunda, Orijinal Ekipman Üreticisi (OEM) firmalar, akustik beklentileri sağlayabilen ve rekabetçi malzemeler sunan tedarikçiler ile çalışmayı tercih etmektedir. Binek araçlarda gürültüyü engelleme ve ses yalıtımı sağlamak amacıyla, kapı paneli, tavan, taban ve motor bölgelerinde çeşitli malzemelerden yapılan iç/dış trim parçalar kullanılmaktadır. Yüksek ses yutumu performansı, düşük maliyet, hafiflik ve kolay işlenebilirlik gibi özelliklere sahip birçok farklı malzemenin, aracın çeşitli bölgelerinde kullanımı söz konusu olabilir. Özellikle, hafif ve geri dönüştürülebilir kompozit malzemeler, yüksek akustik performans

sağlamakla birlikte araç ağırlığını azaltmaya da yardımcı olur. Bu malzemelerin akustik performanslarının ölçülebilmesi için kapsamlı akustik testler gerçekleştirilmektedir. Bu test yöntemleri, numune üretimi, test numunesi hazırlama, testin gerçekleştirilmesi ve test sonuçlarının düzenlenmesi aşamalarını içermesi sebebiyle iş gücü, zaman ve maliyet kaybına yol açmaktadır. Bu nedenle malzemelerin akustik performanslarının tahmin edilmesi için uygulanan geleneksel yöntemlerden ziyade, günümüzde oldukça popülerlik kazanmış olan makine öğrenmesi (MÖ) teknikleri ön plana çıkmaktadır. MÖ algoritmaları, büyük veri kümelerinin analizinde geleneksel istatistiksel yöntemlere göre daha iyi bir performans göstermektedir [1]. Özellikle, denetimli öğrenme algoritmaları, üretim süreçlerindeki dinamik ilişkileri yakalamak için geçmiş verileri kullanmada üstün performans göstermektedir [2].

Literatürde, mikro delikli panellerin ses emilim katsayısının tahminine yönelik çalışmalarda, plaka kalınlığı, boşluk derinliği ve delme oranı gibi geometrik parametreler kullanılarak oluşturulan veri setleri üzerinden makine öğrenmesi modelleri geliştirilmiştir. Bu kapsamda 100–600 Hz frekans aralığında gerçekleştirilen bir çalışmada, destek vektör regresyonu, rastgele orman ve konvolüsyonel sinir ağı modelleri karşılaştırılmış ve destek vektör regresyonu yönteminin en yüksek tahmin doğruluğunu sağladığı raporlanmıştır [3]. Akustik performans tahmini yalnızca emilim özellikleriyle sınırlı kalmayıp, ses yalıtımı problemlerine de uygulanmıştır. Yüksek hızlı trenlerde kullanılan kompozit zemin sistemlerinin ses yalıtım performansını etkileyen temel parametrelerin belirlenmesine yönelik bir çalışmada, deneysel ses yutum katsayısı verileri kullanılarak rastgele orman tabanlı bir model geliştirilmiş; alüminyum profiller, iç paneller ve yüzey yoğunluğunun baskın faktörler olduğu gösterilmiştir [4]. Benzer şekilde, gürültü tahmini problemlerinde de makine öğrenmesi yaklaşımlarının etkinliği ortaya konmuştur. Yankısız ortam koşullarında gerçekleştirilen rüzgar tüneli testlerine dayalı bir çalışmada, yüksek frekanslı geniş bant gürültüsünün tahmini için karar ağacı tabanlı regresyon modeli önerilmiş ve makine öğrenmesi yöntemleriyle gürültü seviyelerinin yüksek doğrulukla öngörülebileceği gösterilmiştir [5]. Akustik metalmalzemelere yönelik çalışmalarda ise makine öğrenmesi, yalnızca tahmin değil aynı zamanda optimizasyon amacıyla da kullanılmıştır. Yarı-Helmholtz akustik metalmalzemelerin ses emilim performansının LightGBM tabanlı bir modelle tahmin ve optimize edildiği bir çalışmada, delik çapı, boyun uzunluğu, duvar kalınlıkları ve iç boşluk derinliği gibi yapısal parametrelerin emilim performansı üzerindeki etkileri analiz edilmiştir. Sonuçlar, makine öğrenmesi yöntemlerinin akustik metalmalzemelerin hızlı ve güvenilir biçimde tasarlanmasına olanak sağladığını göstermiştir [6]. Bir başka çalışmada ise, mikro-delikli plakaların akustik tasarımında geleneksel ileri tasarım yaklaşımları yerine makine öğrenmesine dayalı ters tasarım yöntemi önerilmiştir. Geliştirilen entegre derin sinir ağı modeli, ileri ve ters tahminleri yüksek doğrulukla gerçekleştirmiştir [7].

Mevcut literatürde, özellikle formlu keçe malzemelerin ses yutum katsayısını tahmin etmek için MÖ algoritmalarının uygulanmasına yönelik çalışmalar sınırlı olup, bu alanda daha fazla araştırmaya ihtiyaç duyulmaktadır. Bu çalışmada, akustik performans ölçümü esnasında harcanan maliyet, iş gücü ve süreden kazanç sağlamak amacıyla, iç trim ve dış trim parçalarında kullanılan formlu keçe malzemelerin ses yutum performans tahmini için MÖ teknikleri kullanılmıştır. Bu kapsamda, lineer regresyon (LR), rastgele orman (RO) ve k-en yakın komşu (k-EYK) algoritmaları kullanılarak tahminler yapılmıştır. Çalışmanın amacı bu konuda en iyi performans gösteren yaklaşımın belirlenmesidir.

2. Materyal ve Yöntem

Bu bölümde, veri setinin oluşturulması, uygulanan MÖ algoritmaları ve model performans metrikleri detaylı bir şekilde açıklanmıştır.

2.1. Veri setinin oluşturulması

Çalışma kapsamında hazır bir veri seti kullanılmamış olup, tüm veriler epoksi keçe malzeme ile hazırlanan numuneler üzerinden gerçekleştirilen akustik performans ölçümlerinden elde edilmiştir. Otomotiv sektöründe akustik izolasyon sağlanması amacıyla keçe malzeme kullanımı oldukça yaygındır. Keçe, dokunmamış, genellikle karışık liflerin bir araya getirilmesiyle oluşturulan bir tekstil malzemesi olarak tanımlanmaktadır[8].

Bu çalışmada, Formfleks Otomotiv tarafından üretilen; 700 g/m², 900 g/m² ve 1000 g/m² yoğunluğa sahip üç farklı epoksi keçe tipi seçilmiştir. Epoksi keçe, epoksi reçinesi ile birleştirilmiş, karışık liflerin bir araya gelmesi ile oluşturulan dokusuz bir tekstil malzemesidir [9]. Ham, preslenmemiş keçe numuneleri üzerinde yapılan ölçümler sonucunda, 700 g/m² ham keçe için serbest kalınlık 24.8 ± 1.6 mm, 900 g/m² ham keçe için 26.5 ± 1.5 mm ve 1000 g/m² ham keçe için 28.9 ± 1.7 mm olarak belirlenmiştir.

Akustik izolasyon amacıyla araçların farklı bölgelerinde kullanılan ön göğüs izolatörü, tavan kaplamaları ve kaput altı izolatörü gibi ürünler, tek katmanlı keçe yapıları ile üretilebildiği gibi, iki keçe tabakasının üst üste konumlandırılmasıyla elde edilen çift katmanlı yapılar şeklinde de üretilebilmektedir. Bu nedenle çalışma kapsamında numuneler, tek katmanlı ve birbirleri ile kombinasyon oluşturacak şekilde iki katmanlı konumlandırılarak dokuz farklı numune yapısı oluşturulmuştur. Belirlenen kombinasyonlar Tablo 1’de verilmiştir.

Tablo 1. Numune kombinasyonları

| No | 1. Katman (g/m ²) | 2. Katman (g/m ²) |
|----|-------------------------------|-------------------------------|
| N1 | 700 | - |
| N2 | 900 | - |
| N3 | 1000 | - |
| N4 | 700 | 700 |
| N5 | 700 | 900 |
| N6 | 700 | 1000 |
| N7 | 900 | 900 |
| N8 | 900 | 1000 |
| N9 | 1000 | 1000 |

Ses yutum katsayısındaki değişimleri gözlemlemek ve makine öğrenmesi modeli için girdi verisi oluşturmak amacıyla, her bir numune kombinasyonu için 2 mm, 5 mm, 10 mm, 15 mm, 20 mm ve 25 mm olmak üzere altı farklı bitmiş ürün kalınlığı seviyesi belirlenmiştir.

Bu çalışmada her bir kecenin serbest kalınlığı bilinmekle birlikte, temel parametre olarak bitmiş ürün kalınlığı esas alınmış, presleme sonrası katmanların ayrı ayrı kalınlıklarının belirlenmesi hedeflenmemiştir. Bunun nedeni, presleme sırasında katmanların sıkışma miktarının malzemenin gramajına ve yapısına bağlı olarak değişmesi ve katman bazında kalınlığın pratikte tekrarlanabilir şekilde kontrol edilmesinin zor olmasıdır. Prototip numune üretimleri, Formfleks Otomotiv’de bulunan 250 ton kapasiteli hidrolik pres kullanılarak gerçekleştirilmiştir. Üretilen numuneler Şekil 1’de gösterilmiştir.



Şekil 1. Formlanmış epoksi keçe numuneleri

Tüm numunelerin akustik performans testleri, TS EN ISO 354 standardına uygun olarak Formfleks Ar-Ge laboratuvarında yer alan ve Şekil 2’de gösterilen alfa kabin test cihazı ile 400~10000 Hz ölçüm aralığında gerçekleştirilmiştir. Akustik performans testleri gerçekleştirilirken laboratuvar ortamı, 23 °C sıcaklık ve %45 bağıl neme sahip olacak şekilde şartlandırılmakta ve ortam, test esnasında dış ortamdaki gürültü minimum seviyede olacak şekilde izole edilmektedir.



Şekil 2. Alfa kabin test cihazı

Akustik performans testleri sonrasında, çeşitli frekanslarda ölçülen ses yutum katsayısı değerleri ve malzeme kompozisyonu bilgileri baz alınarak bir veri seti oluşturulmuştur. Veri setinden örnek bir görüntü, Tablo 2’de verilmektedir.

Tablo 2. Veri seti örnek görüntü

| Gramaj 1 (g/m ²) | Gramaj 2 (g/m ²) | Katman Sayısı | Bitmiş Ürün Kalınlığı (mm) | Frekans (Hz) | Ses Yutum Katsayısı (α) |
|---------------------------------|---------------------------------|------------------|-------------------------------------|-----------------|----------------------------------------|
| 900 | 0 | 1 | 5 | 800 | 0.095 |
| 900 | 900 | 2 | 2 | 400 | 0.020 |
| 700 | 700 | 2 | 2 | 4000 | 0.359 |
| 1000 | 700 | 2 | 2 | 2500 | 0.334 |
| 700 | 1000 | 2 | 5 | 1250 | 0.361 |
| 900 | 900 | 2 | 5 | 500 | 0.089 |
| 900 | 0 | 1 | 10 | 6300 | 1.150 |
| 1000 | 0 | 1 | 20 | 4000 | 1.017 |

Çalışmada kullanılan veri seti toplam 690 satırdan oluşmakta olup, deneysel ölçümler sonucunda elde edilen çok değişkenli bir yapıya sahiptir. Veri setinde, beş adet girdi (bağımsız) değişken ve bir adet çıktı (bağımlı) değişken yer almaktadır. Girdi değişkenleri; numunenin yapısal ve fiziksel özelliklerini temsil eden gramaj 1, gramaj 2, katman sayısı, bitmiş ürün kalınlığı ile akustik test koşulunu ifade eden Frekans parametrelerinden oluşmaktadır. Çıktı değişkeni ise malzemenin ses yutum performansını temsil eden ses yutum katsayısıdır. Veri türleri incelendiğinde, katman sayısı kategorik veri olarak sınıflandırılırken diğer tüm veriler sürekli veri niteliğindedir. Veri seti özellikleri Tablo 3'te gösterilmektedir.

Tablo 3. Veri seti özellikleri

| Veri Adı | Açıklama | İstatistiksel Veri Tipi / Alacağı Değer | Veri Yapısı |
|--------------------------|-----------------------------------------------------------------------|---------------------------------------------|----------------|
| Gramaj 1 | İlk katmanın gramajı (g/m ²) | Sürekli veri / 700-1000 arası | Girdi |
| Gramaj 2 | İkinci katmanın gramajı (tek katmanlı numunelerde 0 değeri alır) | Sürekli veri / 0 veya 700- 1000 arasında | Girdi |
| Katman Sayısı | Toplam katman sayısı | Kategorik veri / 1 ya da 2 | Girdi |
| Frekans | Test edilen frekans | Sürekli veri / 400-10.000 arası | Girdi |
| Bitmiş Ürün Kalınlığı | Formlama sonrası kompozisyonun toplam kalınlığı | Sürekli veri / 2-25 mm arası | Girdi |
| Ses Yutum Katsayısı | Bir malzemenin üzerine gelen sesi ne kadar emdiğini gösteren değer | Sürekli veri / 0-1,4 arasında | Çıktı |

Tüm numunelerin ölçüm sonuçları incelendiğinde, bazı frekans bantlarında ses yutum katsayısının (α) 1'in üzerinde değerler aldığı görülmüştür. Teorik olarak ses yutum katsayısı 0 ile 1 arasında olması beklenir; ancak literatürde, özellikle Sabine formülü

kullanılan küçük çınlanım odası (alfa kabin) ölçümlerinde 1'in üzerindeki değerlerin gözlenmesi normal kabul edilmektedir. Bu durum aşağıdaki nedenlerle açıklanmaktadır:

1. Sabine formülü, doğrudan enerji emilim oranı değil, bu oranın logaritmik dönüşümü ile hesaplanır. Enerji emilim oranı (ϵ) yaklaşık %63'ü aştığında formül gereği α değeri 1'i geçebilmektedir. Özellikle yüksek emici malzemelerde bu durum beklenen bir sonuçtur [10].
2. Numune kenarlarında ses dalgalarının kırınımına uğraması ve saçılması, numunenin yalnızca ön yüzeyinin değil kenar yüzeylerinin de etkin soğurucu hale gelmesine yol açar. Bu ilave soğurma alanı hesaplamada ön yüzey alanına bölündüğünden, katsayının 1'i aşmasına neden olabilir [11].
3. Numunenin montaj şekli, arka yüzeyde hava boşluğu bırakılması veya kenarlarda sızıntılar olması ek soğurma etkisi yaratır. Bu etkiler hesaplamada göz önünde bulundurulmadığında katsayılar 1'in üzerine çıkabilir.

Bu nedenlerle, elde edilen ölçümlerde α değerinin 1'i aşması, malzemenin fiziksel olarak enerjiden fazla soğurma yaptığı anlamına gelmemekte; ölçüm yöntemi, oda geometrisi, numune boyutu ve hesaplama yaklaşımı gibi faktörlerin birleşiminden kaynaklanmaktadır.

2.2. Makine öğrenmesi algoritmaları

Makinelerin verilerden öğrenebilmesini sağlayan algoritmalar ve yöntemleri kapsayan MÖ modelleri, karmaşık verilerden doğru tahminler ve analizler çıkarılmasını sağlamaktadır [12]. MÖ tarihi, 1943 yılında Walter Pitts ve Warren McCulloch'un, herhangi bir mantıksal ifadenin, uygun bir sinir ağı modeliyle ifade edilebileceğinin anlatıldığı ilk matematiksel model ile başlamaktadır [13]. MÖ uygulamaları, denetimli öğrenme ve denimsiz öğrenme olmak üzere iki farklı yöntemle çalışabilmektedir. Denetimli öğrenmede, sonuçları bilinen veri setinden yapılan sınıflandırmadan hareketle sonuçları bilinmeyen veri setine dair etkili tahminler yapılırken, denimsiz öğrenmede sadece girdi verileri sisteme verilerek, sistem otomatik olarak keşifler yapar ve ilişki ağını ortaya koymaya çalışır [14].

LR, bağımsız değişkenler ile bağımlı bir değişken arasındaki doğrusal ilişkiyi modelleyen istatistiksel bir tekniktir [15]. Bu yöntem, bağımsız değişkenlerdeki değişimlerin bağımlı değişken üzerindeki etkisini anlamamıza ve gelecekteki değerleri tahmin etmemize olanak tanımaktadır. Bağımlı değişken, diğer değişkenlerin etkisiyle değeri değişen değişken olarak tanımlanmakta ve Y ile gösterilmektedir. Bağımsız değişken ise rastgele koşullara göre değişebilen ve bağımlı değişken üzerinde etkisi olan değişkendir; bağımsız değişken tekse X , birden fazlaysa X_1, X_2, \dots, X_k sembolleriyle gösterilir. Aşağıdaki denklem çoklu doğrusal regresyon modelini temsil eder. Modelin amacı, bağımlı değişken Y 'yi birden fazla bağımsız değişken X_1, X_2, \dots, X_k ile açıklamaktır.

$\beta_1, \beta_2 \dots \beta_k$ değerleri bağımsız değişkenlerin katsayılarını, β_0 ise sabit terim olarak tüm bağımsız değişkenler sıfır olduğunda Y 'nin tahmini değerini gösterir [16]. ϵ_i ise modelin açıklayamadığı rastgele hata veya gürültü olarak adlandırılır.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \epsilon_i \quad (1)$$

Bu çalışmada bağımsız değişkenler X_k ; gramaj, katman sayısı, ürün kalınlığı ve frekanstır. Bağımlı değişken Y ise ses yutum katsayısıdır.

RO, çok sayıda karar ağacının ürettiği tahminleri birleştirerek daha sağlam ve genelleştirilmiş sonuçlar elde eden bir topluluk algoritmasıdır [17]. Verilerdeki tüm olası değişkenliği hesaba katarak, karmaşık veri kümelerini işleyebilen, aşırı öğrenmeye dayanıklı ve daha kesin tahminler elde edilebildiği için yaygın olarak kullanılmaktadır [18]. RO, veri kümesinden rastgele alt kümeler seçerek başlar. Seçilen örneklerden her model, orijinal verilerden değiştirme ile alınan örneklemeler kullanılarak eğitilir; bu işleme önyükleme denir. Her model, kendi örnekleme üzerinde bağımsız olarak eğitilir ve farklı sonuçlar üretir. Daha sonra, tüm modellerin tahminleri aşağıda denklemi verilen çoğunluk oylaması yöntemiyle birleştirilerek nihai çıktı elde edilir. Her karar ağacı $f_b(X)$ bir tahmin yapar, $1(.)$ gösterge fonksiyonu olup, bir ağacın c sınıfını tahmin edip etmediğini kontrol eder. $\arg \max_c$ ise \hat{Y} 'yi belirleyen en sık tahmin edilen sınıfı seçer. Bu sürece toplama adı verilir [19]. Bu yöntem, modelin genelleme kabiliyetini artırarak daha güvenilir tahminler yapmasını sağlamaktadır.

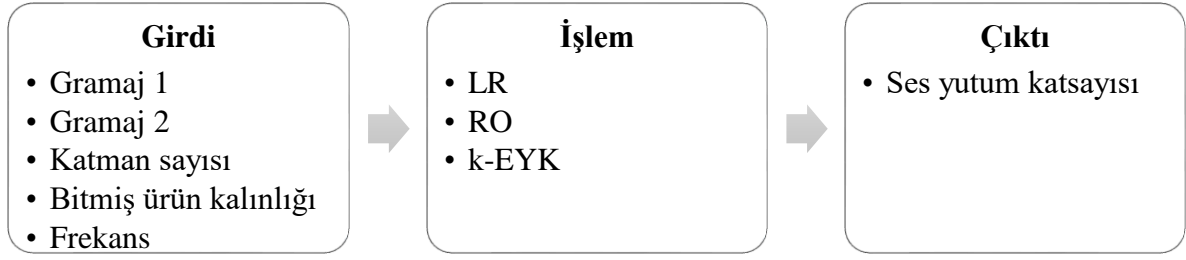
$$\hat{Y} = \arg \max_c \sum_{b=1}^B 1(f_b(X) = c) \quad (2)$$

k-EYK, verileri alt gruplara ayırarak, sınıflandırılmamış verilerin önceden sınıflandırılmış verilere olan benzerliğine göre sınıflandırma işlemi yapan bir algoritmadır [20]. Benzerlik veya yakınlık ilkesine göre çalışır. Tahmin edilmek istenen gözlem birimine en yakın "k" adet farklı gözlem birimi tespit edilir ve bu "k" adet gözlem biriminin bağımlı değişkenleri üzerinden ilgili gözlem için tahminde bulunulur. İlk olarak dikkate alınacak en yakın komşu sayısı olan "k" değeri belirlenir ve hedef veri noktası ile eğitim setindeki tüm noktalar arasındaki mesafe bir uzaklık fonksiyonu kullanılarak hesaplanır., [21]. Uzaklık hesaplamak için genellikle aşağıda verilen "Öklid Fonksiyonu" kullanılmaktadır.

$$\sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Öklid fonksiyonuna alternatif olarak Manhattan, Minkowski ve Hamming fonksiyonları da kullanılabilir. Uzaklık hesaplandıktan sonra sıralanır ve gelen değer uygun olan sınıfa atanır. En kısa mesafeye sahip "k" veri noktası bulunarak, bu komşuların özellikleri analiz edilir ve "k" komşusunun sınıf etiketleri arasında çoğunluk oylaması yapılır. Hedef veri noktasına en sık görülen sınıf etiketi atanır.

Çalışmada, veri setindeki girdilerle çıktılar arasındaki ilişkileri öğrenerek tahmin veya sınıflandırma yapabilmek için denetimli öğrenme algoritmaları tercih edilmiştir. Bu kapsamda değişkenler arası doğrusal ilişki kurabilen, LR modeli ve lineer olmayan RO ve k-EYK modelleri kullanılmıştır. Modellere ait girdi, işlem, çıktı diyagramı Şekil 3'te verilmiştir.



Şekil 3. Girdi, işlem, çıktı diyagramı

Tüm algoritmalar, Python programlama dili 3.13.2 sürümü kullanılarak çalıştırılmıştır. Veri analizi için “Pandas”, sayısal hesaplamalar için “Numpy”, MÖ araçlarının kullanılabilmesi için “Scikit-learn” ve elde edilen sonuçları görselleştirebilmek için “Matplotlib” kütüphaneleri eklenmiştir.

Literatürde genellikle verilerin %80’i eğitim verisi, %20’si ise test verisi olarak kullanılır. Eğitim verisinin kalitesi, ya da anakütleyi temsil kabiliyeti, modelin başarısını belirleyen temel unsurlar arasındadır. Bu çalışmada da verilerin %80’i modeli kurmak ve eğitmek için, kalan %20’lik kısım ise test için kullanılmıştır. Bu işlem scikit-learn kütüphanesinde yer alan, “train_test_split” fonksiyonu kullanılarak gerçekleştirilmiştir. Veriyi bölerken rastgeleliği kontrol etmek amacıyla random_state = 42 komutu kullanılır.

LR modeli, scikit-learn kütüphanesinde yer alan “LinearRegression” sınıfı kullanılarak oluşturulmuştur. Modelde herhangi bir düzenleme uygulanmamış olup, katsayılar en küçük kareler yöntemi ile hesaplanmıştır. LR modeli, doğrusal varsayım altında değişkenler arasındaki temel ilişkiyi incelemek amacıyla referans model olarak kullanılmıştır.

RO modeli, “RandomForestRegressor” sınıfı kullanılarak oluşturulmuştur. Modelde ağaç sayısı (n_estimators) 100 olarak belirlenmiştir. Farklı ağaç sayıları için yapılan ön denemelerde, yaklaşık 100 ağaçtan sonra model performans metriklerinde anlamlı bir iyileşme gözlenmediği belirlenmiştir. Bu nedenle, hem kararlı tahminler elde etmek hem de hesaplama süresini gereksiz yere artırmamak amacıyla ağaç sayısı 100 olarak seçilmiştir. Modelde bootstrap örnekleme yöntemi kullanılmıştır. Her bir karar ağacı, eğitim verisinin rastgele seçilmiş bir alt kümesi ile eğitilmiştir. Her düğümde bölünme yapılırken, dikkate alınan özellik sayısı toplam özellik sayısının karekökü olarak belirlenmiştir. Ağaçlar, saf hale gelene kadar büyütülmüş olup, minimum bölünme ve yaprak örnek sayıları varsayılan değerlerde bırakılmıştır. Bu yapı, modelin aşırı öğrenmeye karşı dayanıklı olmasını ve veri içindeki doğrusal olmayan ilişkileri etkin şekilde yakalamasını sağlamaktadır.

k-EYK modeli, “KneighborsRegressor” sınıfı kullanılarak oluşturulmuştur. Model oluşturulurken, “k” parametresi n_neighbors = 5 olarak belirlenmiştir. Bu, modelin tahmin yaparken kullandığı en yakın beş komşunun değerlerini dikkate alacağı anlamına gelmektedir. Küçük k değerleri modelin aşırı uyum sağlamasına yol açarken, büyük k değerleri ise genelleme gücünü artırır ancak fazla basitleşmesine sebep olabilir. k=5 bu iki uç nokta arasında iyi bir denge sağlamaktadır. Komşu ağırlıkları ise 'uniform' olarak

belirlenmiş olup, modelin her bir komşuya aynı önemi vermesi sağlanmıştır. Algoritmada mesafe ölçümü olarak “Öklid” kullanılmaktadır.

2.3. Model performans metrikleri

MÖ algoritmalarının tahmin başarısının ölçülebilmesi ve modelin doğrulanabilmesi için model performans metriklerinin kullanılması gerekmektedir [22]. Bu çalışmada algoritmaların tahmin performansları hem eğitim hem de test veri setlerinde ayrı ayrı değerlendirilmiştir. Değerlendirme sırasında, belirleme katsayısı (R^2), ortalama mutlak hata (MAE), ortalama karesel hata (MSE), kök ortalama karesel hata (RMSE) ve ortalama mutlak yüzde hata (MAPE) metrikleri ile değerlendirilmiştir. Denklemlerde verilen n toplam gözlem sayısı, y_i gerçek değerlerin i . gözlemi, \hat{y}_i ise tahmin edilen değerlerin i . gözlemi anlamına gelmektedir.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (8)$$

R^2 , bağımlı değişkendeki varyansın bağımsız değişkenlerden tahmin edilebilen oranıdır. 0 ile 1 arasında değer alır; 1'e ne kadar yakınsa, modelin performansı o kadar iyidir [23]. MAE, tahmin edilen ve gerçek değerler arasındaki mutlak farkların ortalamasıdır. Hataların büyüklüğünü doğrudan yansıtır ve tüm hataları eşit oranda değerlendirir. MSE, tahmin edilen ve gerçek değerler arasındaki farkların karelerinin ortalamasıdır. Büyük hataların etkisini daha fazla vurgular. RMSE, MSE'nin kareköküdür ve hataların ortalama büyüklüğünü orijinal birimde ifade eder. MAPE ise tahmin hatalarının yüzde cinsinden ortalamasıdır ve tahminlerin gerçek değerlere göre ne kadar sapma gösterdiğini belirtmektedir [24].

3. Bulgular ve Tartışma

Bu çalışmada formülü keçe malzemelerin ses yutumu katsayısını tahmin etmek için LR, RO ve k-EYK algoritmaları kullanılmış, modelin performans değerlendirilmesi R^2 , MAE, MSE, RMSE ve MAPE gibi temel metrikler kullanılarak yapılmıştır. R^2 değerinin, 1'e yakın olması modelin tahmin ile gerçek veriler arasında güçlü bir uyum olduğunu ve değişkenler arasındaki ilişkileri yakalamada yüksek bir doğruluk oranına sahip olduğunu gösterir [25]. MAE, MSE, RMSE ve MAPE gibi hata metriklerinin sıfıra yakın olması, modelin tahminlerinin minimum hata payına sahip olduğunu ve gözlemlenen değerlerle yüksek düzeyde uyum gösterdiğini gösterir [26].

Algoritmaların performansını kapsamlı bir şekilde değerlendirmek için hem eğitim hem de test veri kümelerinden elde edilen sonuçlar Tablo 4 ve Tablo 5'te verilmiştir. Eğitim kümesi sonuçları, modelin eğitildiği verilerdeki performansını yansıtırken, test kümesi sonuçları, modelin görülmeyen, genel verilerdeki tahmin yeteneğinin bir değerlendirmesini sağlar. Eğitim ve test sonuçları arasındaki minimum farklar, modelin hem eğitim verilerinde hem de genel verilerde iyi performans gösterdiğini ve doğru tahminler yapma yeteneğini gösterdiğini gösterir [27].

Test veri seti performansları incelendiğinde, RO, LR ve k-EYK algoritmaları için R^2 değerleri sırasıyla, 0.9961, 0.6053, 0.4700 olarak hesaplanmıştır.

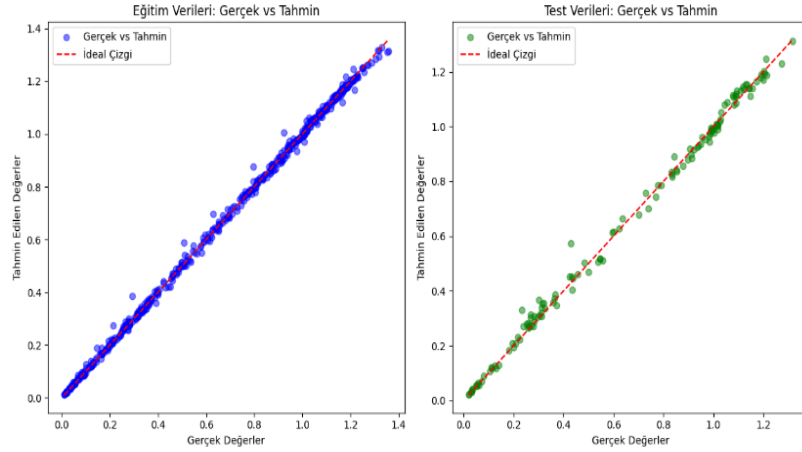
Tablo 4. Eğitim veri seti performans metrikleri

| Algoritma | R^2 | MAE | MSE | RMSE | MAPE |
|------------------|--------|--------|--------|--------|----------|
| Lineer Regresyon | 0.6522 | 0.1966 | 0.0566 | 0.2381 | %76.0608 |
| Rastgele Orman | 0.9988 | 0.0087 | 0.0002 | 0.0138 | %2.4841 |
| k-En Yakın Komşu | 0.6379 | 0.1987 | 0.0590 | 0.2430 | %80.5704 |

Tablo 5. Test veri seti performans metrikleri

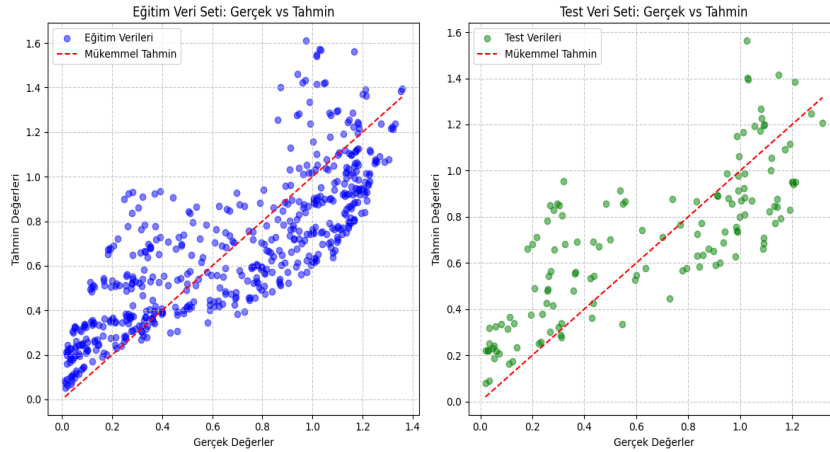
| Algoritma | R^2 | MAE | MSE | RMSE | MAPE |
|------------------|--------|--------|--------|--------|----------|
| Lineer Regresyon | 0.6053 | 0.2085 | 0.0637 | 0.2524 | %90.4648 |
| Rastgele Orman | 0.9961 | 0.0174 | 0.0006 | 0.0251 | %4.5228 |
| k-En Yakın Komşu | 0.4700 | 0.2273 | 0.0855 | 0.2925 | %90.8275 |

En yüksek R^2 değeri, MAE, MSE, RMSE hata metriklerinin 0'a yakın olacak şekilde en düşük olması ve %4.5228 MAPE değeri RO algoritmasıyla elde edilmiş olup, bu durum algoritmanın en yüksek doğruluk oranı ile çalıştığını göstermektedir. RO algoritmasının eğitim ve test veri setlerinde oldukça yüksek performans değerleri sergilemesi modelin aşırı öğrenme riskinin ayrıca değerlendirilmesini gerekli kılmış ve modelin genelleme yeteneğini doğrulamak amacıyla 5-katlı çapraz doğrulama yöntemi uygulanmıştır. Çapraz doğrulama sonuçları, RO modelinin farklı veri alt kümelerinde de benzer doğruluk seviyelerini koruduğunu göstermiş ve modelin yalnızca eğitim verilerine özgü değil, genel veri yapısına uyumlu bir tahmin yeteneğine sahip olduğunu doğrulamıştır. RO modelinin formülü keçe malzemelerin ses yutum katsayısını tahmin etmede güçlü olduğu ve veri seti ile yüksek bir uyum sergilediği Şekil 4'te açıkça görülmektedir.



Şekil 4. RO gerçek-tahmin edilen değer grafiği

LR algoritmasının eğitim ve test veri setinde benzer R^2 değerine sahip olması (~ 0.6) tutarlı olsa da düşük bir açıklama gücüne sahiptir. Bu durum, modelin aşırı öğrenme göstermediğini; ancak açıklama gücünün sınırlı kaldığını göstermektedir. MAPE değerinin %90.4648 olması ile modelin karmaşık doğrusal olmayan ilişkileri yakalamakta zorlandığı Şekil 6’da görülmüştür. Doğrusal regresyon, bağımlı ve bağımsız değişkenler arasındaki ilişkiyi doğrusal olarak varsayan bir modeldir. Bu yöntem, değişkenler arasındaki ilişkinin doğrusal olup olmadığını dikkate almaz ve yalnızca doğrusal bağıntıları yakalayabilir. Literatürde de görüldüğü gibi değişkenler arasındaki ilişki doğrusal olmadığı zaman modelin tahmin performansı olumsuz etkilenmektedir [28].



Şekil 6. LR gerçek-tahmin edilen değer grafiği

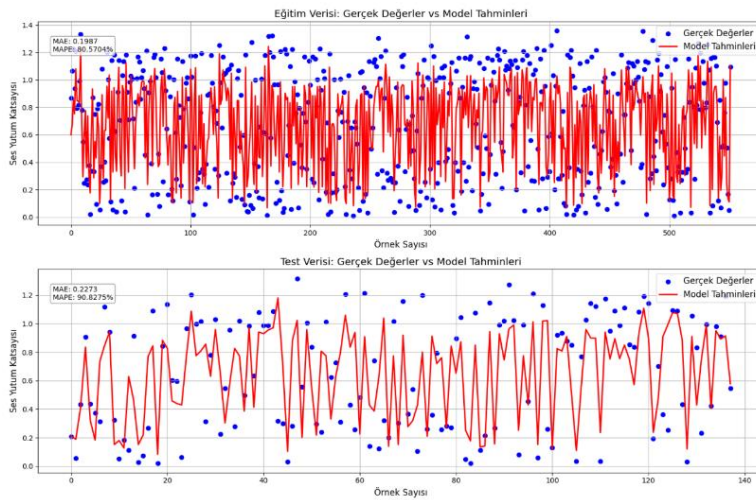
Tablo 6'da verilen LR katsayılarına göre, ses yutum katsayısını en çok etkileyen değişken 0.2066 katsayısı ile katman sayısıdır, bu da literatürde de sıkça vurgulanan çok katmanlı yapıların ses dalgalarını daha etkili sönmüleyebilmesi olgusunu desteklemektedir [29]. Bitmiş ürün kalınlığı (0.0354) da düşük frekansta yutumu artıran önemli bir faktördür [30]. Buna karşın, frekans ve gramajın düşük katsayıları (0.0001), modelin bu değişkenlere duyarlılığının sınırlı olduğunu ve doğrusal olmayan modellerin daha iyi

açıklama yapabileceğini göstermektedir. Gramajın düşük etkisi, lif yoğunluğu, malzeme gözenekliliği gibi ek faktörlerle birlikte ele alınması gerektiğini düşündürmektedir.

Tablo 6. LR girdi katsayıları

| Girdi | Katsayı |
|-----------------------|---------|
| Katman sayısı | 0.2066 |
| Bitmiş ürün kalınlığı | 0.0354 |
| Frekans | 0.0001 |
| Gramaj 1 | 0.0001 |
| Gramaj 2 | 0.0001 |

k-EYK algoritmasında ise, eğitimde $R^2 = 0.6379$, ancak testte $R^2 = 0.4700$ değerine gerileyerek ciddi bir performans düşüşü yaşanmış olup, test veri setinde MAPE: %90.82 ile yüksek hata oranı sergilemiştir. Bu durum Şekil 7’de görülmektedir. Model belirli bir trendi yakalasa da bireysel tahminlerde ciddi sapmalar yaşamaktadır. K-EYK algoritması, mesafe tabanlı yapısı nedeniyle girdi değişkenlerinin ölçeklerine duyarlıdır. Bu çalışmada k-EYK modeli için herhangi bir normalizasyon veya standardizasyon işlemi uygulanmamıştır. Bu durum, modelin özellikle test veri setindeki performansını sınırlamış olabilir. Gelecek çalışmalarda uygun veri ön işleme adımlarının k-EYK performansını artırma potansiyeli bulunmaktadır. Buna rağmen, k-EYK algoritmasının çalışmaya dahil edilmesi, farklı algoritmaların veri yapısına karşı gösterdiği duyarlılığı karşılaştırmalı olarak ortaya koymak açısından önemli bir referans sağlamaktadır.



Şekil 7. k-EYK gerçek-tahmin edilen değer grafiği

Bu çalışma, formlu keçe malzemelerin ses yutum katsayısının makine öğrenmesi yöntemleri ile tahmin edilebilirliğini ortaya koymayı amaçlayan, veri odaklı ve karşılaştırmalı bir yaklaşıma sahiptir. Çalışmada kullanılan veri seti belirli gramaj, katman sayısı ve kalınlık aralıkları ile sınırlandırılmıştır. Ayrıca çalışma kapsamında kullanılan makine öğrenmesi algoritmaları gözetimli öğrenme yöntemleri ile sınırlı olup, derin öğrenme veya hibrit optimizasyon yaklaşımları uygulanmamıştır. Gelecek çalışmalarda belirtilen sınırlılıklar dikkat alınarak, farklı lif türleri veya üretim yöntemleri ile girdiler artırılarak kapsam genişletilebilir. Daha geniş ve çeşitlendirilmiş veri kümeleri ile farklı algoritmalar denenerek modelin genelleme kabiliyeti daha da artırılabilir.

4. Sonuç

Otomotiv endüstrisinde artan konfor beklentileriyle birlikte akustik performans kritik bir odak noktası haline gelmiştir. Bu çalışmada farklı gramaj, katman sayısı, kalınlık ve frekans değerlerine sahip epoksi keçe numunelerin ses yutum katsayısını tahmin etmek için LR, k-EYK ve RO olmak üzere üç farklı MÖ algoritmasının performansı karşılaştırılmıştır. Model performansları, eğitim ve test veri setlerinden elde edilen R^2 , MAE, MSE, RMSE ve MAPE metrikleri ile değerlendirilmiştir. Bulgular, RO algoritmasının diğer algoritmalara kıyasla daha iyi bir performans gösterdiğini, daha yüksek R^2 değeri, MAE, MSE, RMSE hata metriklerinin sıfıra yakın değerler aldığını ve %4.5228 MAPE değeri ile algoritmanın en yüksek doğruluk oranı ile çalıştığını göstermektedir. Literatürde, RO algoritmasının aşırı öğrenmeye yatkın olabileceği öngörülse de bu sonuçlar, yeterli ağaç sayısı ve uygun derinlik parametreleri ile modelin karmaşık desenleri etkin bir şekilde öğrenebildiğini göstermektedir. Sonuç olarak, MÖ modelleri akustik performansın hızlı ve doğru bir şekilde tahmin edilmesini sağlamak ve ampirik ya da analitik yöntemler ile karşılaştırıldığında, zaman, işgücü ve maliyet kaybında belirgin bir tasarruf sağlamaktadır. Bu çalışmanın, otomotiv endüstrisinde akustik performans tahmini alanında, farklı malzemelerden elde edilen verilerle MÖ modellerinin yaygın bir şekilde kullanılmasına yönelik önemli referans kaynak olacağı öngörülmektedir.

Çıkar Çatışması

Bu makale ile ilgili herhangi bir çıkar çatışması bulunmamaktadır.

Yazar Katkısı

Yazarlar eşit katkı sağlamıştır.

5. Kaynaklar

- [1] Yıldız, B. İ. ve Karabağ, K., Prediction of beef production using linear regression, random forest and k-nearest neighbors algorithms, **Kahramanmaraş Sütçü İmam Üniversitesi Tarım ve Doğa Dergisi**, 28(1), 247-255, (2025).
- [2] Ahmed, M. U. ve Hussain, I., Prediction of wheat production using machine learning algorithms in northern areas of Pakistan, **Telecommunications Policy**, 46(6), 102370, (2022).

- [3] Yuan, B., You, T., Jiang, H., Qian, H., Cao, L. ve Zhu, R., Prediction of the sound absorption performance for micro-perforated panel based on machine learning, **Journal of Engineering and Applied Science**, 72(1), 27, (2025).
- [4] Wang, R., Yao, D., Zhang, J., Xiao, X. ve Xu, Z., Identification of key factors influencing sound insulation performance of high-speed train composite floor based on machine learning, **Acoustics**, Vol. 6, No. 1, pp. 1-17, (2023).
- [5] Sasaki, S., Suganuma, S. and Aramaki, T., Development of a low-noise wind tunnel and prediction of broadband noise from a flat plate based on machine learning, **Nagasaki University Graduate School of Engineering Research Reports**, 53 (101), 1-7, (2023).
- [6] Xu, J., Mao, W., Yu, X., and Liu, B., Prediction and optimization of acoustic absorption performance of quasi-Helmholtz acoustic metamaterials based on LightGBM algorithm, **AIP Advances**, 15(1), (2025).
- [7] Zhang, Z., Ye, R., Zhu, G., Huang, X., Qian, Y. And Yurchenko, D., Research on sound absorption performance prediction of microporous structures based on integrated learning, **Materials Today Communications**, 46, 112866, (2025).
- [8] Das, D., Pradhan, A. K., Chattopadhyay, R. ve Singh, S. N. **Composite nonwovens**, Textile Progress, 44(1), 1-84, (2012).
- [9] Xue, B., Xie, L., Bao, Y. ve Zhang, J., Multilayered epoxy/glass fiber felt composites with excellently acoustical and thermal insulation properties, **Journal of Applied Polymer Science**, 136(3), 46935, (2019).
- [10] Polack, J. D., Revisiting reverberation formulae, **In INTER-NOISE and NOISE-CON Congress and Conference Proceedings**, (Vol. 270, No. 8, pp. 3497-3504), Institute of Noise Control Engineering, (2024).
- [11] Peng, L. Sound absorption and insulation functional composites. **In Advanced high strength natural fibre composites in construction**, (pp. 333-373), Woodhead Publishing, (2017).
- [12] Türkmenoglu, C. ve Tantug, A. C., Sentiment analysis in Turkish media, **In International Conference on Machine Learning**, Beijing, (2014)
- [13] McCulloch, W. S. ve Pitts, W., A logical calculus of the ideas immanent in nervous activity. **The Bulletin of Mathematical Biophysics**, 5, 115-133, (1943).
- [14] Korkmaz, G. ve Eroğlu, E., Model karmaşıklığının kontrolü, **İktisadi ve İdari Yaklaşımlar Dergisi**, 2(2), 146-162, (2020).
- [15] Kılıç, S., Doğrusal regresyon analizi, **Journal of Mood Disorders**, 3(2), 90-92, (2013).
- [16] Tonta Y., Regresyon analizi, Ders Notu, Hacettepe Üniversitesi, Ankara, (2008).
- [17] Breiman L., Random forests, **Machine Learning**, 45, 5-32, (2001).
- [18] What is random forest? <https://www.ibm.com/think/topics/random-forest>, (01.04.2025).
- [19] Altman, N. ve Krzywinski, M., Ensemble methods: bagging and random forests, **Nature Methods**, 14(10), 933-935, (2017).

- [20] Taunk, K., De, S., Verma, S. ve Swetapadma, A. A brief review of nearest neighbor algorithm for learning and classification. **International Conference on Intelligent Computing and Control Systems**, p. 1255-1260, Madurai, (2019).
- [21] Taşcı, E. ve Onan, A., K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi, **Akademik Bilişim**, 1(1), 4-18. (2016).
- [22] Kızılkaya, Y. M. ve Oğuzlar, A., Bazı denetimli öğrenme algoritmalarının R programlama dili ile kıyaslanması, **Karadeniz Uluslararası Bilimsel Dergi**, 37(37), 90-98. (2018).
- [23] Figueiredo Filho, D. B., Júnior, J. A. S. ve Rocha, E. C., What is R2 all about?, **Leviathan (São Paulo)**, (3), 60-68. (2011).
- [24] Varoquaux, G. ve Colliot, O., Evaluating machine learning models and their diagnostic value, **Machine Learning for Brain Disorders**, 601-630. (2023).
- [25] Chicco, D., Warrens, M. J., ve Jurman, G., The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation, **Peerj Computer Science**, 7, e623. (2021).
- [26] Bui, D. T., Khosravi, K., Tiefenbacher, J., Nguyen, H. ve Kazakis, N., Improving prediction of water quality indices using novel hybrid machine-learning algorithms, **Science of the Total Environment**, 721, 137612, (2020).
- [27] Prayudani, S., Hizriadi, A., Lase, Y. Y. ve Fatmi, Y., Analysis accuracy of forecasting measurement technique on random K-nearest neighbor (RKNN) using MAPE and MSE, **In Journal of Physics: Conference Series**, Vol. 1361, No. 1, p. 012089, (2019).
- [28] Flom, P., The disadvantages of linear regression (2018) <https://www.sciencing.com/disadvantages-linear-regression-8562780/>, (2.04.2025).
- [29] Yuan, Z. X., Acoustic properties of multilayered structures, **Acoustics Australia**, 48(3), 395-405, (2020).
- [30] Qui, H., ve Enhui, Y., Effect of thickness, density and cavity depth on the sound absorption properties of wool boards, **Autex Research Journal**, 18(2), 203-208, (2018).