

## K En Yakın Komşu Algoritması ile Binalarda Enerji Verimliliği Tahmini

Mehmet HACİBEYOĞLU<sup>1</sup>  Merve ÇELİK<sup>2</sup>  Özlem ERDAŞ ÇİÇEK<sup>3</sup> 

<sup>1</sup> Necmettin Erbakan Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Konya, Türkiye, [hacibeyoglu@erbakan.edu.tr](mailto:hacibeyoglu@erbakan.edu.tr) (Sorumlu Yazar/ Corresponding Author)

<sup>2</sup> Necmettin Erbakan Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Konya, Türkiye

<sup>3</sup> Necmettin Erbakan Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Konya, Türkiye, [ozlem.erdascicek@erbakan.edu.tr](mailto:ozlem.erdascicek@erbakan.edu.tr)

### Makale Bilgileri

### ÖZ

**Makale Geçmişi**  
Geliş: 31.05.2023  
Kabul: 31.07.2023  
Yayın: 31.12.2023

**Anahtar Kelimeler:**  
Enerji verimliliği,  
Lineer regresyon,  
K en yakın komşu,  
Makine öğrenmesi.

Binaların ısıtma ve soğutma yükleri, bir binanın içindeki hava sıcaklığının belirli bir sıcaklıkta tutmak için gereken enerji miktarlarını ifade eder. Bu yükler, binanın yalıtımı, boyutu, şekli, konumu, kullanılan malzemeler ve ısıtma sistemi gibi birçok faktöre bağlıdır. Toplumsal refahın artırılmasında vazgeçilmez bir etken olan enerjinin giderek önem kazandığı günümüzde binalardaki enerji tüketiminin azaltılması hem çevresel sürdürülebilirlik hem de ekonomik açıdan önemlidir. Bu çalışmada denetimli makine öğrenmesi algoritmalarından doğrusal regresyon ve k en yakın komşu algoritmaları kullanılarak bir regresyon problemi olan binalarda enerji verimliliği tahmini yapacak bir model geliştirilmiştir. Deneysel çalışmalar, 12 farklı binadan elde edilen sekiz adet şart özelliğe ve ısıtma yükü ve soğutma yükü olmak üzere iki sınıf değerine sahip toplamda 768 adet veri barındıran enerji verimliliği veri kümesi kullanılarak gerçekleştirilmiştir. Çalışmada k en yakın komşu algoritmasının k parametre değeri optimize edilerek performansı artırılmıştır. Elde edilen deneysel sonuçlara göre k en yakın komşu algoritması doğrusal regresyon algoritmasına göre çok daha başarılı ve %96'lar seviyesinde tahminler gerçekleştirmiştir. Çalışma sonunda elde edilen denetimli makine öğrenmesi modeli ile binaların enerji analizleri kolaylıkla yapılabilecek ve elde edilen analiz sonuçları enerji verimliliğinin artırılabilmesi için geliştirilecek politikalarda kullanılabilir.

## Energy Efficiency Estimation in Buildings with K Nearest Neighbor Algorithm

### Article Info

### ABSTRACT

**Article History**  
Received: 31.05.2023  
Accepted: 31.07.2023  
Published: 31.12.2023

### Keywords:

Energy efficiency,  
Linear regression,  
K nearest neighbor,  
Machine learning.

The quantity of energy needed to maintain a particular temperature inside a structure is referred to as the facility's heating and cooling loads. These loads are influenced by a variety of elements, including the building's insulation, size, shape, location, type of materials used, and heating system. The reduction of energy consumption in buildings is crucial in terms of both the environmental sustainability and the profitability of the building industry. This is because energy is becoming an increasingly important issue, and it is an essential component in improving social welfare. In this study, a model has been developed to predict energy efficiency in buildings with a regression problem by using linear regression and k nearest neighbor algorithms from supervised machine learning algorithms. Experimental studies were carried out using an energy efficiency dataset containing a total of 768 data with eight conditional properties and two class values, heating load and cooling load, obtained from 12 different buildings. The study improved the performance of the k nearest neighbor algorithm by optimizing the k parameter value. According to the experimental results obtained, the k nearest neighbor algorithm was much more successful than the linear regression algorithm and made predictions at the level of 96%. With the supervised machine learning model obtained at the end of the study, energy analyzes of buildings can be easily done and the obtained analysis results can be used in policies to be developed to increase energy efficiency.

**Atıf/Citation:** Hacibeyoğlu, M; Çelik, M. & Erdaş Çiçek, Ö. (2023). K en yakın komşu algoritması ile binalarda enerji verimliliği tahmini, *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 5(2), 65-74.  
<https://doi.org/10.47112/neufmbd.2023.10>



"This article is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) (CC BY-NC 4.0)"

## GİRİŞ (INTRODUCTION)

Makine öğrenmesi, insanların öğrenme şeklini taklit etmek için verilerin ve algoritmaların kullanımına odaklanan bir yapay zekâ ve bilgisayar bilimi dalıdır. Makine öğrenmesi algoritmaları öğrenme yöntemlerine göre temelde denetimli ve denetimsiz olmak üzere iki gruba ayrılırlar. Denetimli makine öğrenmesi algoritmaları, bilgisayarların gerçek dünyada bir probleme çözüm getirmesi için etiketli eğitim kümesinden öğrenen yapay zekânın alt dalıdır. Bu algoritmalar, bir bilgisayar programının insan müdahalesi olmadan öğrenmesine ve gelişmesine olanak sağlar. Denetimli makine öğrenmesi algoritmalarının çalışma mantığında, verileri işleyip onlardan belirli desen ve anlamlar çıkarmak veya belirli modeller geliştirmek vardır [1]. Geliştirilen model ile veriler üzerinde bir görevi yerine getirmek için matematiksel formüller veya algoritmalar kullanılarak öğrenme başarısı artırılmaya çalışılır. Denetimli makine öğrenmesi algoritmalarında genellikle öğrenme süreci algoritmik olarak tekrarlı işlemlerden oluşur. Bu tekrarlı işlemlerde, model veriler üzerinde çalışır ve performansını iyileştirmek için ayarlanır. Daha fazla veri üzerinde uygun parametrelerde ve yeterli süre çalışan denetimli makine öğrenmesi algoritmaları daha iyi sonuçlar elde etmek için sürekli olarak öğrenir ve gelişir. Bu nedenle denetimli makine öğrenmesi algoritmaları, günümüzde teknolojik yeniliklerin bir öncüsü olarak kabul edilmekte ve otonom sürüş sistemleri, sesli asistanlar, nesne tanıma, doğal dil işleme, tıbbi teşhis gibi birçok farklı alanda kullanılmaktadır [2-6]. Denetimli makine öğrenmesi algoritmalarının kullanım alanlarının, her geçen gün artması ile birlikte insanların yapmakta zorlandığı görevler gerçekleştirilebilmekte ve bilgisayarların daha verimli bir şekilde kullanılmasını sağlanmaktadır. Bu görevler verilerin analiz edilmesi, tahmininin yapılması, sınıflandırılması, görüntü tanınması, vb. gerçek dünya problemleridir. Bu problemlerin çözümü için literatürde sıklıkla kullanılan denetimli makine öğrenmesi algoritmaları karar ağaçları, rasgele orman, destek vektör makinesi, naive bayes ve k en yakın komşu algoritmasıdır [7-9]. Denetimsiz makine öğrenmesi algoritmaları da kümeleme problemlerinin çözümünde, veriler arasında ilişkileri keşfetmek ve optimizasyon problemlerinin çözümünde kullanılır [10,11].

Günümüzde dünya genelinde ülkelerin, kurumların ve bireyler enerji ihtiyacı artan nüfus ve teknolojik gelişmeler ile birlikte her geçen gün biraz daha artmaktadır. Ayrıca enerji, endüstriyel alanda birçok ürünün üretiminde de kullanılmaktadır. Artan bu enerji ihtiyacı karşılayabilmek için yeni enerji kaynaklarına finansal yatırımlar yapılmakta ve özellikle ülkemiz için dışarıdan enerji ithal edilmektedir. Bu yüzden enerji ülkelerde ki refahın ve ekonomik gelişmenin sağlanmasında büyük bir öneme sahiptir. Dünya genelindeki enerji tüketiminin yaklaşık %30'u binalarda ısıtma ve soğutma işlemleri için kullanılmaktadır [12]. Binalarda enerji verimliliğinin en yüksek seviyede olması çevresel sürdürülebilirlik hem de ekonomik açıdan önemli bir kriterdir. O yüzden var olan binaların ısıtma ve soğutma enerji verimlilik hızlı bir şekilde hesaplanması ve gerekli önlemlerin alınması enerji tüketimini azaltacaktır. Bu problemin çözümü içinde makine öğrenmesi algoritmalarından faydalanılabilir.

Denetimli makine öğrenmesi algoritmaları genellikle sınıflandırma veya regresyon problemlerinin çözümünde kullanılır. Makine öğrenmesi algoritması şart özelliklerini bildiği problem için sınıf değerini tahmin etmeye çalışır. Eğer tahmin edilen değer kategorik ise sınıflandırma, nümerik ise regresyon problemidir. Bu çalışma da bir regresyon problemi olan binalarda enerji verimliliğinin k en yakın komşu makine öğrenmesi algoritması ile tahmini üzerine bir çalışma gerçekleştirilmiştir. Literatürde enerji verimliliği hakkında son yıllarda yapılan çalışmalar incelendiğinde, 2022 yılında Yıldız ve ark. tarafından yapılan çalışmada [13] kural tabanlı yöntemlerden karar ağaçları kullanılmıştır. Veri kümesine öncelikle belirli veri ön işleme işlemleri yapılmış ve karar ağaçları algoritması ile binaların, duvar alanı, cam yüzey alanı, çatı alanı gibi fiziksel parametrelerinin kullanılarak, ısıtılması ve soğutulması için gerekli enerji yükünün kabul edilebilir hata payları ile öngörülebildiği söylenmiştir. 2022 yılında Canbay ve Taş tarafından yapılan çalışmada [14] yapıların ısıtma ve soğutma yüklerinin yapay zekâ ile tahmini üzerinde durulmuştur. Yapılan çalışmada karar ağaçları, rasgele orman, lineer regresyon, k en yakın komşu ve karar destek sistemleri algoritmaları kullanılarak binaların 8 farklı özelliği üzerinden, binaların ısıtma ve soğutma için ihtiyaç duyacakları enerji tüketimini tahminlemesi gerçekleştirilmiştir. Jia ve ark. 2021 yılında yüksek katlı konut binalarının soğutma yükünü tahmin etmek için 30 tane nörondan oluşan bir

yapay sinir ağı modeli geliştirmişlerdir [15]. Peker ve ark. 2017 yılında yapmış oldukları çalışmada enerji tasarruflu bina tasarımı için ısıtma ve soğutma yüklerinin regresyon tabanlı makine öğrenmesi algoritmaları ile modellenmişlerdir. Destek Vektör Makinesi Regresyonu, Doğrusal Regresyon, Rasgele Orman Regresyonu ve En Yakın Komşu Regresyonu kullanılan çalışmada hem ısıtma ve soğutma için çıkışlar ayrı ayrı hesaplanmıştır ve en iyi sonucu Rasgele Orman Regresyonunun hesapladığı gözlemlenmiştir [16]. Dilber ve ark. 2022 yılında yaptıkları çalışmada enerji verimli binaların ısıtma ve soğutma yüklerini doğrusal regresyon, k en yakın komşuluk, destek vektör makineleri, karar ağaçları, rastgele orman, uç gradyan artırma ve yapay sinir ağları ile tahminlemesini gerçekleştirmişlerdir. Yapılan çalışma ile elde edilen sonuçlara göre tahmin doğruluğu açısından en başarılı algoritmanın uç gradyan artırma algoritmasının olduğu belirtilmiştir [17].

Bu çalışmada binalarda ısıtma ve soğutma işlemleri için enerji verimliliği değerinin makine öğrenmesi yöntemlerinden k en yakın komşu ile tahminlemesi gerçekleştirilmiştir. Enerji verimliliği veri kümesi üzerinde veri ön işleme işlemleri yapılmış ve özgünlük olarak k en yakın komşu algoritmasında parametre ayarı ile performans artışı gerçekleştirilmiştir. Çalışmadan elde edilen makine öğrenmesi modeli ile binaların ısıtma ve soğutma verimlilik değerleri kabul edilebilir hata payı ile çok hızlı bir şekilde hesaplanabilecek ve elde edilen bu değerler yardımıyla binalarda enerji verimliliği artıracak ısı yalıtım vb. ek önlemler anılabilecektir [18].

Çalışmanın geri kalanı şu şekilde düzenlenmiştir. Materyal ve metot kısmında çalışmada kullanılan algoritmalar, performans metrikleri ve enerji verimliliği veri kümesi anlatılmaktadır. Deneysel çalışmalar kısmında çalışmanın genel yapısı kullanılan makine öğrenmesi algoritmaların enerji tahmini problemi için tahminleme sonuçları ve elde edilen en iyi sonuçların literatürdeki diğer çalışmalar ile karşılaştırmalı analizi bulunmaktadır. Sonuç kısmında çalışmadan elde edilen bilgiler paylaşılmakta ve gelecekte yapılabilecek çalışmalardan söz edilmektedir.

## MATERYAL VE METOT (MATERIALS AND METHODS)

### *K En Yakın Komşu Algoritması*

K En yakın komşu (K Nearest Neighbour, KNN) sınıflandırma ve regresyon problemleri için kullanılan bir denetimli makine öğrenmesi algoritmasıdır. KNN algoritması, özellikle küçük boyutlu veri kümelerinde ve sınıflar arasındaki ayrımın belirgin olduğu durumlarda etkili bir şekilde çalışır. KNN algoritması yeni sınıflandırılacak olan örnek ile eğitim kümesindeki örneklerin uzaklıklarını ölçüp en yakın k tane komşunun sınıf değerlerini kullanarak sınıflandırma işlemlerini gerçekleştirir. Sınıflandırılacak veri ve eğitim verileri arasındaki mesafeyi ölçmek için genellikle aşağıdaki denklemlerde sırayla verilen Öklid, Manhattan ve Minkowski uzaklık ölçüm formülleri kullanılır.

$$d(X, Y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

$$d(X, Y) = |x_i - y_i| \quad (2)$$

$$d(X, Y) = \sqrt{\left(\sum_{i=1}^n |x_i - y_i|^p\right)^{1/p}} \quad (3)$$

Yukarıdaki denklemlerde X sınıflandırılacak olan yeni veriyi ve Y ise veri kümesindeki örneği temsil etmektedir. Her iki veride veri kümesi de özellik sayısı kadar eleman içeren vektör şeklinde tutulmaktadır.

$$X = \{x_1, x_2, \dots, x_n\} \quad (4)$$

$$Y = \{y_1, y_2, \dots, y_n\} \quad (5)$$

KNN algoritması regresyon problemlerinde uzaklık ölçütü belirlendikten sonra veri noktasına en yakın

komşuları hesaplayıp komşulara ait hedef değerlerinin interpolasyonlarını hesaplayarak tahminde bulunur [19]. Bu noktada k sayısının optimum değerde seçilmesi en önemli adımlar biridir. k'nın optimum değerde seçilebilmesi için k değerinin belirli aralıklarda testleri gerçekleştirilir [20]. Bu çalışmada k değerinin optimum değerde seçilebilmesi için gerekli test işlemleri yapılmıştır ve sonuçlar deneysel çalışmalar bölümünde verilmektedir.

### **Regresyon Analizi**

Regresyon analizinde amaç belli aralıklarda bulunan bağımsız x değişkenlerini kullanarak bağımlı y değişkenini tahmin etmektir. Bunun için kullanılan pek çok yöntem bulunmaktadır. Bu çalışmada doğrusal regresyon yöntemi kullanılmıştır.

### **Doğrusal Regresyon**

Doğrusal regresyon, gözetimli öğrenmede kullanılan basit ama etkili bir araçtır. Tek bağımsız değişken bulunduğu durumda doğrusal regresyon  $y = ax + b$  denklemini çözmeye çalışır. Ancak günümüzde makine öğrenmesi problemlerinde çok sayıda bağımsız olduğu düşünülen değişken bulunmaktadır. Bu durumda çok değişkenli doğrusal regresyon yöntemi bağımsız değişkenleri ile bağımlı y değişkeni arasındaki ilişkiyi modellemeyi amaçlar [21]. Basit bir çok değişkenli doğrusal regresyon modeli y değişkenini doğrusal bir düzlemde düşünür ve  $y = f(x_1, x_2, \dots, x_n) = a_0 + a_1x_1 + a_2x_2 + \dots + a_nx_n$  denklemini çözmeye çalışır ve bunu yaparken eğitim kümesindeki verileri kullanır. Burada  $a_i$ 'ler,  $i=0, 1, \dots, n$  olmak üzere, doğrusal denklemin bilinmeyenlerdir ve gerçek y değerleri ile modelin tahmin ettiği y değerleri arasındaki artık kareler toplamını (AKT) en küçük değerinde tutmayı hedefler. Elimizdeki veriyi X matrisi, tahmin edilecek değerleri y vektörü ve öğrenilecek parametreleri de a vektörü olarak düşünecek olursak problem  $AKT(a) = (y - aX)^T(y - aX)$  denklemini minimize etme problemine dönüşür. Bunun için de AKT'nin a vektörüne göre alınan diferansiyelini sıfırlayan  $a_i$  değerlerinin  $\hat{a} = (X^T X)^{-1} X^T y$  sisteminin çözülerek bulunması gerekir [22].

### **Ortalama Kare Hatası (Mean Squared Error, MSE)**

Ortalama kare hatası bir regresyon probleminde regresyon çizgisinin bir dizi veri noktasına ne kadar yakın olduğunu ölçen metriktir. Karesi alınmış hata kaybının beklenen değerine karşılık gelen bir risk fonksiyonudur. Ortalama kare hatası değerinin büyük olması regresyon problemlerinde tahmin edilen değerlerin istenen noktalardan uzak olduğu manasına gelir ve istenmeyen bir durumdur. Ortalama kare hatası ne kadar küçük olursa tahmin edilen değerlerin veri noktalarına yakın bir şekilde dağıldığını gösterir.

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

Burada y değeri olması gereken değer iken  $\hat{y}$  değeri de makine öğrenmesi algoritmasının tahmin değeridir.

### **Ortalama Mutlak Yüzde Hatası (Mean Absolute Percentage Error, MAPE)**

Ortalama mutlak yüzde hatası bir tahmin sisteminin doğruluğu yüzde olarak ölçer. Ortalama yüzde hatası muhtemelen değişkenin birimlerinin anlaşılmasını kolaylaştıran yüzde birimlerine ölçeklendirildiğinden, hatayı tahmin etmek için kullanılan en yaygın ölçüdür [23].

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \frac{y - \hat{y}}{y} \quad (5)$$

Burada y değeri olması gereken değer iken  $\hat{y}$  değeri de makine öğrenmesi algoritmasının tahmin değeridir.

### **Ortalama Mutlak Hata (Mean Absolute Error)**

Ortalama mutlak hata, gerçek değerler ile model çıktısı olan tahminler arasındaki ortalama farktır.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (6)$$

Burada y değeri olması gereken değer iken  $\hat{y}$  değeri de makine öğrenmesi algoritmasının tahmin değeridir.

### R-Kare (R-Squared)

R-Kare hatası regresyon çizgisinin gerçek değerlere göre toplam varyasyon değerini gösterir yani regresyon modelleri için bir uygunluk ölçüsüdür. En iyi R-kare değeri birdir ve ne kadar yüksek olursa makine öğrenme algoritmasının o kadar iyi tahmin ettiğini gösterir.

$$R - Kare = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (7)$$

Burada y değeri olması gereken değer iken  $\hat{y}$  değeri de makine öğrenmesi algoritmasının tahmin değeri ve  $\bar{y}$  değeri y değerlerinin ortalamasıdır.

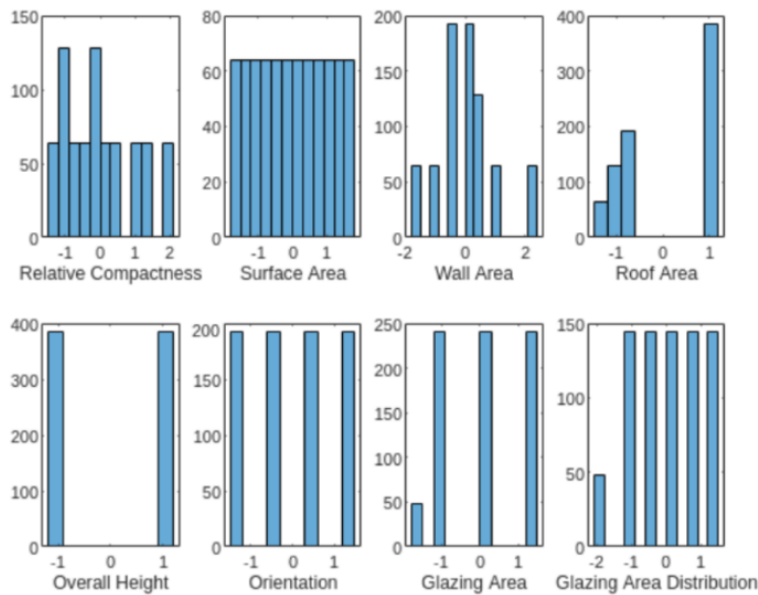
### Enerji Verimliliği Veri Kümesi

Bu çalışmada Kaggle makine öğrenimi ve veri bilimi topluluğundan elde edilen enerji verimliliği veri seti kullanılarak python programlama dili ile tahminleme gerçekleştirilmiştir [24]. Çalışmada kullanılan veri setinde 12 farklı binadan elde edilen toplamda 768 adet veri içermektedir. Veri setinin 8 adet şart niteliği ve ısıtma yükü ve soğutma yükü olmak üzere iki farklı sınıf değeri niteliği bulunmaktadır. Deneysel çalışmalar kısmında şart özellikleri üzerinde herhangi bir özellik seçme yapılmış olup normalizyon veri önışlemi sonucunda elde edilen yeni veriler makine öğrenmesi regresyon verilmiştir. Bütün öznitelikler sayısal değerdedir. Veri kümesine ait olan şart öznitelikleri ve açıklamaları aşağıdaki tabloda gösterilmektedir.

**Tablo 1.** Enerji verimliliği veri kümesine ait öznitelikler

İsim	Açıklama
Bağlı Kompaktlık	Yapının dış mimarisinin bütünlüğünü belirtir
Yüzey Alanı	Yapının dış yüzey alanı
Duvar Alanı	Yapının toplam duvar alanı
Çatı Alanı	Yapının toplam çatı alanı
Toplam Yükseklik	Yapının toplam yüksekliği
Cephe Yönelim	Yapının cephe yönelmesi
Cam Alanı	Yapının toplam cam alanı
Cam Alanı Dağılımı	Yapının cam alanlarının toplam dış alanlarına oranı

Veri Setinin şart özniteliklerinin değer dağılımı Şekil 1'de gösterilmektedir. Veri setinde eksik veri bulunmamaktadır.



**Şekil 1.** Şart özniteliklerin değer dağılımı

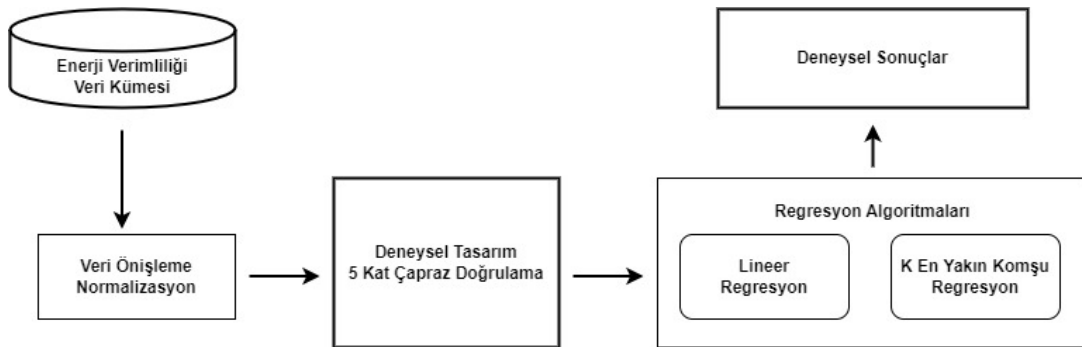
## DENEYSEL ÇALIŞMALAR (EXPERIMENTAL STUDIES)

Deneyisel çalışmalarda KNN algoritmasının performansını karşılaştırmak amacıyla lineer regresyon yöntemi seçilmiştir. Tahminleme işlemi öncesinde enerji verimliliği veri kümesinin şart özellikleri aşağıdaki denklemde verilen Min-Max normalizasyon yöntemine göre [0,1] aralığında normalize edilmiştir.

$$x' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (8)$$

Bu denklemde  $x'$  girdi değerinin 0-1 aralığında normalize edilmiş yeni değerini,  $x_{min}$  enerji verimliliği veri kümesindeki ilgili şart özelliğinin aldığı en küçük değeri ve  $x_{max}$  ise enerji verimliliği veri kümesinde ilgili şart özelliğinin aldığı en büyük değeri göstermektedir. Normalizasyon veri önışlemi sonucunda enerji verimliliği veri kümesindeki bütün şart özelliklerinin değerleri [0,1] arasında normalize edilerek eğitim işlemine bütün şart özelliklerinin aynı oranda etki etmesi sağlanmıştır. Bu sayede denetimli öğrenme algoritmaları daha tutarlı ve sağlam bir şekilde eğitilmişlerdir.

Regresyon algoritmalarının eğitim ve test işlemleri 5 çapraz doğrulama ile yapılmıştır. Eğitim işlemi öncesinde veri kümesi 5 eşit parçaya bölünerek deneyisel çalışmanın her bir iterasyonunda beş eşit parçanın bir tanesi test kümesi ve kalan dört tanesi eğitim kümesi olarak seçilmiştir. 5 iterasyon sonucunda elde edilen performans metriklerinin ortalamaları alınarak deneyisel çalışma tamamlanmıştır. Deneyisel çalışmalarda python programlama dilinde sklearn kütüphanesi kullanılarak gerçekleştirilmiştir. Ayrıca KNN regresyon algoritmasında optimum k değerini belirlemek için 3-11 arasında tek sayılar ile KNN regresyon algoritması çalıştırılmış ve ne başarılı sonucu veren k değeri seçilmiştir. Önerilen sistemin blok diyagramı aşağıdaki şekilde gösterilmektedir.

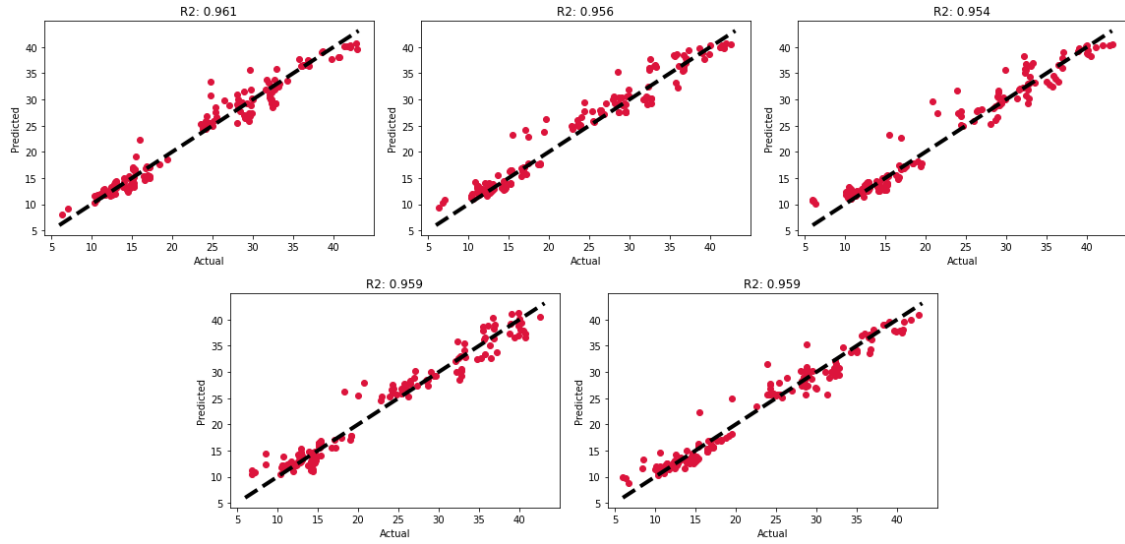


Şekil 2. Önerilen sistemin blok diyagramı

Yapılan deneyisel çalışmalarda ısıtma yükü ve soğutma yükü sınıf değerleri ayrı ayrı tahmin edilmiş olup performans metrik sonuçları ve tahmindeki hataların şekilsel görünüşleri aşağıda gösterilmektedir.

Tablo 2. Isıtma yükü sınıf değeri için deneyisel çalışma sonuçları

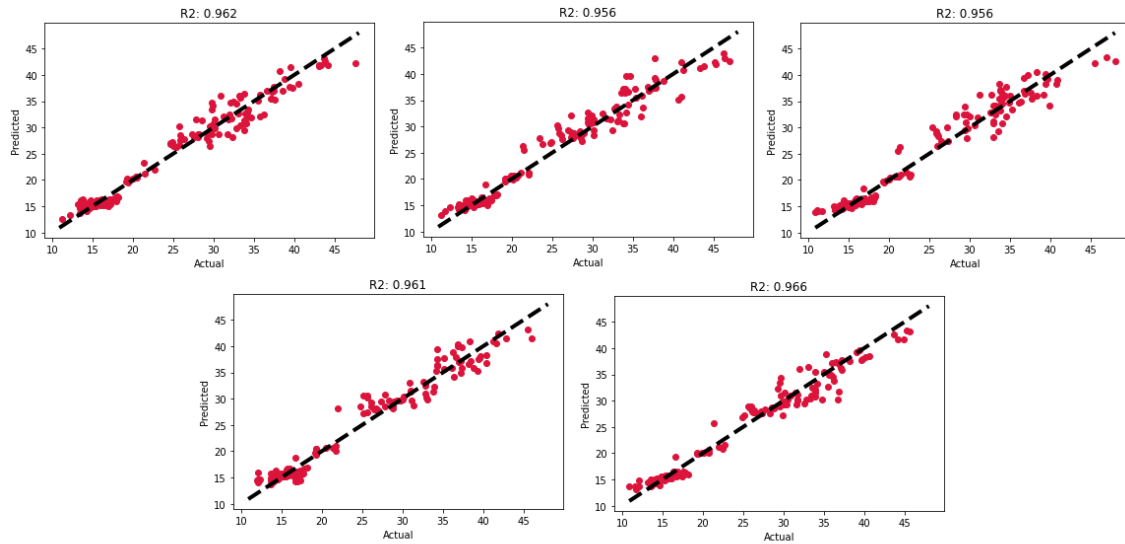
Algoritma	MSE	MAPE	MAE	R-Kare
Lineer Regresyon	8,654	0,099	2,09	0,914
KNN Regresyon (k=3)	4,882	0,092	1,717	0,952
KNN Regresyon (k=5)	<b>4,301</b>	<b>0,084</b>	<b>1,529</b>	<b>0,958</b>
KNN Regresyon (k=7)	5,553	0,096	1,792	0,945
KNN Regresyon (k=9)	6,734	0,108	2,003	0,934
KNN Regresyon (k=11)	6,94	0,109	2,009	0,932



Şekil 3. Isıtma yükü sınıf değeri için  $k=5$  değeri ile 5 çapraz doğrulama tahmin hatalarının şekilsel görünümü

Tablo 3. Soğutma yükü sınıf değeri için deneysel çalışma sonuçları

Algoritma	MSE	MAPE	MAE	R-Kare
Linear Regresyon	10,259	0,089	2,253	0,886
KNN Regresyon (k=3)	3,836	0,064	1,514	0,957
KNN Regresyon (k=5)	<b>3,569</b>	<b>0,06</b>	<b>1,426</b>	<b>0,96</b>
KNN Regresyon (k=7)	4,721	0,069	1,642	0,948
KNN Regresyon (k=9)	5,488	0,075	1,801	0,939
KNN Regresyon (k=11)	5,776	0,077	1,843	0,936



Şekil 4. Soğutma yükü sınıf değeri için  $k=5$  değeri ile 5 çapraz doğrulama tahmin hatalarının şekilsel görünümü

DeneySEL sonuçlar incelendiği zaman hem ısıtma yükü hem de soğutma yükü tahminleri için en iyi performansı KNN algoritmasının  $k=5$  değeri için gerçekleştirdiği görülmektedir. Isıtma yükü ve soğutma yükü sınıf değerleri için KNN algoritması sırasıyla 0.958 ve 0.96 olmak üzere çok yakın değerler elde etmiş ve tahminleme açısından başarılı bir performans sergilemiştir. Ayrıca enerji tahmini problemi için KNN algoritması ile lineer regresyon algoritması karşılaştırıldığında ısıtma yükü ve soğutma yükü sınıf değerleri için

KNN algoritmasının net bir şekilde daha başarılı olduğu gözlenmektedir. Gerçekleştirilen çalışmadan elde edilen KNN modelinin literatürdeki benzer çalışmaların KNN algoritması sonuçları ile performans karşılaştırılması aşağıdaki tabloda gösterilmektedir.

**Tablo 4.** Isıtma ve soğutma yükü sınıf değeri deneysel çalışma sonuçlarının literatür ile karşılaştırılması

Çalışma	Isıtma				Soğutma			
	MSE	RMSE	MAE	R-Kare	MSE	RMSE	MAE	R-Kare
Bu çalışma	<b>4,301</b>	<b>0,084</b>	1,529	<b>0,958</b>	<b>3,569</b>	<b>1,889</b>	1,426	<b>0,96</b>
[16]	4,665	2,160	1,512	-	5,189	2,278	<b>1,313</b>	-
[18]	5,360	2,310	<b>1,460</b>	0,92	6,250	2,500	1,610	0,930

Yapılan çalışmada geliştirilen modelin sonuçları literatürde yapılan benzer çalışma sonuçları ile karşılaştırıldığında MSE, RMSE ve R-Kare değerlerinde performans olarak daha başarılı olduğu görülmektedir. Bunun gerçekleşmesinde hem veri ön işleme aşamasında veri kümesine normalizasyon işleminin yapılması hem de KNN regresyon algoritmasının doğru parametreler ile çalıştırılması yer almaktadır.

### SONUÇ (CONCLUSION)

İnsan yaşamının vazgeçilmez bir parçası olan enerjinin verimli kullanılması günümüzde üzerinde en çok çalışma yapılan güncel konulardandır [25]. Günümüzde binaların ısıtılması ve soğutulması için büyük miktarlarda enerji tüketilmektedir. Bir binanın ısıtma veya soğutma yükü, genellikle birim alana (metrekare) düşen ısıtma veya soğutma ihtiyacı olarak ifade edilir. Özellikle soğuk ve sıcak iklimlerde ısıtma ve soğutma yükleri çok önemli olup doğru bir şekilde hesaplanması, enerji tasarrufu ve maliyet azaltımı açısından büyük önem taşır. Bu çalışmada binaların ısıtma yükü ve soğutma yükü tahmini için denetimli makine öğrenmesi algoritmalarından KNN üzerine bir çalışma gerçekleştirilmiştir. Yapılan deneysel çalışmalar sonunda k=5 değeri için KNN algoritması yaklaşık %96 oranında ısıtma ve soğutma yüklerini doğru tahmin etmiştir. Tahmin edilen bu değerler geliştirilen modelin gerçek dünya ortamında kullanılabilmesi için kabul edilebilir bir hata oranına sahiptir. Bundan sonraki çalışmalarda daha kapsamlı bir veri kullanılarak farklı makine öğrenmesi algoritmalarının veya algoritmaları hibrit kullanımlarının gerçekleştirilmesi önerilmektedir.



**KAYNAKÇA (REFERENCES)**

- [1] C. Janiesch, P. Zschech, K. Heinrich, Machine learning and deep learning. *Electron Market*. 31 (2021), 685–695. doi:10.1007/s12525-021-00475-2
- [2] D. van de Sande, , et al. Developing, implementing and governing artificial intelligence in medicine: a step-by-step approach to prevent an artificial intelligence winter, *BMJ Health & Care Informatics*. 29(1) (2022). doi:10.1080/13645706.2019.1575882
- [3] H. Hozhabr Pour, et al. A machine learning framework for automated accident detection based on multimodal sensors in cars, *Sensors*. 22(10) (2022). doi:10.3390/s22103634
- [4] I. Lauriola, A. Lavelli, F. Aioli, An introduction to deep learning in natural language processing: Models, techniques, and tools. *Neurocomputing*. 470 (2022), 443-456. doi:10.1016/j.neucom.2021.05.103
- [5] X. Liang, J.A. Batsis, Y. Zhu, T.M. Driesse, R.M. Roth, , D. Kotz, B. MacWhinney, Evaluating voice-assistant commands for dementia detection. *Computer Speech & Language*. 72 (2022). doi:10.1016/j.csl.2021.101297
- [6] Y. Gong, Y. Otomo, H. Igarashi, Sensorless metal object detection for wireless power transfer using machine learning. *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*. 41(3) (2022), 807-823. doi:10.1108/COMPEL-03-2021-0069
- [7] M. Bansal, A. Goyal, A. Choudhary, A comparative analysis of K-Nearest Neighbour, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning. *Decision Analytics Journal*. 3 (2022). doi:10.1016/j.dajour.2022.100071
- [8] H. Dabiri, V. Farhangi, M.J. Moradi, M. Zadehmohamad, M. Karakouzian,. Applications of Decision Tree and Random Forest as Tree-Based Machine Learning Techniques for Analyzing the Ultimate Strain of Spliced and Non-Spliced Reinforcement Bars. *Applied Sciences*. 12(10) (2022). doi:10.3390/app12104851
- [9] C. Bogdal, R. Schellenberg, O. Höpli, M. Bovens, M. Lory , Recognition of gasoline in fire debris using machine learning: Part I, application of random forest, gradient boosting, support vector machine, and naïve bayes. *Forensic Science International*. 331 (2022). doi:10.1016/j.forsciint.2021.111146
- [10] A. Pektaş, O. İnan, Ağaç Tohum Algoritmasının Kümeleme Problemlerine Uygulanması, *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*. 4(1) (2022), 1-10.
- [11] M. Karakoyun, A. Özkış, Transfer Fonksiyonları Kullanarak İkili Güve-Alev Optimizasyonu Algoritmalarının Geliştirilmesi ve Performanslarının Karşılaştırılması, *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 3(2) (2021), 1-10.
- [12] <https://www.dunyaenerji.org.tr/wp-content/uploads/2019/11/21112019Sunum.pdf> (erişim 28 Temmuz 2023).
- [13] M. E. Yıldız, F. Beyhan, M.K. Uçar, Enerji verimli bina tasarımında kural tabanlı yöntem yardımıyla ısıtma ve soğutma yüklerinin tahminlemesi, *Online Journal of Art and Design*, 10(4) (2022).
- [14] P. Canbay, H. Taş, Yapıların Isıtma ve Soğutma Yükünün Yapay Zeka ile Tahmini, *International Journal of Pure and Applied Sciences*. 8(2) (2022), 478-489. doi:10.29132/ijpas.1166227
- [15] B. Jia, D. Hou, A. Kamal, I.G. Hassan, L. Wang, Developing machine-learning meta-models for high-rise residential district cooling in hot and humid climate. *Journal of Building Performance Simulation*. 15(4) (2022), 553-573. doi:10.1080/19401493.2021.2001573
- [16] M. Peker, O. Özkaraca, , B. Kesimal, Enerji tasarruflu bina tasarımı için ısıtma ve soğutma yüklerini regresyon tabanlı makine öğrenmesi algoritmaları ile modelleme. *Bilişim Teknolojileri Dergisi*. 10(4) (2017), 443-449. doi:10.17671/gazibtd.310154
- [17] B. Dilber, A.F. Özdemir, Enerji Verimli Binaların Isıtma ve Soğutma Yüklerinin Düzenli Modeller ile Tahmin Edilmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*. 10(3) (2022), 1091-1106. doi:10.29130/dubited.939933
- [18] D. Bayraktar, E. Bayraktar, Mevcut Binalarda Isı Yalıtımı Uygulamalarının Değerlendirilmesi. *Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi*. 7(1) (2016), 59-66.

- [19] S. Shabani, S. Samadianfard, M. T. Sattari, A. Mosavi, S. Shamshirband, T. Kmet, A. R. Várkonyi-Kóczy, Modeling pan evaporation using Gaussian process regression K-nearest neighbors random forest and support vector machines; comparative analysis. *Atmosphere*. 11(1) (2020). doi:10.3390/atmos11010066
- [20] G.A. Seber, A. J. Lee, Linear regression analysis John Wiley & Sons, 329, (2012).
- [21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, E. Duchesnay, Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research*. 12 (2022), 2825-2830.
- [22] T. Hastie, R. Tibshirani, J. H. Friedman, Friedman, The elements of statistical learning: data mining, inference, and prediction, New York: springer, 2 (2009), doi:10.1007/978-0-387-21606-5
- [23] [https://docs.oracle.com/en/cloud/saas/planning-budgeting-cloud/pfusu/insights\\_metrics\\_MAPE.html](https://docs.oracle.com/en/cloud/saas/planning-budgeting-cloud/pfusu/insights_metrics_MAPE.html) (erişim 31 Mayıs 2023)
- [24] <https://www.kaggle.com/datasets/elikplim/eergy-efficiency-dataset> (erişim 31 Mayıs 2023)
- [25] S. Ata, M.E. Boyacıoğlu, R. Şahin, A. Kahraman, ORÇ ile Düşük Sıcaklıklı Isı Kaynaklarından Elektrik Üretilmesinde Islak ve Yeni Nesil Akışkanların Çevresel ve Termodinamik Performanslarının Karşılaştırılması. *Necmettin Erbakan Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 2(1) (2021), 1-13.