

## Makine Öğrenmesi Algoritması ile Çanakkale İlinin Güneş Işınım Şiddetinin Analizi

Necla TEKTAŞ<sup>1\*</sup>, Ebru KORKMAZ<sup>2</sup>

<sup>1</sup>Yöneylem Anabilim Dalı/İİBF, Bandırma Onyedi Eylül Üniversitesi, Türkiye

<sup>2</sup>Ekonometri/Sosyal Bilimler Enstitüsü, Bandırma Onyedi Eylül Üniversitesi, Türkiye

(\*ntektas@bandirma.edu.tr)

**Özet** – Geçmişten günümüze ülkelerin gelişimi için en önemli etmenlerden biri enerjidir. Yenilenebilir ve yenilenemez enerji kaynakları bakımından zengin olan ülkeler hem sanayi de hem de turizm de gelişmişlerdir. Ülkemiz güneş, rüzgâr ve biyokütle enerjisi gibi yenilenebilir enerji kaynaklarına uygun bir bölgede yer almaktadır. Bu nedenle son yıllarda güneş enerjisi, rüzgâr enerjisi, biokütle enerjisi birçok yatırımcının ilgi odağı olmuştur. Yatırımların doğru yerde olması için araştırma, tahmin, sınıflandırma gibi istatistiksel analizler yaparak uygun bölge ve illeri belirlenmesi önemlidir. Bu çalışmada Çanakkale iline ait güneş enerjisine ait güneş ışınım şiddeti için makine öğrenmesi yöntemleri ile yatırımcılar için ön çalışma amaçlanmıştır. Bu nedenle 2015-2019 yılları için güneş enerjisinden elde edilen güneş ışınım (günlük/saatlik) verileri kullanılarak Matlab programı aracılığı ile makine öğrenmesi algoritmasında sınıflama analizi dört temel algoritma ile Karar Ağaçları, Standartlaştırılmış Değişken Mesafeler, En Yakın Komşu Algoritması, Naive Bayes yapılmıştır. Yapılan analiz sonucuna göre güneş ışınım verilerinde makine öğrenmesi algoritmasının sınıflama yöntemlerinden karar ağaçları algoritması ile daha iyi sonuçlar elde edilmiştir.

*Anahtar Kelimeler – Yenilenebilir Enerji, Güneş Işınım Miktarı, Makine Öğrenmesi, Sınıflama.*

### I. GİRİŞ

Yenilenebilir enerji kaynakları, sürdürülebilirliği ve doğada sınırsız varlığı nedeniyle günümüzde en çok araştırılan çalışmalar yapılan konulardan olmuştur. Yenilenebilir enerji kaynakları başlıca; güneş enerjisi, jeotermal enerji, rüzgâr enerjisi, hidrojen enerjisi, hidrolik enerji, deniz kökenli enerjiler ve biyokütle enerji gibi enerji çeşitleri altında incelenmektedir [1].

Güneş enerjisi ile ilgili çalışmalar çok eski yıllardan başlamış ve teknolojik gelişmelerle birlikte hız kazanmıştır. Güneş ışınım miktarı dünyamızın şekinden dolayı coğrafi şekiller ve konumlar yerkürenin her bölgesine ve her ülkesine aynı açı ve düzlemde gelmediğinden farklılık göstermektedir. Dünyanın bazı bölgeleri Afrika'nın büyük bir kısmı, Avustralya ve Amerika'nın orta

bölgesi en çok güneşi gören bölge olarak bilinmektedir ve dünyanın diğer bölgelerine göre güneş bakımından en bol olan bölgedir. Ülkemizde m<sup>2</sup>'ye yıllık 1500–2000 kWh civarında güneş ışınımın ulaştığı ve ortalamanın üstünde şanslı yerlerdendir. Dünyada güneş ışınımı kullanımlarında birinci sırada Almanya, ikinci sırada İtalya gelmektedir. Ülkelerin güneş enerjilerine yönelmelerinin en önemli nedenlerinden bir diğeri sera gazlarının salınımını azaltmaktır. Bu konuda çalışma yapan ülke Japonya'dır. Ülkenin genelinde pek çok fabrikada güneş enerji üretimi yapılmaktadır. Japonya, İspanya ve Amerika Birleşik Devletleri takip etmekte, yaptıkları yatırımlar ve çok çalışmaları öne çıkmaktadırlar [2].

Bu çalışmada da 2015-2019 yıllarına ait Çanakkale ilinin (günlük/saatlik) güneş verileri Meteoroloji Genel Müdürlüğü'nden alınarak Makine Öğrenmesi Algoritması ile sınıflama analizi MATLAB programı ile yapılmıştır.

## II. MATERYAL VE YÖNTEM

Makine öğrenmesi (MÖ), bilgisayarlara açık bir biçimde programlamadan öğrenme yeteneği kazandıran bir yapay zekâ türüdür. Matematiksel ve istatistiksel yöntemleri kullanarak mevcut verilerden çıkarımlar yapan ve bu çıkarımlarla bilinmeyene ilişkin tahminleme, kümeleme ve sınıflama yapan metod paradigmasıdır [4].

MÖ algoritmaları temel olarak beş farklı başlık altında değerlendirilmektedir. Bunlar;

- Danışmanlı Öğrenme (Supervised Learning)
- Danışmansız Öğrenme (Unsupervised Learning)
- Yarı Danışmanlı Öğrenme (Semi-Supervised Learning)
- Takviyeli Öğrenme (Reinforcement Learning)
- Yoğun Öğrenme (Intensive Learning)

Danışmanlı öğrenme yönteminde, bir uzman tarafından oluşturulan örnek girdiler ve buna karşılık beklenen çıktılar modele sunulur. Modelin eğitim süreci, eğitim verilerinde istenen doğruluk seviyesine ulaşana kadar devam eder. Daha sonra modele yeni bir girdi sunulur tahmin edilmesi sağlanır. Danışmansız öğrenme yönteminde, modelin öğrenme sürecinde belirlenen örnek girdiler sunulur fakat herhangi bir beklenen çıktı(etiket) verilmez. Bu sebeple, girdi verisinin hangi sınıfa ait olduğu belli değildir ve girdi parametreleri arasındaki ilişkileri bizatihi öğrenmesi beklenir. Yarı danışmanlı öğrenme yöntemi ise danışmanlı ve danışmansız öğrenme yöntemlerinin arasındadır. Çünkü yöntem eğitim için hem etiketli hem de etiketlenmemiş verileri kullanır. Bu yöntemi kullanan sistemler öğrenme doğruluğunu büyük bir oranla iyileştirebilir. Takviyeli öğrenme yönteminde, eylemler üreterek çevresiyle etkileşime geçen olumlu veya olumsuz durumları keşfeden bir yöntemdir. Bu yöntemde bir danışman/eğitici bulunur fakat danışmanlı öğrenme modellerindeki gibi sisteme çok detay vermez veya müdahale etmez öğrenme olarak da bilinmektedir [4]. Hiyerarşik öğrenme olarak da bilinmektedir. Bu öğrenme yöntemi derin grafiklerde birçok doğrusal ve doğrusal olmayan dönüşümlerden ve çoklu işlem

katmanlarından oluşturulmuş verilerde. üst düzey soyutlamalar kullanılarak elde edilen model girişlerine dayalı bir dizi algoritmalarla geliştirilmiş makine öğrenmesidir [5].

Makine Öğrenimindeki sınıflandırma, kesikli veya sürekli özellik değerlerinin bir vektörünü girdi (özellik) ve tek bir ayrık değer olan sınıfı çıktı(etiket) bir sistemdir. Bir dizi öğrenim kategorisini veya sınıfını tahmin etmek için sınıflandırma algoritmaları kullanılır. Sınıflandırma algoritmaları:

- K-Nearest Neighbors (K-En Yakın Komşular)
- Naive Bayes
- Decision Trees (Karar Ağaçları)
- Standardized Variable Distances (SVD)(Standartlaştırılmış Değişken Mesafeler)
- Support Vector Machine (Destek Vektör Makinesi)
- Random Forest(Rastgele Orman) gibi algoritmalarıdır.

Bu çalışmada beş katlımanlı çapraz doğrulama tekniği kullanılmıştır. Bu teknik, literatür de k-fold cross validation olarak bilinmektedir. Makine öğrenme modellerinin sınıflandırma başarılarının değerlendirilmesi amacıyla kullanılır.

### Makine Öğrenmesi Doğrulama Yöntemleri

- Çapraz Doğrulama Yöntemi

Çapraz doğrulama, bir MÖ modelinde yapılan testin hatasını daha iyi tahmin edebilmek için model yapısı seçiminde kullanılan bir tekniktir. Çapraz doğrulamanın ana yapısı, eğitim verileri setinden doğrulama yani test kümeleri olarak bilinen örnek gözlem kümelerini oluşturmaktır. Modelde eğitim verilerini yerleştirdikten sonra performans her yeni test kümesinin nasıl performans göstereceğine ilişkin daha iyi bir sonuç elde etmektir.

Çapraz doğrulama adını aynı veri üzerinde örneklerin farklı şekillerde gruplanmasından yani çapraz bir şekilde alınarak grupların oluşmasından alır. Çaprazlanan gruplar üzerinde yapılan sınıflama sonuçları birleştirilerek modelin genel tahmin başarısı ölçülür. Çapraz doğrulama yöntem olarak ikiye ayrılmaktadır.

- i. K-Parçalı Çapraz Doğrulama

K-parçalı doğrulama yöntemi, toplam n örneğin bulunduğu bir veri setinde her birinde n/k örneğin bulunduğu k adet farklı kümeye parçaya ayrılır. Her

defasında farklı bir veri seti kümesi test için ayrılarak geriye kalan diğer k-1 veri seti eğitim için kullanılır. Her defasında test kümesi değiştirilerek sınıflandırıcı k adet eğitilir. Böylelikle elde edilen k adet hatanın ortalaması ile sınıflandırıcı performansı tahmin edilmiş olur [6].

## ii. Birisi Dışarıda Çapraz Doğrulama Yöntemi

Bu yöntemde, n adet örneğin bulunduğu veriler n parçaya kadar bölünür, n parçaya bölünmüş veri setinde her seferinde bir örnek dışarıda kalacak biçimde geriye kalan tüm veriler (n-1) ile sınıflandırıcı eğitilir ve bu eğitime tüm örnekler için tekrarlanır. Gerçekte örnek sayısı kadar k parçaya ayrılan k-Parçalı Çapraz Doğrulama Yöntemi gibi düşünülmektedir [6].

### III. BULGULAR

Bu çalışmada kullanılan veriler Ankara Meteoroloji Genel Müdürlüğü'nden Bandırma Onyedü Eylül Üniversitesi Sosyal Bilimler Enstitüsü'nün onayı ve izni ile alınmıştır. Veri seti Türkiye'nin Marmara Bölgesi'nde bulunan güneş ışınımı ölçümü yapan istasyonların verisinden oluşmaktadır. Çanakkale iline ait olan veriler üzerinde analiz yapılmıştır. Çalışmada, 2015-2019 yıllarına ait günlük/saatlik güneş ışınım, toplam yağış, nispi nem, buharlaşma, saatlik güneşlenme süresi ve aktüel basınç değişkenleri üzerinde analiz yapılmıştır. Çalışmanın modelinde bağımlı değişken(Y) olarak günlük/saatlik güneş ışınımı, bağımsız değişken olarak da toplam yağış (X1), nispi nem(X2), buharlaşma (X3), saatlik güneşlenme süresi (X4) ve aktüel basınç(X5) verileri setleri kullanılmıştır.

Meteoroloji Genel Müdürlüğü'nden alınan veriler ilk olarak veriler analize uygun duruma getirilip düzenlenmiştir. Sonra verilerin normal dağılıma uygunluğu araştırılmıştır. Sonuç olarak verilerin normal dağılmadığı tespit edilmiş ve verilere normalizasyon işlemi uygulanmıştır.

Güneş ışınımı, ampirik ölçümlerde 0 ile 1 arasında değişmektedir. Bu veri setlerini bir sınıflandırma problemine dönüştürmek için belirli ölçüm aralıklarının sınıf etiketlerine dönüştürülmesi gerekmektedir. Çanakkale veri setlerindeki güneş ışınımının (y), sınıf etiketlerine ayrıştırılması Tablo 1'de görülmektedir.

Tablo 1. Güneş Radyasyonunun Sınıf Etiketlerine Ayrıştırılması

Sınıf ID	Radyasyon Aralığı
1	$0 \leq y < 0.1$
2	$0.1 < y < 0.2$
3	$0.2 < y < 0.3$
4	$0.3 < y < 0.4$
5	$0.4 < y < 0.5$
6	$0.5 < y < 0.6$
7	$0.6 < y < 0.7$
8	$0.7 < y < 0.8$
9	$0.8 < y < 0.9$
10	$0.9 < y \leq 1.0$

Çanakkale ilinin makine öğrenmesinin sonuçlarına göre Tablo 2'de gösterilmiştir.

Tablo 2. Çanakkale Makine Öğrenmesi Algoritmasının Karşılaştırması

Yöntemler	Doğruluk
Karar Ağaçları	0,8585**
Standartlaştırılmış Değişken Mesafeler	0,8368
En Yakın Komşu Algoritması	0,8545
Naive Bayes	0,8517
	<b>Duyarlılık</b>
Karar Ağaçları	0,2885**
Standartlaştırılmış Değişken Mesafeler	0,2134
En Yakın Komşu Algoritması	0,2686
Naive Bayes	0,2305
	<b>Belirginlik</b>
Karar Ağaçları	0,9202**
Standartlaştırılmış Değişken Mesafeler	0,9095*
En Yakın Komşu Algoritması	0,9183*
Naive Bayes	0,9159*
	<b>Hassaslık</b>
Karar Ağaçları	0,3063**
Standartlaştırılmış Değişken Mesafeler	0,1714
En Yakın Komşu Algoritması	0,2742
Naive Bayes	0,249
	<b>F-skor</b>
Karar Ağaçları	0,2842**
Standartlaştırılmış Değişken Mesafeler	0,1684
En Yakın Komşu Algoritması	0,2691
Naive Bayes	0,1866

\*Makine Öğrenmesi Ölçeklendirme

\*\* Makine Öğrenmesi Sınıflama

Tablo 2'de Çanakkale veri setinin MÖ'si analizleri görülmektedir. Analiz sonuçları birbirlerine yakın çıkmıştır. En iyi sınıf modellemesini karar ağaçları algoritması ile ulaşılmıştır.

Tablo 3:  $0 \leq y < 0.1$  Sınıf Aralığının Çanakkale İlinin Makine Öğrenmesinin Karşılaştırılması

Yöntemler /Metrikler	Karar Ağaçları	Standartlaştırılmış Değişken Mesafeler	En Yakın Komşu Algoritması	Naive Bayes
Doğruluk	<b>0,8966</b>	0,8471	0,8545	0,8517
Duyarlılık	0,2683	<b>0,3127</b>	0,2686	0,2305
Belirginlik	<b>0,9416</b>	0,8856	0,9183	0,9159
Hassaslık	0,2453	0,1609	<b>0,2742</b>	0,2490
F-Skor	0,2456	0,2114	<b>0,2692</b>	0,1866

Tablo 1’de verilen etiketlerin herbiri için MÖ yöntemleri analiz yapılmış tüm analizlerde aynı sıralama elde edilmiştir. Burada sadece örnek olması nedeniyle  $0 \leq y < 0.1$  sınıf aralığı için olan analiz sonucu yeralmaktadır.

Doğruluk ölçütünün sonuçları incelendiğinde karar ağaçları 0,8966, duyarlılık ölçütü için 0,3127 Standartlaştırılmış Değişken Mesafeler, belirginlik ölçütü sonuçları 0,9416 değeri karar ağaçları hassaslık ve F-skor ölçütü için en iyi sonuçları sonuçları 0,2742 ve 0,2692 değer ile En Yakın Komşu Algoritması ile elde edilmiştir.

#### IV. SONUÇLAR

Güneş enerjisi, evrende en çok bulunan yenilenebilir enerji kaynağı olmasının yanı sıra çevre dostu, temiz, sınırsız ve işletme maliyetleri düşük bir enerji kaynağıdır. Ülkelerin gelişimi, enerjiye olan arz ve talebin artışıyla yenilenemez enerji kaynaklarının tükenmesinden dolayı güneş enerjisi kullanımı gün geçtikçe rağbet gören bir enerji türü konumundadır. Bu kapsamda yapılan girişimler, çalışmalar insan yararına olan atılımlar güneş enerjisinin vazgeçilemez bir konumda olduğunu göstermektedir. Güneş enerjisinin potansiyelini belirlemek için yapılan çalışmalarda güneş ışınım şiddetinin miktarı önemli bir parametre olarak karşımıza çıkmaktadır. Günümüz yaşam koşullarının geliştiği ve enerji kaynaklarının tüketildiği göz ardı edilemeyecek bir gerçektir. Bu farkındalıkla birlikte insanoglu sürdürülebilir nitelikte yenilenebilir enerji kaynakları arayışlarına girmektedir.

Enerji kaynakları değerlendirilirken, çeşitli kriterler göz önünde bulundurulmalıdır. Enerji teknolojilerinin değerlendirilmesinde, ülke yapısına ilişkin kriterler, teknolojik kriterler, ekonomik kriterler ve sosyo-ekonomik kriterler ayrı ayrı incelenmelidir.

Bu çalışmada, Çanakkale ilinin sahip olduğu güneş enerji potansiyeli araştırılmış ve güneş ışınım miktarının tahminlenmesi Makine Öğrenmesi (MÖ) ile yapılmıştır. Analizde Ankara Meteoroloji Genel Müdürlüğü’nden Çanakkale iline ait 2015-2019 yıllarına ait beş yıllık günlük/saatlik güneş ışınım miktarını tahminlemek için gerekli olan verileri kullanılmış ve modele dâhil edilen 1818 adet verinin istatistik analizleri MATLAB paket programı aracılığıyla yapılmıştır. Burada değişkenler toplam yağış, nispi nem, buharlaşma saatlik, güneşlenme süresi ve aktüel basınç olarak belirlenmiştir.

Güneş ışınım miktarının ölçülmesi ile ilgili deneye dayalı ölçümlerde 0 ile 1 arasında değişmektedir. Bu veri setlerini bir sınıflandırma problemine dönüştürmek için belirli ölçüm aralıklarının sınıf etiketlerine dönüştürülmesi gereklidir. Bu nedenle MÖ algoritması ile analiz edebilmek için ilk olarak ölçüm aralıkları sınıf etiketlerine çevrilmiştir. Dönüştürülen sınıf etiketine göre sınıflama analizlerinden K-Nearest Neighbors (K-En Yakın Komşular, Naive Bayes, Standardized Variable Distances (SVD)(Standartlaştırılmış Değişken Mesafeler ve Decision Trees (Karar Ağaçları) yöntemleri kullanılmıştır. Analiz sonuçlarına göre. Çanakkale verilerinin sınıflandırılmasında belirginlik ölçütü ile daha iyi sınıflama yapılabildiği görülmüştür. Bu sonuçlar doğrultusunda, Çanakkale güneş ışınım verilerinde makine öğrenmesi algoritmasının sınıflama yöntemlerinden Decision Trees (Karar Ağaçları) algoritmasıyla en iyi sonuca ulaşılmıştır.

#### TEŞEKKÜR

Çalışmamın hazırlanma sürecinde desteğini esirgemeyen Doç. Dr. Necla TEKTAŞ ve Prof. Dr. Mehmet TEKTAŞ ve Doç.Dr. Abdullah ELLEN’e teşekkür ve saygılarımı sunarım.

#### KAYNAKLAR

- [1] Önal, E., & Yarbay, R. Z. (2010). Türkiye’de Yenilenebilir Enerji Kaynakları Potansiyeli ve Geleceği. In İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi (Vol. 9, Issue 18).
- [2] URL1. (2021). Türkiye’de Güneş Enerjisi Potansiyeli. <https://www.incitas.com.tr/bilgi-merkezi/blog/turkiyede-gunes-enerjisi-potansiyeli>
- [3] Özen, İ. (2021). Dünyada Güneş Enerjisi Üretimi ve Kullanımı. <https://www.enerji.gen.tr/dunyada-gunes-enerjisi.html>
- [4] Avuçlu, E., & Elen, A. (2020). Makine Öğrenme Algoritmasına Genel Bakış. Geleceğin Dünyasında

Bilimsel ve Mesleki Çalışmalar 2020 Bilgisayar Mühendisliği (Birinci Ba, pp. 36–37).

- [5] Alkan, M. A. (2017). Makine Öğrenimi Nedir? 2017.
- [6] Narin, A., İşler, Y., & Özer, M. (2014). Çapraz Doğrulama Yöntemlerinin Sınıflandırıcı Karşılaştırılması (Comparison of the Effects of Cross Validation Methods on Determining Performances of Classifiers Used in Diagnosing Congestive Heart Failure). 1–8.