Uluslararası Mühendislik Araştırma ve Geliştirme Dergisi
International Journal of Engineering Research and Development
UMAGD, (2019) 11(3), 769-779.
10.29137/umagd.514933
Cilt/Volume:11 Sayı/Issue:3 Aralık/December 2019
Araştırma Makalesi / Research Article

PV Güç Santrallerinden Elde Edilecek Enerjinin Makine Öğrenmesi Metotları Kullanılarak Tahmin Edilmesi

Estimation of Energy to be Obtained from PV Power Plants Using Machine Learning Methods

Sinan Uğuz[1], Okan Oral [2], Nuri Çağlayan [2]

[1]Isparta Uygulamalı Bilimler Üniversitesi, Bilgisayar Mühendisliği Bölümü, 32260, Isparta, TÜRKİYE
[2] Akdeniz Üniversitesi, Mekatronik Mühendisliği Bölümü, 07070, Antalya, TÜRKİYE

***Başvuru / Received**: 18/01/2019* ***Kabul / Accepted**: 28/09/2019* ***Son Versiyon/Final Version**: 31/12/2019*

Öz

Elektrik enerjisi ihtiyacının artması, ülkeleri güvenilir, ucuz ve temiz enerji teminine yöneltmiştir. Son zamanlarda bu enerji kaynakları arasından fotovoltaik (PV) güç sistemleri öne çıkmaktadır. Güneş enerjisi potansiyelinin yüksek olduğu Türkiye'de devlet teşvikleriyle birlikte PV güç santrallerine olan yatırımların sayısı artmaktadır. PV santrallerin kuruluş yeri seçimi için fizibilite çalışmalarının yapılması ve buna bağlı olarak sistemlerin tasarlanması, yapılacak yatırımların ekonomikliğinin belirlenmesi açısından önemli bir konu olarak görülmektedir. Santral kurulmadan önce ışınım enerjisine göre elde edilebilecek elektrik enerjisinin hesaplanması için eşitlikler ve yöntemler geliştirilmiştir. Bu yöntemlerden biri de makine öğrenme modellerinin geliştirilmesi ve simülasyon sonuçlarının elde edilmesidir. Bu çalışmada; Türkiye'de 125 farklı bölge için kurulması planlanan PV santrallerinin üreteceği elektrik gücünün, makine öğrenmesi modelleri ile tahmin edilmesi amaçlanmıştır. Bu amaç doğrultusunda PV sistemler için güneş ışınımı tahmin edilmesinde Yapay Sinir Ağları (YSA), Çoklu Doğrusal Regresyon (ÇDR) ve k-En Yakın Komşu Regresyon (k-EYKR) makine öğrenimi metodolojileri kullanılmıştır. Bu metodolojilerin performansını analiz etmek amacıyla bir dizi deneysel değerlendirmeler yapılmıştır. Değerlendirmeler için veri seti, Numpy, Pandas, Scipy gibi temel python kütüphanelerinin yanı sıra makine öğrenmesi uygulamaları için geliştirilmiş olan scikit-learn kütüphanesinde test edilmiştir. Deneysel sonuçlar, girdi olarak kullanılmış yedi adet bağımsız değişkenin, makine öğrenimine dayalı tahmin algoritmalarının çalıştırılmasıyla PV tarafında üretilen elektrik gücünü tahmin edebildiğini göstermiştir.

Anahtar Kelimeler

"Yenilenebilir enerji, makine öğrenmesi, çoklu doğrusal regresyon, k-en yakın komşu regresyon, yapay sinir ağları"

Abstract

The increase in the need for electricity has led the countries to provide reliable, inexpensive and clean energy. Recently, among those energy sources, those based on photovoltaic (PV) power systems have come forward. Solar energy potential in Turkey is high, the number of investments in PV power plants with government incentives are increasing. The feasibility studies for the selection of the location of PV plants and the design of the systems are considered as an important issue in terms of determining the economic value of the investments to be made. Equations and methods have been developed to calculate the electrical energy that can be obtained according to the radiant energy before the plant is established. One of these methods is to develop machine learning models and to obtain simulation results. In this study; The establishment of the 125 planned for different areas of electric power to be generated by the PV power plant in Turkey, aimed to estimate with machine learning models. For this purpose, in the estimation of solar radiation for PV systems; artificial neural networks (ANN), multiple linear regression (MLR) and k-nearest neighbors regression (KNNR) machine learning methodologies were used. In order to analyze the performance of these methodologies, a series of experimental evaluations were made. The data set for evaluations has been tested in basic python libraries such as Numpy, Pandas, Scipy, as well as in the scikit-learn library developed for machine learning applications. Experimental results have shown that seven independent variables used as input can predict the electrical power produced by PV based on the study of prediction algorithms based on machine learning.

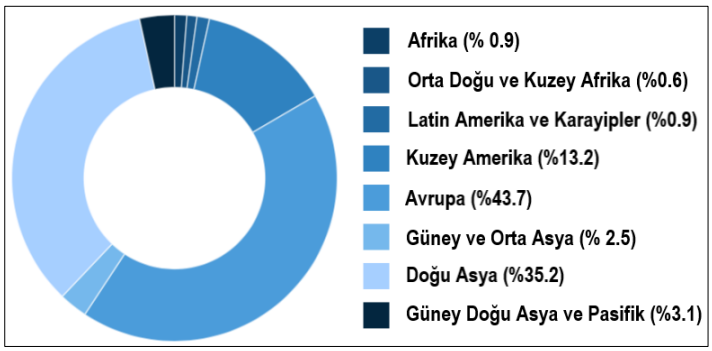
Key Words

"Renewable energy, machine learning, multiple linear regression, k-nearest neighbors regression, neural networks"

***Sorumlu Yazar:** okan@akdeniz.edu.tr*

1. Giriş

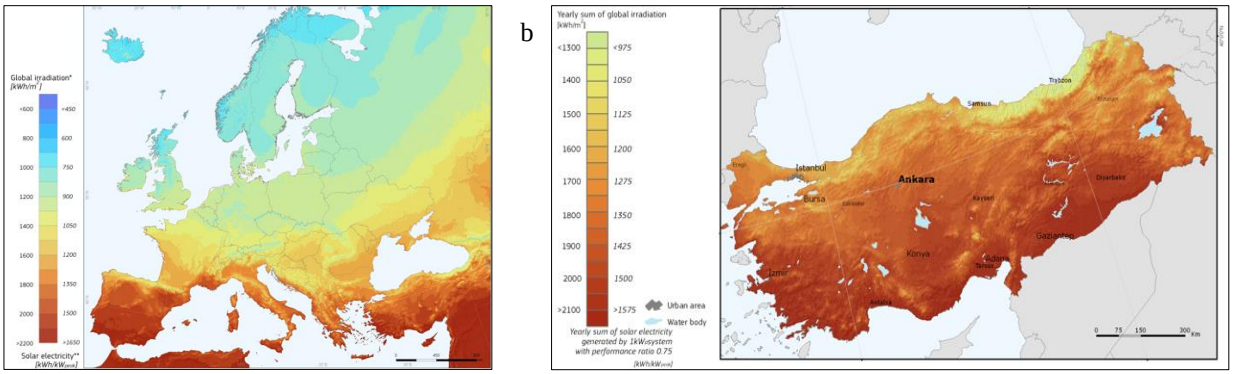
Ülkelerin nüfus artışına bağlı olarak enerji talebi artışı yüksek oranlarda olmaktadır. Bu artış karşısında enerji arz güvenliği ile ilgili endişeler yaşamaları ve iklim değişikliği konusu üzerinde yapılan yoğun tartışmalar; dünya genelinde olduğu gibi Türkiye'de de yenilenebilir enerji kaynaklarına olan ilgiyi önemli ölçüde artırmıştır (Yurdadoğ ve Tosunoğlu, 2017). Özellikle güneş enerjisi, elektrik üretimi için fosil yakıtlara göre sürdürülebilir bir alternatif sunan en umut verici yenilenebilir enerjilerden biridir (Koca vd., 2018). 2017 yılı itibarıyla dünya güneş enerjisi santrali kurulu gücü toplam 227 GW'tır ve bu kurulu gücün %43.7'si Avrupa'da (World Energy Council, 2018) bulunmaktadır (Şekil 1). Bu kurulu gücün 2020 yılına kadar yenilenebilir enerjiden elektrik üretim kapasitesi artışının yaklaşık 700 GW olması ve tahmin edilen küresel toplam kapasite artışının %26'sını oluşturması beklenmektedir (IEA, 2015).



Şekil 1. Bölgelere göre güneş enerjisi kurulu gücü (World Energy Council, 2018).

Küresel kurulu güç oranının en yüksek olduğu Avrupa'da en fazla güneş alan ülke Türkiye'dir (Şekil 2. a, b). Türkiye'nin hem coğrafi olarak büyük bir alana yayılmış hem de güneşli gün sayısı yüksek bir ülke olması bu durumun sebebidir. Türkiye'nin yeryüzündeki coğrafi konumu gereği yıl içerisindeki güneşli gün sayısının fazla olması, teknik açıdan bu denli yüksek bir potansiyele sahip olmasında en büyük etkenlerin başında gelmektedir. Türkiye'nin yıllık toplam güneşlenme süresi 2741 saat, yıllık toplam gelen güneş enerjisi 1527 kWh m^{-2}'dir (Enerji ve Tabi Kaynaklar Bakanlığı, 2018). Öte yandan, Türkiye elektrik üretiminde güneş enerjisi teknik potansiyeli 189 GWh $yıl^{-1}$ olarak hesaplanmıştır (Karagöl ve Kavaz, 2017). Bu alanda Türkiye kendisine en yakın ülkeler olan İspanya ve Fransa'dan yaklaşık %30 daha fazla potansiyele sahiptir. Bu da ülkenin mevcut elektrik talebinin %75'inin güneş enerjisiyle karşılanabileceğini göstermektedir.

a b



Şekil 2. (a) Avrupa ve **(b)** Türkiye fotovoltaik güneş enerjisi potansiyeli (PVGIS, 2018).

Türkiye'nin elektrik enerjisi kurulu gücü, 2017 Yılı üçüncü çeyrek sonu itibariyle 81647 MW'a yükselmiştir. Bu kurulu gücün kaynaklara göre dağılımı; %33.2'si hidrolik, %32.6'sı doğal gaz, %22.1'i kömür, %7.7'si rüzgâr, %1.2'si jeotermal, %2.2'si güneş ve %1'i ise diğer kaynaklardan oluşmaktadır. Türkiye'de yatırımcılar 1 MW'ye kadar olan PV santrallerini maliyetli bir lisanslama sürecine dâhil olmadan kurabilmektedirler. Bu kapsamda, 2017 yılı üçüncü çeyrek sonu itibariyle 1753 MW kurulu güce sahip 2104 adet lisanssız güneş santrali devrededir (Enerji ve Tabii Kaynaklar Bakanlığı, 2018). 2023 yılı brüt elektrik talebinin 500 bin MW olacağı öngörüsü altında Türkiye tüm güneş potansiyelini kullanmak koşuluyla 2023 yılına gelindiğinde elektrik talebinin tamamını sadece güneş enerjisinden karşılayabilecektir. Türkiye' de güneş enerjisinden elektrik üretiminin 2023 yılında 5 GW'a ulaşacağı tahmin edilmektedir (Karagöl ve Kavaz, 2017). Bu nedenle, Türkiye'nin sadece güvenli enerji arzını artırmak için değil, aynı zamanda ekonomik kalkınmayı artırmak için de güneş enerjisinden elektrik üretimini geliştirmesi çok önemlidir.

Türkiye'de güneş enerjisinden optimum düzeyde faydalanarak, PV santralleri sayesinde elektrik üretebilmek için, PV santrallerinin kurulacağı konumların doğru belirlenmesi önem kazanmaktadır. Santralin kurulacağı bölgenin rakımı, enlemi, ortam sıcaklığı, güneşlenme süresi ve PV sisteminin donanımsal özellikleri, PV santrallerinin kurulacağı konumların doğru belirlenmesi için önem arz eden parametreler arasındadır. Deneysel çalışmalar için belirli maliyetler oluşması bu analizlerin çeşitli yazılımlar ile ya da makine öğrenmesi algoritmaları gibi tahmin yöntemleri ile gerçekleştirilebilir. Bu çalışmada da Türkiye'de güneş potansiyeli

yüksek 125 farklı konum için kurulması planlanan PV santrallerinin üretebileceği elektrik gücünün, makine öğrenmesi modelleri kullanılarak tahmin edilmesi amaçlanmıştır.

2. İlgili Çalışmalar

Makine öğrenim yaklaşımı ile güneş ışınımının tahmin yöntemleri arasında kullanım açısından pek çok yöntem kullanılmıştır. Bu yöntemler arasında genellikle YSA ve ARIMA (Autoregressive Integrated Moving Average) modelleri kullanılmaktadır. İlave olarak, Destek vektör makineleri, k-ortalamalar yöntemleri de artık daha sık kullanılmaya başlandığı görülmektedir. Bunların dışındaki Boosting, Regression ağaçları, Rastgele ormanlar vb. yöntemler ise daha nadiren kullanılmaktadır (Voyant vd., 2017). Kullanılan yöntemler, performans sıralaması, veri kümesinin çeşitliliği, zaman aşaması, tahmin görüşleri ve performans göstergeleri nedeniyle karmaşıktır. Fakat genel olarak, tahmin hatası eşdeğer seviyededir. Tahmin performansının iyileştirilmesi için bazı araştırmacılar ise hibrit modellerin kullanımını önermektedirler. Örneğin, Saklı Markov modeline ve Destek Vektör Regresyonuna (DVR) dayanan kısa vadeli güneş ışınımı tahmin algoritmaları kullanılarak farklı hava koşullarında gelecek 5-30 dakikalık güneş ışınımını doğru şekilde tahmin edilebildiği gösterilmiştir (Jiaming vd., 2016).

Literatürde PV sistem projelerinin çeşitli parametrelerinin tahmin edilmesi ile ilgili farklı makine öğrenmesi algoritmaları üzerinde çalışılmıştır. Son yıllarda farklı enlem ve iklime sahip yerler için güneş radyasyonu modelleme çalışmasında özellikle YSA modellerinin sıklıkla kullanıldığı görülmektedir (Elizondo vd., 1994). YSA kullanılarak küresel güneş ışınımının tahmini üzerine yapılan çalışmada (Khatib vd., 2012), Malezya'daki 28 siteden toplanan veriler kullanılmış ve gökyüzü için açıklık (Berraklık) indeksleri tahmin edilmiştir. Bu indeksler daha sonra küresel güneş ışınımını tahmin etmek için kullanılmıştır. Ayrıca, Malezya için yaygın güneş radyasyonu tahminleri için bir denklem geliştirilmiştir. Bu denklem ile yaygın ve küresel güneş ışınımının ve berraklık endeksinin bir fonksiyonu olarak hesaplanabilmektedir. Önerilen yaklaşımı test etmek için Malezya'daki beş ana saha kullanılmıştır. Tahmin edilen küresel güneş ışınımı için performans ölçütleri MAPE (Ortalama mutlak yüzde hata), MBE (Ortalama bias hata) ve RMSE (Ortalama kare hata karekök) değerleri sırasıyla %5.92, %1.46 ve %7.96 olarak elde edilmiştir. Yaygın güneş ışınımı için ise MAPE %9.8 olarak bulunmuştur.

Yapılan diğer bir çalışmada farklı hava koşullarında PV güç üretimini doğru olarak tahmin etmek için genelleştirilmiş bir Destek vektör regresyon tabanlı tahmin modeli önerilmiştir. Önerilen model üç farklı PV güç istasyonuna uygulanmış ve daha sonra performansı analizleri yapılmıştır. Her belirli noktada gerçek ölçülen ve tahmin edilen PV güçlerindeki sapmalar kabul edilebilir bir aralıkta (%10'dan fazla olmamak üzere) olduğu ve herhangi bir noktada ciddi dalgalanmaların gözlemlenmediği bildirilmiştir. Modelin performans ölçtleri ise RMSE, MAE (Ortalama mutlak hata) ve MBE değerleri sırasıyla %34.57 ve %11.34 ve %3.08 olarak elde edilmiş (Das vd., 2017).

PV sistemlerinin güç çıkışını tahmin etmek için farklı makine öğrenim modellerinin performansını değerlendirildiği bir çalışmada (Theocharides vd., 2018), YSA, DVR ve regresyon ağaçları dahil olmak üzere çeşitli yöntemler araştırılmıştır. Her bir modelin çıkış gücü tahmin performansı, bir yıl boyunca edinilen ve mevcut bir sürekli modele kıyasla gerçek PV üretim veri setleri üzerinde testler yapılmıştır. Optimal olarak tasarlanan farklı modeller arasındaki karşılaştırmalı analiz, YSA'nın diğer modellerden daha iyi performans gösterdiğini, MAPE ve RMSE performans ölçüt değerleri sırasıyla %0.6 ve %0.76 olarak ifade edilmiştir.

Gerçek uygulamada, eğitim verileri zamana duyarlı olduğundan, eğitim verilerinin gerçek zamanlı olarak güncellenmesi gerekir. Giriş değişkenleri hava durumuna göre sınıflandırılan 24 saatlik online bir tahmin modeli (Chen vd., 2011) kullanan bazı araştırmacılar, güneş enerjisinin istatistiksel olarak standartlaştırılmasıyla elde edildiği iki aşamalı bir yöntemi tarif etmişlerdir. Standart hale getirilmiş güneş enerjisi veri tahminleri ise başka bir araştırmada PV gücünü tahmin etmek için kullanılmıştır (Bacher vd., 2009). Diğer bir çalışmada (Duffie ve Beckman, 2013), genel olarak önceki tarihlere ait meteorolojik veriler (sıcaklık, bulutlanma durumu, küresel radyasyon) ile tesisin bulunduğu yere ait konum verilerinden yola çıkılarak PV tesisinden elde edilebilecek gücün hesaplanmasında YSA yönteminden yararlanmışlardır. Yazarlar, PV güç tahmininde en iyi YSA düzeninde kullanılabilecek prosedürü ve ayarları belirleyebilmek için ayrıntılı bir analiz yapmışlardır. Bu sonuçlara ilave olarak Makine öğrenmesi tekniklerinin tahmin performansının, veri setinin düzeninden ve girdi seçiminden etkilendiğini göstermiştir (Netsanet vd., 2016; Panapakidis ve Christoforidis, 2017).

DVR, Eğim Artan Regresyon (EAR), Rastgele orman regresyon ve bunların yanı sıra alt ölçeklere birleştirmek ve Nümerik Hava Tahmini (NHT) tarafından sağlanan 3 saatlik kümülatif güneş ışınımı tahminlerini geliştirmek için hibrit bir yöntem uygulanmıştır (Gala vd., 2016). Ayrıca, açık gökyüzü teorik ve deneysel radyasyon modellerine dayanan interpolasyonu kullanarak 3 saatlik tahminlerden saatlik tahminlerin elde edilmesine yönelik çalışma yapılmıştır. Makine öğrenmesi yöntemlerinin veya daha genel olarak hibrit yapay zeka sistemlerinin oldukça etkili olduğunu ve dolayısıyla güneş ışınımının tahmini için uygun olduğunu göstermektedir.

Kaliforniya'da faaliyet gösteren 1 MW_p, tek eksenli bir izleme, fotovoltaik santralin güneş enerjisi çıkışını tahmin etmek için harici girişler kullanmayan birkaç tahminleme tekniği değerlendirilip karşılaştırılmıştır (Pedro ve Coimbra, 2012). Çalışmada 2 yıllık veriler kullanılmış olup, bir ve iki saat ileri saatlik ortalama güç çıkışı için ARIMA, k-EYK, YSA ve Genetik Algoritma (GA) tarafından optimize edilmiş YSA modelleri oluşturulmuştur. Bulgular, YSA temelli tahmin modellerinin diğer tahmin tekniklerinden daha iyi performans gösterdiğini, YSA parametrelerinin GA optimizasyonu ile önemli iyileştirmelerin elde edilebileceğini ve tüm modellerin doğruluğunun güneş değişkenliğinin mevsimsel özelliklerine güçlü bir şekilde bağlı olduğunu göstermiştir.

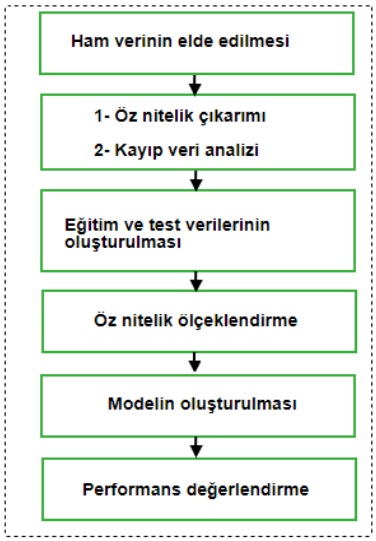
YSA, bulanık mantık (BM) ve hibrit sistemler (GA/YSA, YSA/BM) yataya gelen güneş ışınımını belirleyen temel fiziksel süreçlerin karmaşık doğasını modellemek için çok uygundur (PV tesislerinin güç çıkışı gibi). Genellikle doğrusal olmayan atmosferik olayları tanımlamak için ARMA yöntemi ve ARIMA gibi onun varyasyonlarını içeren diğer regresyon yöntemleri kullanılır (Gordon, 2009).

Çoğu PV uygulamasında modül uyumsuzluklarından dolayı enerji üretim kayıpları ortaya çıkmaktadır. Temelde dizilerin kısmi gölgelenmesi ve PV modüllerindeki farklılıklardan kaynaklanan güç kayıpları, bir güneş dizisinin modül ara bağlantılarının değiştirilmesiyle azaltılabilir. Bunun yerine çeşitli çevresel koşullardaki mevcut PV dizisi üretimini tahmin etmek ve kayıpları belirleyebilmek için yeni bir yöntem sunmuşlardır (Picault vd., 2010). Bu yaklaşımda, saha ölçüm verileri kullanılmıştır. Yöntemde, uyumsuzluk kayıplarını azaltmak için uyarlanabilir modül ara bağlantı şemaları ile PV dizilerini simüle etmektedir. Model, kısmen gölgeli ve normal çalışma koşullarında güvenilir güç üretim tahmin hassasiyeti gösteren üç farklı ara bağlantı şeması ile 2.2 kWp'lik bir tesis üzerinde yapılan deneysel sonuçlar ile doğrulanmıştır. Güneş radyasyonu tahmin hesaplamaları için doğrusal regresyon ve Gauss işlem regresyon modellerinin performansı MAE, MSE ve RMSE gibi istatistiksel hata kriterleri açısından değerlendirilmiştir. En iyi sonucun elde edildiği model ve model parametreleri için MAE değeri 0.016620, MSE değeri 0.000514 ve RMSE değeri 0.022674 olarak hesaplanmıştır (Karasu vd., 2017).

Güneş enerjisi potansiyelini tahmin etmek için yapılan bir çalışmada DVM ve coğrafi bilgi sistemlerinin kombinasyonu kullanılarak İsviçre'deki 2477 komün topluluğunun toplam 1901'i için çatı güneş enerjisi potansiyeli tahmin edilmiştir. Sonuçlar, her binanın toplam zemin kat alanının ortalama %81'inin PV kurulumu için mevcut çatı alanına karşılık geldiğini göstermektedir. Çalışmada, veri seti k-kat çapraz doğrulama yöntemi ile eğitim ve test kümelerine ayrılmıştır. Ayrıca, farklı DVM modelleri kurularak performansları RMSE değerlendirme ölçütü ile karşılaştırılmıştır (Assouline vd., 2017). Fotovoltaik gücün tahmini için YSA da kullanılabilir. Giriş olarak, ortam ve panel yüzey sıcaklıkları ($^{\circ}$C), toplam günlük enerji (J), ışınım enerjisi (W/m^2) ve elektrik gücü (W) değişkenleri kullanılarak yapılan bir çalışmada (Teo vd., 2015), 30 gün için üç veri seti oluşturulmuştur.

3. Çalışmanın Tanımı ve Yöntemleri

Bu çalışmada gerçekleştirilen makine öğrenmesi sürecine ait adımlar Şekil 3'deki diyagramda görülmektedir.



Şekil 3. Bir makine öğrenmesi sürecinin şematik gösterimi.

3.1. Ham Verinin Elde Edilmesi

Bu çalışmada, Türkiye'deki 10 konum için 1, 3, 5 ve 10 MW güçlerindeki PV santrallerinden elektrik üretimine yönelik yapılan analizler kullanılmıştır (Caglayan vd., 2014). Analizlerin yapılmasında RETScreen Clean Energy Project Analysis (RETScreen, 2011) yazılımı kullanılmıştır. Bu 10 konum için elde edilen analiz sonuçları dört farklı makine öğrenmesi modeli kullanılarak Türkiye'deki 125 konum için çıkarımlar yapılmıştır. PV santrallerinde kullanılan modüller sabit, tek ve iki eksende hareketli olarak planlanmıştır. Önerilen PV santrallerindeki modül özellikleri, Tablo 1'de verilmiştir.

RETScreen, temiz enerji teknolojisi modelleri kullanarak karar vermeyi kolaylaştıran, güvenilir bir sonuç veren yaklaşım izlemektedir. Ayrıca, RETScreen fizibilite çalışmalarının zamanını ve ilgili maliyetleri azaltmaya da yardımcı olur (Salata vd., 2014). Bu yazılıma yatay yüzeye gelen global solar radyasyon, hava sıcaklığı ve güneş izleme durumuna göre veriler girilerek önerilen güçlerdeki PV santralinden elde edilebilecek minimum, maksimum ve ortalama spesifik verim (SY) değerleri hesaplanabilir. SY, önerilen PV santrallerin ürettiği elektriğin bir yıl boyunca sabit, tek ve çift eksenli güneş izleme sistemlerinin her biri için PV dizi alanına oranı olarak hesaplanır (Caglayan vd., 2014).

Tablo 1. Fotovoltaik modül özellikleri.

Özellikler	Değer
Üretici	BP Solar
PV Modül Tipi	Mono-si
Modül Numarası	BP 250 F
Nominal verimlilik (%)	11.4
PV modül derecesi (W)	50
Voltaj [@ Zayıf Güç] (V)	17
Akım [@ Zayıf Güç] (A)	2.94
Açık akım voltajı (V)	21.2
Kısa Devre Akımı (V)	3.22
Çerçeve alanı (m^2)	0.44
Raf montaj alanı ($W.m^{-2}$)	100
İki modül arası mesafe (m)	1.25
Montaj boyut kalınlığı (mm)	38.6
Genişlik (mm)	530
Yükseklik (mm)	825
Ağırlık (kg)	6

Analizlerde girdi olarak kullanılan iklimsel veriler, NASA Surface Meteorology and Solar Energy (2019)'den sağlanan konuma göre uzun süreli iklim veri kayıtlarından oluşmaktadır (Tablo 2). Bu verilerden, aylık ortalama günlük küresel güneş ışınımı ve güneşlenme süreleri elde edilmiş ve Türkiye üzerindeki küresel radyasyonun (GSR), güneşlenme süresinin (SSD) ve hava sıcaklıklarının (T) dağılımını içeren veri seti oluşturulmuştur. Elde edilen veri seti, önerilen PV santrallerinden elde edilebilecek elektrik güç verileri için bir girdi oluşturmaktadır.

Tablo 2. Veri setine ait öznitelikler

Öznitelik	Açıklama
ALT	Ölçüm yapılan bölgenin rakımı
LAT	Ölçüm yapılan bölgenin enlemi
T	Aylık ortalama dış sıcaklık (°C),
GSR	Güneşten gelen ışınım (MWh m^{-2} $yıl^{-1}$)
SSD	Güneşlenme süresi (h $yıl^{-1}$)
POW	Planlanan PV sisteminin gücü (1, 3, 5 ve 10 MW)
AX	Planlanan PV sistemi (Sabit, Tek eksenli, İki eksenli)
EPOW	Planlanan PV sisteminin yıllık ürettiği elektrik enerjisi (GWh $yıl^{-1}$)

3.2. Model Oluşturma Süreci

Yapılan çalışmada öncelikle Türkiye'deki 10 konumda kurulduğu varsayılan 1, 3, 5 ve 10 MW güçlerindeki PV santrallerinden elde edilen elektrik enerjisi RETScreen Clean Energy Project Analysis yazılımı ile analiz edilmiştir. Adından 125 farklı konum için toplam 1620 adet ham veri elde edilmiştir. Ham veriler toplandıktan sonra makine öğrenmesi uygulamasında kullanılacak olan öz nitelikler, öz nitelikler çıkarımı işlemi ile seçilerek model için istatistiksel öneme sahip olmayanlar veri setinden çıkarılmıştır. Makine öğrenmesinde oluşturulan bir modeli, veri setinde olmayan yeni veriler ile sınamak modelin geçerliliği için önemlidir. Bunu gerçekleştirmek için veri setini, eğitim ve test seti olarak bölmek gerekmektedir. Veri setinin eğitim ve test seti olarak ayrılması esnasında veri setinin dağılımında oluşabilecek düzensizlikler, makine öğrenmesi modelinin performansını olumsuz etkileyebilir. Bu problem k-kat çapraz doğrulama (k-fold cross validation) yöntemi (Stone, 1974) ile çözüme kavuşturulabilir. Bu çalışmada veri seti k-kat çapraz doğrulama yöntemi kullanılarak %70 eğitim seti ve %30 test seti olacak şekilde bölünmüştür. Makine öğrenmesi algoritmalarında çok sayıda öznitellik ile çalışmak mümkündür ve veri setindeki tüm özniteliklerin aynı değer aralığında olmasına ihtiyaç duyulur (Bollegala, 2017). Bu sebeple, özniteliklerin tümünün aynı değer aralığında ifade edilmesi işlemi (öznitelik ölçeklendirme) yapılmalıdır.

Bu çalışmada, veri ön işleme aşamasında veri setindeki tüm veriler öznitelik ölçeklendirme yapılarak aynı değer aralığında (-1 ile +1) ifade edilmiştir. Bir sonraki aşama olan makine öğrenmesi modelerinin oluşturulması sürecinde YSA, ÇDR ve K-EYKR kullanılmıştır. Bu çalışmada kurulan makine öğrenmesi modelleri için, Numpy, Pandas, Scipy gibi temel python kütüphanelerinin yanı sıra makine öğrenmesi uygulamaları için geliştirilmiş olan scikit-learn kütüphanesi de kullanılmıştır.

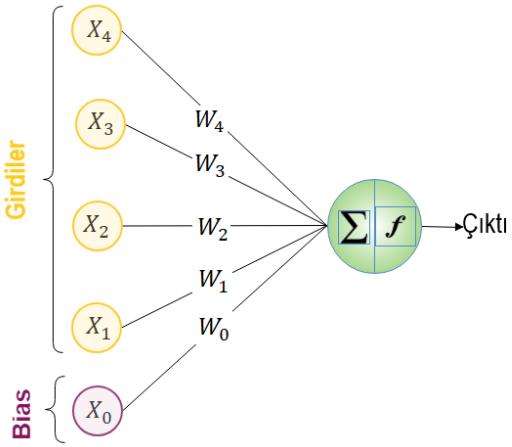
3.3. Performans Değerlendirme Kriterleri

Bir makine öğrenmesi modeli kurduktan sonra modelin performansının değerlendirilmesi gerekmektedir. Regresyon analizinde performans değerlendirmesi yapılırken R^2, MAE, MSE ve MedAE (Medyan ortalama hata) gibi çeşitli performans ölçütleri kullanılabilir (Zheng, 2015; Al-Ghobari vd., 2018) R^2, bir saçılım grafiği ile gösterilebilen veri seti örneklerinin regresyon çizgisine yakınlığı hakkında bilgi veren bir performans ölçütüdür. Veri setindeki örnekler regresyon çizgisine yaklaştıkça R^2'nin değeri artar dolayısı ile daha uygun bir model elde edilir. R^2, 0 ile 1 arasında bir değer alır (Grégoire, 2014). MAE, MSE ve MedAE hata değerlerinin ise olabildiğince düşürülmesi hedeflenir.

3.4. Yapay Sinir Ağları (YSA)

Bir YSA mimarisinde girdi katmanı, gizli katman ve çıktı katmanı olmak üzere üç katmanlı bir yapı söz konusudur. Katmanlar içerisinde yer alan nöronları birbirine bağlayan ağırlıklar YSA'da bilgi saklanması için kullanılan yapılardır. YSA'nın davranışında nöron sayısı, katman sayısı ve katmanlar arasındaki bağlantıların tipi önemli bir yer tutmaktadır (Patterson ve Gibson, 2017). Basit bir YSA mimarisinde sadece girdi ve çıktı katmanları yer alır ve bu mimari tek katmanlı YSA olarak adlandırılır. Şekil 4'de, dört girdi ve bir çıktıya sahip basit bir YSA modeli görülmektedir. Girdilerin ve bias değerinin kendi ağırlık değerleri ile çarpımlarının birleştirme fonksiyonu ile toplamı, bir aktivasyon fonksiyonundan (Sigmoid, hiperbolik tanjant, doğrusal fonksiyon vb.) geçirilerek çıktı değeri elde edilmektedir.

Genel olarak bu işlem denklem (1) ile ifade edilebilir.



Şekil 4. Basit YSA Modeli

$$\text{Çıkış} = \text{f}\left(\sum_{\text{i=0}}^{4} \text{w}_\text{i}\ \text{x}_\text{i}\right) \tag{1}$$

Burada i, nöronun girişlerindeki indeks, x_i nöronun girişidir, w_i o girişe eklenmiş ağırlık faktörüdür ve w_0 nöronun yanlılığıdır.

Giriş ve çıkış katmanları arasına gizli katmanlar eklendiğinde çok katmanlı YSA mimarisi oluşturulabilir. çok katmanlı YSA yapısında ağ, geri yayılım algortiması gibi çeşitli öğrenme algoritmaları ile eğitilir. Geri yayılım algortiması aşağıdaki temel adımları içerir. Bu maddelerdeki olaylar bir döngü şeklinde devam eder ve belirlenen iterasyon sayısına ya da hata değerine ulaşıldığında sona erer.

- Ağ girdisinin girdi katmanından çıktı katmanına doğru ilerlemesi
- Hatanın hesaplanması ve geriye doğru hatanın yayımı.
- Katmanlar arasındaki ağırlıkların ve bias değerlerinin güncellenmesi.

Bir YSA modelinin performansını etkileyen unsurlardan başlıcaları arasında gizli katman sayısı, katmanlardaki nöron sayısı, aktivasyon fonksiyonunun türü, öğrenme algoritmasının türü sayılabilir. Dolayısıyla en iyi YSA performansını elde etmek bu sayılan faktörleri değiştirerek denemeler yapmayı gerektirir. Bu çalışmada, PV sistemlerinden üretilen güç değerlerinin tahmin edilmesi amacıyla geri yayılım algoritması ile eğitilen ağda iki gizli katman üzerinde çalışılmıştır. En iyi modelin seçimi için gizli katmanlardaki nöron sayıları, 1'den 10'a kadar bir döngü kurularak denemeler gerçekleştirilmiştir. Algoritmanın sonlandırılması için kriter olarak iterasyon sayısı 1000 olarak belirlenmiştir.

3.5. Çoklu Doğrusal Regresyon (ÇDR)

ÇDR, bir veri setine ait bağımlı değişken ile birkaç bağımsız değişken arasındaki en iyi ilişkiyi tahmin etmek için kullanılabilen ve makine öğrenmesi altında uygulanan regresyon modellerinden birisidir. ÇDR modeli gözlenen ve tahmin edilen değerlerin farklarının karelerinin toplamını minimize eden en küçük kareler yöntemine dayanır (Tiryaki ve Aydın, 2014). Bir ÇDR modeli, Denklem 2'deki çıktı bağımlı değişkeninin hesaplanmasında kullanılan fonksiyon ile ifade edilebilir.

Denklemde *n* tane olmak üzere, *x* değerleri bağımsız değişkenleri ifade etmektedir. ε ise hata terimidir. W_i parametreleri ise ağırlık değerlerini oluşturmaktadır.

$$\text{Çıktı} = w_0 + w_1x_1 + w_2x_2 + \cdots + w_nx_n + \varepsilon \quad (1)$$

ÇDR uygulamalarında bazı varsayımlar üzerinden model oluşturulur. Regresyon analizinin sağlıklı olarak gerçekleştirilmesi için bu varsayımların ihlal edilmemesi gerekir. Bu varsayımlardan birisi çoklu bağlantıdır. Veri setinde iki veya daha fazla bağımsız değişkenin birbiriyle korelasyon içinde olması çoklu bağlantı oluşmasına neden olur. Bu durumunda, katsayıların standart hatası artarak bağımsız değişkenlerin istatistiksel açıdan anlamsız olarak değerlendirilmesi söz konusu olur. Varyansın yükseliş miktarı, Varyans etki faktörü (VIF) ile tespit edilebilir. Bulunan VIF değeri, eğer 10'dan daha yüksek bir değer ise bağımsız değişkenler arasında yüksek korelasyondan yani çoklu bağlantıdan söz edilebilir.

Bu çalışmada çoklu bağlantı analizi yapılarak bağımsız değişkenlerin VIF değerleri hesaplanmış ve veri setinde çoklu bağlantı olmadığı sonucuna varılmıştır. İhlal edilmemesi gereken bir diğer varsayım ise hata terimlerinin birbiriyle ilişkili olması durumu olarak ifade edilebilen oto korelasyonun oluşmamasıdır. Bu çalışmada, bağımsız değişkenler arasında oto korelasyon olmadığı sonucuna Durbin-Watson testi gerçekleştirilerek ulaşılmıştır. Gerçek hayatta bağımlı değişkenleri etkileyen birçok faktör bulunmaktadır. Bu faktörlerin sayısının çok fazla olması modelin karmaşıklığını artıracağı gibi kullanışlılığını da olumsuz yönde etkileyecektir. Bu yüzden bağımlı değişken üzerinde daha az etkili olan öz niteliklerin modelden çıkartılması gerekir. Bağımsız değişkenlerin modelden çıkartılması adımsal regresyon teknikleri ile gerçekleştirilebilir (John vd., 1994). Bu çalışmada ÇDR modelinin eğitimi gerçekleştirilirken bu tekniklerden geriye doğru eleme yöntemi kullanılmıştır. Bu yöntemde tüm değişkenler algoritmanın başlangıcında modele dahil edilirler. Daha sonra algoritma, belirlenen anlamlılık düzeyinden daha büyük olan (Birden fazla ise en büyük olan) öznitelik modelden çıkarılacak şekilde çalıştırılır. Bu işlem belirlenen anlamlılık düzeyinden daha büyük özellik kalmayana kadar devam eder (Narin vd., 2014).

3.6. k-En Yakın Komşu Regresyon (k-EYK)

Sınıflandırma uygulamaları için ilk olarak geliştirilen k-EYK algoritması (Cover ve Hart, 1967) etkili ve basit olmasından dolayı popüler bir makine öğrenmesi algoritması olarak nitelendirilebilir. Her makine öğrenmesi algoritmasında parameter optimizasyonu, performansın geliştirlmesi için önemli bir kriterdir. k-EYK algoritmasında ise k parametresi optimize edilerek veri setindeki özniteliklere ait örneklerin dahil olacağı sınıf belirlenebilmektedir. Regresyon uygulamalarında k-EYKR olarak adlandırılan algoritma temel olarak veri setindeki örneklerin birbirine olan uzaklıklarının Euclidean, Manhattan ve Minkowski gibi uzaklık ölçüm yöntemleri ile hesaplanması prensibine dayanır. Hesaplama sonucunda k parametresine bağlı olarak birbirine mesafe olarak yakın örnekler aynı sınıfa dahil edilirler.

4. Sonuç ve Tartışma

Bu çalışmada yedi adet bağımsız değişken girdi olarak kullanılmış ve PV tarafında üretilen elektrik gücü tahmin edilmiştir. Tablo 3'de girdi ve çıktı değişkenlerine ait veri aralıkları ve toplam örnek sayısı görülmektedir.

Tablo 3. Makine öğrenim modellerinde kullanılan veri setinin özeti

Girdi Değişkenleri								**Çıktı Değişkeni**
TÖ	ALT	LAT	SSD	T	GSR	POW	AX	EPOW
1260	0 - 2173	33.5 - 42	3.8 - 8.4	6 - 18.5	1.3 - 2	1-10	1-3	1.27- 26.19

TÖ= Toplam Örnek

Tablo 4'de veri setinin eğitim ve test kümesine ayrıldıktan sonra bağımsız değişkenlerin aldığı tanımlayıcı istatistiksel parametreler olan minimum, maksimum, ortalama, standart sapma, skewness ve kurtosis değerleri görülmektedir. Bir veri setinin normal dağılıma sahip olduğunu test etmek için skewness ve kurtosis katsayılarına bakılabilir. Normal dağılıma sahip bir veri seti için skewness katsayısının 3'den küçük ve kurtosis katsayısının -2 ile 2 arasında olması beklenir (Chemingui ve Ben lallouna, 2013). Tablo 5 incelendiğinde, veri setinin normal dağılıma sahip verilerden oluştuğu yorumu yapılabilir.

Tablo 4. Eğitim ve test sürecinde kullanılan değişkenlerin tanımlayıcı istatistikleri

	Eğitim Seti							**Test Seti**						
	ALT	LAT	SSD	T	GSR	POW	AX	ALT	LAT	SSD	T	GSR	POW	AX
X_{Min}	0	33.5	3.8	6	1.3	1	1	0	33.5	3.8	6	1.3	1	1
X_{Mak}	2108	42.0	8.4	18.5	2	10	3	2108	42.0	8.4	18.5	2	10	3
X_{Ort}	864	38.8	6.9	12.8	1.6	4.7	2	888	38.9	6.9	12.6	1.6	4.6	2
X_{Std}	577	1.5	0.9	3.3	0.1	3.3	0.8	592	1.5	0.9	3.2	1.3	3.3	0.8
X_{Skew}	0.4	-0.1	-0.8	-0.1	0.09	0.5	0.01	0.3	0.05	-0.8	-0.1	-0.05	0.6	-0.03
X_{Kr}	-0.6	-0.2	0.2	-0.9	0.06	-1.0	-1.4	-0.6	-0.9	0.06	-0.7	-0.3	-1.0	-1.4

Geliştirilen YSA modellerinin performans değerlendirme kriterlerinin sonuçları Tablo 5'de görülmektedir. Her bir model iki gizli katmandan ve her gizli katmanda çeşitli sayıda nörondan oluşmaktadır. Her gizli katmanda 1'den 10'a kadar çeşitli varyasyonlar denenerek en iyi performansı veren nöron sayısına sahip beş model Tablo 6'da görülmektedir. Buna göre tüm denemeler neticesinde birinci gizli katmanda 8 ve ikinci gizli katmanda 10 nöronun seçildiği model en iyi performansı vermiştir.

Tablo 5. YSA modelinin gizli katmanlardaki çeşitli nöronlarla istatistiksel performansı

YSA Modelleri		**Performans Ölçütleri**			
GK1	GK2	R2	MAE	MSE	MedAE
8	10	0.979	0.025	0.001	0.015
9	7	0.945	0.048	0.003	0.042
10	9	0.943	0.049	0.003	0.041
2	8	0.938	0.050	0.004	0.040
5	5	0.929	0.051	0.004	0.033

GK1= Gizli katman 1'deki nöron sayısı, GK2= Gizli katman 2'deki nöron sayısı

ÇDR modeli için adımsal regresyonu gerçekleştirdikten sonra, bağımlı değişken olan EPOW'u etkileyen bağımsız değişkenlerin, standart hata, t-istatistik ve p-value değerleri Tablo 6'da görülmektedir. p değeri regresyon katsayılarının istatistiksel olarak önemini ifade etmektedir. Bu çalışmada güven aralığı %95 olarak belirlendiği için, p-değeri 0.05'den daha düşük ya da eşit olan bağımsız değişkenlerin regresyon modelinde önemli olduğu düşünülebilir.

T ve GSR bağımsız değişkenlerinin p-value değeri, adımsal regresyon sonunda 0.05'den büyük çıktığı için modele dahil edilmemiştir ve tablolarda katsayıları boş bırakılmıştır.

Tablo 6. ÇDR katsayısının standart hatası, t istatistiği ve bağımsız değişkenlerin olasılığı

					Bağımsız Değişkenler[a]			
	Kesen	ALT	LAT	SSD	T	GSR	POW	AX
SH	2.086	5.12e-05	0.037	0.438	-	-	0.009	0.036
t-stat	-7.859	3.850	3.990	11.269	-	-	211.2	34.298
p-value	0.000	0.000	0.000	0.000	-	-	0.000	0.000

[a]SH=Standart Hata, t-stat= t istatistik, p-value=olasılık

ÇDR analizi neticesinde, yedi bağımsız değişken için türetilen model, denklem 3'de görülmektedir. T ve GSR bağımsız değişkenleri adımsal regresyon sonunda modelden çıkartıldıkları için denklemde yer almamaktadır.

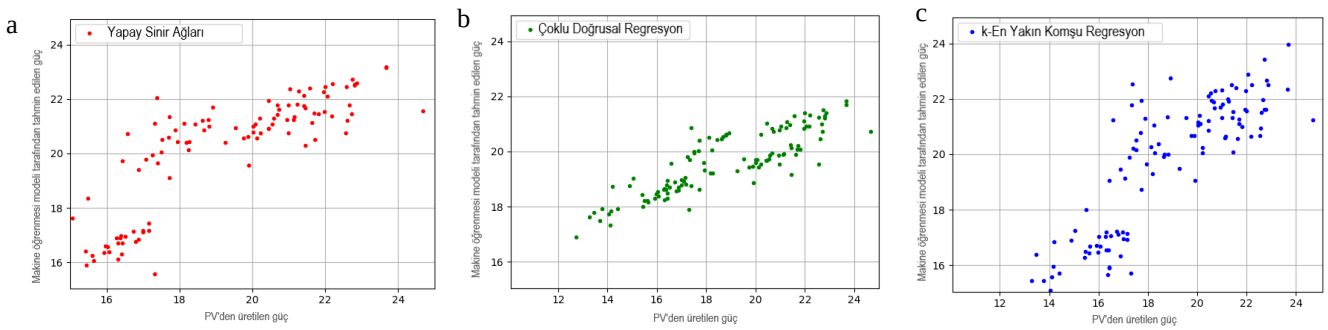
$$\mathrm{EPOW} = -16.39 + 0.0002\mathrm{ALT} + 0.147\mathrm{LAT} + 4.939\mathrm{SSD} + 1.872\mathrm{POW} + 1.245\mathrm{AX} \quad (3)$$

Tablo 7'de çalışmada kullanılan üç farklı makine öğrenmesi modelinin performansları görülmektedir. En yüksek sonuç YSA modeli ile elde edilmiş olup diğer modellerde YSA'ya yakın sonuçlar vermiştir.

Tablo 7. Makine öğreniminin test sırasındaki istatistiksel performansı

Modeller	**Değerlendirme ölçütleri**			
	R^2	MAE	MSE	MedAE
ÇDR	0.959	0.039	0.002	0.031
k-EYKR	0.971	0.030	0.001	0.020
YSA	0.979	0.025	0.001	0.033

Şekil 5'de, PV sistemi tarafından üretilen güç değerleri ile makine öğrenmesi modelinin tahmin ettiği sonuçlara ait grafiklerler görülmektedir. Veri setindeki örnek sayısının, verilerin grafikler üzerinde görselleştirilmesini zorlaştıracak derecede çok olmasından dolayı, 16-24 arasındaki bölgeye ait veriler grafiklerde verilmiştir.



Şekil 5. PV sistemin ürettiği güç ile makine öğrenmesi modelinin tahmin ettiği sonuçlar, **(a)** YSA; **(b)** ÇDR; **(c)** k-EYKR

Türkiye'de PV güç santrallerinin kurulum çalışmaları son yıllarda hız kazanmıştır. Bu durum, santrallerin kurulacağı yerler için fizibilite çalışmalarının yapılması ve kurulması planlanan santralin ekonomikliğinin araştırılması konularının önemini arttırmaktadır. Bu amaçla yapılan çalışmada, daha önce Türkiye'deki 10 konumda kurulduğu varsayılan PV güç santralleri için RETScreen yazılımı ile yapılmış analiz sonuçları makine öğrenmesi modelleri için girdi olarak kullanılmıştır. Bu modeller için geliştirilen algoritmaların çalıştırılmasıyla da 125 farklı konum için PV gücü tahmin edilmiştir. Deneysel verilerin % 30'u makine öğrenme yöntemleri kullanılarak elde edilen modelleri test ederken, verilerin % 70'i modelleri eğitmek için kullanılmıştır. YSA modellerinin eğitim aşamasında en yüksek R^2 değeri 0.959 iken, en düşük hata değerlerinin MAE, MSE ve MedAe için sırasıyla 0.039, 0.002 ve 0.031 olduğu görülmüştür. Üç modelin en iyi performansı, YSA modeliyle elde edilmiştir. Sonuçlar, makine öğrenimine dayalı tahmin algoritmaları kullanılarak PV elektrik gücünün yüksek oranda doğrulukla tahmin edilebileceğini göstermiştir.

Referanslar

Al-Ghobari, H. M., El-Marazky, M. S., Dewidar, A. Z., & Mattar, M. A. (2018). Prediction of wind drift and evaporation losses from sprinkler irrigation using neural network and multiple regression techniques. Agricultural Water Management, 195, 211-221.

Assouline, D., Mohajeri, N., & Scartezzini, J. L. (2017). Quantifying rooftop photovoltaic solar energy potential: a machine learning approach, Solar energy 141, 278-296.

Bacher, P., Madsen, H., & Nielsen, H. A. (2009). Online short-Term solar power forecasting. Sol. Energy, 83, 1772-1783.

Bollegala, D. (2017). Dynamic feature scaling for online learning of binary classifiers. Knowledge-Based Systems, 129, 97-105.

Caglayan, N., Ertekin, C., & Evrendilek, F. (2014). Spatial viability analysis of grid-connected photovoltaic power systems for Turkey, International journal of electrical power & energy systems, 56, 270-278.

Chemingui, H., & Ben lallouna, H. (2013). Resistance, motivations, trust and intention to use mobile financial services. International Journal of Bank Marketing, 31(7), 574-592.

Chen, C., Duan, S., Cai, T., & Liu, B. (2011). Online 24-h solar power forecasting based on weather type classification using artificial neural network. Sol. Energy, 85, 2856-2870.

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.
Das, U., Tey, K., & Seyedmahmoudian, M. (2017). SVR-based model to forecast PV power generation under different weather conditions. Energies, 10(7), 876.

Duffie, J.A., & Beckman,W. A. (2013). Solar engineering of thermal processes. Retrieved from https://www.amazon.com/Solar-Engineering-Thermal-Processes-Duffie/dp/0471698679.

Elizondo, D., Hoogenboom, G., & McClendon, R. W. (1994). Development of a neural network model to predict daily solar radiation, Agricultural and Forest Meteorology, 71(1-2), 115-132.

Enerji ve Tabii Kaynaklar Bakanlığı 2018 Yılı Bütçe Sunumu (2018). Strateji Geliştirme Başkanlığı Retrived from http://www.enerji.gov.tr/File/?path=ROOT%2f1%2fDocuments%2fB%c3%bct%c3%a7e%20Konu%c5%9fmas%c4%b1%2f2018_Butce_Sunus_Kitabi.pdf.

Gala, Y., Fern andez, A., Díaz, J., & Dorronsoro, J. R. (2016). Hybrid machine learning forecasting of solar radiation values, Neurocomputing, 176, 48-59.

Gordon, R. (2009). Predicting solar radiation at high resolutions: a comparison of time series forecasts. Solar Energy 83, 342-349.

Grégoire, G. (2014). Multiple linear regression. European Astronomical Society Publications Series, 66, 45-72.

International Energy Agency (IEA), (2015). The Medium-Term Renewable Energy Market Report. Retrieved from https://www.iea.org/publications/freepublications/publication/MTRMR2015.pdf .

Jiaming, L., John, W., Taffy, T., Lyle, C., & Glenn, P. (2016). Machine Learning for Solar Irradiance Forecasting of Photovoltaic System. Renewable Energy, 90(1), 542–553.

John, G. H., Kohavi, R., & Pfleger, K. (1994). Irrelevant features and the subset selection problem. In Machine Learning Proceedings , 121-129.

Karagöl, T. E., & Kavaz, İ. (2017). Dünyada ve Türkiye'de Yenilenebilir Enerji. Siyaset, Ekonomi ve Toplum Araştırmaları Vakfı (SETA) Analiz Yayını 197. Retrieved from https://setav.org/assets/uploads/2017/04/YenilenebilirEnerji.pdf.

Karasu, S., Altan, A., Sarac, Z., & Hacioglu, R. (2017). Prediction of solar radiatıon based on machine learning methods, The journal of cognitive systems, 2(1), 16-20.

Khatib, T., Mohamed, A., Sopian, K., & Mahmoud, M. (2012). Solar energy prediction for malaysia using artificial neural networks, International journal of photoenergy, 2012(419504), 16.

Koca, A, Oztop, H. F., Varol, Y, & Koca G. O. (2018). Estimation of solar radiation using artificial neural networks with different input parameters for Mediterranean region of Anatolia in Turkey. Expert Syst Appl 2011; 38(7), 8756–62.

Ministry of Energy and Natural Resources (2018). Turkey has high solar energy potential due to its geographical location. Retrieved from http://www.enerji.gov.tr/en-US/Pages/Solar.

Narin, A., Isler, Y., & Ozer, M. (2014). Investigating the performance improvement of HRV Indices in CHF using feature selection methods based on backward elimination and statistical significance. Computers in biology and medicine, 45, 72-79.

NASA Surface Meteorology and Solar Energy. Redrived from https://power.larc.nasa.gov.

Netsanet, S., Zhang, J., Zheng, D., & Hui, M. (2016). Input parameters selection and accuracy enhancement techniques in PV forecasting using Artificial Neural Network. In Proceedings of the 2016 IEEE International Conference on Power and Renewable Energy (ICPRE), Shanghai, China, pp. 565-569.

Panapakidis, I. P., & Christoforidis, G. C. (2017). A hybrid ANN/GA/ANFIS model for very short-term PV power forecasting. In Proceedings of the 2017 11th IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), Cadiz, Spain, 412-417.

Patterson, J., & Gibson, A. (2017). Deep Learning: A Practitioner's Approach. O'Reilly Media, Inc.

Pedro, H. T. & Coimbra, C. F. (2012). Assessment of forecasting techniques for solar power production with no exogenous inputs. Solar Energy, 86, 2017-2028.

Photovoltaic Geographical Information System (2018). Country and regional maps. European Commission Joint Research Centre Energy Efficiency and Renewables Unit (PVGIS team). Retrieved from http://re.jrc.ec.europa.eu/pvg_download/map_index_c.html.

Picault, D., Raison, B., Bacha, S., de la Casa, J., & Aguilera, J. (2010). Forecasting photovoltaic array power production subject to mismatch losses. Solar Energy 84, 1301-1309.

RETScreen, (2011). Retrived from https://www.nrcan.gc.ca/energy/software-tools/7465.

Salata F., Vollaro, Andrea. L., & Vollaro, R. L. (2014). A case study of technical and economic comparison among energy production systems in a complex of historic buildings in Rome. 68th Conference of the Italian Thermal Machines Engineering Association, ATI2013. Energy Procedia 45, 482-491.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society. Series B (Methodological), 111-147.

Teo, T. T., Logenthiran, T., & Woo, W. L. (2015). Forecasting of photovoltaic power using extreme learning machine, in: IEEE Innovative smart grid technologies - Asia (ISGT ASIA), 1-6.

Theocharides, S., Makrides, G., Georghiou, G. E., & Kyprianou, A. (2018). Machine learning algorithms for photovoltaic system power output prediction, 2018 IEEE International Energy Conference (ENERGYCON), Limassol, 2018, pp. 1-6.

Tiryaki, S., & Aydın, A. (2014). An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. Construction and Building Materials, 62, 102-108.

Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review, Renewable Energy, 105, 569-582.

World Energy Council (2018). Energy Resources (Solar). Retrieved from https://www.worldenergy.org/data/resources/resource/solar/

Yurdadoğ, V., & Tosunoğlu, Ş. (2017). Renewable energy support policies in Turkey. Eurasian academy of sciences eurasian business & economics journal, 9, 1-21.

Zheng, A. (2015). Evaluating Machine Learning Models, Farnham, U.K.:O'Reilly Media, Inc.