



Kahramanmaraş Sütçü İmam University Journal of Engineering Sciences



Geliş Tarihi : 20.10.2024
Kabul Tarihi : 17.01.2025

Received Date : 20.10.2024
Accepted Date : 17.01.2025

NEHİR TİPİ SANTRALLERDEKİ IZGARA KİRLİLİK ORANLARININ MAKİNE ÖĞRENME YÖNTEMLERİ İLE TAHMİNİ

PREDICTION OF GRID POLLUTION RATES IN RIVER-TYPE POWER PLANTS USING MACHINE LEARNING METHODS

Kağan KONU^{1}, (ORCID: 0009-0002-6466-0799)*
Prof. Dr. Osman BİLGİN², (ORCID: 0000-0002-4576-0113)
Doç. Dr. Hakan AÇIKGÖZ³, (ORCID: 0000-0002-6432-7243)

1 Gaziantep İslam Bilim ve Teknoloji Üniversitesi, Elektrik-elektronik Mühendisliği Bölümü, Gaziantep, Türkiye

2 Gaziantep İslam Bilim ve Teknoloji Üniversitesi, Elektrik-elektronik Mühendisliği Bölümü, Gaziantep, Türkiye

3 Kahramanmaraş İstiklal Üniversitesi Yazılım Mühendisliği Bölümü, Kahramanmaraş, Türkiye

*Sorumlu Yazar / Corresponding Author: Kağan KONU, kagankonu@gmail.com

ÖZET

Izgara kirlilik değeri farklı konumlara yerleştirilen sensörlerin ölçtüğü basınç değerlerinden elde edilmektedir. Sensörlerin kalibrasyonunun bozulması veya çevresel faktörlerle hatalı ölçümler yapılması ünitelerin verimsiz ve sağlıklı çalışmalarına neden olmaktadır. Bu çalışmada, ilk olarak Karkamış hidroelektrik santralının SCADA sisteminden alınan örneklerden oluşan bir veri seti toplanmıştır. Bu veri seti sensörlerden alınan debi, güç, düşü ve ızgara kirliliğine ait ölçümlerden oluşmaktadır. Izgara kirlilik tahminini gerçekleştirmek için toplanan veri seti eğitim, doğrulama ve test alt setlerine ayrılmıştır. Eğitim verileri ile makine öğrenme yöntemleri için eğitim çalışması gerçekleştirilmiştir. Daha sonra ise ızgara kirlilik tahmini için makine öğrenme yöntemlerinin performansları test veri seti üzerinde farklı ünitelerde karşılaştırılmıştır. Gerçekleştirilen çalışmalardan, ince ağaç modeli ile en iyi korelasyon katsayısı (R) değerlerini ünite-1 için 0.9833, ünite-2 için 0.9716, ünite-4 için 0.9792 ve ünite-5 içinde 0.9810 ile sağlamıştır. Elde edilen sonuçlardan makine öğrenme yöntemleri ile ızgara kirlilik tahmininin etkin bir şekilde yapılabileceği açıkça gözlemlenmiştir.

Anahtar Kelimeler: Izgara kirlilik tahmini, makine öğrenmesi, karar ağaçları, destek vektör makineleri

ABSTRACT

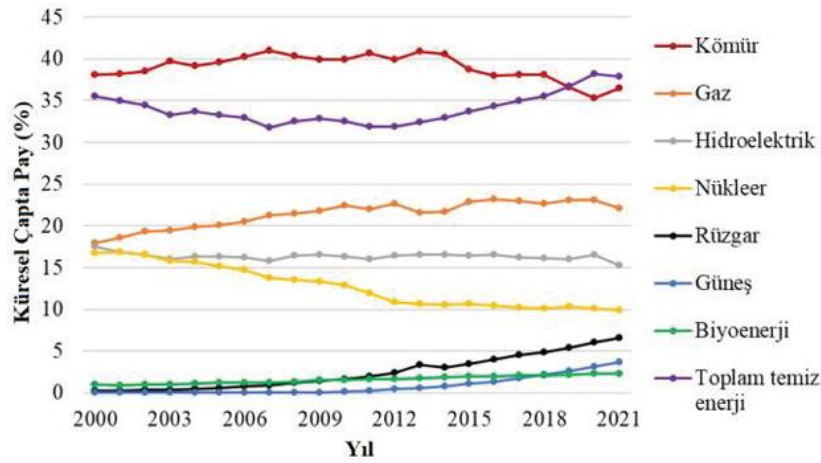
The grid pollution value is obtained from the pressure values measured by sensors placed at different locations. Deterioration of the calibration of the sensors or incorrect measurements due to environmental factors cause the units to operate inefficiently and unhealthily. In this study, firstly, a dataset consisting of samples taken from the SCADA system of Karkamış hydroelectric power plant are collected. This dataset consists of flow, power, head and grid pollution measurements taken from the sensors. The collected dataset is divided into training, validation and test subsets to perform grid pollution prediction. Training study is realized for machine learning methods with training data. Then, the performances of machine learning methods for grid pollution prediction are compared on different units on the testing dataset. From the studies, the best correlation coefficient (R) values are provided with the thin tree model as 0.9833 for unit-1, 0.9716 for unit-2, 0.9792 for unit-4 and 0.9810 for unit-5. It is clearly observed from the obtained results that grid pollution prediction can be done effectively with machine learning methods.

Keywords: Grid pollution prediction, machine learning, decision trees, support vector machines

ToCite: : KONU, K., & BİLGİN, O., AÇIKGÖZ, H., (2025). NEHİR TİPİ SANTRALLERDEKİ IZGARA KİRLİLİK ORANLARININ MAKİNE ÖĞRENME YÖNTEMLERİ İLE TAHMİNİ. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik Bilimleri Dergisi*, 28(2), 613-629.

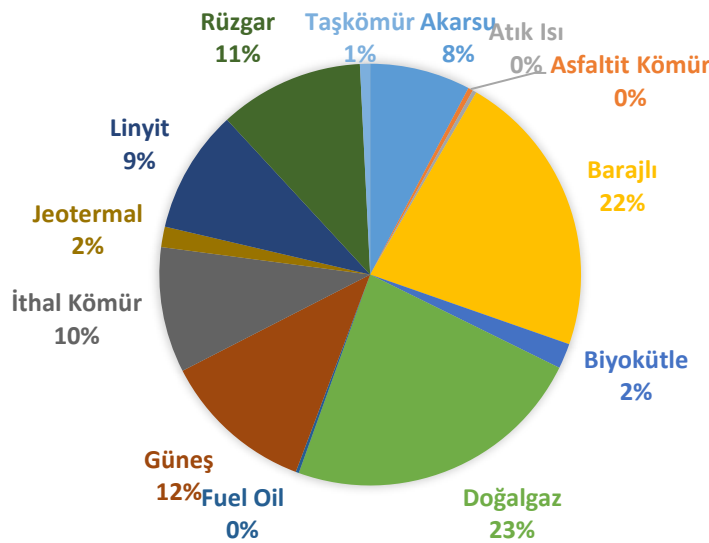
GİRİŞ

Enerji insanlığın hayatını devam ettirebilmesi için en temel unsurlardan biridir. Pamir, enerjiyi “işe dönüştürülen şey” olarak tanımlamaktadır. Enerjinin fizikteki kullanımı ise bir fiziksel sistemin ne kadar iş yapabileceğini ya da ne kadar ısı transferi yapabileceğini göstermektedir (Pamir, 2017). Tüm maddelerin bir enerjisi mevcuttur. Maddelerin enerjisi fiziksel veya kimyasal olarak dönüştürülüp ekonomik olarak kullanılabiliriyorsa bu maddeler enerji kaynakları tanımına girmektedir (Coşkun, O & Doğanay, H, 2020). Enerji kaynakları kullanışlarına ve dönüştürülebilmelerine göre yenilenebilir ve yenilenemez enerji kaynakları olmak üzere iki ana başlık altında ele alınmaktadır (Koç & Şenel, 2013). Yenilenemez enerji kaynaklarının çevreye verdiği zarar çok fazladır. Bu zararı azaltmak ve önlemek için yenilenebilir enerji kaynaklarının kullanımı her geçen gün hızlanarak artmaktadır. Şekil 1’de küresel çapta enerji kaynaklarından elde edilen elektrik enerjisine ait üretiminin değişimi yıllara göre sunulmuştur (Özbektaş vd., 2023).



Şekil 1. Küresel çapta yıllara göre farklı enerji kaynaklarından elde edilen elektrik enerjisi üretim miktarları

Hidroelektrik santraller (HES’ler), suyun hidrolik akım enerjisini elektrik enerjisine dönüştürmek ve sudan elektrik enerjisi üretmek amacı ile kurulan tesislerdir. Santraller kurulacağı yerin topografik durumuna göre çeşitli şekillerde yapılabilmektedir. Bununla birlikte hidrolik santraller; düşü, güç, havza, yapı, çalışma ve depolama özelliklerine göre sınıflandırabilmektedir (Başesme, 2003). 2024 yılının mart ayı verilerine göre ülkemizin elektrik üretimi kurulu gücünün 107.959 MW’ a ulaştığı görülmektedir.



Şekil 2. Türkiye elektrik üretimi Mart ayı sonu itibarıyla kurulu gücümüzün kaynaklara göre dağılımı

2024 yılı mart ayı sonu itibarıyla kurulu gücümüzün kaynaklara göre dağılımı ayrıntılı olarak Şekil 2 de verilmiştir. Şekilde de görüleceği gibi kaynakların yaklaşık %29,6'sı hidrolik enerji, %23,2'si doğal gaz, %20,2'si kömür, %11,2'si rüzgâr, %11,7'si güneş, %1,6'sı jeotermal ve %2,5'i ise diğer kaynaklar şeklinde sıralanmaktadır. Ayrıca ülkemizde elektrik enerjisi üretim santrallerinin sayısı, 2024 yılı mart ayı sonu itibarıyla 18.231'e (Lisanssız santraller dâhil) ulaşmıştır. Mevcut santrallerin 757 adedi hidroelektrik, 69 adedi kömür, 364 adedi rüzgâr, 63 adedi jeotermal, 363 adedi doğal gaz, 16.144 adedi güneş, 471 adedi ise diğer kaynaklı santrallerdir (T.C. Enerji ve Tabii Kaynaklar Bakanlığı, 2024). Bu doğrultuda nehir tipi santrallerin genelinde yapılacak verim artışı toplam üretime büyük katkı sağlayacaktır. Böylece daha az fosil yakıt kullanılmış olacaktır.

Literatürde HES'lerde verimliliği ele alan ve yapay zekâ ile analiz edilen pek çok çalışma vardır. Ülkemizde yapılan çalışmalarda suyun mevsimsel değişimleri araştırılarak üretilecek enerji ve verim tahmini yapılmıştır. Benzer olarak iklimsel değişikliklerin verime olan etkisinin incelenmesine dair çalışmalarda bulunmaktadır. Çalışmaların ortak özelliği çalışma koşulları veri olarak girilip üretim ve verimlilik tahmininin yapılmasıdır. Çalışmalarda genellikle göl havza kapasitesi ve üretim ilişkisi ele alınmaktadır (İnalı vd., 2014; Karakus ve Altun, 2014; Makas ve Karaatlı, 2016; Emeç vd., 2019). Literatürde yapılan çalışmaların bir kısmında ise ızgara kirliliğinin verimle olan ilişkisinin ele alındığı görülmüştür. Bununla birlikte ızgara tasarımı, kirliliğin akışa ve düşüye olan etkisi üzerinden verimlilik çalışmaları gerçekleştirilmiştir (Hamid ve ark., 2016; Walczak vd, 2020; Bragalli vd., 2023). Bu çalışmalarda hem deneysel hem de sahada uygulamalı olarak kirliliğin oluşturduğu düşü kayıplarına odaklanılmıştır.

Izgara kirliliği verim ilişkisine dair yapılan çalışmalarda Hribernik yıl içerisindeki kirlilik durumlarını takip ederek kirlilik tahmin çalışması gerçekleştirmiştir. Yapmış olduğu çalışmada kirliliğin 20 MWh enerji kaybına yol açtığını tespit etmiştir. Bu çalışmada Hribernik tarafından kirliliğin mevsimsel olarak değiştiği de tespit edilmiştir. Ayrıca yazar, ızgara kirliliğinin temizlik periyotlarını değiştirerek ızgara kirlilik durumlarına göre optimum verimi sağlamaya çalışmıştır. Çalışmada yapılan kirlilik tahmin modelinin %97 doğruluk oranı ile yüksek performanslı bir sonuç verdiği tespit edilmiştir (Hribernik, 2020).

Walczak vd., hidrolik kanalda ızgaraya yerleştirilmiş bir kuvvet ölçer ile su üzerinde bulunan buzların su akışını etkilediğine dair farklı bir çalışma yapmıştır. Yapılan araştırmalar, küçük hidrolik santrallerde ızgarada biriken buzun üretimi olumsuz etkileyebileceğini, buna bağlı olarak verimin sürüklenme kuvvetinin artmasıyla (düşü kayıplarının artmasıyla) ilişkili olduğunu göstermiştir. Çalışmada ayrıca ızgaralardaki buz görünümünün düşü kayıplarını %14'e kadar arttırdığı tespitinde bulunulmuştur (Walczak vd, 2020).

Himreen yaptığı çalışmada hidroelektrik santralinde hidrolik net düşü, verimlilik, su debisi gibi değişkenler üzerinden santralin üretim değerlerini yapay sinir ağları (ileri yayılım-geri beslemeli) yardımıyla tahmin etmiştir. Buna bağlı olarak üretim talebinin karşılanmasında YSA modelinin kullanılması üretimin daha doğru ve etkili şekilde yapılmasını sağlamıştır (Hamid vd., 2016).

Karakaya Hidroelektrik Santralinin verim ve elektrik üretimi değerlerinin tahmini yapay sinir ağları ile çalışılmıştır. Yapay sinir ağlarının girdileri olarak; ortalama nem (%), yağış miktarı (mm), açık yüzey buharlaşma miktarı, ortalama basınç (hPa), ortalama bulutluluk ve ortalama sıcaklık (°C), özgül su sarfıyatı (m^3/kWh), ortalama net düşü (m), gelen suyun debisi (m^3/sn) değerlerini kullanmışlardır. Yapay sinir ağlarının çıkış verileri ise, yüzdelik verim ve üretim değeri olarak alınmıştır. Yapay sinir ağlarının oluşturulmasında MATLAB programı kullanılmıştır. Çalışma sonucunda Karakaya Hidroelektrik Santralinin aktif elektrik üretim değeri ve santralin enerji verimi yüksek bir performans ve çok düşük hatalarla tahmin edilmiştir (İnalı vd., 2014).

2014 yılında Karakus ve Altun, Hirfanlı barajı örneğinde yapay sinir ağları ile hidroelektrik santralinin verim- üretim tahminini yapmışlardır. Bu araştırmada meteorolojiden alınan buharlaşma, yağış miktarı vb. veriler alınarak verim ve üretim tahmini yapılmıştır. Çalışmada sonuç olarak hidroelektrik santralinin verimlilik tahmini için iklimsel faktörlerin etkisi olduğu tespitinde bulunulmuştur (Karakus ve Altun, 2014).

2016 yılında ise Makas ve Karaatlı yapmış oldukları çalışmada, Isparta ili Aksu ilçesinde yer alan kaynağını Yenice deresinden alan bir hidroelektrik santralin bir yıllık elektrik üretimi tahminini yapmışlardır. Aylık ortalama olarak

alınan yağış, nem ve debi verileri Yapay Sinir Ağlarına (YSA) girdi olarak alınmıştır. 2007-2013 tarihlerinde alınan veriler arasındaki ilişki ileri beslemeli geri yayılım sinir ağı yöntemi ile analiz edilip 2014 yılı için yüksek doğrulukta üretim tahmininde bulunmuşlardır. Sonrasında aynı çalışmaya 2014 verileri eklenerek aynı YSA modeli ile 2015 yılı elektrik üretimi tahmin edilmiştir (Makas ve Karaatlı, 2016).

Emeç vd., Türkiye’deki Hidroelektrik Santrallerin Etkinliklerinin Veri Zarflama Analizi ile Değerlendirilmesi adlı çalışmada 51 HES’in analizini yapmışlardır. Çalışmada kurulu güç, üretim kapasitesi (yıllık), harcanması gereken su miktarı girdi olarak alınmıştır. Ne kadar üretim yapıldığı ve üretim ile kaç insan-hanenin elektrik ihtiyacının karşılandığı bilgisi çıktı olarak alınmıştır. Buradan hareketle etkin santral oranı %19.1 olarak tespit edilmiştir (Emeç vd., 2019).

Diniz ve ark., Brezilya’da bulunan iki hidroelektrik santral için, ortalama birim verimlilik faktörü ve elde edilen üretim ile ünitenin çalışma koşullarına göre verimlilik değişimi üzerinden yapılan üretim arasında bir değerlendirme yapmışlar. Bu çalışmada iki farklı matematiksel model geliştirilmiş ve türbin generatör verimi analiz edilmiştir. Çalışmada ünite verimlilik eğrilerine bağlı kalınarak üretimin planlı şekilde yapılması ve düşünün dikkate alınmasıyla çalışmanın fayda vereceği öngörülmüştür (Diniz vd., 2007).

Bragalli vd., net düş ve verimlilik dalgalanmalarının birleşik etkisini göstermek için küresel verimlilik kavramını kullanmışlardır. Deşarj alanındaki kaybolan enerjinin işlenmiş akış hızıyla ilişkilendirilmesi yaklaşımını benimsemişlerdir. Bu yaklaşım, Bologna’nın (İtalya) merkezinde yer alan mevcut Cavaticcio nehir tipi santralinde uygulanmıştır. Sonuçlar, deşarj sırasında kaybolan enerjideki dalgalanmaların, net yük değerini yukarı akış kayıplarından daha fazla etkileyebileceğini göstermektedir (Bragalli vd., 2023).

Alatas vd., Mikro hidroelektrik santrali girişinde çöp tıkanmasını en aza indirmek için yatay çöp rafı yönlendirici sistem çalışması ile ilgili bir çalışma yapmışlar. Bu çalışmada Endonezya da Yogyakarta da bulunan 8 adet mini HES’in ortak problemi olan ızgara kirliliğinin santral çalışmasına olumsuz etkisini problem olarak belirlemiştir. Problem çözümü için ızgara tasarımı değiştirilerek bu kirliliğin üretime olan olumsuz etkisini mümkün olduğunca ortadan kaldırılması amaçlanmıştır. Çalışmada kirliliğin etkisini azaltmak için Yatay Çöp Rafı Yönlendirici Çöp Kutusu ile uygun teknoloji yeniliğine ihtiyaç olduğunu belirtmişlerdir. (Alatas vd., 2022).

Nehir tipi hidrolik santraller; nehir üzerinde kurulu, göl kapasitesi olmayan ya da çok az olan yapılardır. Su tutma kapasitelerinin az olması nedeni ile gelen suyun bir an önce üretime çevrilmesi gerekmektedir. Bu tip santrallerde genellikle ızgara kirlilik problemi yaşanmaktadır ve bu durum üretim kaybına yol açmaktadır. Bu çalışmada, ızgara kirliliğinin tespiti veya tespitindeki hataların ortadan kaldırılması için makine öğrenmesi temelli bir tahmin sistemi geliştirilmiştir. İlk olarak, çalışma sahası olarak 189 MW kurulu güce sahip olan Karkamış HES seçilmiştir. Santralden alınan güç, debi, düşü ve ızgara kirlilik örneklerinden oluşan bir veri seti toplanarak makine öğrenme yöntemlerine giriş olarak uygulanmıştır. Daha sonra ise farklı üniteler için test çalışmaları gerçekleştirilerek detaylı bir analiz gerçekleştirilmiştir. Çalışmanın diğer kısımları ise kısaca şu şekilde özetlenebilir. 2. Bölüm Karkamış HES’ten elde edilen veri seti hakkında genel bilgileri içermektedir. 3. Bölümde kullanılan makine öğrenme yöntemlerinin detaylı açıklamaları sunulmuştur. 4. Bölümde ise deneysel çalışmalar sunulmuştur. Son olarak, elde edilen sonuçlar ışığında genel bir değerlendirme ise 5. Bölümde gerçekleştirilmiştir.

VERİ SETİNİN TOPLANMA AŞAMALARI

Karkamış Hidroelektrik Santrali; Gaziantep ilinin Karkamış ilçesine 14 km mesafede yer almaktadır. Güneydoğu Anadolu Projesi (GAP) kapsamında Fırat Nehri üzerinde inşa edilmiş son barajdır. Karkamış HES 6 adet ünitesinin herbiri 31,5 MW enerji kapasitesi ile toplam 189 MW kurulu güce sahiptir. Yıllık üretim kapasitesi 652 milyon kWh olarak hesaplanmaktadır. Fırat Havzası’nda bulunan Karkamış Barajı ve Hidroelektrik Santrali’nin inşaatına 07.12.1995 tarihinde başlanmış ve 1999 yılında tamamlanarak işletmeye alınmıştır. Kıl çekirdekli dolgu tipinde inşa edilen barajın gövde hacmi 2.100.000 m³, dere yatağından yüksekliği 29 m, gövde uzunluğu ise 1.608 m’dir. Gölün hacmi 156,9 hm³, normal su seviyesindeki gölalanı ise 28,4 km²’dir (EÜAŞ,2024). Karkamış HES’e ait fotoğraf ise Şekil 3’te verilmiştir.



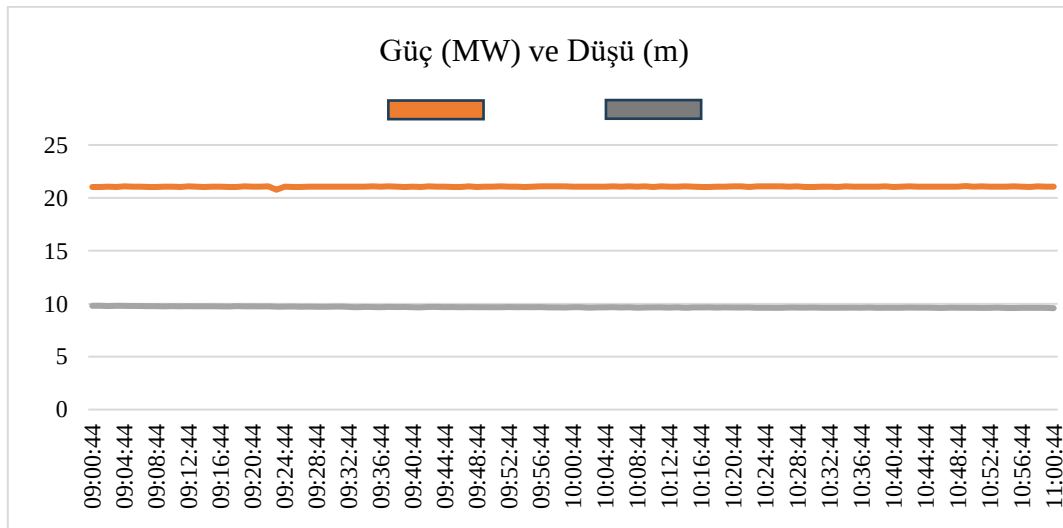
Şekil 3. Karkamış Hidroelektrik Santrali

Nehir tipi ve yatay kaplan bir santral olan Karkamış HES’inde kronik problemleri arasında ızgara kirliliği yer almaktadır. Yosun ve çevresel atıklar ızgarada birikip su akışını engellemektedir. Bu durum; ızgara temizken üretilen enerjiyi, ızgara kirliyken üretmeye çalıştığımızda çok daha fazla su kullanmamıza ve verimsiz çalışmaya neden olmaktadır. Santralin ızgara temizliğinin yapılmasına ait çalışma ise Şekil 4’te gösterilmiştir.

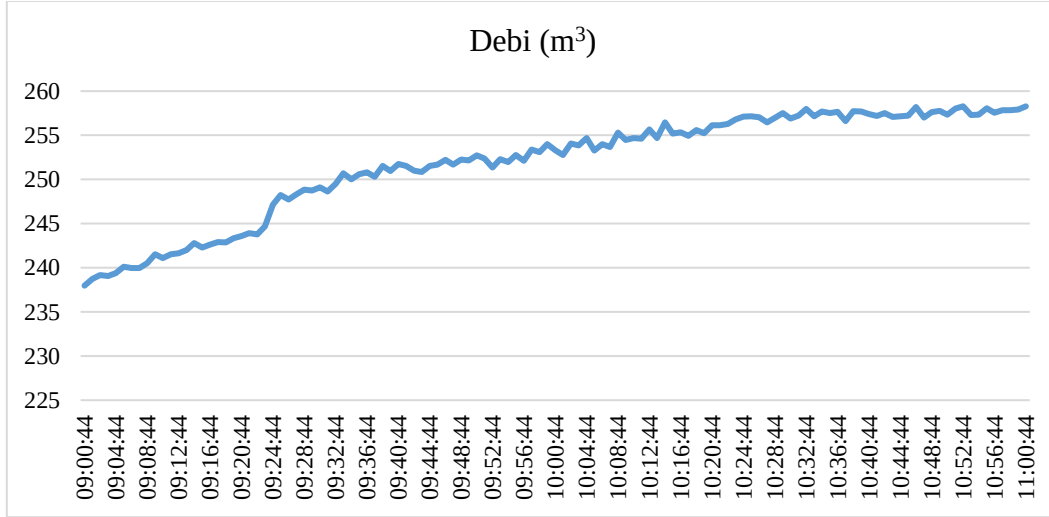


Şekil 4. Karkamış HES ızgara temizlik çalışması

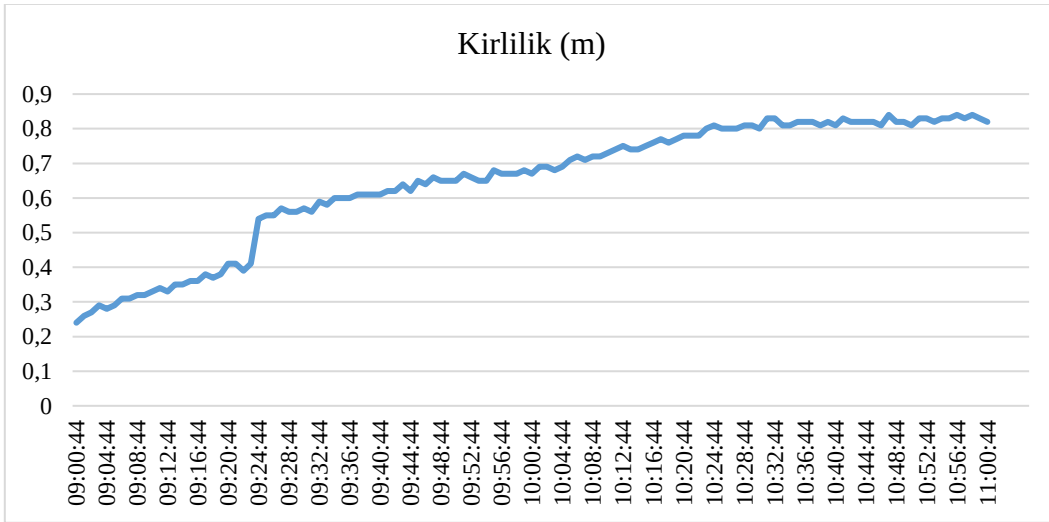
Karkamış HES’in 5. Ünitesinde bu durumun örneği aşağıdaki grafiklerle anlatılmıştır. Karkamış HES Ünite-5 için 2 saat içerisinde ızgara kirliliğinin artması ile debi miktarındaki değişim örneğini aşağıda verilmiştir.



Şekil 5. Karkamış HES Ünite-5’e ait güç ve düşü grafiği

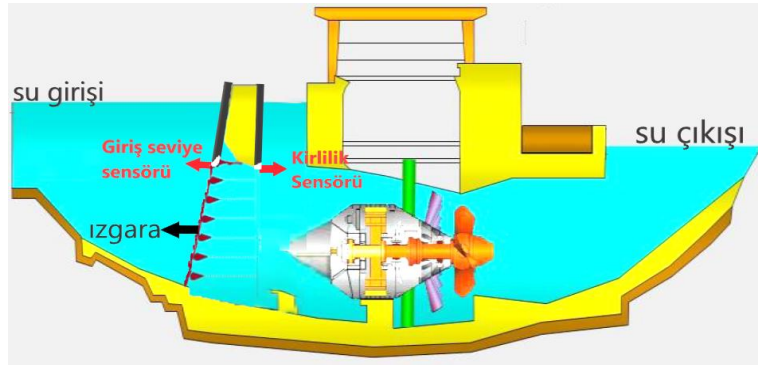


Şekil 6. Karkamış HES Ünite-5 için debi grafiği



Şekil 7. Karkamış HES Ünite-5 kirlilik grafiği

Bu örnekte üretim değerlerinin sabit tutulması için artan debi miktarı görülmektedir. Bu durum verimsiz çalışmayı göstermektedir. Çalışma başlangıcından sonuna kadar 21 MW üretim yapıldığı ve düşü miktarının da değişmediği Şekil 5' te görülmektedir. Ancak üretim ve düşü sabitken kullanılan su miktarı ile ızgara kirliliğinin doğru orantılı olarak arttığı Şekil 6-7'de görülmektedir. Izgara kirliliğindeki artış su kullanımında %19 artış yaşanmasına neden olmuştur. Bu durum ızgara kirliliğinin oluşturduğu verimsiz çalışmaya bir örnektir.



Şekil 8. Kirlilik ölçüm sensörlerinin santraldeki konumları

Izgara kirlilik tespiti, Şekil 8’de gösterilen ızgara girişine ve ızgara sonrasında dinlenme havuzu olarak adlandırılan konuma yerleştirilen basınç sensörleri ile yapılır. Bu basınç ölçüm değerleri SCADA sisteminde kot bilgisi olarak gösterilir. İki sensör arasındaki fark kirliliği verir. Bu sistemde daldırma tip basınç sensörleri kullanılmaktadır. Basınç sensörleri statik basınç değişimlerini ölçerler. Kablolar yardımıyla basınç değişimleri kumanda odasındaki SCADA sistemine aktarılır. Sensörün sabitlendiği yerin kot bilgisi yazılıma girilir. Sensör su yüksekliğine göre basınç bilgisini okur ve sisteme bir değer olarak aktarır. Bu değer yazılımda mevcut kot değerleri ile toplanır ve su seviye bilgisi elde edilir. Kullanılan sensöre ait teknik özellikler Tablo 1’de verilmiştir. Şekil 9’da arızalı ölçüm yapan bir sensör fotoğrafı verilmiştir.

Tablo 1. Seviye ölçüm sensör özellikleri

Özellikler	Teknik Açıklama
Kulandan kısım	Izgara kirliliği seviye ölçüm sensörleri
Tipi	Hidrostatik daldırma tip basınç sensörü
Çalışma aralığı	0-3 bar - 0-30 metre su sütunu
Hassasiyet	$\pm \% 0,1$
Çalışma akımı	4-20 mA

**Şekil 9.** Arızalı seviye ölçüm sensörü**MAKİNE ÖĞRENME YÖNTEMLERİ**

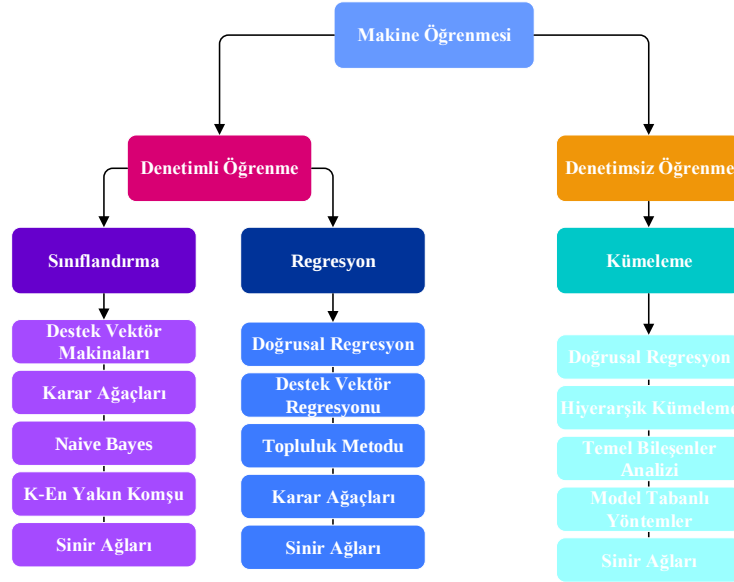
Bilgisayarda bir problemi çözmek için uygun bir algoritmaya ihtiyaç duyulmaktadır. Algoritma, sonuca ulaşabilmek için mevcut verilere kural ve formüllerin uygulanmasıdır. Veri setinde bulunan girdi-çıkı arasında bağlantının karmaşık olmaması durumlarında algoritma ile problem çözülür. Ancak algoritmanın oluşturulamadığı veya yeterli olmadığı durumlarda örnek veriler kullanılarak makine öğrenmesi ile sonuca ulaşılması hedeflenir. Makine öğrenimi ile eğitim gerçekleştirmek için istatistik teorisinden yararlanarak matematiksel modeller oluşturularak örneklerden çıkarım yapılmaktadır (Alpaydın, 2014). Bu çalışmada ise makine öğrenim programlarının eğitimi için MATLAB’da bulunan Regression Learner kütüphanesinden yararlanılmıştır.

Makine öğrenme yöntemleri sınıflandırılması; denetimli öğrenme, denetimsiz öğrenme ve destekleyici öğrenme olmak üzere üç başlık altında yapılmaktadır. Literatürde genellikle öğrenme yöntemleri denetimli ve denetimsiz olmak üzere iki grupta sunulmaktadır. Denetimli öğrenme modellerinde sınıflandırma ve regresyon kullanılırken denetimsiz öğrenme modelleri ise kümeleme uygulamalarında kullanılmaktadır (Metlek ve Kayaalp, 2020).

Denetimli öğrenme, girdi ve çıktı verilerine dayalı bir tahmin modelidir. Denetimli öğrenme algoritmaları; sınıflandırma ve regresyon algoritmalarını içermektedir. Denetimli öğrenmede çıktı, sınırlı bir değer kümesiyle sınırlandırıldığında sınıflandırma ile öğrenme yapılmaktadır. Ancak çıktı bir aralık içinde herhangi bir sayısal değere sahip olduğunda ise regresyon algoritmaları kullanılmaktadır (Ray, 2019).

Denetimsiz (gözetimsiz, eğitici-siz) öğrenme, istenen çıktı sonucu için özel bir tanımlama/etiketleme/sınıflandırma yapılmadığında veya belirsizlik durumu söz konusu olduğunda veri kümesindeki gizli kalmış ilişki ve örüntüleri

tanımlamak için algoritmaların kullanılması anlamına gelmektedir. Diğer bir deyişle, bir hedef değişken (çıktı değeri) bulunmayıp verileri ilişkisel yapıya göre kümeye ayırma hali söz konusu olmaktadır. Algoritmaların dışardan bir rehberliğe sahip olmadan, kendi kendine veri noktalarını gruplamasına olanak tanınır. Denetimsiz yöntemler veriyi tanımlamak amacıyla kullanılır. Denetimsiz makine öğrenmesi kümeleme yöntemleri olarak incelenir. Regresyon, üzerinde durulan yanıt değişkeni ile ilişkili olabileceği düşünülen açıklayıcı değişkenlerin bir fonksiyonu olarak ifade edilmektedir. Değişkenler arasındaki ilişkinin fonksiyonel şekli regresyon modelleri ile incelenmektedir (Arı ve Önder, 2013). Makine öğrenme modellerinin sınıflandırılması Şekil 10'daki gibidir.

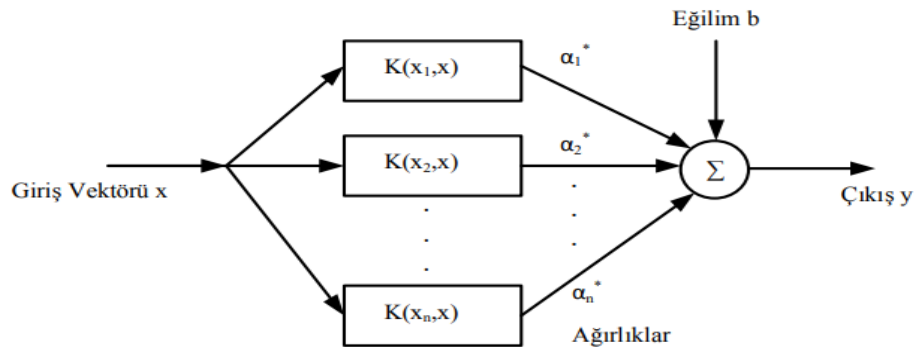


Şekil 10. Makine Öğrenme Modelleri

Bu çalışmada kullanılan veri setindeki girdi- çıktılar ilişkili olması ve çıktılar sayısal bir değere sahip olması gibi etkenlerden dolayı denetimli öğrenme modeli kullanılmıştır. Denetimli öğrenme içerisinde yüksek boyutlu veri setinde başarılı olan destek vektör makineleri ve yüksek öğrenim performansı gösteren karar ağaç modelleri ile çalışılmıştır.

Destek Vektör Makineleri

Makine öğrenim yöntemi modellerinden destek vektör makineleri (DVM), sınıflandırma ve regresyon problemlerinde çokça kullanılır. DVM, ilk olarak Vladimir Vabnik ve Alexey Shervonenikis'in çalışmaları ile ortaya çıkmıştır (Daş vd., 2019). Destek vektörü regresyonu, veri setindeki etiketli değerler ile sürekli yeni değer üretimini sağlayan bir model kullanan denetimli öğrenme algoritmasıdır. Geleneksel olarak yapılan regresyon modellerinde hedef hatayı en aza indirecek bir fonksiyonla algoritmayı oluşturmaktır. Bunun yanı sıra DVM regresyonu, veri setindeki değerlerden çıkan sonuçların marjı yüksek olanlar için toleranslı bir algoritma da oluşturur (Bekçioğulları vd., 2021). Destek vektör makinalarının genel yapısı Şekil 11'de gösterildiği gibidir.



Şekil 11. Destek vektör makinalarının genel yapısı

Bu durum destek vektör regresyonlarının diğer yöntemlere nazaran genel hata ve uygulama hatalarının üst sınırı aşağı çekerek hata riskini düşürür. Veri setinde bulunan ve doğrusal olmayan giriş-çıkış verilerinin kümesi $T = \{(x_i, y_i) : i = 1, 2, \dots, N\}$ gösterilmektedir. Tahmin değerini gösterecek fonksiyon ise $f(x) = wx + b$ şeklinde ifade edilir. Burada kullanılan w ve b değerleri hata oranını minimize etmek için kullanılan çeşitli fonksiyonlarla hesaplanır. Bu fonksiyonlar gerçek değer ile hesaplanan tahmin değeri arasındaki farkı azaltacak yöntemleri içerir (Yoldaş, 2023).

$$|y - f(x)| = \max\{0 |y - f(x)| - \varepsilon\}, \varepsilon > 0 \quad (1)$$

$$\min_{w, b} \frac{1}{2} w^T \cdot w + c \sum_{i=1}^i (\xi_i + \xi_i^*) \quad (2)$$

Denklem 2’de c sabit olup dengeleme ve düzgünleştirme yapar. Bu formülde kısıt, hata değerinin gerçek değerle hedef değer arasındaki farktan mutlak manada küçük olmasıdır. Doğrusal olmayan durumlarda uygulanan denklem ise aşağıda verilmiştir.

$$\max_s \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (a_i^+ - a_i^-) (a_j^+ - a_j^-) k(x_i^+ \cdot x_j^-) - \varepsilon \sum_{i=1}^N (a_i^+ - a_i^-) - \sum_{i=1}^N y_i (a_i^+ - a_i^-) \quad (3)$$

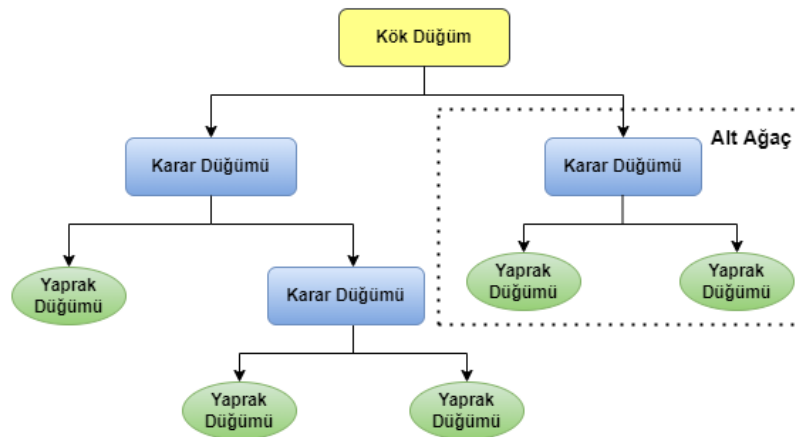
Denklem 4 ile ana fonksiyon denklemleri gösterilmiştir. Denklem 5’te ise ağırlık çarpan değerlerini elde edilmektedir.

$$f(x) = \sum_{i=1}^N (a_j^+ - a_j^-) k(x_i x_j) + b \quad (4)$$

$$w = \sum_{i=1}^N (a_j^+ - a_j^-) \phi(x_i) \quad (5)$$

Karar Ağaçları

Karar ağaçları, çok değişkenli analizin basit ama güçlü bir şeklidir. Verileri dal benzeri bölümlere ayırmanın çeşitli yollarını tanımlayan algoritmalar tarafından üretilirler. Karar ağaçları, verileri girdi değişkenlerinin kategorilerine göre alt kümelerle bölerek birinin karar yolunu anlamaya yardımcı olur (Karakuş, 2001). Karar ağaçları veri setindeki ilişkileri çözerek ağaç diyagramları ile gösterebilmektedir. Karar ağaçlarının kolay ve anlaşılabilir olmasından dolayı, birçok makine öğrenme uygulamasında kullanılmaktadır. Girdi ve çıktı veri setlerinden oluşan KA denetimli bir öğrenme yöntemi ile çalışmaktadır. Bu denetimli yapıda KA, girdi ve çıktı arasındaki bağlantıyı tespit etmektedir (Metlek ve Kayaalp, 2020).



Şekil 12. Karar Ağaç Grafiği

Şekil 12’de karar ağaç modelinin oluşmasına ait bir grafik verilmiştir. Grafikte görüleceği üzere modelin oluşturduğu ağacın tepesinde yer alan eleman kök düğüm olarak isimlendirilmektedir. Kök düğümün altında yer alan her alt düğümün oluşması belirgin bir özellik değerinin farklılığına bağlıdır. Ağaç son düğümlerin oluşmasıyla

tamamlanacak olup bu son düğümlere yaprak adı verilir. Ana kök düğüm ile yaprak arasında kalan kısma dal denilir. Tüm veriler için model uygulanıp yaprak düğümlerinin bitimi ile sonlanır (Daş ve Türkoğlu, 2014).

Karar ağacı ile yapılan regresyon, bir bağlı değişkenin diğer değişkenlerin ilişkisine göre değerlendirip olası değişiklikleri tahmini için uygulanan bir algoritmadır (Kingsford ve Salzberg, 2008). Karar ağacı oluşturmak için kullanılan AID (Automatic Interaction Detector - Otomatik Değişken Seçimi Algoritması), ID3 (Iterative Dichotomiser 3), C4.5, C5, CART (ya da C&RT), CHAID, MARS, SLIQ, SPRINT ve QUEST algoritmalarıdır. Karar ağaç modellerinde en çok kullanılan modellerden CART ve CHAID algoritmaları Denklem 6 ve 7’de verilen eşitlik ile ifade edilebilir. CART algoritmasında her veri, dal dediğimiz iki kısma ayrıldıktan sonra ilk kısımdakiler kendi içerisinde ikinci kısımdakiler ise kendi içerisinde ayrılır. CART algoritmasında tüm değişkenler iki dalın bir yanında yer alırken soldaki grup a ve sağdaki b olarak gösterilir.

$$Gini_a = 1 - \sum_{i=1}^k \left(\frac{a_i}{Toplam_a} \right)^2 \quad Gini_b = 1 - \sum_{i=1}^k \left(\frac{b_i}{Toplam_b} \right)^2 \quad (6)$$

Sonrasında aşağıda verilen formül ile her bir değişken için ağırlıklandırılma yani $Gini_j$ değerleri elde edilir.

$$Gini_{ij} = \frac{1}{n} (Toplama * Gini_a + Toplamb * Gini_b) \quad (7)$$

Gini değeri en düşük olan değer düğüme yerleştirilir ve bölünmeye başlanır. Yeni düğümlerle tekrar eden bu süreç sonlanıp yapraklar bitene kadar devam eder. CHAID (Ki-Kare Otomatik Etkileşim Dedektörü) algoritması CART’ tan farklı olarak ikiden fazla dal üretebilmektedir. Eğitim verileri üzerinden hedef değişken ile bağımsız değişkenler arasında ikili olarak çapraz tablolar oluşturulur. Oluşturulan tüm ikili değerler için Denklem 8 kullanılır ve değerler hesaplanır. Bu değerler içerisinde en büyük değere sahip hedef değerle bağlantısı en kuvvetli olan değişken kök düğüme yerleştirilir. Sonrasında ayrılacak dal kalmayana kadar bölünme devam eder (Gacar, 2023).

$$X^2 = \sum_{i=1}^k \left(\frac{(Gözlenen\ değer - Beklenen\ değer)^2}{Beklenen\ değer} \right) \quad (8)$$

DENEYSSEL ÇALIŞMALAR

Bu bölümde, makine öğrenme yöntemleri ile ızgara kirlilik tahmini için deneysel çalışmalar gerçekleştirilmiştir. İlk olarak, Karkamış HES SCADA sisteminden 2022 ve 2023 yıllarına ait güç, debi, düşü ve ızgara kirlilik örneklerine ait sayısal değerler toplanmıştır. Toplanan veri setindeki ölçüm ve haberleşme hatalarının oluşturduğu anlamsız değerler tespit edilip filtreden geçirilerek veri setinden çıkarılmıştır. Böylece veri seti hatalardan ve anlamsız değerlerden arındırılmış olarak deneysel çalışma için hazır hale getirilmiştir. Veri setindeki tüm örneklerin %70’i eğitim, %15’i doğrulama ve %15’i de test için ayrılmıştır. Eğitim için ayrılan veri seti makine öğrenme yöntemlerine giriş olarak uygulanarak sonuçlar elde edilmiştir. Makine öğrenme yöntemi olarak ise Karar Ağacı ve DVM tercih edilmiştir. Çalışmada kullanılan makine öğrenme yöntemleri MATLAB/Regression Learner kütüphanesinde bulunan hazır modellerden elde edilmiştir. Tahmin işlemi için yazılan tüm kodlar MATLAB ortamında yazılarak Intel(R) Core(TM) i7-7700HQ CPU@ 2.80GHz, NVIDIA GeForce GTX 1050 ve 16 GB RAM özellikli bir iş istasyonu üzerinden gerçekleştirilmiştir. Ayrıca, deneysel çalışmalardan elde edilen sonuçların değerlendirilebilmesi için, korelasyon katsayısı (R), hataların karesinin ortalamasının karekökü (RMSE) ve hataların mutlak değerlerinin ortalaması (MAE) gibi performans ölçütleri kullanılmıştır. Bu ölçütler kısaca şu şekilde açıklanabilir:

RMSE, uygulanan bir model sonucunda elde edilen tahminler ile gerçek değerler arasındaki sapmayı yani hatayı gösterir. RMSE, tahmin ile gerçek değer arasındaki farkları değerlendirmek için çok yaygın olarak kullanılan hata ölçüm tekniğidir. RMSE değerlerinin sıfıra yakın olması uygulanan modelin performansının yüksek olduğunu gösterir. RMSE, Denklem 9’daki gibi hesaplanabilir (Artuğer & Özkaynak, 2022).

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (P_k^p - \hat{P}_k)^2} \quad (9)$$

R ise bir modelin veri setindeki değişkenliği ne kadar iyi açıkladığını ölçer. 1'e yaklaştıkça, model veriyi daha iyi açıklar. R değeri, verilerin regresyon hattına ne kadar yakın olduğunun istatistiksel bir ölçütüdür ve aşağıdaki gibi matematiksel olarak ifade edilebilir (Zilyas & Yılmaz, 2023).

$$R = \frac{\sum_{k=1}^n [(P_k^p - \hat{P}_k^p)(P_k - \hat{P}_k)]}{\sqrt{\sum_{k=1}^n (P_k^p - \hat{P}_k^p)^2 \sum_{k=1}^n (P_k - \hat{P}_k)}} \quad (10)$$

MAE (Ortalama Mutlak Hata), hatanın mutlak ortalamasıdır. Gerçek değerler ile tahmin sonucu bulunan değerlerin farklarının ortalaması alınarak hesaplanmaktadır. Tahminler ile gerçek değerler arasındaki mutlak farkların ortalamasıdır. MAE değerinin formülü Denklem 11'deki gibi verilebilir.

$$MAE = \frac{1}{n} \sum_{k=1}^n (P_k^p - P_k) \quad (11)$$

Yukarıda verilen denklemlerde, P_k ve P_k^p ölçülen ve tahmin edilen güçtür. $\hat{P}_k$ ve $\hat{P}_k^p$ ise ölçülen ve tahmin edilen gücün ortalamasıdır.

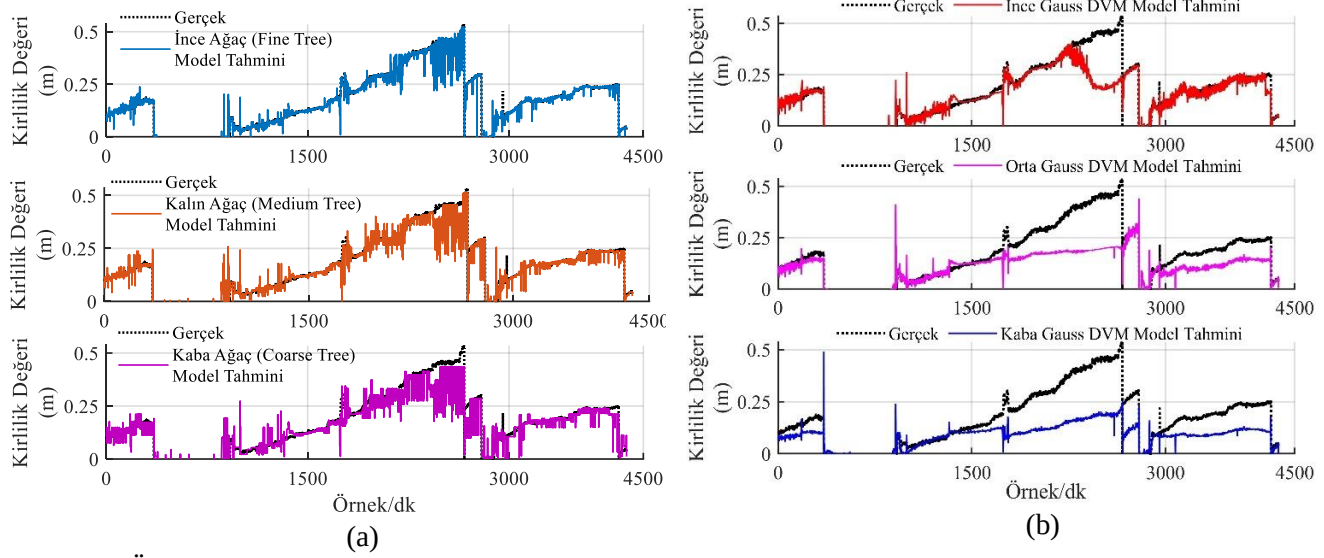
Deneyisel Çalışma-1: Ünite-1 İçin Izgara Kirlilik Tahmin Performanslarının İncelenmesi

Bu bölümde, Karkamış HES'te bulunan Ünite-1 için tahmin çalışması gerçekleştirilmiştir. Veri seti olarak üniteye ait güç, debi, düşü ve ızgara kirliliğinden oluşan 90.000 örnek ile bir veri toplanmıştır. Bu veriler üzerinde yapılan filtrelemeler sonucu 16.376 örnek içeren veri seti elde edilmiştir. Ünite-1'in üretim yaptığı zamanlarda ölçülen ızgara kirlilik değerlerinin en değişken olduğu aralık veri setini belirlemede diğer bir etken olmuştur. Yapılan çalışmalar sonucundaki ölçüt değerleri Tablo 2'de listelenmiştir. Tabloda yer alan en iyi sonuçlar koyu renk ile belirtilmiştir. Ünite-1 için yapılan çalışmada kullanılan tüm makine öğrenim modellerine ait tahmin sonuçları Şekil 13'de sunulmuştur.

Tablo 2. Makine öğrenme yöntemlerinin Ünite-1'deki ölçüt sonuçları

Metot	R	RMSE	MAE
İnce Ağaç Modeli	0.9833	0.0266	0.0103
Kaba Ağaç Modeli	0,9691	0,0365	0,0170
Kalın Ağaç Modeli	0,9405	0,0512	0,0283
İnce Gauss DVM Modeli	0,8803	0,0734	0,0327
Orta Gauss DVM Modeli	0,6738	0,0680	0,1016
Kaba Gauss DVM Modeli	0,6042	0,0920	0,1220

Tablo 2'deki sayısal sonuçlar incelendiğinde, Ünite-1 için en iyi R değerini 0.9833 ile İnce Ağaç modeli verirken, en kötü R değerini 0.6042 ile Kaba Gauss DVM modeli vermiştir. Diğer performans ölçütü olan RMSE değeri kıyaslandığında aynı şekilde 0.0266 değeri ile ince ağaç modeli en iyi sonucu vermiştir. Diğer modellerin R performans ölçüt değerleri ise kaba ağaç modeli için 0.9691, kalın ağaç modeli için 0.9405, İnce Gauss DVM modeli için 0.8803 ve Orta Gauss DVM modeli için 0.6738 olarak hesaplanmıştır. Ünite-1 için yapılan çalışmada en iyi tahmin sonuçlarını veren model ince ağaç ile en kötü tahminde bulunan model Kaba Gauss DVM modeli R performans ölçütü açısından kıyaslandığında %62.74 daha iyi performans göstermiştir. Şekil 13'te görüleceği üzere Karar ağaçları ile elde edilen sonuçlar hem daha iyi hem de daha güvenilirdir. Burada, Karar ağaçlarının ızgara kirlilik tahmininde daha iyi bir genelleme yeteneğine sahip oldukları ifade edilebilir.



Şekil 13. Ünite-1 için karşılaştırma sonuçlarının gösterimi a) Karar Ağaç Modelleri b) Destek Vektör Makineleri

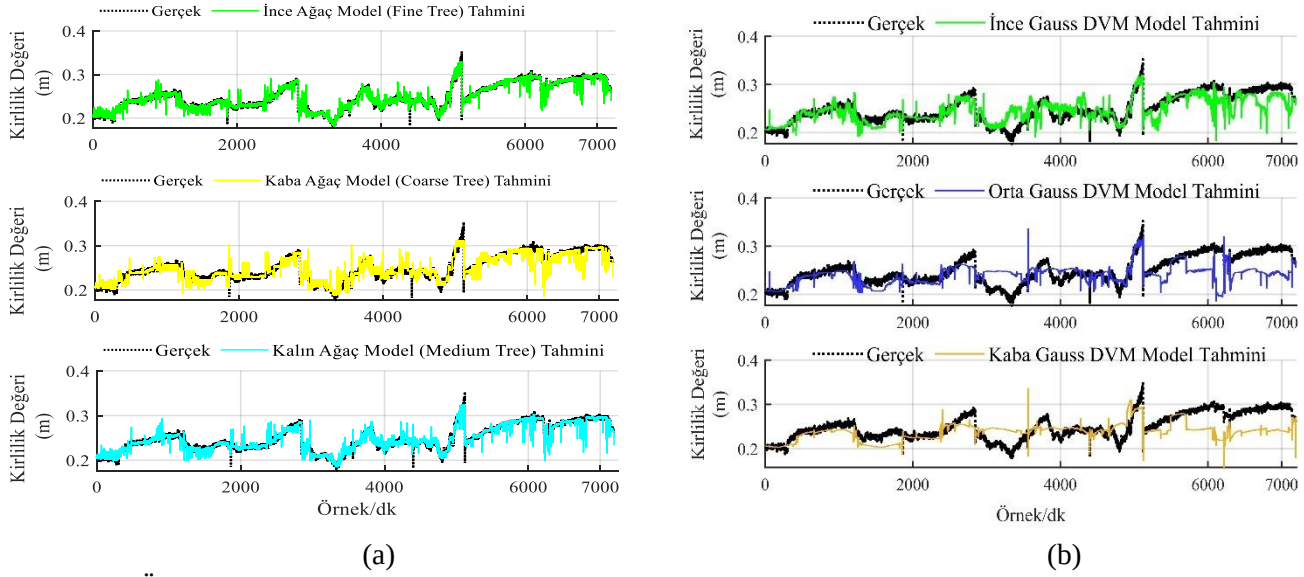
Deneyisel Çalışma-2: Ünite-2 İçin Izgara Kirlilik Tahmin Performanslarının İncelenmesi

Karkamış HES'in 2. Ünitesi için ızgara kirlilik tahmin çalışması gerçekleştirilmiştir. Veri seti için Karkamış HES SCADA sisteminden alınmış 78.962 örnek filtrelenmiş 47.000 satırlık veri seti elde edilmiştir. Bu veriler güç, debi, düşü ve kirlilik değerlerinden oluşmaktadır. Ünite-2'nin ölçülen ızgara kirlilik değeri ünite 1'e göre daha düşük ancak veri setindeki örnek sayısı daha yüksektir. Yapılan çalışmalar sonucundaki ölçüt değerleri Tablo 3'te listelenmiştir. Listede yer alan en iyi sonuçlar koyu renk ile belirtilmiştir. Ünite-2 için yapılan çalışmada kullanılan tüm makine öğrenme modellerine ait tahmin sonuçları Şekil 14'te sunulmuştur

Tablo 3. Makine öğrenme yöntemlerinin Ünite-2'deki ölçüt sonuçları

Metot	R	RMSE	MAE
İnce Ağaç Modeli	0,9716	0,0069	0,0040
Kaba Ağaç Modeli	0,9071	0,0123	0,0078
Kalın Ağaç Modeli	0,9486	0,0092	0,0056
İnce Gauss DVM Modeli	0,7751	0,0186	0,0129
Orta Gauss DVM Modeli	0,4746	0,0276	0,0200
Kaba Gauss DVM Modeli	0,4200	0,0292	0,0222

Ünite-2'ye ait çalışmalar için performans değerlerini içeren Tablo 3 incelendiğinde ince ağaç modeli R (0.9716), RMSE (0.0069) ve MAE (0.0040) değerleri ile tüm ölçütlerde en iyi sonucu vermiştir. Çalışmanın en kötü sonucu Kaba Gauss DVM Modeli; R için 0.4200, RMSE için 0.0292 ve MAE için 0.0222 değerleriyle vermiştir. R ölçüt değerleri Kaba ağaç modelinde 0.9071, Kalın ağaç modelinde 0.9486, İnce Gauss DVM Modeli modelinde 0.7751 ve Orta Gauss DVM Modeli modelinde 0.4746 olarak hesaplanmıştır. Çalışmanın en iyi sonuç veren modeli ile en kötü modeli Ünite-1 ile aynı modellerdir. Ünite-2 için yapılan çalışmada en iyi tahmin sonuçları veren model ince ağaç ile en kötü tahminde bulunan model Kaba Gauss DVM Modeli R performans ölçütü açısından kıyaslandığında %231.3 daha iyi performans elde edilmiştir. Şekil 14'te görüleceği üzere Karar ağaçları ile elde edilen sonuçlar daha başarılıdır.



Şekil 14. Ünite-2 için karşılaştırma sonuçlarının gösterimi a) Karar Ağaç Modelleri b) Destek Vektör Makineleri

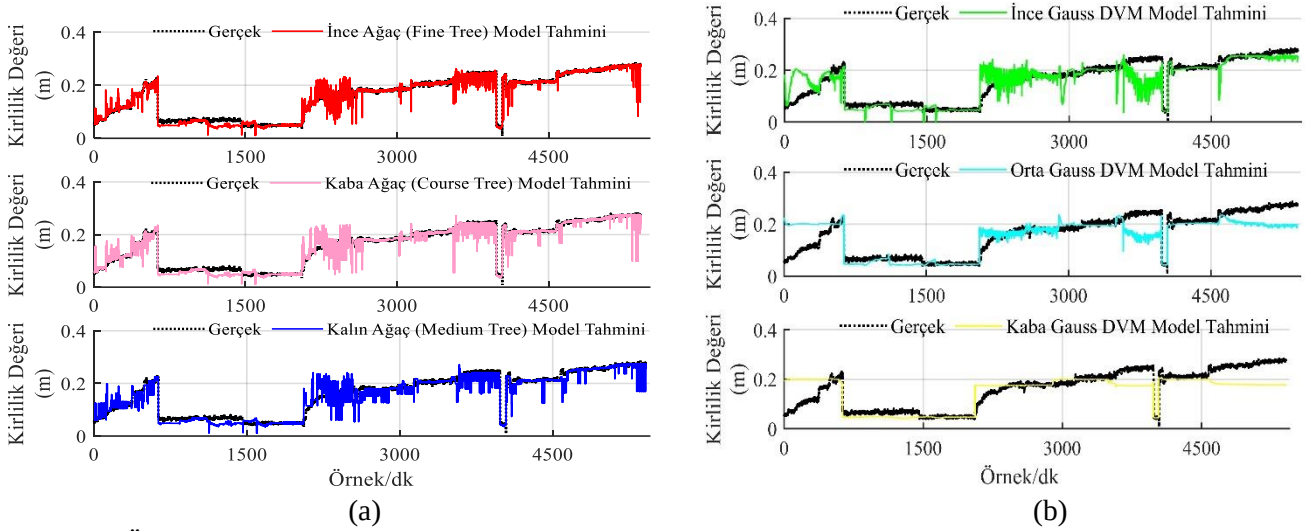
Deneysel Çalışma-3: Ünite-4 İçin İzgara Kirlilik Tahmin Performanslarının İncelenmesi

Ünite-4 için yapılan çalışmanın verileri SCADA sisteminden alınan güç, debi, düşü ve kirlilik değerlerinden oluşan örneklerdir. Bu örnekler incelenip filtreden geçirildikten sonra 36.087 örnek içeren veri seti elde edilmiştir. Ünite 4 veri seti örnek miktarı Ünite-1'den fazla Ünite-2'den düşüktür. Çalışma neticesinde elde edilen performans ölçüt değerleri liste şeklinde Tablo 4'te verilmiştir. Listede yer alan en iyi değerler koyu renk ile gösterilmiştir. Makine öğrenme yöntemlerinin sonuçları Şekil 15'te sunulmuştur.

Tablo 4. Makine öğrenme yöntemlerinin Ünite-4'teki ölçüt sonuçları

Metot	R	RMSE	MAE
İnce Ağaç Modeli	0,9792	0,0164	0,0088
Kaba Ağaç Modeli	0,9673	0,0203	0,0114
Kalın Ağaç Modeli	0,9563	0,0234	0,0140
İnce Gauss DVM Modeli	0,8612	0,0403	0,0246
Orta Gauss DVM Modeli	0,8015	0,0470	0,0323
Kaba Gauss DVM Modeli	0,7799	0,0497	0,0351

Tablo 4 verilen sayısal sonuçlar incelendiğinde, Ünite-4 için en iyi R değerini 0.9792 ile İnce Ağaç modeli verirken, en kötü R değerini 0.7799 ile Kaba Gauss DVM Modeli vermiştir. Diğer performans ölçütü olan RMSE değeri kıyaslandığında aynı şekilde 0.0164 değeri ile ince ağaç modeli en iyi sonucu vermiştir. Diğer modellerin R performans ölçüt değerleri ise 0.9673 ile kaba ağaç modeli, 0.9563 değeri ile kalın ağaç modeli, 0.8612 değeri ile İnce Gauss DVM Modeli ve 0.8015 değeri ile Orta Gauss DVM Modeli olarak hesaplanmıştır. En iyi sonuçların alındığı ince ağaç modelinin cevapları Kaba Gauss DVM Modeli ile kıyaslandığında %25.55 oranında daha iyi bir performans göstermiştir. DVM modellerinin ve karar ağaç modellerinin performans ölçütleri Ünite-2'den daha başarılıdır.



Şekil 15. Ünite-4 için karşılaştırma sonuçlarının gösterimi a) Karar Ağaç Modelleri b) Destek Vektör Makineleri

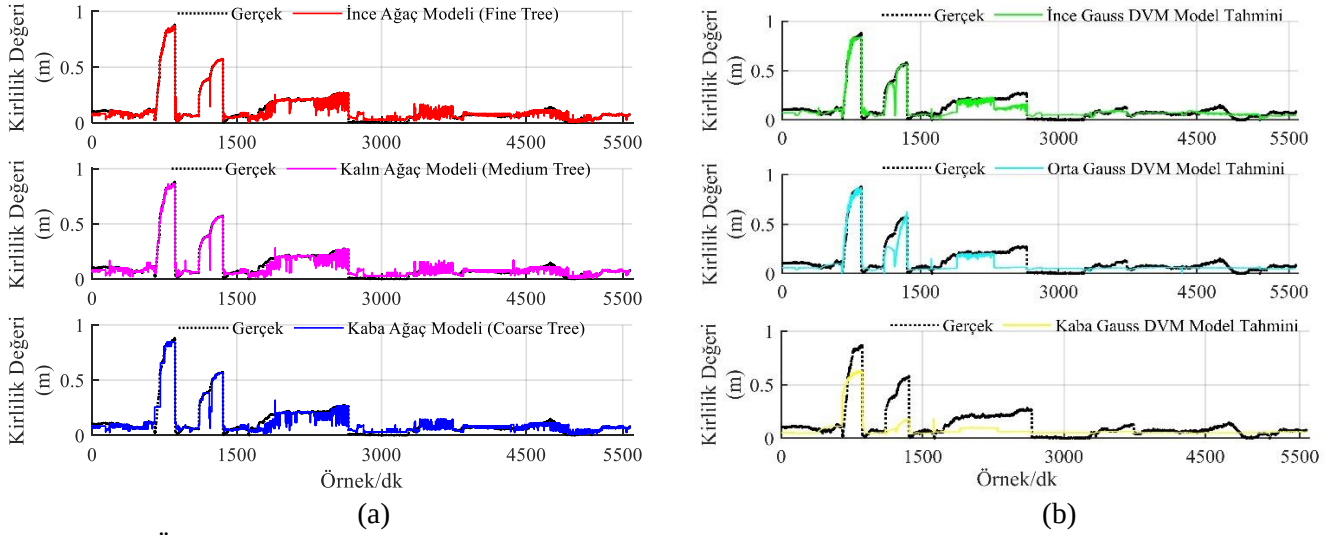
Deneysel Çalışma-4: Ünite-5 İçin İzgara Kirlilik Tahmin Performanslarının İncelenmesi

Bu bölümde; ünite 5 için yapılan tahmin çalışmada Ünite-5'e ait güç, debi, düşü ve ızgara kirlilik değerlerini içeren 144.000 örneğin toplandığı bir veri havuzu oluşturulmuştur. Bu veri miktarı filtrelenerek 36285 örneği içeren veri setine dönüşmüştür. Ünite-5 için hazırlanan veri seti Ünite-4'ün veri setine sayısal olarak benzerlik göstermektedir. Veri seti oluşturulurken ızgara kirliliğinin en değişken olduğu aralık önemli bir etken olmuştur. Elde edilen sayısal sonuçlar Tablo 5'de listelenmiştir. Tablo 5'de en iyi performans ölçütü koyu renk ile gösterilmiştir. Ayrıca, makine öğrenme yöntemlerinin tüm çeşitlerine ait tahmin cevapları Şekil 16'da sunulmuştur.

Tablo 5. Makine öğrenme yöntemlerinin Ünite-5'deki ölçüt sonuçları

Metot	R	RMSE	MAE
İnce Ağaç Modeli	0,9810	0,0296	0,0018
Kaba Ağaç Modeli	0,9769	0,0326	0,0203
Kalın Ağaç Modeli	0,9627	0,0414	0,0241
İnce Gauss DVM Modeli	0,9525	0,0480	0,0328
Orta Gauss DVM Modeli	0,9119	0,0675	0,0448
Kaba Gauss DVM Modeli	0,7839	0,1063	0,0669

Tablo 5'deki istatistiksel sonuçlar incelendiğinde, Ünite-5 için en iyi R değerini 0.9810 ile ince ağaç modeli verirken aynı ölçütteki en kötü sonucu 0.7839 değeri ile Kaba Gauss DVM Modeli vermiştir. Diğer performans ölçütü olan RMSE için ise yine en iyi sonucu ince ağaç modeli 0.0296 değeri ile sağlamıştır. Kaba ağaç modeli, kalın ağaç modeli, İnce Gauss DVM Modeli ve Orta Gauss DVM Modeli için R değerleri ise sırasıyla 0.9769, 0.9627, 0.9525 ve 0.9119 olarak hesaplanmıştır. Burada, en iyi sonucu veren ince ağaç modelinin R değeri en kötü performansa sahip Kaba Gauss DVM Modelinden %25.14 daha iyi sonuçlanmıştır. Grafikte görüldüğü gibi kirlilik değeri 1 m'ye yaklaşmıştır. Diğer ünitelerdeki tahmin çalışmalarının performans ölçütleri ile kıyaslandığında DVM modellerinin en başarılı sonucu bu çalışmada alınmıştır.



Şekil 16. Ünite-5 için karşılaştırma sonuçlarının gösterimi a) Karar Ağaç Modelleri b) Destek Vektör Makineleri

SONUÇ

Dünyada ve ülkemizde elektrik tüketimi ve buna bağlı olarak elektrik ihtiyacı sürekli artmaktadır. Enerji ihtiyacı için kullanılan fosil enerji kaynakları çevre kirliliğinin ve küresel ısınmanın artmasına sebep olmaktadır. Fosil enerji kaynaklarının kullanımının azaltılıp yerine çevre dostu yenilenebilir enerji kaynaklarının kullanımı yaygınlaştırılmalıdır. Yenilenebilir enerji kaynaklarının etkin ve verimli şekilde kullanılması ihtiyaç olan enerjide fosile olan bağımlılığı azaltacaktır. Bu sebeple yenilenebilir enerji kaynaklarının küresel olarak üretim kapasitesinin en büyük payını oluşturan Hidrolik kaynaklarda yapılacak verimlilik çalışmaları büyük katkı sağlayacaktır. Bu çalışmada HES’lerde makine öğrenme metotları kullanılarak arızaları tespit etmek, daha verimli üretim yapabilmek ve mevcut sistemlere alternatif yöntemler oluşturmak amaçlanmıştır. İzgara kirliliğinin, sensörlerin basınç farkı ile tespitindeki hataların üretime olan olumsuz etkilerini tespit etmek önemlidir. Bu sebeple Karkamış HES özelinde; ızgara kirliliğinin, üretim ve üretime bağlı diğer değerlerden oluşan veri seti kullanılarak farklı makine öğrenme modelleri ile tahminler yapılmıştır. Yapılan tahminlerin performansları detaylı olarak analiz edilmiştir.

Karkamış HES verileri ile gerçekleşen makine öğrenmesinde kullanılan Karar Ağaç modelleri yüksek performans göstermiş olup en başarılı modeli 0.9833 doğruluk değeri ile İnce Ağaç modelidir. Diğer ağaç modellerinde de başarılı bir öğrenme gerçekleşmiş olup doğruluk oranları genellikle 0.90 üzerinde görülmüştür. Hata oranları genellikle çok küçük değerlerde çıkmıştır. Bu durum karar ağaçlarının veri setlerindeki değişimin kirlilik oranlarındaki farklılıkları tespit ettiğini ve ayrılmış geniş test verileri için gerçekçi tahminler yaptığını göstermektedir. DVM modeli ile yapılan makine öğrenmelerinde genel olarak Karar Ağaç modellerine nazaran başarısız bir eğitim gerçekleşmiştir. DVM modelleri içerisinde en başarılı olanı İnce Gauss modelidir. En başarılı tahmin performansını 0.9525’lik R değeri ile Ünite-5 için sağlamıştır. En başarısız tahmin performansı ise Ünite-2’de görülmüştür. Sonuç olarak, makine öğrenme yöntemleriyle ızgara kirlilik tahmininin başarılı bir şekilde gerçekleştirilebileceği ve yüksek başarımın elde edilebileceği yapılan deneysel çalışmalarla doğrulanmıştır. İleriki çalışmalarda ise, derin öğrenme modellerinin kullanılması düşünülmektedir.

KAYNAKLAR

- Acar, G. (2022). Zaman serisi yöntemleri ve xgboost algoritması ile kömür satış tahmini: türkiye tarım kredi kooperatifleri uygulaması.Yüksek Lisans Tezi. Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilişim Sistemleri Mühendisliği Ana Bilim Dalı, Sakarya 72s.
- Alatas, M.,Susilowati E.,Theresia M., Budiastuti S., Gunawan T., Setyono P. (2022). Horizontal Trash Rack Diverter Trash (HTDT) to Minimize Trash Clogging at the Intake of Micro-Hydro Power Plant. *International Journal of Sustainable Development and Planning*, Vol.17,No.6,October.1713-1720. <https://doi.org/10.18280/ijstdp.170604>
- Artuğer, F & Özkaynak, F (2022). Görüntü Sıkıştırma Algoritmalarının Performans Analizi İçin Değerlendirme Rehberi.*Munzur Üniversitesi Dergisi*, 8(1).102-110. <https://doi.org/10.29132/ijpas.1012013>
- Alpaydin, E. (2014). Introduction to machine learning . The MIT press. (4th ed.)
- Arı, A & Önder, H. (2013). Farklı Veri Yapılarında Kullanılabilecek Regresyon Yöntemleri. *Anadolu Tarım Bilim Dergisi*, 28 (3), 168-174. <https://doi.org/10.7161/anajas.2013.28.3.168>
- Bragalli C. , Micocci D. , Naldi G. (2023).On the influence of net head and efficiency fluctuations over the performance of existing run-of-river hydropower plants, *Volume 206*, April 2023. 1170-1179.
- Başeşme, Hidayet (2003). Hidroelektrik santraller ve hidroelektrik santral esisleri, EÜAŞ Genel Müdürlüğü Hidrolik Santraller Dairesi Başkanlığı.
- Bekçioğulları, M. F., Dikici, B., Açıkgöz, H., Keçecioğlu, Ö. F. (2021). Güneş Enerjisinin Kısa-Dönem Tahmininde Farklı Makine Öğrenme Yöntemlerinin Karşılaştırılması. *EMO Bilimsel Dergi*, 11(22), 37-45.
- Coşkun, O & Doğanay,H (2020). Enerji kaynakları. İstanbul: Pegem Akademi.
- Daş, M., Balpetek, N., Kavak Akpınar, E., & Akpınar, S. (2019). Investigation of wind energy potential of different provinces found in Turkey and establishment of predictive model using support vector machine regression with the obtained results. *Journal of the Faculty of Engineering and Architecture of Gazi University*. 34(4). 2203-2212. <https://doi.org/10.17341/gazimmfd.432590>
- Daş, B & Türkoğlu İ. (2014 Kasım). DNA dizilimlerinin sınıflandırılmasında karar ağacı algoritmalarının karşılaştırılması, Eleco 2014 Elektrik- Elektronik Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu.
- Diniz, A.L, Esteves P. P. and Sagastizabal C. A. (2007). "A Mathematical Model for the Efficiency Curves of Hydroelectric units," 2007 IEEE Power Engineering Society General Meeting, Tampa, FL, USA. 1-7.
- Emeç, Ş., Adar, T., Akkaya, G. & Delice, E. K. (2019). Efficiency Assessment of Hydroelectric Power Plant in Turkey by Data Envelopment Analysis (DEA). *Avrupa Bilim ve Teknoloji Dergisi*, (Özel Sayı), 34-45. <https://doi.org/10.31590/ejosat.594716>
- EÜAŞ (2024). Karkamış HES. (Erişim Tarihi:18.12.2023) <https://www.euas.gov.tr/en-US/power-plants/karkamis-hepp>
- Gacar, Kocarik, B. (2023). Fonksiyonel veri analizinin karar ağaçlarında kullanımı. Doktora Tezi. Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı. İzmir 186s.
- Hribernik, A. (2020).Evaluation of Clogged Hydropower Plant Trash Rack Losses. *Journal of Mechanical Engineering*. 66(2020)2,142-152 <http://doi.org/10.5545/sv-jme.2019.6313>
- Hammid, A. T, Sulaiman, M. H.,Abdalla, A. N., (2017). Prediction of Small Hydropower Plant Power Production in Himreen Lake dam (HLD) using artificial neural network, *Alexandria Engineering Journal*, 57 (1). <https://doi.org/10.1016/j.aej.2016.12.011>
- İnalı, K., Işık, E. & Dağtekin, İ. (2014). Karakaya HES’de verim ve üretim parametrelerinin yapay sinir ağı ile tahmini. *Dicle Üniversitesi Mühendislik Dergisi*. 5(1),59-68.

- Koç, E., Şenel, M. C.(2013). Dünyada ve Türkiye’de Enerji Durumu - Genel Değerlendirme. *Mühendis ve Rea*, 54(639), 32-44.
- Kingsford, C., & Salzberg, S. L. (2008). *What are decision trees?*. *Nature biotechnology*. 26(9).1011-1013
- Karakuş, C.(2001). Makine Öğrenmesi Algoritmaları <https://ckk.com.tr/ders/ML/ML%2001%20Makine%20%C3%96%C4%9Frenmesi%20Algoritmalar%C4%B1.pdf>
- Makas, Y & Karaatlı, M. (2016). Yapay Sinir Ağlarıyla Hidroelektrik Enerji Üretiminin Çok Dönemli Tahmini, *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 23(3).757-772.
- Metlek, S & Kayalp, K. (2020). Makine Öğrenmesinde Teoriden Örnek Matlab Uygulamalarına Kadar Destek Vektör Makineleri, Ankara: İksad Yayınları.
- Özbektaş, S., Şenel M.C., Sungur B. (2023). Dünya’da ve Türkiye’de Yenilenebilir Enerji Durumu ve Kurulum Maliyetleri, *Mühendis ve Makina/ Engineer and Machinery*, 64(711), 317-351.
- Pamir, N. (2017). Enerjinin iktidarı. İstanbul: Hayy Kitap Yayınları
- Ray, S. (2019, February). A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon). 35-39).
- T.C. Enerji ve Tabii Kaynaklar Bakanlığı (2024). <https://enerji.gov.tr/bilgi-merkezi-enerji-elektrik/> Erişim Tarihi 18.03.2024.
- Walczak, N& Walczak, Z& Niec J. (2020), Assesment of the Resistance Value of Trash Racks at a Small Hydropower plant operating at low temperature energies. *Energies* 2020, 13(7),1775 <https://doi.org/10.3390/en13071775>
- Yoldaş, Y.B. (2023). Fotovoltaik güç santrali ile desteklenen güç sisteminde zayıf kaynak besleme koşullarında destek vektör makinalarıyla mesafe korumanın geliştirilmesi. Yüksek Lisans Tezi. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Elektrik Mühendisliği Ana Bilim Dalı, İstanbul 105s.
- Zilyas, D. & Yılmaz, A. (2023). Makine Öğrenmesi Yöntemleri ile Eğitim Başarısının Tahmini Modeli. *Dicle University Journal of Engineering Dergisi*, 14 (3), 437-447. <https://doi.org/10.24012/dumf.1322273>