

MÜHENDİSLİK BİLİMLERİNDE GÜNCEL ARAŞTIRMALAR

Araştırma, Metodoloji ve Yenilik

Editörler

Duygu KAVAK

Ali KILIÇER



LIVRE DE LYON

2024

Mühendislik Bilimleri

MÜHENDİSLİK BİLİMLERİNDE GÜNCEL ARAŞTIRMALAR

Araştırma, Metodoloji ve Yenilik

Editörler

Duygu KAVAK & Ali KILIÇER



LIVRE DE LYON

Lyon 2024

MÜHENDİSLİK BİLİMLERİNDE GÜNCEL ARAŞTIRMALAR

Araştırma, Metodoloji ve Yenilik

Editörler

Duygu KAVAK & Ali KILIÇER



LIVRE DE LYON

Lyon 2024

**Mühendislik Bilimlerinde Güncel Araştırmalar: Araştırma,
Metodoloji ve Yenilik**

Editors • Prof. Dr. Duygu KAVAK • Orcid: 0000-0002-1189-3110
Assoc. Prof. Dr. Ali KILIÇER • Orcid: 0000-0002-1745-854X

Cover Design • Motion Graphics
Book Layout • Motion Graphics
First Published • December 2024, Lyon

e-ISBN: 978-2-38236-767-4
DOI: 10.5281/zenodo.14537399

copyright © 2024 by Livre de Lyon

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission from the Publisher. The author or authors of the relevant section are responsible for any copyright infringement that may occur due to the images and graphics used in the book. The editor or publisher does not assume responsibility in this regard.



Publisher • Livre de Lyon
Address • 37 rue marietton, 69009, Lyon France
website • <http://www.livredelyon.com>
e-mail • livredelyon@gmail.com



ÖNSÖZ

Değerli okurlar;

Livre de Lyon Yayınevi tarafından hazırlanan “MÜHENDİSLİK BİLİMLERİNDE GÜNCEL ARAŞTIRMALAR: Araştırma, Metodoloji ve Yenilik” isimli bu kitap farklı alanlarda yapılan mühendislik, bilim ve teknoloji konularından oluşan araştırma ve incelemeleri bir araya getirmektedir.

Kitaba çalışmalarını sunarak destek veren tüm bölüm yazarlarına, eser değerlendirmelerini gerçekleştiren değerli hakemlere ve bu çalışmalarını bir kitaba dönüştüren Livre de Lyon Yayınevi çalışanlarına teşekkürlerimizi sunarız. Bu kitabın başta akademisyenler, araştırmacılar ve öğrenciler olmak üzere tüm okuyucular için faydalı olmasını ve çalışmalarına katkı sağlamasını dileriz.

Editörler

Prof. Dr. Duygu Kavak

Doç. Dr. Ali Kılıçer

İÇİNDEKİLER

	ÖNSÖZ	I
BÖLÜM I.	BİYOKÜTLE PİROLİZİ İLE BİYO-YAĞ ÜRETİMİ: DEĞERLENDİRME VE ÖNERİLER	1
	<i>Ahmet Rasim GİRİŞEN & Gürkan KARA & Hakan ÖZCAN</i>	
BÖLÜM II.	BİYOKÜTLE PİROLİZİNDE KULLANILAN REAKTÖR TÜRLERİ	29
	<i>Gürkan KARA & Ahmet Rasim GİRİŞEN & Hakan ÖZCAN</i>	
BÖLÜM III.	BİYOGAZ SANTRALLERİ ÜZERİNE BİR ARAŞTIRMA	49
	<i>Emine Merve ERDAL & Âdem YILMAZ</i>	
BÖLÜM IV.	SİĞ JEOTERMAL ENERJİ: TEMEL KAVRAMLAR, TERMAL ETKİLEŞİM VE TERMAL BOZUNUM	67
	<i>Ali Samet ÖNGEN</i>	
BÖLÜM V.	HİBRİT CNN-GRU MODELİ İLE GÜÇ TÜKETİMİ TAHMİNİ: TETOUAN ŞEHİRİ	83
	<i>Ardan Hüseyin EŞLİK & Emre AKARSLAN & Fatih Onur HOCAOĞLU</i>	
BÖLÜM VI.	SİMEKTİT GRUBU KİL MİNERALLERİNİN PETROL JEOLJİSİNDEKİ ÖNEMİ	93
	<i>Berna Yavuz PEHLİVANLI</i>	
BÖLÜM VII.	KİL MİNERALLERİNİN OLUŞUMU, SINIFLANDIRILMASI VE ÖZELLİKLERİ	117
	<i>İsmail DEMİR</i>	
BÖLÜM VIII.	DOĞADA YER ALAN REAKTİF SİLİS İÇERİKLİ AGREGALARIN BETON İÇERİSİNDE KULLANILABİLİRLİĞİ	147
	<i>Demet DEMİR ŞAHİN & Hasan EKER</i>	
BÖLÜM IX.	KARAYOLLARINDA YATAY İŞARETLEME: BOYA TÜRLERİNİN UYGULAMALARI VE PERFORMANS KRİTERLERİ	181
	<i>Murat KARACASU & Kadir Berkhan AKALIN</i>	
BÖLÜM X.	KCL ÇAMURUNDA KAYIP ÖNLEYİCİ MALZEME OLARAK YERFİSTİĞİ KABUĞUNUN KULLANIMININ ARAŞTIRILMASI	199
	<i>Ebrahim Thabit Mohammed ALADRAN & İnci TÜRK TOĞRUL</i>	
BÖLÜM XI.	ÇAPRAZ AKIŞLI NANOFİLTASYON YÖNTEMİYLE PLASTİK ENDÜSTRİSİ ATIK SUYUNDAN ASKIDA KATI MADDE GİDERİMİ	211
	<i>Büşra Elif TAŞ & Duygu KAVAK</i>	

- BÖLÜM XII.** TOPRAKLARDA MİKROPLASTİK KİRLİLİĞİ,
TAŞINIMI VE MODELLENMESİ 219
Olca GÜLÇİÇEK UYSAL
- BÖLÜM XII.** SÜTÇÜLER İLÇE MERKEZİ (ISPARTA) SU
KAYNAKLARININ SU KALİTE İNDEKS (WQI)
YÖNTEMİ İLE DEĞERLENDİRMESİ 233
*Simge VAROL & Nedime GÜRELİ & Füsün KILÇIK &
Mehmet ULUSOY*
- BÖLÜM XIV.** DRONE VE YAPAY ZEKA DESTEKLİ GÖRÜNTÜ
İŞLEME TEKNİKLERİ KULLANILAN BİTKİ TANIMA
SİSTEMLERİ 265
Burak Efe KILIÇ & Etem Berat GÜNEŞ & Kamil ORMAN
- BÖLÜM XV.** ELEKTRİK ARIZALARININ KESTİRİMİNİN ARTTIRILMASI:
OTOMATİK BİR MAKİNE ÖĞRENMESİ YAKLAŞIMI 279
*Mohammed Wadi & Wisam Elmasry &
Furkan Ahmet TAMYİĞİT & İsmail KÜÇÜK*
- BÖLÜM XVI.** Wi-Fi TEKNOLOJİSİNDE SON GELİŞMELERE
GENEL BAKIŞ: Wi-Fi 6/6E ve Wi-Fi 7 311
Hakkı SOY
- BÖLÜM XVII.** YAPAY ZEKA DESTEKLİ DİL ÖĞRENME UYGULAMASI
LANG-AI 341
Emirhan ALICI & Kamil ORMAN
- BÖLÜM XVIII.** NESNELERİN YAPAY ZEKASI TABANLI AKILLI SOKAK
AYDINLATMA UYGULAMASI 355
Semih AŞÇI & Emre KOÇAK & Hakkı SOY
- BÖLÜM XIX.** PLC PROGRAMLAMA 365
Yunus AKALTUN
- BÖLÜM XX.** AYAKTA ÇALIŞMA MASALARININ OFİS
ÇALIŞANLARI ÜZERİNDEKİ ETKİLERİ 375
*Gülberg ERGİN & Ash Ece ACAR FİLİZCİ &
Ergün ERASLAN & Dilek ÖZTAŞ*
- BÖLÜM XXI.** İŞ EKİPMANLARININ GÜVENLİ KULLANIMI 397
*Hamza TOPGÜL & Gülberg ERGİN & Ash Ece ACAR FİLİZCİ &
Cihat ÖZTÜRK & Dilek ÖZTAŞ & Ergün ERASLAN*
- BÖLÜM XXII.** KİTAB-I BAHRİYE VE KIBRIS ADASI 409
Mustafa YILMAZ
- BÖLÜM XXIII.** HFORMASYONU BİTÜMLÜ ŞEYLLERİN İÇERİKLERİNİN
JEOİSTATİKSEL YORUMU (BEYPAZARI, ANKARA,
TÜRKİYE) 427
Berna Yavuz PEHLİVANLI

BÖLÜM XV

ELEKTRİK ARIZALARININ KESTİRİMİNİN ARTTIRILMASI: OTOMATİK BİR MAKİNE ÖĞRENMESİ YAKLAŞIMI

*Enhancing Prediction of Electrical Faults:
An Automated Machine Learning Approach*

**Mohammed Wadi¹ & Wisam Elmasry² & Furkan Ahmet
TAMYİĞİT³ & İsmail KÜÇÜK⁴**

¹(Doç. Dr.) İstanbul Sabahattin Zaim
Üniversitesi, Mühendislik ve Doğa Bilimleri
Fakültesi, Elektrik – Elektronik Mühendisliği
Bölümü, 34303, Halkalı, İstanbul, Türkiye.
E-mail: mohammed.wadi@izu.edu.tr,
ORCID: 0000-0001-8928-3729

²(Dr.) İstanbul Kültür Üniversitesi,
Mühendislik Fakültesi, Bilgisayar
Mühendisliği Bölümü, 34158, Bakırköy, İstanbul, Türkiye.
E-mail: w.elmasry@iku.edu.tr,
ORCID: 0000-0002-0234-4099

³(M.Sc. & Ph.D. Öğrencisi) İstanbul Sabahattin
Zaim Üniversitesi, Mühendislik ve Doğa
Bilimleri Fakültesi, Elektrik – Elektronik
Mühendisliği Bölümü, 34303, Halkalı, İstanbul, Türkiye.
E-mail: furkan.tamyigit@izu.edu.tr,
ORCID: 0000-0001-9873-0877

⁴(Prof. Dr.) İstanbul Sabahattin Zaim
Üniversitesi, Mühendislik ve Doğa Bilimleri
Fakültesi, Elektrik – Elektronik Mühendisliği
Bölümü, 34303, Halkalı, İstanbul, Türkiye.
E-mail: ismail.kucuk@izu.edu.tr,
ORCID: 0000-0003-3071-0612

1. Giriş

Son yıllarda, elektrik tüketimine yönelik talepteki artış eğrisi, elektrik enerjisi şebekelerinin katlanarak genişlemesine yönelik dikkate değer bir odak kayması olmuştur. Yukarıda belirtilen yukarı yönlü eğri, modern enerji altyapılarının her yönü için önemli riskler ve ortaya çıkan yeni karmaşıklıklar sunar (Raza ve ark., 2020). Dağıtım şebekesi, (Prasad ve ark., 2018b) tarafından vurgulandığı gibi bu tür tehlikelere ve engellere karşı en hassas kısımdır. Elektrik güç sistemleri bağlamında dağıtım şebekeleri, insan vücudundaki sinir sistemi ile karşılaştırılabilir temel bir bileşen olarak hizmet etmektedir. İnsan vücudunda elektrik sinyalleri, sistem içindeki çeşitli uç noktalara taşınmasından ve iletilmesinden sorumludur. Sonuç olarak, herhangi bir kapsamlı şebeke planlama sürecinin kalıcı hedefi, dağıtım sisteminin sürekliliğini ve güvenilirliğini korumaktır. Bu, yalnızca elektrik gücünün kalitesini artırmakla kalmaz, aynı zamanda kullanıcıların memnuniyetini de artırır (Wadi ve ark., 2021). Böylece, elektrik güç kalitesi önemli ölçüde yükselmektedir. Bununla birlikte, müşterilerin memnuniyeti de yukarı doğru tırmanmaktadır (Wadi ve ark., 2021).

Elektrik dağıtım şebekelerinde istikrarsız bir duruma neden olan başlıca faktörlerden biri, yaygın olarak bilinen bir olgu olan iletim veya dağıtım hatlarında aniden ortaya çıkan elektrik arızalarıdır (Elmasry ve Wadi, 2022). (Chen ve ark., 2016)'e göre bir elektrik arızası, akım ve gerilim seviyelerinde sıra dışı bir sapma olarak bahsedilir. Özellikle elektrik arızası, standart çalışma koşulları sırasında tipik olarak beklenen büyüklüğün üzerinde bir yükseklik ile karakterize edilmektedir. Mevcut gerilim ve akım değerlerinin nominal değerlerden sapmalarına; insan hataları, yıkıcı çevre koşulları veya ekipman arızaları gibi birden fazla faktör sebep olabilir (Prasad ve ark., 2018). Ayrıca, bir elektrik arızası durumunda, çeşitli aparat ve ekipmanlara zarar verebilecek aşırı akımlar şebeke boyunca akmak durumunda kalmaktadır. Bu tip durumlarda, güç sisteminin genel güvenilirliği tehlikeye girebilir. Bunun sonucunda, bakım giderlerinde artış gözlemlenebilir (Elmasry ve Wadi, 2022). Elektrik dağıtım

şebekelerindeki arızaların erken ve kesin tespiti; ekipman hasarını, hizmet aksamasını ve insan ve hayvan yaşamının kaybını azaltmak için gerekli hale gelmektedir (Elmasry ve Wadi, 2022).

Dağıtım şebekelerindeki elektriksel arıza tespit süreci, yaygın olarak, arızaların tanımlanması ve yerleştirilmesi için önemli bir sistem olarak kabul edilmektedir (Elmasry ve ark., 2022). İstenen amaca ulaşmak için arıza tespit sistemi, özellikle bir elektrik sinyalinin normal ve anormal olanı arasında ayırım yapmaya çalışmaktadır. Sonuç olarak, algılama konusu, elde edilen sinyallerin "normal" olanları gösteren negatif örnekler veya "hatalı" olanları temsil eden pozitif örnekler olarak kategorize edilebildiği bir ikili sınıflandırma görevi olarak çerçevelenebilir (Wadi ve ark., 2023). İkili sınıflandırma görevlerinin otomasyonu için makine öğrenimi tekniklerinin kullanılması giderek daha da yaygın hale gelmektedir. Bu eğilim, belirli bir veri setinden öğrenmek ve ardından iki sınıfın etkili bir şekilde ayrılmasını sağlamak için makine öğrenimi algoritmaları yöntemlerine atfedilebilir. Metin girmek için buraya tıklayın veya dokununuz. (Wadi ve Elmasry, 2022; Wadi ve Elmasry, 2023). Son yıllarda, elektrik arızalarının saptanması için makine öğrenimi tabanlı yaklaşımların geliştirilmesine yönelik önemli ölçüde araştırmalar yapılmıştır (Dong ve ark., 2019; Eboule ve ark., 2018) (Gururajapathy ve ark., 2018). Bununla birlikte, bu artan ilgiye rağmen mevcut literatür içerisinde bu alanın işleme (Huang ve ark., 2024; Jamil ve ark., 2015) (Koley ve ark., 2015; Mbamaluikem ve ark., 2018), doğrulama (Qu ve ark., 2020; Uzubi ve ark., 2017) ve otomasyon (Wadi, 2021; Wadi ve Elmasry, 2021) gibi çeşitli yönlerine ilişkin anlayışımızda önemli boşlukların kaldığını göstermektedir (Chen ve ark., 2016; Prasad ve ark., 2018b, 2018a; Raza ve ark., 2020).

Mevcut bağlamda işleme, eğitim sınıflandırma yöntemleri etkin bir şekilde kullanılmadan önce ham verilerin etkin bir şekilde ön işlenmesi anlamına gelmektedir (Wadi ve Elmasry, 2021). Alternatif olarak doğrulama, verilerin ve yöntemin gerçek dünya uygulama alanlarında elektrik güç şebekelerinde yaklaşık kullanma derecesi ile ilgilidir (Elmasry ve Wadi, 2022). Ek olarak, çalıştırma parametrelerini seçme ve eylemleri gerçekleştirme süreci, çalışma (Elmasry ve ark., 2022)'de belirtildiği gibi otomasyon tarafından belirlenmektedir. Elektrik dağıtım şebekelerindeki arıza olaylarının kapsamlı bir analizinin ardından, önceki çalışmaların sınırlamalarının dikkatli bir şekilde incelenmesiyle birleştiğinde, güç şebekelerindeki arızaları tespit etmek için gelişmiş bir sistem önerilmiştir. Önerilen sistem, gerçek güç sistemlerinde optimum performans gösterecek şekilde tasarlanmıştır. Somut olarak, önceden var olan bir veri işleme aşaması, işlenmemiş verileri etkili bir şekilde incelemek

ve iyileştirmek için hata tespit modellerini konuşlandırmadan önce dahil edilmektedir. Ayrıca, doğrulama görevini etkili bir şekilde ele almak için yaygın olarak kullanılan makine öğrenimi sınıflandırıcılarıyla birlikte gerçek zamanlı veriler uygulanır. Ayrıca, optimum özellikleri ve hiper parametreleri seçme sürecini kolaylaştırmak ve otomatikleştirmek için çok sayıda optimizasyon tekniği kullanılır. Önerilen sistemin yukarıda belirtilen faydaları, bu çalışmaya önemli katkıları temsil etmektedir.

Çalışmamızın sonraki bölümleri, sıradaki cümlelerdeki gibi sistematize edilmiştir. Bölüm 2, VSB veri setinin özniteliklerini ve veri setini dengelemek için kullanılan teknikleri açıklamaktadır. Bölüm 3, önerilen makine öğrenimi yöntemine ilişkin ayrıntılı bir açıklama sağlar. Bölüm 4, sekiz değerlendirme ölçüsünü ve bunlara karşılık gelen matematiksel formülleri tanıtmaktadır. Bölüm 5, çeşitli analitik teknikler kullanılarak deneysel sonuçların kapsamlı bir analizini sunmaktadır. Sonuç bölümü, potansiyel ilerleme alanlarına ilişkin değerli içgörüler sağlarken, araştırma bulgularına kapsamlı bir genel bakış sağlamaktadır. Ayrıca, gelecekteki araştırmalar için olası yollar, Bölüm 6'da tartışılmaktadır.

2. VSB Veri Seti

Ostrava Teknik Üniversitesi (VSB) veri seti, 2018'de Kaggle Yarışması web sitesinde yayınlanan modern bir veri setidir (*VSB Power Line Fault Detection*, 2018). Araştırma için değerli olan bu veri seti, VSB tarafından oluşturulmuştur. Ayrıca, incelenmekte olan arıza tespit veri seti, VSB'deki ENET Merkezinden alınmıştır ve bu veri seti, elektrik hatları üzerinden iletilen elektrik sinyallerini yakalayan yeni bir cihaz tarafından üretildiği için pratiklik göstermektedir (ENET Centre, t.y.).

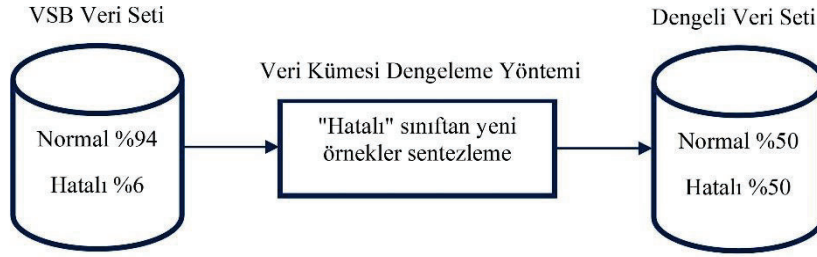
VSB veri setinin mimarisi ile ilgili olarak, her bir örnek, yalnızca tamsayı değerlerinde saklanan 800.000 gerilim ölçümünü kapsayan tek bir elektrik sinyalidir ve toplamda bu veri setinde 8.712 adet örnek vardır. Burada açıklanan sinyaller, 50 Hz frekanslı gerçek bir üç fazlı elektrik güç şebekesinden elde edilmiştir. Bahsedilen sinyallerin tamamı, 20 milisaniyelik bir süreyi kapsayan tek bir şebeke döngüsü üzerinden alınmıştır. Ek olarak, VSB veri setinde, "Sınıf" özelliği olarak adlandırılan belirgin bir özelliğin, burada yer alan çeşitli sinyal türlerini tanımladığı ve bunlar arasında ayırım yaptığı belirtilmelidir. Daha spesifik olarak, normal aralıkta olduğu kabul edilen sinyaller "0" olarak belirtilmektedir. Aynı zamanda, arıza veya hata belirtisi gösterenler "1" olarak etiketlenir. VSB veri setindeki örneklerin çoğu normal sinyallerle ilgilidir ve toplam 8.187 örnek normal sinyalleri barındırırken, geri kalan 525 örnek hatalı

sinyallerden oluşmaktadır. Sınıflandırıcılar, baskın sınıf olma eğiliminde olabileceğinden, VSB veri setindeki normal ve hatalı örneklerin miktarındaki önemli tutarsızlık, yetersiz sınıflandırma sonuçlarına neden olabilir. Bu nedenle, herhangi bir veri seti ön işlemeden önce, veri setinin doğal dengesizliğini ele almak zorunludur.

Genellikle aşırı örnekleme olarak adlandırılan dengesiz bir veri setini dengeleme süreci kuşkusuz olarak, sadece yeterince temsil edilmeyen sınıfa ait örneklerin sayısında bir artışa yol açmaz, aynı zamanda modellerin tahmin yeteneklerinde dikkate değer bir artışa da yol açmaktadır (He ve Mu, 2013). Dengesiz bir veri setinde denge sağlamaya yönelik geleneksel yaklaşım, yetersiz temsil edilen sınıfın örneklerini çoğaltmayı gerektirir. Bununla birlikte, bu yaklaşımı uygulamak potansiyel olarak dengeleme sorununu çözebilirken, örneklerin çoğaltılması modele yeni içgörüler kazandırmaz ve bu nedenle performans geliştirmesini engeller (Weiss, 2013). Sınıf dengesizliği sorununu ele almak için alternatif bir yaklaşım, yeni sentetik örnekler üretmek azınlık sınıfı içindeki örneklerin yüksek oranda örneklemesini içerir. Sentetik Azınlık Yüksek Örnekleme Tekniği (SAYÖT), yeni sentetik örneklerin oluşturulduğu üstün performanslı bir teknik olarak bilinmektedir (Chawla ve ark., 2002).

Önceki araştırmalarda alıntılanan kaynaklara göre, SAYÖT pratik bir yaklaşımdır (Chawla ve ark., 2002). Bunun başlıca nedeni, yeni üretilen örneklerin yüksek derecede makullük sergilemesidir. Spesifik olarak bu örnekler, öznelik alanı içindeki orijinal azınlık sınıfı örneklerine yakın bir yerde bulunur. Aksine, SAYÖT özellikle sınıfların arasında önemli çakışmaların olduğu durumlarda, çoğunluk sınıfı için yeterli değerlendirmeyi yapmaksızın yetersiz temsil edilen sınıftan yeni örnekler üretmesi nedeniyle bir dezavantajla karakterize edilir (Chawla ve ark., 2002). Bu uzantıların temelini oluşturan temel kavram, örneklerin rastgele seçilmesi yerine azınlık sınıfından hassas örneklerin seçilmesini içerir. Mevcut hassas örnekler, yalnızca belirli bir sınıflandırıcının hatalı bir şekilde sınıflandırdığı, azınlık sınıfına ait örneklerin bir alt kümesine atıfta bulunur. (Han ve ark., 2005)'ün bulgularına göre, azınlık sınıfıyla ilgili yanlış sınıflandırılan örneklerin belirlenmesi için k-en yakın komşu (K-EYK) sınıflandırıcı kullanılması önerilmiştir. Bununla birlikte, (Nguyen ve ark., 2011) içinde vurgulanan alternatif bir öneri, bu görevi gerçekleştirmek için sınıflandırıcı olarak bir Destek Vektör Makinesi (DVM) modeli uygulamaktır. Sınır bölgesindeki azınlık sınıfı içindeki hassas örneklerin yüksek örneklemesinin önemi, doğru sınıflandırma için bu örneklerin yeniden değerlendirilmesiyle modellerin etkinliğinin artırılmasında yatmaktadır (Han ve ark., 2005; Nguyen ve ark., 2011).

Uyarlanabilir Sentetik Örneklem (USÖ) yöntemi, azınlık sınıfı içinde aynı sınıfa ait bölgelerin yoğunluğuyla negatif ilişkili olan sentetik örnekler üretmektedir. Spesifik olarak, o sınıfın düşük yoğunluğu ile karakterize edilen özellik uzaylarında sentezlenen azınlık örneklerinin sayısında bir artış meydana gelecektir. Buna karşılık, yoğunluğun yüksek olduğu bölgelerde bunun tersi olacaktır. Bu yaklaşım, literatürde sunulmaktadır (He ve ark., 2008). Ayrıca, Yerelliğe Duyarlı Karma SAYÖT (YDK-SAYÖT) yaklaşımı, geleneksel SAYÖT'ün değiştirilmiş bir yinelenmesini temsil etmektedir (Hassib ve ark., 2018). Geleneksel SAYÖT ve YDK-SAYÖT yöntemleri arasındaki birincil zıtlık, önceki (Hassib ve ark., 2018)'da kullanılan k-en yakın komşu yaklaşımının aksine, ikincisinin seçilen örneğin en yakın komşularını belirlemek için bölgeye duyarlı karma tekniğini kullanmasıdır. Veri seti dengeleme prosedürünü tamamladıktan sonra, VSB veri seti, her bir ilgili sınıfta mevcut 8.187 numune ile 1:1 oranına ulaşarak, "normal" ve "hatalı" numunelerin dengeli bir dağılımını sergileyecektir. Şekil 1'de gösterilen mevcut çizim, VSB veri seti ile ilgili veri seti dengeleme kavramını göstermektedir.



Şekil 1. Veri seti dengeleme kavramı.

3. Önerilen Yaklaşım

Önerilen makine öğrenimi yöntemi, açık ve anlaşılır olacak şekilde tasarlanmıştır. Süreç dört ardışık aşamadan oluşur: veri ön işleme, ön eğitim, eğitim ve test aşamalarıdır. Bu aşamalar sonraki alt bölümlerde açıklanacaktır. Önerilen yaklaşımın diyagramı Şekil 2'de gösterilmektedir.

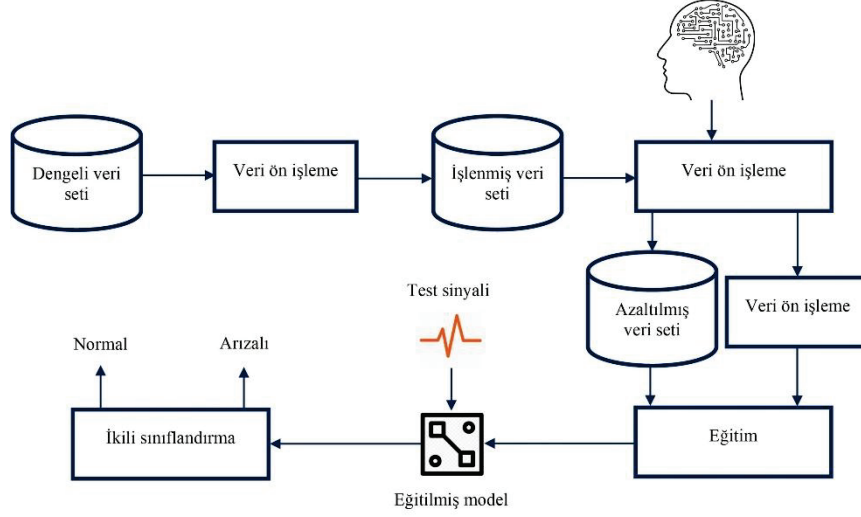
3.1. Veri Ön İşleme Aşaması

Adından da anlaşılacağı gibi, başlangıç aşaması öngörülen sistem için giriş noktası görevi görür. Veri ön işleme aşaması, verileri çeşitli makine öğrenimi modelleriyle kullanıma uygun hale getirmek için ilk elektrik sinyali verilerini,

özellikle gerilim sinyallerini, iyileştirmeyi amaçlar. Mevcut aşama, yakalanan analog sinyali bir dizi sayısal özelliğe dönüştürme yeteneği göz önüne alındığında, genel sistem içinde önemli bir öneme sahiptir. Bu özellikler ise dağıtım şebekelerinde meydana gelebilecek arızaların tanımlayıcı özelliklerini kavramalarını sağlayarak sınıflandırıcıların işini kolaylaştırır. Veri ön işleme aşaması, sorumluluğunu etkin bir şekilde yerine getirmek için dengeli veri seti üzerinde, aşağıda ana hatlarıyla belirtildiği gibi, sinyal filtreleme, sinyal ayrıştırma, özellik çıkarma ve veri normalleştirmeyi içeren dört sıralı işlemi yürütülmesini içerir.

3.1.1. Sinyal Filtreleme

VSB veri setinde denge sağlandıktan sonra, gerilim sinyalleri kaydedilirken harici gürültü ölçümlerinin yakalanması mümkün olduğundan, her örnekteki gerilim ölçümleri için bir temizleme işlemi yapılması gerekli hale gelir. Ölçüm mekanizmalarının arızaları, öngörülemeyen koşulların ortaya çıkması ve tesadüfi çevresel faktörler dahil ancak bunlarla sınırlı olmamak üzere, elde edilen verilerde gürültünün varlığına birden fazla faktör katkıda bulunur. Makine öğrenimi sınıflandırıcılarının eğitiminde veya testinde gürültü karışmış verilerin kullanılması, sonuçta hata algılama alanında yetersiz sonuçlara yol açacaktır. Mevcut fenomen, incelenen verilerde kullanılan tanımlayıcı istatistiklerin güvenlik açıklarıyla, özellikle de istatistiksel analizlerde, önemli gözlemlerin potansiyel olarak ihmâl edilmesine veya gerçek sonuçların saptırılmasına yol açan, konu dışı ince ayrımlara veya gürültüye karşı hassasiyetleri ile açıklanabilir. Sonuç olarak, dengelenmiş veri seti üzerinde bir sinyal filtreleme işlemi gerçekleştirmek, herhangi bir dış gürültü değerini ortadan kaldırmak için zorunludur (Naveed ve ark., 2021).



Şekil 2. Elektriksel arıza tespit sistemine yönelik önerilen genel makine öğrenimi yaklaşımı blok diyagramı.

Veri setlerindeki gürültü olgusunun daha iyi anlaşılması, daha etkili veri temizleme prosedürlerini kolaylaştıracaktır. Gürültü, genellikle bir veri setindeki aykırı değerlerin tezahürü olarak adlandırılır. Operasyonel olarak, veri popülasyonunun geri kalanından önemli ölçüde sapan atipik gözlemler olarak tanımlanabilir. Genel olarak gürültünün tanımlanmasının daha kesin olması gerekir. Yukarıda belirtilen gürültü kavramının kullanılmasıyla, potansiyel gürültü kaynaklarını belirlemek için çeşitli istatistiksel teknikler kullanılabilir. Çeyrekler Açıklığı (ÇA) tekniğini kullanan bu çalışma, gürültünün etkisini azaltmak için bir filtreleme stratejisi uygulamayı amaçlamaktadır. Araştırma tasarımı ve yöntemi de yazarlar tarafından kapsamlı bir şekilde gözden geçirilmiş ve geçerlilik ve güvenilirliği sağlamak için değerlendirilmiştir. Nihayetinde bu konu, ilgili alandaki önemine ve bilgiyi ilerletme potansiyeline göre seçilmiştir. İncelenmekte olan teknik, Çalışma (García-Gil ve ark., 2019)'de özetlendiği gibi, dikkate değer basitlik, genellik ve sağlamlık niteliklerini sergilemektedir.

ÇA tekniği, beş sıralı adım kullanarak dengeli bir veri kümesindeki her bir örneğe etkili bir şekilde bir filtreleme işlemi uygular. Başlangıçta süreç, daha önce bahsedildiği gibi, gerilim ölçümlerini artan bir sırada düzenlemektedir. Ardından gerilim ölçümlerinin birinci çeyrek (Ç1) ve üçüncü çeyrek (Ç3) değerleri arasındaki sayısal fark belirlenerek çeyrekler arasındaki açıklığın (ÇA) hesaplanması gerçekleştirilir. Bu durum, Eşitlik (1)'de gösterilmektedir.

Prosedürün üçüncü aşamasında, eşik değerinin belirlenmesi, çeyrekler arası açıklığın (ÇA) 3 ile çarpılmasını içerir. Bu hesaplama, Eşitlik (2) ile gösterilmektedir. Bu işlem, Çalışma (García-Gil ve ark., 2019)'de bildirildiği gibi bir düzeltme faktörü olarak kabul edilir. Bu faktörü kullanmanın amacı, aykırı değerlerin sağlamlığını tespit etmektir. Elektrik arıza tespitinde, aykırı değerleri iki farklı kategoride sınıflandırmak mümkündür. İlk kategori, normal değerlerden küçük bir sapma sergileyen arıza ölçümleriyle ilgilidir. İkinci kategori, normdan aşırı bir sapma gösteren gürültü ölçümlerini kapsar. Sonuç olarak, arıza tespitini etkili bir şekilde yürütmek için, gürültü ölçümlerini atarken arıza ölçümlerini korumak zorunludur. Ardından bir çift sınır değeri, alt (AS) ve üst sınır ($ÜS$) değerleri olmak üzere hesaplanmaktadır. AS değeri için *eşik* değerinin Ç1 değerinden çıkarılması ve $ÜS$ değeri için *eşik* değerinin, Ç3 değerine eklenmesiyle türetilir. Eşitlik (3) ve (4)'te, alt ve üst sınır değerlerinin hesaplanması gösterilmektedir.

ÇA tekniği, sistemdeki herhangi bir gerilimi etkili bir şekilde ortadan kaldırır. Kaldırılan ölçüm değerleri, AS değerinin altında veya $ÜS$ değerinin üzerindedir. Filtreleme işlemi bir sonraki sinyale ilerler ve eşitlenen veri setindeki tüm sinyaller, filtrelenceye kadar aynı eylemi yürütmeye devam eder.

$$\text{ÇA} = \text{Ç3} - \text{Ç1} \quad (1)$$

$$eşik = 3 \times \text{ÇA} \quad (2)$$

$$AS = \text{Ç1} - eşik \quad (3)$$

$$ÜS = \text{Ç3} + eşik \quad (4)$$

3.1.2. Sinyal Ayrıştırma

Sinyal filtreleme prosedürü tamamlandıktan sonra, dengeli veri seti içindeki i . sinyalinden çıkarılan gerçek gerilim ölçümleri, $(800.000-l)$ 'nin çıkarılmış değerine eşdeğerdir. Burada l , i . sinyali içinde gürültü olarak algılanan ve daha sonra atılan gerilim ölçümlerinin sayısını temsil eder. Sınırlı hatalı ölçümlerin normal sinyal ölçümleri alanına yerleştirilmesi nedeniyle, her bir sinyalin artık gerilim ölçümlerindeki arızaların tespit edilmesi hala bir zorluk teşkil etmektedir. Yukarıda bahsedilen sorunun aşılmasında vazgeçilmez olan sinyal ayrıştırma işlemi çok önemlidir (Yue, 2019).

Sinyal ayrıştırma, sinyal işleme alanında kullanılan ve her bir sinyalin artık gerilim ölçümlerinin daha küçük birimlere bölüldüğü bir tekniktir. Bu yaklaşım, sinyalin bu ayrı parçaları (*chunk*) içindeki hataların tanımlanmasını ve

çözümlemesini kolaylaştırmayı amaçlamaktadır. Bu yöntemin uygulanmasının, her grup içindeki ölçüm aralığını azaltması ve sonuç olarak karşılık gelen normal ölçümlerin spektrumu içindeki herhangi bir arızanın daha iyi anlaşılmasını sağlaması beklenmektedir (Yasir ve Koh, 2018). Parça sayısındaki bir artış, gelişmiş bir performans düzeyi ile sonuçlanacaktır. Bu ampirik araştırma kapsamında performanstaki farklılıkları araştırmak için, bireysel sinyal verileri 1, 2, 4 ve 8 parçayı kapsayan farklı deneysel parçalara ayrılmıştır. M parça sayısını temsil etmektedir. Sinyal ayrıştırma sürecinde, dengeli veri setindeki her sinyalin kalan gerilim ölçümleri M parçalarına bölünür. Bu durum, Eşitlik (5)'te gösterilmektedir.

$$parça_boyutu = yuvarla\left(\frac{Sinyalin_kalan_ölçümleri}{M}\right) \quad (5)$$

Burada $parça_boyutu$, parçaların boyutudur ve $yuvarla()$, bir sayıyı bir tamsayı değerine yuvarlayan bir işlevdir.

$$Parça_j^i = X_{[(j-1) \times parça_boyutu] + 1}^i, X_{[(j-1) \times parça_boyutu] + 2}^i, \dots, X_{[(j-1) \times parça_boyutu] + parça_boyutu}^i, \quad j = 1, 2, \dots, M \quad (6)$$

Burada X_d^i , i sinyalin d . gerilim ölçümüdür ve $Parça_j^i$, i sinyalin j . parçasıdır.

3.1.3. Özellik Çıkarma

Sinyal filtreleme ve ayrıştırma tekniklerinin uygulanmasının ardından her bir sinyal için gerilim ölçümü sayısındaki azalmaya rağmen, sinyalin her bir parçasının hala on veya yüzbinlerce gerilim ölçümü içerdiği tahmin edilmektedir. Bu gözlem, sinyal analizi ve işleme faaliyetleri sırasında dikkate alınması gereken gerilim ölçümlerinin önemli hacminin altını çizmektedir. Bunun nedeni, dengeli veri setindeki her sinyalin orijinal olarak 800.000'i bulan önemli miktarda gerilim ölçümünden oluşmasıdır. Mevcut senaryoda, sinyalin artık gerilim ölçümleri, yüksek boyutlu bir girdi uzayında çalışan uygulamalı modellere girdi görevi görecektir. Çoğu makine öğrenimi modeli yüksek boyutluluk için pratik olmadığından ve boyutlardaki bir artış şüphesiz zayıf model performansına ve başarısızlığa neden olacağından, yüksek boyutlu bir girdi uzayının ortaya çıkışı makine öğreniminde önemli bir zorluk teşkil eder.

Bu sorunu azaltmak için, özellik uzayının boyutsallığını azaltmak amacıyla bir özellik çıkarma prosedürü kullanılır. Mevcut gerilim ölçümlerine dayalı olarak her bir sinyal parçasından on dokuz farklı özellik çıkarılır. Bu işlem, sinyalin her parçası için ayrı ayrı uygulanır. Tüm sinyal parçalarından özellik

çıkarmasının ardından, çıkarılan özellikler, ilgili sinyal için bir veri demeti veya kaydı oluşturmak üzere karşılık gelen sinyalin "Sınıf" etiketi ile birleştirilir. Özellik çıkarma aşaması, tüm sinyallerin işlenmesinin tamamlanmasının ardından sona erer ve ardından elde edilen tüm veri kayıtları, işlenmiş veri seti olarak adlandırılan yeni bir veri setine entegre edilir. Eşit derecede dikkate değer bir yön, işlenen veri setindeki özelliklerin sayısındaki, bölünme için belirlenen parçaların sayısına bağlı olarak değişikliklerle ilgilidir. Spesifik olarak, sayı $(19 \times M) + 1$ 'e eşdeğerdir, burada M bu tür parçaların sayısını gösterir.

Sinyalin bireysel parçalarından elde edilen özellikler, gerilim ölçümlerinin dağılımı ve altında yatan modeller hakkında değerli bilgiler sağlama potansiyeline sahip kapsamlı bir istatistiksel ölçüm aralığına sahiptir. Çıkarılan tüm özellikler, ortalamayı, standart sapmayı ve maksimum değeri açıkça belirten sayısal değerler içerir. Çalışma, kaydedilen en düşük ölçümü %1, %25, %50, %75 ve %99 için yüzdeleri incelemiştir. Bu çalışma, belirli bir bağlamda çeşitli yüzdelerin (%0, %1, %25, %50, %75, %99 ve %100) görece değerlerini araştırmaktadır. Yüzdeler, alt ve üst bantları ve boyu kapsayan metrikler akademik olarak ifade edilebilir. Eşitlik (7), (8), (9) ve (10), fiziksel niceliklerin ve ilişkili birimlerin orantılı oranını ifade ederek görece nicelikler arasındaki ilişkiyi göstermektedir. Hesaplanan değerler arasında yüzde dilim, alt sınır, üst sınır ve büyüklük bulunur. Buna karşılık, kalan değerler temel bilgi olarak kabul edilir.

$$\%P_{bağlı\ yüzde} = \%P_{yüzdeler} - Ortalama_j \quad (7)$$

Burada $\%P_{bağlı\ yüzde}$, sinyalin j . parçasının $\%P$ bağlı yüzdeliğidir ve $\%P_{yüzdeler}$, sinyalin j . parçasının $\%P$ yüzdeliğidir.

$$Alt_Sınır_j = Ortalama_j - Standart_Sapma_j \quad (8)$$

$$Üst_Sınır_j = Ortalama_j + Standart_Sapma_j \quad (9)$$

Burada sırasıyla $Alt_Sınır_j$ ve $Üst_Sınır_j$, sinyalin j . parçasının alt ve üst sınırlarıdır.

$$Yükseklik_j = Maksimum_j + Minimum_j \quad (10)$$

Burada j . sinyale ait $Yükseklik_j$, $Maksimum_j$ ve $Minimum_j$ parametreleri isimleri ile aynı işlevi görmektedir.

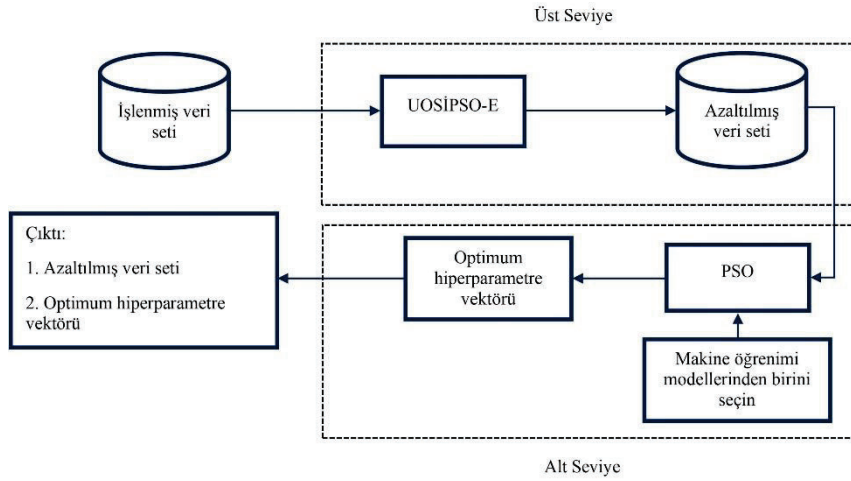
3.1.4. Veri Normalleştirme

İşlenen veri setindeki her bir veri kaydı için, özellikler ("Sınıf" özelliği hariç), aşağıdaki gibi min-maks dönüşümü kullanılarak [0,1]'e normalleştirilir. Eşitlik (11)'de bu durum gösterilmektedir.

$$x_i = \frac{x_i - Min}{Maks - Min} \quad (11)$$

Burada x_i , işlenen veri setindeki i . veri kaydına ait sayısal karşılığı özellik olarak vermektedir. Min ve $Maks$, her bir sayısal özelliğe ait en düşük ve en yüksek değerlerini sırasıyla temsil etmektedir.

3.2. Ön Eğitim Aşaması



Şekil 3. Çift PSO ile hiper parametre ve özelliklerin optimizasyonu blok şeması.

Ön eğitim aşaması, analiz için rafine etmek üzere ön işleme tabi tutulan veri setinin alınmasıyla başlar. İşlenen veri seti, kullanılan modellerin eğitimi ve test edilmesi amacıyla iki farklı nedenden dolayı yetersiz kalmaktadır. En önemli dayanaklardan biri, çıkarılan özniteliklerin kendi aralarında veya "Sınıf" etiketi ile ilgili korelasyonu ile ilgilidir. Bu çalışma, çıkarılan özniteliklerin elektriksel arıza tespit problemini ne ölçüde etkili bir şekilde temsil edebileceğini araştırmayı amaçlamaktadır. Varlıkları, doğruluk ve karmaşıklık açısından yetersiz sınıflandırma sonuçlarına yol açabileceğinden, özellik seti içinde gereksiz ve yinelenen özelliklerin varlığını dikkate almak önemlidir (Elmasry ve ark., 2020). İkincil gerekçe, ağırlıklı olarak hiper parametrelerin ilgili

değerlerinden etkilenen, kullanılan modellerin etkinliği ile ilgilidir. Bir modelin hiper parametreleri, eldeki belirli sınıflandırma görevini yerine getirmek için modelin ayarlanmasını sağlayan yapılandırılabilir parametreler olarak hizmet eder. Değerleri belirlemek, eğitim aşamasına başlamadan önce ele alınması gereken önemli bir husustur (Elmasry ve ark., 2020). Sonuç olarak, ön eğitim aşaması, eldeki belirli görevde kullanılan modellerin genel etkinliğini artırmayı amaçlayan en avantajlı özellik alt setini ve hiper parametreleri belirlemeyi amaçlar (Elmasry ve ark., 2020).

Bu çalışmada, özellikleri ve hiper parametreleri tek bir işlemde eş zamanlı olarak optimize etmek için Çift Parçacık Sürü Optimizasyonu (PSO) algoritması kullanılmıştır (Elmasry ve ark., 2020). Çift PSO algoritması iki farklı seviyeden oluşur. İlk seviye, özellik seçimi için Uygunluk Orantılı Seçim İkili Parçacık Sürü Optimizasyonu ve Entropi (UOSİPSO-E) algoritmasını kullanır. Bu çalışma, daha önceki bir çalışma (Elmasry ve ark., 2020) tarafından bildirildiği gibi, daha düşük bir düzeyde hiper parametre seçimi için PSO tabanlı (Elmasry ve ark., 2018) bir algoritma kullanır. Bu çalışma, temel optimizasyon algoritması olarak UOSİPSO-E (Zhou ve ark., 2014) ve optimal özellik alt kümesini belirlemek için tamamlayıcı bir araç olarak tek amaçlı filtre tabanlı entropi yönteminden (Cervante ve ark., 2012) oluşan UOSİPSO-E algoritmasının tanıtılmasıyla başlamaktadır. Algoritmanın birincil amacı, en etkili özellik kümesini elde etmek için fazlalığı en aza indirmek ve seçilen özellik alt kümesi ile sınıf kümesi arasındaki alaka düzeyini en üst düzeye çıkarmaktır. Tersine, PSO'ya dayalı algoritma, belirli bir eğitim setini referans olarak kullanarak belirli bir makine öğrenimi modeline uygulandığında maksimum doğrulukla sonuçlanan hiper parametreleri belirleme nihai hedefiyle birden çok yinelemeden geçer. Sonuç olarak, çift PSO algoritması iki temel çözüm üretir: kesilmiş veri seti ve optimal hiper parametre vektörü. Şekil 3, çift PSO tabanlı algoritmanın mekanizmasını göstermektedir.

3.3. Eğitim ve Test Aşamaları

Eğitim aşamasındaki ilk adım, azaltılmış veri setinin eğitim seti ve test seti olmak üzere iki otonom bileşene bölünmesini içerir. Verileri bölme işlemi katmanlara ayrılmıştır. Böylece ortaya çıkan her set, 1:1 oranını koruyarak, indirgenmiş veri setiyle eşdeğer sınıf dağılımı sergiler. Bölme prosedürü, eğitim seti için 13.100 örneği (%80) rastgele seçmek ve kalan 3.274 örneği (%20) test kümesine tahsis etmek için değiştirmeden uzatma örneklemesini kullanır. Eğitim ve test veri kümelerinin adil bileşiminin, çoğunluklu kategoriye yönelik model yanlılığı sorununu etkili bir şekilde ele alması beklenmektedir. Tablo 1'de

sunulan tablo, küçültülmüş veri setinin temel özelliklerini ve ilgili eğitim ve test alt kümelerini özetlemektedir.

Ardından, en uygun hiper parametreler kullanılarak özel makine öğrenimi modeli kurulur ve ardından tüm eğitim verisi seti ile tam eğitime tabi tutulur. Sonraki bir adım olarak test aşaması, belirlenmiş test setinin değerlendirilmesi yoluyla eğitilen modelin etkinliğinin değerlendirilmesini gerektirir. Ortaya çıkan ikili sınıflandırma sonuçları daha sonra sonraki analiz ve manipülasyon için korunur. Ampirik deneylerimizi oluştururken kullanılan yöntem, Şekil 4'te gösterilmektedir.

4. Değerlendirme Ölçütleri

Bulgular, test aşaması tamamlandıktan sonra ikili bir sınıflandırma çerçevesi aracılığıyla elde edilmiştir. İkili sınıflandırma bağlamında, dört farklı hücreden oluşan bir karışıklık matrisi, gelen bir veri setini doğru bir şekilde sınıflandırmada belirli bir modelin etkinliğini değerlendirmek için tipik bir şekilde bir araç olarak kullanılır. Bu ölçümler; Doğru Negatif (DN), Yanlış Negatif (YN), Doğru Pozitif (DP) ve Yanlış Pozitif (YP). Buradaki terimler, istenen ve istenmeyen örneklere dair olarak sınıflandırmaların doğruluklarını ve yanlışlıklarını nicelleştirerek göstermektedir. Ek olarak, yukarıda bahsedilen dört ölçüm, Doğruluk, Kesinlik, Duyarlılık, F1 puanı, Yanlış Alarm Oranı (YAO), Özgünlük, Yanlış Negatif Oranı (YNO), Matthews Korelasyonu Katsayı (MKK) gibi yaygın olarak kullanılan sekiz değerlendirme metriğini hesaplamak için anahtar faktörler olduğu kabul edilmektedir. Bu çalışma, kullanılan modellerin performansını değerlendirmek için Bölüm 5'te özetlenen ölçütleri kullanacaktır. Bu değerlendirme ölçütlerinin hesaplanmasını açıklayan eşitlik listesi (Eşitlik (12)-(19)) aşağıda sunulmaktadır.

$$\text{Doğruluk} = \frac{DP+DN}{DP+DN+YP+YN} \quad (12)$$

$$\text{Kesinlik} = \frac{DP}{DP+YP} \quad (13)$$

$$\text{Duyarlılık} = \frac{DP}{DP+YN} \quad (14)$$

$$\text{F1 puanı} = \frac{2 \times \text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (15)$$

$$\text{YAO} = \frac{YP}{YP+DN} \quad (16)$$

$$\text{Özgünlük} = \frac{DN}{DN+YP} \quad (17)$$

$$\text{YNO} = \frac{YN}{YN+DP} \quad (18)$$

$$\text{MKK} = \frac{(DP \times DN) - (YP \times YN)}{\sqrt{(DP+YN) \times (DP+YN) \times (DN+YP) \times (DN+YN)}} \quad (19)$$

Tablo 1. VSB veri setinin azaltılması ile özellik dağılımı.

Karakteristik	Veri Seti
Örnek sayısı	16.374
Sınıf sayısı	2
Özelliklerin tipi	Sayısal
Tam özellik seti	1 parça için 20 2 parça için 39 4 parça için 77 8 parça için 153
Optimum özellik alt kümesi	1 parça için 12 2 parça için 24 4 parça için 47 8 parça için 91
Eğitim seti dağılımı	Normal=6.550 Hatalı=6.550 Toplam=13.100
Test seti dağılımı	Normal=1.637 Hatalı=1.637 Toplam=3.274

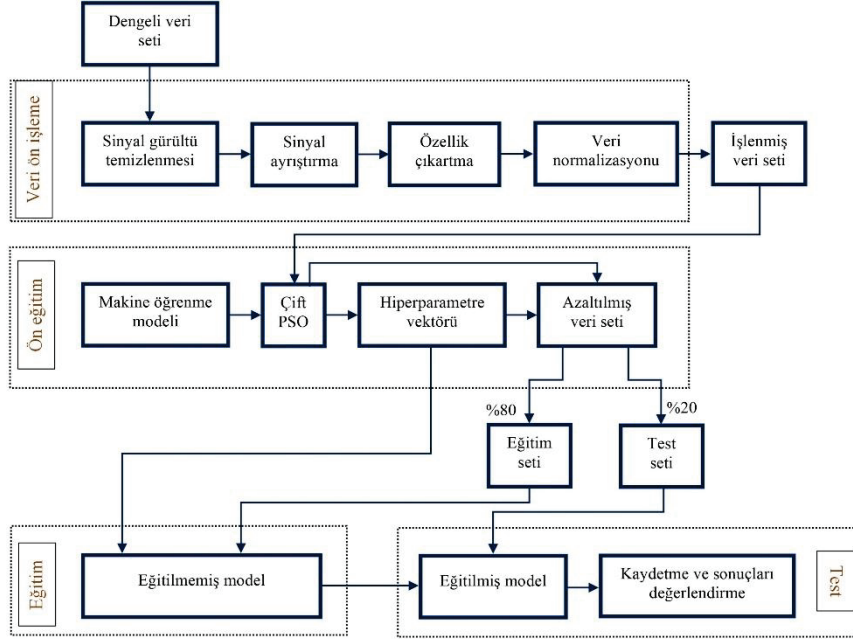
5. Sonuçlar ve Tartışma

NumPy kütüphanesi ile ilişkili Python programlama dili sürüm 3.9.5, hesaplamalı analiz için verileri gerekli niteliklerle donatmayı ve ön model eğitimini uygulamayı kolaylaştırmaktadır. Bunun aksine, eğitim ve test aşamalarını oluşturmak ve uygulamak için Azure Machine Learning (AML) Studio (*Azure Machine Learning Studio*, 2020) kullanılır. Bu süreç içerisinde AML Studio'da altı yerleşik ikili sınıflandırma modelleri olan Yapay Sinir Ağları (YSA), Destek Vektör Makinesi (DVM), Naive Bayes (NB), Arttırılmış Karar Ağacı (AKA), Karar Ormanı (KO) ve Karar Sık Ormanı (KSO) (Decision Jungle (DJ)) kullanılır. AML, kullanıcılara birçok avantajlı özelliği ve işlevi ücretsiz olarak sunan bulut tabanlı bir platformu temsil etmektedir. Bu teknoloji, (*Azure Machine Learning Studio*, 2020)'de bildirildiği gibi, kapsamlı bilgi işlem kaynaklarını kullanarak makine öğrenimi deneylerini tasarlamak ve yürütmek için iş birliğine dayalı bir platformu sunmaktadır.

5.1. Performans Analizleri

Bir modelin elektrik arızalarını ne kadar iyi tespit ettiğine karar vermek, değerlendirme ölçütlerinin değerlendirilmesine bağlıdır. Sonraki alt bölümlerde

deneysel sonuçların analitik bir açıklamasını sağlamak için altı farklı durum dikkate alınacaktır. Alt bölümler, deneysel bulguların altı senaryoda incelenip elde edilmesi ile açıklamaktadır.



Şekil 4. Deneylerimizin yöntemi.

5.1.1. Senaryo 1: Veri Setini Dengelemenin Etkisi

Bu senaryo ile veri seti dengeleme aşaması ön plana çıkarılmak üzere yazarlar tarafından kurgulanmıştır. En yakın komşu sayısının (k) belirlenmesi, Bölüm 2'de açıklanan veri seti dengeleme yöntemlerini kullanmaktan daha zorlu bir görevdir. En yakın komşu sayısının (k) geleneksel değeri, genellikle 5 olarak alınsa da önceki araştırma (Chawla ve ark., 2002) tarafından önerildiği gibi, belirli değerlerin veri setine ve eldeki göreve son derece duyarlı olabileceğine dikkat edilmelidir. Şekil 5'te gösterilen diyagram, k parametresinin 1 ila 10 arasında değişen farklı varyasyonları altında kullanılan modellerin ortalama doğruluk ölçüsünü göstermektedir. Bu nedenle, (k) parametresi için uygun değerlerin seçilmesi, optimum performansla sonuçlanabileceği ve oluşturulan dengeli veri setinde uygulanan modeller arasında olası aşırı uyum sorununu azaltabileceği için büyük önem taşımaktadır (Chawla ve ark., 2002). Şekil 5'te sunulan bulgulara göre, k parametresi için en uygun değer 6 olduğu sonucuna

varılabilir. Çünkü bu iddia, k 'nın değeri 6'dan büyük olduğunda kullanılan modellerin ortalama doğruluğu önemli ölçüde azalır.

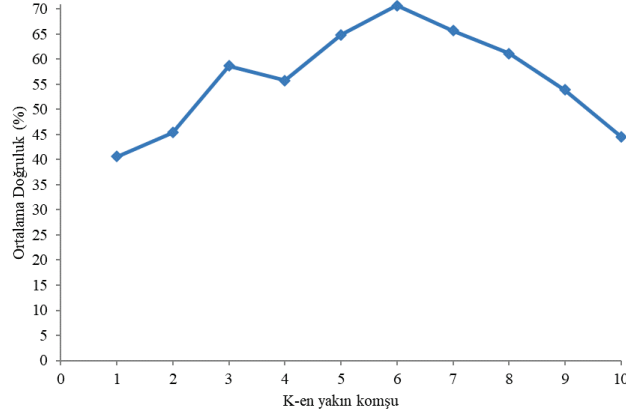
Öte yandan, veri setlerini dengelemek için uygun bir yaklaşımın belirlenmesi de önemlidir. Bölüm 2'de özetlenen veri seti dengeleme tekniklerinin etkinliğini araştırmak için kullanılan modeller, veri ön işleme ve ön eğitim aşamaları uygulanmadan eğitilmiş ve değerlendirilmiştir. Ek olarak, veri seti dengeleme yaklaşımları tarafından üretilen dengelenmiş veri setleri için bölümlenme prosedürü, 3.3 alt bölümünde açıklanan prosedüre benzemektedir. Eğri Altındaki Alan (EAA) metriği, bu yöntemlerin etkisini nicelleştirmek adına önemli bir araçtır. Hesaplaması, Denklem (20)'ye dayalıdır (Sokolova ve Lapalme, 2009).

$$EAA = \frac{1}{2} \times (\text{Duyarlılık} + \text{Özgünlük}) \quad (20)$$

Tablo 2. Veri seti dengeleme yöntemleri kullanılırken kullanılan modellerin $k=6$ ile EAA değerlerinin ortalaması.

Veri seti dengeleme yöntemleri	EAA (%)
Veri seti dengelemeden	40,04
SAYÖT	56,37
Sınır çizgisi-SAYÖT K-EYK	58,89
Sınır çizgisi-SAYÖT DVM	61,74
USÖ	66,23
YDK-SAYÖT	70,65

Tablo 2, çeşitli veri seti dengeleme tekniklerinde kullanılan modellerden elde edilen ortalama EAA değerlerini göstermektedir. Tablo 2, veri seti dengeleme tekniklerinin uygulanmasıyla ortalama EAA değerinde önemli bir artış olduğunu göstermektedir. YDK-SAYÖT yöntemi diğer yöntemlere göre en yüksek EAA değeri ile ortaya çıkmıştır. Sonuç olarak, YDK-SAYÖT yaklaşımı ile elde edilen dengeli veri seti, bir sonraki alt bölümde uygulanmak üzere uygun görülmüştür. Mevcut çalışma, YDK-SAYÖT yönteminin parametrelerinin, çalışma (Hassib ve ark., 2018)'da belirtildiği üzere bire bir değerleri üzerinden ayarlanmıştır.



Şekil 5. En yakın k komşu sayısını değiştirirken kullanılan modellerin ortalama doğruluğu.

5.1.2 Senaryo 2: Veri Ön İşleme Etkisi

İkinci senaryoda önerilen deneyin birincil amacı, veri ön işleme aşamasının oynadığı zorunlu rolü göstermektir. Eğitilen ve değerlendirilen modeller, veri ön işleme işlemlerinden geçen dengeli bir veri seti kullanır. Ancak, önceki senaryodaki ön eğitim aşamasını içermezler. Tablo 3, iki farklı senaryo için kullanılan modellerle elde edilen ortalama Eğri Altındaki Alan (EAA) değerlerini göstermektedir. EAA değeri ortalamasının, veri ön işleme aşamasının uygulanmasının ardından, veri ön işleme yapılmadan elde edilen karşılık gelen değere kıyasla %7 arttığı gözlemlenmiştir. Bu olgu; veri temizleme, özellik çıkarma ve veri normalleştirme mekanizmalarının toplu uygulamasına bağlanabilir.

Tablo 3. Senaryo 2'nin deneyi bittikten sonra kullanılan modellerin ortalama EAA değerleri.

Durumlar	EAA (%)
Veri ön işleme aşaması olmadan	70,65
Veri ön işleme aşaması ile (bir parça)	77,23

5.1.3 Senaryo 3: Ön Eğitimin Etkisi

Bir önceki açıklama, önerilen yöntemin ikinci aşaması ile ilgilidir. Ön eğitim aşamasında, çift Parçacık Sürü Optimizasyonu (PSO) tabanlı algoritma,

önceki senaryodan türetilen önceden işlenmiş veri setini alır ve belirli modelin optimal hiper parametre vektörü ile optimum özellik alt kümesini üretmektedir. Çift PSO tabanlı algoritmanın çalışma parametreleri, (Elmasry ve ark., 2020) tarafından yürütülen araştırmada açıklanan yöntem altında oluşturulmuştur. Daha sonra, geliştirilmiş modeller, indirgenmiş veri seti kullanılarak eğitime ve değerlendirmeye tabi tutulur. Dört farklı senaryo için kullanılan modellerin ortalama EAA değerleri, Tablo 4'te sunulmuştur. Tablo 4'te gösterildiği gibi ikinci senaryo, kullanılan modellerin, varsayılan hiper parametreler kullanılarak ayarlamaya tabi tutulduğunu ve indirgenmiş bir veri seti üzerinde eğitildiğini ima eder. Özellikle, Tablo 4'te gösterilen üçüncü örnek, kullanılan modellerin tek bir değer ayarlanmasını göstermektedir. Bunun ile, en uygun hiper parametre değeri ile önemli ayarlamalar yapılmaktadır. Bunun için özellik seti kullanılarak eğitildiği gözler önüne serilmektedir. Tablo 4'teki bulgulara dayanarak çift PSO tabanlı algoritma, ortalama EAA değerini ön eğitim fazının kullanılmamasına karşılık gelen değere göre %10 artırmıştır. Ek olarak, kullanılan modellerin etkinliği, hem veri ön işleme hem de ön eğitim aşamalarının kullanılmasıyla %16 oranında artırılmıştır.

Tablo 4. Senaryo 3'ün deneyi bittikten sonra kullanılan modellerin ortalama EAA değerleri.

Durumlar	EAA (%)
Ön eğitim fazı olmadan	77,23
Sadece özellik seçimi ile	83,12
Sadece hiper parametre seçimi ile	81,55
Ön eğitim fazı ile	86,49

5.1.4 Senaryo 4: Parça Sayısının Etkisi

Dördüncü senaryo, özellikle bir, iki, dört ve sekiz olmak üzere değişen sayıda parçaya uygulandığında kullanılan modellerin sınıflandırma doğruluğunu araştırmak için tasarlanmıştır. Bu çalışma, dikkate alınan önceden var olan modellere bir makine öğrenimi yöntemi uygular ve ardından Bölüm 4'te belirtildiği gibi belirlenmiş değerlendirme metriklerini hesaplar. Tablo 5, kullanılan parça sayısına göre her bir model için yüzde olarak ifade edilen değerlendirme ölçümlerini göstermektedir. Ayrıca, ilgili denemede ve tüm denemelerde kullanılan modellerin optimum sonuçları Tablo 5'te sunulan vurgulanan parametreler aracılığıyla gözlemlenebilir.

Tablo 5. Farklı parça sayısına sahip kullanılan modellerin sonuçları.

Parça sayısı	Modeller	Değerlendirme Ölçütleri (%)					
		Doğruluk	Kesinlik	F1 puanı	YAO	Özgünlük	MKK
1	YSA	83,75	89,32	82,51	9,16	90,84	68,19
	NB	84,94	89,72	83,98	9,04	90,96	70,40
	DVM	85,80	90,30	84,96	8,61	91,39	72,05
	AKA	88,67	91,59	88,26	7,82	92,18	77,53
	KO	90,01	92,53	89,71	7,03	92,97	80,17
	KSO	90,71	93,03	90,46	6,60	93,40	81,55
2	YSA	87,08	92,87	86,14	6,17	93,83	74,85
	NB	88,52	93,45	87,82	5,80	94,20	77,53
	DVM	90,87	95,70	90,36	3,85	96,15	82,20
	AKA	92,97	96,31	92,71	3,42	96,58	86,17
	KO	93,95	96,81	93,76	2,99	97,01	88,07
	KSO	95,24	97,56	95,12	2,32	97,68	90,58
4	YSA	89,46	94,86	88,79	4,52	95,48	79,50
	NB	90,75	96,00	90,18	3,54	96,46	82,03
	DVM	91,84	96,54	91,41	3,12	96,88	84,12
	AKA	94,53	97,22	94,37	2,63	97,37	89,21
	KO	95,82	98,14	95,71	1,77	98,23	91,74
	KSO	96,67	98,42	96,61	1,53	98,47	93,40
8	YSA	90,71	95,87	90,16	3,67	96,33	81,95
	NB	91,63	96,71	91,15	2,93	97,07	83,76
	DVM	93,68	97,79	93,39	2,02	97,98	87,68
	AKA	96,06	98,21	95,97	1,71	98,29	92,21
	KO	96,82	98,54	96,77	1,41	98,59	93,71
	KSO	97,46	98,87	97,43	1,10	98,90	94,97

Varsayıldığı gibi, parça sayısı arttıkça tüm modellerin değerlendirme metriklerinde önemli bir iyileşme olduğu gözlemlenmiştir. Örneğin, sinyal verilerinin sekiz bölüme ayrılması, doğruluk ölçüsünde %7'lik bir marjla kayda değer bir iyileşme, duyarlılık ölçüsünde %8'lik bir artış ve bununla ilişkili olarak tek bir segment kullanılırken elde edilen performans ölçümleriyle karşılaştırıldığında, YAO değerinde %5,5'lik bir azalma ile sonuçlanmıştır. Bununla birlikte, sinyallerin daha fazla sayıda parçaya ayrışması, karmaşıklıkta bir artışa neden olur. Alternatif olarak belirtildiği gibi, bir model içindeki parçaların sayısındaki artış, modelin yapısal karmaşıklığında buna eşlik eden bir

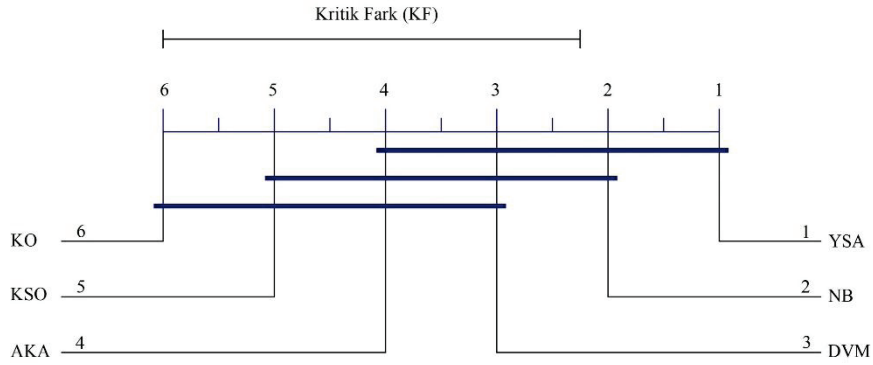
artışa neden olur. Parçaların sayısı ile sonuçta elde edilen öznelilikler arasındaki korelasyon aşamalı olarak artar ve sonuç olarak model için girdi değişkenlerinin sayısı artar. Bu sonuç, parça sayısının çıkarılan özelliklerin genel miktarı üzerindeki doğrudan etkisine bağlanabilir. Yüzeysel makine öğrenimi modellerinin güvenilirliği, yüksek boyutlu özellik uzaylarında azalır. Bu nedenle, derin öğrenme modellerini kullanmak bu özel bağlamda daha etkili bir yaklaşım olarak görünmektedir (Elmasry ve ark., 2018, 2020).

Kullanılan modellerin performansı ile ilgili olarak, AKA, KO ve KSO modellerinin tüm örneklerde tüm metrikler açısından YSA, NB ve DVM modellerine göre üstünlük gösterdiği söylenebilir. Bu olgu için, AKA, KO ve KSO'nun topluluk modelleri, geri kalan modellerin ise modelin mimarisi çerçevesinde tekil modeller olması ayırımından kaynaklanmaktadır (*Azure Machine Learning Studio*, 2020). Buna rağmen, modeller elektrik arızalarını tespit etmek için uygundur. Bu çalışmaya yönelik olarak kullanılan modellerin performansı, Kritik Fark Diyagramı (KFD) vasıtası ile görselleştirilmektedir (Demšar, 2006). Çeşitli sayıda parçada kullanılan modeller için KFD'nin grafiksel gösterimi Şekil 6'da gösterilmektedir. İlişkili şeklin üzerine yerleştirilmiş bir çubukla temsil edilen ve Kritik Fark (KF) olarak belirtilen değerın sayısal karşılığı 3,7702'dir.

Önerilen yöntemin kararlılığını doğrulamak ve elde edilen sonuçların tutarlılığını sağlamak için bazı istatistiksel testlerin yapılması gerekli görülmektedir. Uzamsal kısıtlamalar nedeniyle, Tablo 5'te sunulan bulgularla ilgili olarak yalnızca Friedman testi uygulanmıştır. Friedman testi, çok sayıda tekrarlanan işlemler arasındaki farkları açıklamayı amaçlayan, yaygın olarak bilinen bir istatistiksel testtir (Friedman, 1937). Friedman testi, parametrik olmayan bir test olduğu için uygulama kolaylığı, her şeyi kapsayıcılık ve herhangi bir özel dağılım varsayımına dayanmama gibi çeşitli avantajlara sahiptir. Tablo 5, her biri kullanılan parça sayısına göre değişen dört farklı deneysel koşulu içermektedir. Pratik bir bakış açısından, bu deneyler özellikle önemlidir. Ayrıca, her deneyin kullanılan modellerle ilişkili altı konuyu içerdiğine dikkat edilmelidir. Friedman testinin sıfır hipotezi, deneylerin sayısından bağımsız olarak hepsinde reaksiyona sahip olduğunu varsayar. Bu, soru testinin altında yatan varsayımı temsil eder. Sıfır hipotezini reddetmek için iki temel kriterin karşılanması gerekir. Göze çarpan birinci kriter, hesaplanan istatistiğin ($F\hat{I}$) kritik değeri (FK) aşmasıdır. İkinci kriter, anlamlılık seviyesinin (α) hesaplanan olasılık değerini ($P_{değer}$) aşmasını gerektirir. Bu araştırmada α değerinin 0,05 olarak seçilmesi literatürdeki yaygın kullanımına dayanmaktadır.

Tablo 6'da Friedman testinin DP, DN, YN ve YP çıktıları uygulandığındaki sonuçları görülmektedir.

Tablo 6'dan, Friedman testinin sıfır hipotezi reddedilir. Çünkü, iki koşul da tüm durumlarda sağlanır ($FK < F\dot{I}$ ve $\alpha > P_{değer}$). Bu nedenle, kullanılan modellerin çıktıları anlamlı ve birbirinden farklıdır.



Şekil 6. Senaryo 4'ün tüm deneyleri için kullanılan modellerin kritik fark diyagramı.

Tablo 6. Friedman testinin sonuçları.

Çıktılar	FK	$F\dot{I}$	α	$P_{değer}$
DN	7,6	18	0,05	0,00044
YP	7,6	18	0,05	0,00044
DP	7,6	18	0,05	0,00044
YN	7,6	18	0,05	0,00044

5.1.5 Senaryo 5: Özellik Seçim Yönteminin Etkisi

Bu senaryoda kullanılan UOSİPSO-E öznitelik seçim yöntemi, Stokastik Fraktal Arama tabanlı Kılavuzlu Balina Optimizasyon Algoritması (SFA-Kılavuzlu BOA) (El-Kenawy ve ark., 2020), Hibrit Gri Kurt Optimizasyonu ve Parçacık Sürüsü Optimizasyonu (GKO-PSO) (Friedman, 1937), Hibrit Gri Kurt Optimizasyonu ve Genetik Algoritma (GKO-GA) (Şenel ve ark., 2019), Biyocoğrafya Tabanlı Optimize Edici (BTOE) (Simon, 2008), Ateşböceği Algoritması (AA) (Fister ve ark., 2012) ve Saten Çardak Kuşu Optimize Edici (SÇKOE) (Samareh Moosavi ve Khatibi Bardsiri, 2017) gibi diğer güncel özellik seçim yöntemleriyle, EAA ve Özellik Azaltma Oranı (ÖAO) (Elmasry ve ark., 2020) metriklerinin ortalaması açısından performansını doğrulamak için karşılaştırılır. ÖAO metriği, seçilen özelliklerin tüm özellik setine oranının tamamlayıcısıdır ve Eşitlik (21)'deki gibi hesaplanabilir.

$$\text{ÖAO} = 1 - \frac{\text{Seçilen özellik sayısı}}{\text{Tüm özelliklerin sayısı}} \quad (21)$$

Deney, tüm özellik setinin 20 bileşenden oluştuğu tek bir parça üzerinde gerçekleştirilmiştir. Ayrıca kullanılan modeller, öznitelik seçme teknikleriyle küçültülmüş veri setleri kullanılarak eğitim ve değerlendirme süreçlerinden geçirilmektedir. Beşinci senaryoda gerçekleştirilen çalışma sonuçları Tablo 7'de gösterilmektedir. UOSİPSO-E algoritmasının, hem performansta hem de minimum özellik alt kümesi seçiminde alternatif özellik seçim tekniklerinden daha iyi performans gösterdiği sonucuna varılmaktadır.

Tablo 7. UOSİPSO-E ile diğer öznitelik seçim yöntemlerinin karşılaştırılması.

Özellik seçim yöntemleri	Seçilen Özellik Sayısı	ÖAO (%)	EAA (%)
SFA-Kılavuzlu BOA	15	25	83,54
GKO-PSO	17	15	82,37
GKO-GA	17	15	82,64
BTOE	19	5	80,46
AA	16	20	82,95
SÇKOE	18	10	81,71
UOSİPSO-E	12	40	86,49

5.1.6. Senaryo 6: Hiper parametre Optimizasyon Yönteminin Etkisi

Kullanılan PSO tabanlı hiper parametre optimizasyon yönteminin performansı, orijinal Genetik Algoritma (GA) (Holland, 1992), Çekirge Optimizasyon Algoritması (ÇOA) (Saremi ve ark., 2017), Balina Optimizasyon Algoritması (BOA) (Mirjalili ve Lewis, 2016), Gri Kurt Optimizasyonu (GKO) (Mirjalili ve ark., 2014), Yarasa Algoritması (YA) (Karakonstantis ve Vlachos, 2021) ve Çoklu Evren Optimizasyonu (ÇEO) (Mirjalili ve ark., 2016) gibi diğer popüler optimizasyon algoritmaları ile karşılaştırılarak kullanılan modellerin ortalama EAA değerleri açısından incelenmektedir. Tablo 8, hiper parametrelerin en uygun değerinin bulunması üzerine kurgulanan senaryo 6'ya ait deney bulgularını paylaşmaktadır. Hiper parametre seçimi açısından kullanılan optimizasyon yöntemleri arasında en iyi performansı PSO algoritması vermiştir.

Tablo 8. Hiper parametre seçimi için PSO ile diğer optimizasyon algoritmaları arasındaki karşılaştırma.

Optimizasyon Algoritmaları	EAA (%)
GA	84,09
ÇOA	81,22
BOA	83,65
GKO	84,41
YA	80,84
ÇEO	83,38
PSO	86,49

5.3 Önceki Çalışmalarla Karşılaştırma

Bölüm 1'de açıklandığı gibi, elektrik arıza tespitindeki önceki çalışmaların çoğu, gerçek elektrik şebekelerindeki elektrik arızalarının modelini yansıtamayan simüle edilmiş veri setleri kullanmıştır. Önerdiğimiz yaklaşım, sadece (Dong ve ark., 2019; Wadi ve Elmasry, 2021)'deki çalışmalarla karşılaştırılmıştır. VSB veri seti, (Dong ve ark., 2019; Wadi ve Elmasry, 2021) çalışmalarında da kullanılmıştır. Böylece okuyucular sonuçları kolayca karşılaştırabilir. Tablo 9, bulgularımız diğer bulgular arasındaki farkları belirlemek için oluşturulmuştur. Açıkça görülmektedir ki, aynı veri setine kullanan bu çalışmalara göre çalışmamızdaki model performansları, önerdiğimiz yöntem ile yükselmiştir. Böylece, elektriksel arıza tespitine yönelik olarak makine öğrenimi yaklaşımı etkili bir sistemdir. Tablo 9 ile bu durum kanıtlanmaktadır.

Tablo 9. Önceki çalışmalarla karşılaştırma.

Çalışma	Teknik	Doğruluk	Kesinlik	F1 puanı
(Dong ve ark., 2019)	STL+DVM (Lineer Çekirdek)	-	0,72	0,7
	STL+DVM (Derece 6 ile Polinom Çekirdek)	-	0,67	0,45
	STL+DVM (RTF Çekirdek)	-	0,75	0,72
	STL+DVM (Sigmoid Çekirdek)	-	0,53	0,5
(Wadi ve Elmasry, 2021)	Anomali tespiti+DVM (RBF Çekirdek)	0,7984	0,6984	0,6842
	Anomali tespiti+TBA	0,7928	0,6673	0,6953
Çalışmamız	Çift PSO+YSA	0,9071	0,9587	0,9016
	Çift PSO+NB	0,9163	0,9671	0,9115
	Çift PSO+DVM (RTF Çekirdek)	0,9368	0,9779	0,9339
	Çift PSO+AKA	0,9606	0,9821	0,9597
	Çift PSO+KO	0,9682	0,9854	0,9677
	Çift PSO+KSO	0,9746	0,9887	0,9743

STL: Seasonal-Trend decomposition using LOESS, RTF: Radyal Temelli Fonksiyon, TBA: Temel Bileşen Analizi

6. Sonuç

Bu çalışma, elektriksel arıza tespiti konusundaki araştırmaların mevcut durumunu eleştirel bir şekilde incelemekte ve bu alandaki önemli araştırma boşluklarını tespit ederek bu alandaki araştırmacıları aydınlatmaktadır. Elektrik enerjisi endüstrisinde etkili arıza tespit önlemlerinin uygulanmasının hizmet kalitesini artırması ve ekonomik kayıpların azaltılmasına katkıda bulunması beklenmektedir. Sonuç olarak, mevcut sistemlerin eksikliklerini ve sınırlamalarını ele almak için otomatik elektrik arıza tespiti için önerilen bir yöntem tanıtılmaktadır. Önerilen makine öğrenimi yaklaşımı; dengeli veri setinin 4 fazda gerçekleştirilmesi ile tamamlanır. Bunlardan ilk ikisi veri ön işleme ve ön eğitim aşamalarıdır. Bu aşamalar, eğitim ve test aşamalarına veri seti sunmak için gereklidir. İlk ve ikincil aşamalar, temel görevlerin yerine getirilmesinde önemli aşamalar olarak kabul edilir. Veri hazırlamanın ilk aşaması, gürültü giderme, ayrıştırma, özellik çıkarma ve veri normalleştirme dahil olmak üzere çeşitli prosedürler yoluyla ham elektrik sinyallerinin uygun şekilde önceden işlenmesini gerektirir. Buna karşılık olarak ön eğitim

aşamasında, en uygun özelliklerin seçilmesi ve hiper parametre seçimleri tespit edilmelidir. Bunun için çift PSO algoritması uygulanmıştır. Önerilen yöntemi doğrulamak için altı makine öğrenimi modeli kullanılmıştır: Yapay Sinir Ağları, Naive Bayes, Destek Vektör Makinesi, Artırılmış Karar Ağacı, Karar Ormanı ve Karar Ormanı. Bu modellerin kullanılma amaçları, VSB veri setindeki hataları tespit etmektir. Deneysel bulgular, analitik araçlar olarak AÇK analizinin ve Friedman testinin konuşlandırılmasıyla birlikte, deneylerin sistematik olarak altı benzersiz senaryoya bölünmesiyle derinlemesine incelenir. Elde edilen bulgulara göre, çalışmamız önemli bir gelişmeyi göstermektedir. Elektrik arıza tespiti için önerilen yaklaşımın uygulanması, özellikle doğrulukta ve duyarlılıkta olmak üzere sınıflandırma ölçütlerinin hassasiyetinde %15'lik fark edilebilir bir artışa yol açmıştır. Ayrıca, önerilen yöntemi içermeyen önceki modellere kıyasla yanlış sınıflandırma ölçütlerinde, özellikle YAO'da %10'luk önemli bir azalma gözlemlenmiştir. Bununla birlikte, parça sayısını artırmak, daha önce kullanılan modellere potansiyel olarak zarar verebilecek özellik uzayının boyutsallığını çoğaltma potansiyeline sahiptir. Bu nedenle, elektrik arıza tespitini artırmak için derin öğrenme modellerini kullanan önerilen yaklaşımın uygulanması, gelecekteki araştırmalar için uygun bir yol olarak kabul edilir.

Kaynakça

- Azure Machine Learning Studio*. (2020). Microsoft. <https://studio.azureml.net/>. Yayın tarihi 2020. Erişim tarihi November 2022.
- Cervante, L., Xue, B., Zhang, M., ve Shang, L. (2012). *Binary Particle Swarm Optimisation for Feature Selection: A Filter Based Approach*.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., ve Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).
- Chen, K., Huang, C., ve He, J. (2016). Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. In *High Voltage* (Vol. 1, Issue 1, pp. 25–33). Institution of Engineering and Technology. <https://doi.org/10.1049/hve.2016.0005>
- Demšar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. In *Journal of Machine Learning Research* (Vol. 7).
- Dong, M., Sun, Z., ve Wang, C. (2019). A pattern recognition method for partial discharge detection on insulated overhead conductors. *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, 1–4.

Eboule, P. S. P., Pretorius, J. H. C., Mbuli, N., ve Leke, C. (2018). Fault detection and location in power transmission line using concurrent neuro fuzzy technique. *2018 IEEE Electrical Power and Energy Conference (EPEC)*, 1–6.

El-Kenawy, E. S. M., Ibrahim, A., Mirjalili, S., Eid, M. M., ve Hussein, S. E. (2020). Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images. *IEEE Access*, 8. <https://doi.org/10.1109/ACCESS.2020.3028012>

Elmasry, W., Akbulut, A., ve Zaim, A. H. (2018). Deep Learning Approaches for Predictive Masquerade Detection. *Security and Communication Networks*, 2018. <https://doi.org/10.1155/2018/9327215>

Elmasry, W., Akbulut, A., ve Zaim, A. H. (2020). Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic. *Computer Networks*, 168. <https://doi.org/10.1016/j.comnet.2019.107042>

Elmasry, W., ve Wadi, M. (2022a). Detection of Faults in Electrical Power Grids Using an Enhanced Anomaly-Based Method. *Arabian Journal for Science and Engineering*, 47(11), 14899–14914. <https://doi.org/10.1007/s13369-022-07030-x>

Elmasry, W., ve Wadi, M. (2022b). EDLA-EFDS: A Novel Ensemble Deep Learning Approach For Electrical Fault Detection Systems. *Electric Power Systems Research*, 207. <https://doi.org/10.1016/j.epsr.2022.107834>

Elmasry, W., ve Wadi, M. (2022c). Enhanced Anomaly-Based Fault Detection System in Electrical Power Grids. *International Transactions on Electrical Energy Systems*, 2022, 1–19. <https://doi.org/10.1155/2022/1870136>

Elmasry, W., Wadi, M., ve Shahinzadeh, H. (2022). Two-Tier Cascaded Classifiers to Improve Electrical Power Quality. *2022 26th International Electrical Power Distribution Conference, EPDC 2022*, 96–101. <https://doi.org/10.1109/EPDC56235.2022.9817239>

ENET Centre. (t.y.). VSB. Erişim tarihi Aralık, 2020.

Fister, I., Yang, X.-S., Fister, I., ve Brest, J. (2012). *Memetic firefly algorithm for combinatorial optimization*. <http://arxiv.org/abs/1204.5165>

Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675–701.

García-Gil, D., Luengo, J., García, S., ve Herrera, F. (2019). Enabling Smart Data: Noise filtering in Big Data classification. *Information Sciences*, 479, 135–152. <https://doi.org/10.1016/j.ins.2018.12.002>

Gururajapathy, S. S., Mokhlis, H., ve Illias, H. A. Bin. (2018). Classification and regression analysis using support vector machine for

classifying and locating faults in a distribution system. *Turkish Journal of Electrical Engineering and Computer Sciences*, 26(6), 3044–3056. <https://doi.org/10.3906/elk-1711-194>

Han, H., Wang, W.-Y., ve Mao, B.-H. (2005). Borderline-SMOTE: A New Over-Sampling Method in Imbalanced Data Sets Learning. In *LNCS* (Vol. 3644).

Hassib, E. M., Eldesokey, A., Labib, L., ve Elghamrawy, S. (2018). LSH-SMOTE: a modified SMOTE algorithm for imbalanced data-sets. *Ciência Técnica Vitivinícola*, 33(4), 50–65.

He, H., Bai, Y., Garcia, E. A., ve Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>

He, H., ve Mu, Y. (2013). *Imbalanced Learning: Foundations, Algorithms, and Applications*. John Wiley ve Sons.

Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.

Huang, C., Ding, S., Li, S., ve Liu, R. (2024). LMFE: Learning-Based Multiscale Feature Engineering in Partial Discharge Detection. *IEEE Transactions on Neural Networks and Learning Systems*, 35(5), 5848–5856. <https://doi.org/10.1109/TNNLS.2022.3222671>

Jamil, M., Sharma, S. K., ve Singh, R. (2015). Fault detection and classification in electrical power transmission system using artificial neural network. *SpringerPlus*, 4(1). <https://doi.org/10.1186/s40064-015-1080-x>

Karakonstantis, I., ve Vlachos, A. (2021). Bat algorithm applied to continuous constrained optimization problems. *Journal of Information and Optimization Sciences*, 42(1), 57–75.

Koley, E., Verma, K., ve Ghosh, S. (2015). An improved fault detection classification and location scheme based on wavelet transform and artificial neural network for six phase transmission line using single end data only. *SpringerPlus*, 4(1). <https://doi.org/10.1186/s40064-015-1342-7>

Mbamaluikem, P. O., Awelewa, A. A., ve Samuel, I. A. (2018). An Artificial Neural Network-Based Intelligent Fault Classification System for the 33-kV Nigeria Transmission Line. In *International Journal of Applied Engineering Research* (Vol. 13, Issue 2). <http://www.ripublication.com>

Mirjalili, S., ve Lewis, A. (2016). The Whale Optimization Algorithm. *Advances in Engineering Software*, 95, 51–67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>

Mirjalili, S., Mirjalili, S. M., ve Hatamlou, A. (2016). Multi-Verse Optimizer: a nature-inspired algorithm for global optimization. *Neural Computing and Applications*, 27(2), 495–513. <https://doi.org/10.1007/s00521-015-1870-7>

Mirjalili, S., Mirjalili, S. M., ve Lewis, A. (2014). Grey Wolf Optimizer. *Advances in Engineering Software*, 69, 46–61. <https://doi.org/10.1016/j.advengsoft.2013.12.007>

Naveed, K., Akhtar, M. T., Siddiqui, M. F., ve ur Rehman, N. (2021). A statistical approach to signal denoising based on data-driven multiscale representation. *Digital Signal Processing: A Review Journal*, 108. <https://doi.org/10.1016/j.dsp.2020.102896>

Nguyen, H. M., Cooper, E. W., ve Kamei, K. (2011). Borderline over-sampling for imbalanced data classification. *International Journal of Knowledge Engineering and Soft Data Paradigms*, 3(1), 4–21. <https://doi.org/10.1504/IJKESDP.2011.039875>

Prasad, A., Belwin Edward, J., ve Ravi, K. (2018a). A review on fault classification methodologies in power transmission systems: Part-II. *Journal of Electrical Systems and Information Technology*, 5(1), 61–67. <https://doi.org/10.1016/j.jesit.2016.10.003>

Prasad, A., Belwin Edward, J., ve Ravi, K. (2018b). A review on fault classification methodologies in power transmission systems: Part—I. *Journal of Electrical Systems and Information Technology*, 5(1), 48–60. <https://doi.org/10.1016/j.jesit.2017.01.004>

Qu, N., Li, Z., Zuo, J., ve Chen, J. (2020). Fault Detection on Insulated Overhead Conductors Based on DWT-LSTM and Partial Discharge. *IEEE Access*, 8, 87060–87070. <https://doi.org/10.1109/ACCESS.2020.2992790>

Raza, A., Benrabah, A., Alquthami, T., ve Akmal, M. (2020). A review of fault diagnosing methods in power transmission systems. In *Applied Sciences (Switzerland)* (Vol. 10, Issue 4). MDPI AG. <https://doi.org/10.3390/app10041312>

Samareh Moosavi, S. H., ve Khatibi Bardsiri, V. (2017). Satin bowerbird optimizer: A new optimization algorithm to optimize ANFIS for software development effort estimation. *Engineering Applications of Artificial Intelligence*, 60, 1–15. <https://doi.org/10.1016/j.engappai.2017.01.006>

Saremi, S., Mirjalili, S., ve Lewis, A. (2017). Grasshopper Optimisation Algorithm: Theory and application. *Advances in Engineering Software*, 105, 30–47. <https://doi.org/10.1016/j.advengsoft.2017.01.004>

Şenel, F. A., Gökçe, F., Yüksel, A. S., ve Yiğit, T. (2019). A novel hybrid PSO–GWO algorithm for optimization problems. *Engineering with Computers*, 35(4), 1359–1373. <https://doi.org/10.1007/s00366-018-0668-5>

Simon, D. (2008). Biogeography-based optimization. *IEEE Transactions on Evolutionary Computation*, 12(6), 702–713. <https://doi.org/10.1109/TEVC.2008.919004>

Sokolova, M., ve Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>

Uzubi, U., Ekwue, A., ve Ejiogu, E. (2017). Artificial neural network technique for transmission line protection on Nigerian power system. *2017 IEEE PES PowerAfrica*, 52–58.

VSB Power Line Fault Detection. (2018). <https://www.kaggle.com/c/vsb-Power-Line-Fault-Detection/Data>. Erişim tarihi Aralık, 2020.

Wadi, M. (2021). Fault detection in power grids based on improved supervised machine learning binary classification. *Journal of Electrical Engineering*, 72(5), 315–322. <https://doi.org/10.2478/jee-2021-0044>

Wadi, M., ve Elmasry, W. (2021, March 23). An Anomaly-based Technique for Fault Detection in Power System Networks. *2021 International Conference on Electric Power Engineering - Palestine, ICEPE-P 2021*. <https://doi.org/10.1109/ICEPE-P51568.2021.9423479>

Wadi, M., ve Elmasry, W. (2022). Comparison of five different distributions based on three metaheuristics to model wind speed distribution. *Journal of Electrical Systems*.

Wadi, M., Elmasry, W., ve Tamyiğit, F. A. (2023). Important considerations while evaluating wind energy potential. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 38(2), 947–962. <https://doi.org/10.17341/gazimmfd.1066351>

Weiss, G. M. (2013). Foundations of Imbalanced Learning. In *Imbalanced Learning* (pp. 13–41). John Wiley ve Sons, Ltd. <https://doi.org/https://doi.org/10.1002/9781118646106.ch2>

Yasir, M. N., ve Koh, B. H. (2018). Data decomposition techniques with multi-scale permutation entropy calculations for bearing fault diagnosis. *Sensors (Switzerland)*, 18(4). <https://doi.org/10.3390/s18041278>

Yue, X. (2019). Data decomposition for analytics of engineering systems: Literature review, methodology formulation, and future trends. *International Manufacturing Science and Engineering Conference*, 58745, V001T02A011.

Zhou, Z., Liu, X., Li, P., ve Shang, L. (2014). Feature selection method with proportionate fitness based binary particle swarm optimization. *Simulated Evolution and Learning: 10th International Conference, SEAL 2014, Dunedin, New Zealand, December 15-18, 2014. Proceedings 10*, 582–592.