ENDÜSTRİYEL KONTROL SİSTEMLERİNDE ANOMALİ TESPİTİ

Emine Hümeyra DEMİRCİOĞLU

T.C.
BURSA ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ENDÜSTRİYEL KONTROL SİSTEMLERİNDE ANOMALİ TESPİTİ

Emine Hümeyra DEMİRCİOĞLU
0000-0003-4296-5532

Doç. Dr. Ersen YILMAZ
(Danışman)

DOKTORA TEZİ
ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

BURSA – 2024
Her Hakkı Saklıdır

ÖZET

Doktora Tezi

ENDÜSTRİYEL KONTROL SİSTEMLERİNDE ANOMALİ TESPİTİ

Emine Hümeyra DEMİRCİOĞLU

Bursa Uludağ Üniversitesi
Fen Bilimleri Enstitüsü
Elektronik Mühendisliği Anabilim Dalı

Danışman: Doç. Dr. Ersen YILMAZ

Endüstriyel sistemlerde oluşan büyük veri bu sistemler için kontrol ve izleme araçlarının geliştirilmesi süreçlerinde önemli zorluklar oluşturmaktadır. Son yıllarda, endüstriyel sistemlerde oluşan büyük verinin analizi gelişen teknolojiye paralel olarak yapay zekâ temelli yöntemler ile gerçekleştirilmektedir. Endüstriyel kontrol sistemlerinin görevleri arasında sistemin sağlık durumunun izlenmesi önemli bir yer tutmaktadır. Sistem sağlık durumunun izlenmesi için kullanılan yaklaşımlar arasında ise anomali tespiti temelli yöntemler yoğun ilgi çekmektedir. Sistem bileşenlerinde oluşan anomaliler tüm sistemi etkileyebilmekte, zaman ile sistemin başarımını düşürebilmekte ve hatta sistemin arızalanmasına sebep olabilmektedir. Arızaların oluşma nedeninin, zamanın ve süresinin tespit edilmesi zaman alıcı ve aynı zamanda zor olabilmektedir. Bu zorlukların aşılmasında yapay zekâ temelli yöntemlerin kullanımı önemli bir avantaj sağlamaktadır.

Bu tez çalışmasında endüstriyel sistemlerin bileşenlerinde oluşan anomalilerinin tespitinde otokodlayıcıların başarım analizlerinin gerçekleştirilmesi amaçlanmıştır. Bu nedenle endüstriyel sistemlerin önemli bileşenleri arasında yer alan ve bu sistemlerde sıklıkla karşımıza çıkan DC motorlarda oluşabilecek anomalilerin tespiti üzerine yoğunlaşılmıştır. DC motor üzerinde yer alan sensörler aracılığı ile toplanan ve motorun sağlık izlemesinde temel gösterge olarak kullanılan gövde sıcaklığı verisindeki anomalilerin tespiti otokodlayıcılar aracılığı ile gerçekleştirilmiştir. Gövde sıcaklığı verisinde oluşabilecek noktasal ve toplu anomali tipleri için otokodlayıcıların başarım analizleri yapılmıştır. Başarım analizleri yapılırken karşılaştırma yöntemi olarak noktasal anomaliler için 3-SgOut Aykırı Değer yöntemi kullanılırken toplu anomaliler için ise İzolasyon Ormanı yöntemi kullanılmıştır.

Başarım analizleri yapılırken bu çalışmaya özel olarak geliştirilen bir deney düzeneği aracılığı ile toplanan veri seti kullanılmıştır. Yapılan deneyler sonucunda Otokodlayıcı mimarisinin noktasal ve toplu anomalilerin tespit edilmesinde yüksek başarım oranlarına sahip olduğu ve karşılaştırma amacıyla kullanılan yöntemlerden daha iyi başarım gösterdiği görülmüştür.

Anahtar Kelimeler: Otokodlayıcı, anomali tespiti, endüstriyel sistemler, yapay sinir ağları, DC motor

2024, xvii + 68 sayfa.

ABSTRACT

PhD Thesis

ANOMALY DETECTION IN INDUSTRIAL CONTROL SYSTEMS

Emine Hümeyra DEMİRCİOĞLU

Bursa Uludağ University
Graduate School of Natural and Applied Sciences
Department of Electronics Engineering

Supervisor: Assoc. Prof. Dr. Ersen YILMAZ

Big data generated in industrial systems poses significant difficulties in the development of control and monitoring tools for these systems. In recent years, analysis of big data occured in industrial systems has been carried out with artificial intelligence based methods in parallel with developing technology. Monitoring the health status of the system has an important place among the tasks of industrial control systems. Among the approaches used to monitor system health, anomaly detection-based methods attract great attention. Anomalies occurring in system components can affect the entire system, reduce the performance of the system over time, and even cause it to malfunction. Determining the cause, time and duration of malfunctions can be time-consuming and also be difficult. The use of artificial intelligence-based methods provides a significant advantage in overcoming these difficulties.

In this thesis, it is aimed to perform performance analysis of autoencoders in detecting anomalies in the components of industrial systems. For this reason, we focused on detecting anomalies in DC motors, which are important components of industrial systems and are frequently used in these systems. Detection of anomalies in the body temperature data collected by the sensors on DC motor, which is used as a basic indicator in the health monitoring of the motor, was carried out with autoencoders. Performance analyzes of autoencoders were performed for point and collective anomaly types in body temperature data. While performing performance analyses, the 3-SgOut Outlier method was used as a comparison method for point anomalies, while the Isolation Forest method was used for mass anomalies.

While performing performance analyses, the data set obtained through the experimental setup developed specifically for this study was used. It was seen that the Autoencoder architecture used in the thesis study had high performance rates in detecting point and collective anomalies and showed better performance than the methods used for comparison purposes.

Key words: Autoencoder, anomaly detection, industrial systems, artificial neural networks, DC motor

2024, xvii + 68 pages.

ÖNSÖZ ve/veya TEŞEKKÜR

Doktora sürecimde emeğini, desteğini esirgemeyen, yaptığım araştırmalara ve çalışmalara yön veren kıymetli danışmanım Doç. Dr. Ersen Yılmaz hocama teşekkürlerimi sunarım.

Doktora sürem boyunca desteğini esirgemeyen eşim Tayfur Kerem Demircioğlu'na, annem Nur Metin'e, babam Osman Metin'e, kardeşim Süleyman Metin'e teşekkürlerimi sunarım.

Doktora çalışmalarımda bana şans getirdiği için kızım Çiğdem Nur Demircioğlu'na teşekkürlerimi sunarım.

Tez çalışmasında kullanılan deney düzeneğinin geliştirilmesi sürecinde sağladığı desteklerden dolayı Yasin AYGÖR hocama teşekkürlerimi sunarım.

Bu tez çalışması 118C157 numaralı TÜBİTAK BİDEB 2244-Sanayi Doktora Programı'nın "Endüstriyel Gerçek Zamanlı Akıllı Kontrol Sistemlerinin Geliştirilmesi" projesi kapsamında tamamlanmıştır. Bu kapsamda sağlamış olduğu destekten dolayı TÜBİTAK'a ve projenin sanayi ortağı EMKO Elektronik A.Ş.'nin Yönetim Kurulu Başkanı Ayhan ISPALAR ile Arge Merkezi Yöneticisi Erkan EGELİ'ye teşekkürlerimi sunarım.

Emine Hümeyra DEMİRCİOĞLU
1/11/2024

İÇİNDEKİLER

Sayfa

ÖZET.....viii
ABSTRACT.....ix
ÖNSÖZ ve/veya TEŞEKKÜR.....x
SİMGELER ve KISALTMALAR DİZİNİ.....xiii
ŞEKİLLER DİZİNİ.....xv
ÇİZELGELER DİZİNİ.....xvii
1. GİRİŞ.....1
1.1. Endüstriyel Sistemler.....1
1.2. Anomali Tespiti.....2
1.3. Tezin Amaç ve Kapsamı.....3
1.4. Tezin Katkısı.....4
1.5. Tezin Organizasyonu.....5
2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI.....6
2.1. Yapay Zekâ.....18
2.2. Makine Öğrenmesi.....19
2.2.1. Denetimli Öğrenme.....19
2.2.2. Denetimsiz Öğrenme.....20
2.2.3. Yarı Denetimli Öğrenme.....21
2.2.4. Takviyeli Öğrenme.....21
2.3. Derin Öğrenme.....22
2.3.1. Yapay Sinir Ağları.....22
2.4. Anomali Nedir?.....25
2.4.1 Anomali Tipleri.....27
3. MATERYAL ve YÖNTEM.....29
3.1. DC Motor Deney Düzeneği ve Anomali Veri Seti.....29
3.1.1. DC Motor Deney Düzeneği.....29
3.1.2. Anomali Veri Seti.....33
3.2. Anomali Tespit Yöntemleri.....33
3.2.1. Otokodlayıcı.....34
3.2.2. 3-Sigma Aykırı Değeri (3-SgOut).....36
3.2.4. Karar Ağacı (KA).....37
3.2.4.1. İzolasyon Ormanı.....37
3.2.5. Başarım Değerlendirmesi.....40
4. BULGULAR ve TARTIŞMA.....42
4.1. Bulgular.....42
4.1.1. Noktasal Anomali Tespiti.....42
4.1.1.1 Otokodlayıcı ile Anomali Tespiti.....46
4.1.1.2. 3-SgOut ile Anomali Tespiti.....49
4.1.2. Toplu Anomali Tespiti.....50
4.1.2.1. Otokodlayıcı ile Anomali Tespiti.....52
4.1.2.2. İzolasyon Ormanı ile Anomali Tespiti.....55
4.2. Tartışma.....56
4.2.1. Noktasal Anomali Tespiti.....56
4.2.2. Toplu Anomali Tespiti.....57

4.2.3. Pencere Boyutu Seçimi ve Başarım Sonuçlarına Etkisi.....57
5. SONUÇ.....62
KAYNAKLAR.....63
ÖZGEÇMİŞ.....68

SİMGELER ve KISALTMALAR DİZİNİ

Simgeler	Açıklama
$\hat{f}$	Makine öğrenmesinde öğretilmiş fonksiyon
w	Sinir Ağının Ağırlığı
t	Zaman
b	Önyargı-Bias
α_f	Aktivasyon Fonksiyonu
z	Giriş Ağırlıkları
s	Algılayıcıların Durumu
η	Öğrenme Oranı
y	Nöron Çıkışı
N	Yapay Sinir Ağında Bulunan Birim Küme
y_u	u Biriminin Çıktısı
$W_{[v,u]}$	v Birimi u Birimi Arasındaki Ağırlık
$X_{[v,u]}$	v Biriminden u Birimine Gelen Giriş
x_u	u Biriminin Girişi
z_u	u Biriminin Giriş Ağırlığı
b_u	u Biriminin Bias'ı
s_u	u Biriminin Durumu
f_u	u Biriminin Sıkıştırma Fonksiyonu
e_u	u Biriminin Hatası
ϑ_u	u Biriminin Hata Sinyali
y_u	u Biriminin Çıkışı
x	Giriş Vektörleri
h	Kodlar
$\hat{x}$	Yeniden Oluşturulmuş Vektörler
x_t	Girdilerin Vektörel Gösterimi
K_t	Motor Tork Sabiti
J	Rotorun Eylemsizlik Momenti
i	Armatür Akımı
b	Motor Viskoz Sürtünme Sabiti
R	Elektrik Direnci
K_e	Elektromotor Kuvvet Sabiti
L	Elektrik Endüktansı
E	Ağın Genel Hatası

Kısaltmalar	Açıklama
KNN	K Nearest Neighborhood
FFNN	Feed-Forward Neural Networks
NTS	Normal Test Numunesi
BPTT	Backpropagation Through Time
RTRL	Real-Time Recurrent Learning
ATS	Anomali Test Numunesi
TV	Test Verisi
OKT	Otokodlayıcı Tahmini
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
MSE	Mean Squared Error
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
3-SgOut	3-Sigma Outlier Detection
LVQ	Learning Vector Quantization
ART	Adaptip Rezonans Teori Ağları
SPI	Serial Peripheral Interface
ATmega2560	Arduino Mega 2560
NTÖ	Normal Test Örneği
ATÖ	Anomali Test Örneği
KOH	Karesel Ortalama Hata
TP	True Positive (Gerçek Pozitif)
FP	False Positive (Yanlış Pozitif)
TN	True Negative (Gerçek Negatif)
FN	False Negative (Yanlış Negatif)
Logsig	Log-Sigmoid
Pureline	Linear Transfer Function

ŞEKİLLER DİZİNİ

		Sayfa
Şekil 1.1.	Endüstri 4.0 Bileşenleri.....	2
Şekil 2.1.	Yapay Zekâ için Venn Şeması	18
Şekil 2.2.	Denetimli Öğrenme Algoritmasının Şeması	20
Şekil 2.3.	Takviyeli Öğrenme Algoritmasının Şeması	21
Şekil 2.4.	Yapay Sinir Ağı (Algılayıcı) Yapısı	23
Şekil 2.5.	ÇKYSA Mimarisi	24
Şekil 2.6.	Veri Madenciliğini Yapısındaki Disiplinler	26
Şekil 2.7.	Anomalilik Tespiti ile İlgili Ana Yapının Bileşenleri.....	27
Şekil 3.1.	Oluşturulan DC Motor Deney Düzeneği	31
Şekil 3.2.	Oluşturulan DC Motor Sisteminin Akım, Hız, Sıcaklık Verilerinin İzlenebildiği Panelin Görüntüsü	31
Şekil 3.3.	Oluşturulan DC Motor Sisteminin Ayrıntılı Çizimi	32
Şekil 3.4.	Veri Değerlerinin (Sıcaklık, Hız, Akım) İzlenmesi	33
Şekil 3.5.	Normal ve Hatalı Veri Örneklerinin Görselleştirilmesi	34
Şekil 3.6.	Tipik bir otokodlayıcı mimarisi	35
Şekil 3.7.	3-Sigma Aykırı Değeri için Grafiksel Gösterim.....	36
Şekil 4.1.	Sıcaklık Sensörü Zaman Serisinde Yer Alan Noktasal Anomaliler	43
Şekil 4.2.	Noktasal Anomali Tespiti Görsel Anlatımı	44
Şekil 4.3.	Otokodlayıcı ile Noktasal Anomali Tespiti için Akış Diyagramı	45
Şekil 4.4.	3-Sigma Aykırı Değeri ile Noktasal Anomali Tespiti için Akış Diyagramı	46
Şekil 4.5.	Otokodlayıcı ile Noktasal Anomali Tespitinde; Normal Test Örneği için (a) NTÖ (b) NTÖ Kodu (c) Yeniden Oluşturulan NTÖ (d) NTÖ Hatası.....	47
Şekil 4.6.	Otokodlayıcı ile Noktasal Anomali Tespitinde; Anomali Test Örneği için (a) ATÖ (b) ATÖ Kodu (c) Yeniden Oluşturulan ATÖ (d) ATÖ Hatası.....	48
Şekil 4.7.	Toplu Anomali Tespiti Görsel Anlatımı.....	50
Şekil 4.8.	İzolasyon Ormanı ile Toplu Anomalilerin Tespiti için Akış Diyagramı	51
Şekil 4.9.	Otokodlayıcı ile Toplu Anomali Tespitinde; Normal Test Örneği için (a) NTÖ (b) NTÖ Kodu (c) Yeniden Oluşturulan NTÖ (d) NTÖ Hatası.....	53
Şekil 4.10.	Otokodlayıcı ile Toplu Anomali Tespitinde; Anomali Test Örneği için (a) ATÖ (b) ATÖ Kodu (c) Yeniden Oluşturulan ATÖ (d) ATÖ Hatası.....	54
Şekil 4.11.	Otokodlayıcının Doğruluk Yüzdesinin Pencere Boyutu ile Değişimi.....	58
Şekil 4.12.	Otokodlayıcının Geri Çağırma Yüzdesinin Pencere Boyutu ile Değişimi.....	58
Şekil 4.13.	Otokodlayıcının Kesinlik Yüzdesinin Pencere Boyutu ile Değişimi.....	59
Şekil 4.14.	İzolasyon Ormanının Doğruluk Yüzdesinin Pencere Boyutu ile Değişimi.....	60

Şekil 4.15. İzolasyon Ormanının Geri Çağırma Yüzdesinin Pencere Boyutu ile Değişimi..... 60
Şekil 4.16. İzolasyon Ormanının Kesinlik Yüzdesinin Pencere Boyutu ile Değişimi..... 61

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 2.1. Biyolojik Sinir Ağları Terimlerine Karşılık Gelen Yapay Sinir Ağları Terimleri 23
Çizelge 3.1. Hata Matrisi..... 41
Çizelge 4.1. Otokodlayıcı için Hata Matrisi..... 48
Çizelge 4.2. Otokodlayıcı Başarım Sonuçları..... 49
Çizelge 4.3. 3-SgOut için Hata Matrisi 49
Çizelge 4.4. 3-SgOut Başarım Sonuçları 49
Çizelge 4.5. Sıcaklık Sensörü Zaman Serilerinde Yer Alan Toplu Anomaliler 51
Çizelge 4.6. Otokodlayıcı için Hata Matrisi Çizelgesi..... 54
Çizelge 4.7. Otokodlayıcı için Başarım Sonuçları 55
Çizelge 4.8. İzolasyon Ormanı için Hata Matrisi Çizelgesi 55
Çizelge 4.9. İzolasyon Ormanı için Başarım Sonuçları 55
Çizelge 4.10. Otokodlayıcı ve 3-SgOut için Başarım Sonuçları 56
Çizelge 4.11. Otokodlayıcı ve İzolasyon Ormanı için Başarım Sonuçları..... 57

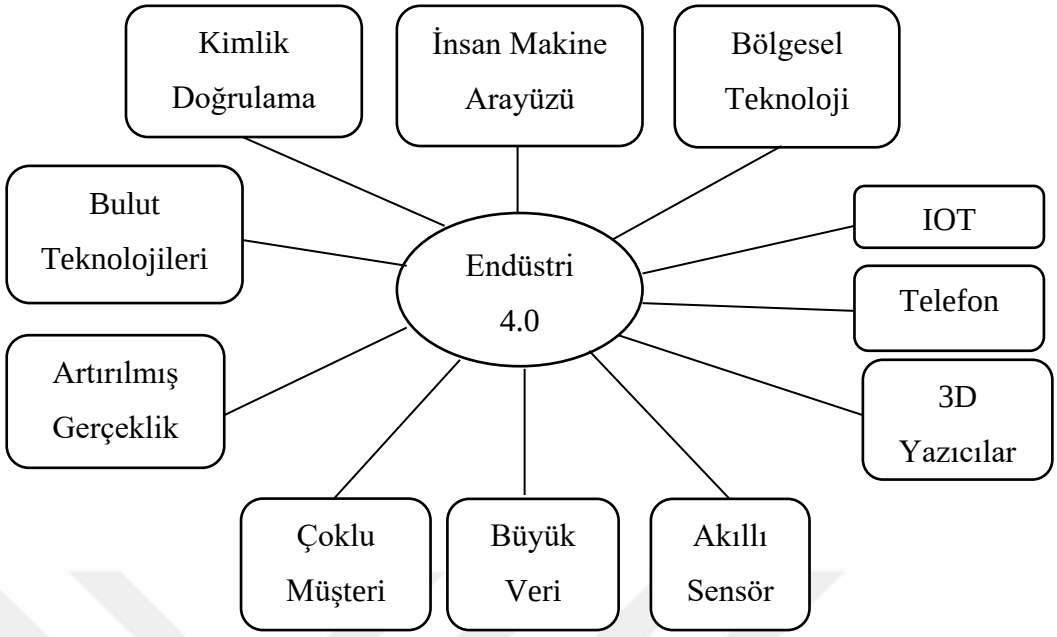
1. GİRİŞ

1.1. Endüstriyel Sistemler

Endüstriyel sistemler sürekli bir biçimde değişmekte ve gelişmektedirler. Özellikle 18. yüzyılın sonlarında, 1970 yılından sonra, sanayi devrimi ile başlayan gelişimini hızlanarak sürdürmüşlerdir [(Kagermann vd., 2011), (Kılıç., 2023)].

Endüstri devriminden sonra 1980'li yıllarda beyaz yakalı çalışan kavramı gelişirken, Endüstri 4.0'ın hayatımıza girmesiyle birlikte altın yakalı çalışan kavramı oluşmaya ve gelişmeye başlamıştır. Altın yakalı çalışan kavramı teknolojik araçları yetkin bir biçimde kullanabilen insan kaynağını tanımlamaktadır. Bu araçlar, robotik teknolojiler, artırılmış gerçeklik, nesnelerin interneti, siber güvenlik, üç boyutlu yazıcı teknolojileri, büyük veri analizleri, siber güvenlik sistemleri vb. teknoloji alanlarında kullanılan araçlar olarak geniş bir pencerede karşımıza çıkmaktadır [(Tüsiad, 2016), (Kılıç., 2023)].

Endüstri 4.0 ile başlayan teknolojik gelişim yapay zekâ ile daha da hız kazanmıştır. Günümüz endüstriyel sistemleri akıllı sistemler ve bilgi teknolojileri içeren pek çok alt bileşene sahiptir. Endüstri 4.0 için özet bir bileşen şeması Şekil 1.1.'de sunulmuştur [(Ötleş ve Özyurt, 2016), (Kılıç., 2023)].



Şekil 1.1. Endüstri 4.0 Bileşenleri [(Ötleş ve Özyurt, 2016), (Kılıç., 2023)]

Endüstriyel sistemlerde oluşan büyük veri bu sistemler için kontrol ve izleme araçlarının geliştirilmesi süreçlerinde önemli zorluklar oluşturmaktadır. Son yıllarda, büyük verinin analizi gelişen teknolojiye paralel olarak yapay zekâ yöntemleri ile gerçekleştirilmektedir. Endüstriyel kontrol sistemlerinin görevleri arasında sistemin sağlık durumunun izlenmesi önemli bir yer tutmaktadır. Sistem sağlık durumunun izlenmesi için kullanılan yaklaşımlar arasında anomali tespiti temelli yöntemler yoğun ilgi çekmektedir. Sistem bileşenlerinde oluşan anomaliler tüm sistemi etkileyebilmekte, zamanla sistemin başarımını düşürebilmekte ve hatta arızalanmasına sebep olabilmektedirler. Arızaların oluşma nedeninin, zamanın ve süresinin tespit edilmesi zaman alıcı ve zor olabilmektedir. Bu zorlukların aşılmasında yapay zekâ temelli yöntemlerin kullanımı önemli bir avantaj sağlamaktadır.

1.2. Anomali Tespiti

Endüstriyel sistemlerde anomali tespiti literatürde yoğun ilgi çeken önemli bir çalışma alanıdır. Anomaliler sistemlerin başarımlarını düşürebilmekte ve belirli bir zaman sonra ise sistemlerde arızaların oluşmasına sebep olabilmektedirler. Anomali tespitinde yapay zekâ temelli yöntemler başarı ile uygulanmaktadır. Literatürde, farklı sektör ve uygulama alanları için gerçekleştirilmiş çok sayıda anomali tespiti çalışması yer almaktadır,

örneğin; [(Bayraktar ve Gökçen, 2020), (Zhao vd., 2014), (Ferreira vd., 2007), (Das vd., 2020), (Himeur vd., (2021), (Theumer vd., (2021), (Lu vd., (2017), (Hiruta vd., 2021), (Hızarcı vd., 2021), (Jami ve Heyns, 2018), (Takiddin vd., 2022), (Givnan vd., 2022), (Homayouni vd., 2020), (Kalıpcıoğlu vd., 2019), (Zavrak, 2020), (Seitz vd., 2023), (Bingchen vd., 2023), (Yuying vd., 2023), (Mahsan vd., 2023), (Armin Masoumian vd., 2023), (Brummer ve Vleeschouwer, 2023), (Antonenko ve Read 2023), (Martino vd., 2022), (Jamjoom vd., 2023), (Ye ve Bors, 2023), (Lopes ve Pedronette 2023), (Tavares vd., 2023), (Lambrecht, 2023), (Volya ve Mishra, 2023), (Rahman vd., 2023), (Sauvalle ve Fortelle, 2023), (Hassanaly vd., 2023), (Zhu vd., 2023), (Li vd., 2019), (Maasoum vd., 2020), (Roy vd., 2018), (Ammann vd., 2020), (Ahmad vd., 2020), (Borghesi vd., 2019), (Meire ve Karsmakers, 2019), (Ferreira vd., 2019), (Tolpin, 2019), (Zhang vd., 2023), (Kirchgässner vd., 2020)].

Anomali tespit yöntemleri belirli bir tür için geliştirilmiş olabileceği gibi genel amaçlı yöntemlerde bulunmaktadır. Sahip olduğu avantaj ve dezavantajlar dikkate alınarak anomali tespit yönteminin seçimi problemin başarıyla çözülmesinde ilk adımı oluşturmaktadır (Chandola vd., 2009).

Anomali, bir sistemde bulunan aykırı değerler, sapmalar, farklılıklar, uyumlu olmayan değerler olarak ifade edilebilir. Anomali tespiti, siber güvenlik ve dolandırıcılık önlemeden, sağlık hizmetleri ve öngörücü bakım uygulamalarına uzanan çok geniş bir yelpazede karşımıza çıkmaktadır. Anomali tespiti bu uygulamalar için hayati önem taşımaktadır. Bir dolandırıcılık saldırısının erken tespit edilmesi, sağlık hizmetlerinde ölümcül hastalıkların ya da acil durumların erken teşhis edilmesi, siber güvenlik saldırılarının önüne geçilmesi ve endüstriyel sistemlerde bileşen arızalarının önlenmesi önemli örnekler olarak verilebilir (Chandola vd., 2009).

1.3. Tezin Amaç ve Kapsamı

Bu tez çalışmasında endüstriyel sistemlerin bileşenlerinde oluşan anomalilerinin tespitinde otokodlayıcıların başarım analizlerinin gerçekleştirilmesi amaçlanmıştır. Bu kapsamda endüstriyel sistemlerin önemli bileşenlerinden olan ve bu sistemlerde sıklıkla

yer verilen DC motorlardaki anomalilerin tespiti üzerine yoğunlaşılmıştır. DC motor üzerinde yer alan ve motorun sağlık izlemesinde temel gösterge olarak kullanılan gövde sıcaklığındaki anomalilerin tespiti otokodlayıcılar ile gerçekleştirilmiştir. Gövde sıcaklığı verisinde noktasal ve toplu anomali türleri için otokodlayıcıların başarım analizleri yapılmıştır. Başarım analizleri yapılırken karşılaştırma yöntemi olarak noktasal anomaliler için 3-SgOut Aykırı Değer yöntemi kullanılırken toplu anomaliler için ise İzolasyon Ormanı yöntemi kullanılmıştır.

Bu tez çalışması 118C157 numaralı TÜBİTAK BİDEB 2244-Sanayi Doktora Programı'nın "Endüstriyel Gerçek Zamanlı Akıllı Kontrol Sistemlerinin Geliştirilmesi" projesi kapsamında yürütülmüştür.

1.4. Tezin Katkısı

Doktora tezi kapsamında gerçekleştirilen çalışmaların ve bu çalışmaların sonuçlarını içeren yayınların katkısı iki ana başlık altında özetlenebilir.

- Deney Düzeneği ve Veri Seti

  Doktora tezi kapsamında DC motor gövde sıcaklık verisini izlemek, anomali oluşturmak ve veri kaydını gerçekleştirmek amacıyla özel bir deney düzeneği geliştirilmiştir. Bu deney düzeneği ile otokodlayıcıların başarım analizlerinde kullanılan veri seti elde edilmiştir. Oluşturulan veri setinin planlanan yayınlar tamamlandıktan sonra açık bir veri seti ambarına bağışlanması hedeflenmektedir.

- Endüstriyel Sistem Bileşenlerinin Anomali Tespitinde Otokodlayıcıların Uygulanabilirliği

  Yapılan deneyler sonucunda Otokodlayıcı mimarilerin DC motor gövde sıcaklığı verisinde yer alan noktasal ve toplu anomalilerin tespitinde başarıyla kullanılabileceği gösterilmiştir (Demircioğlu ve Yılmaz 2023). Bununla birlikte, tez çalışmasında kullanılan Otokodlayıcı mimarisinin endüstriyel sistemlerin

diğer bileşenlerinde oluşabilecek anomalilerin tespitinde de başarıyla uygulanabileceği sonucu literatürde paylaşılmıştır.

1.5. Tezin Organizasyonu

Doktora tezinin 1. Bölüm sonrasına ait organizasyonu aşağıda verildiği biçimde gerçekleştirilmiştir.

2. Bölümde; literatür taraması gerçekleştirilmiştir. Bununla birlikte, yapay zekâ, makine öğrenmesi, derin öğrenme, yapay sinir ağları ile anomali tipleri hakkında temel bilgiler sunulmuştur.

3. Bölümde; Tez çalışmasında kullanılan DC Motor deney düzeneği ile anomali veri seti detaylandırılmıştır. Ayrıca, Otokodlayıcı, 3-Sigma Aykırı Değeri ile İzolasyon Ormanı yöntemlerinin temelleri anlatılmıştır.

4. Bölümde; Otokodlayıcı, 3-SgOut Aykırı Değer ve İzolasyon Ormanı ile yapılan noktasal ve toplu anomalilerin tespitine ait sonuçlar sunulmuştur.

5. Bölümde; Noktasal ve toplu anomaliler için Otokodlayıcı ve karşılaştırma yöntemleri olarak seçilen 3-SgOut Aykırı Değer ve İzolasyon Ormanı yöntemlerinin başarım sonuçları karşılaştırılmıştır.

2. KURAMSAL TEMELLER ve KAYNAK ARAŞTIRMASI

Literatürde, anomali tespitini hedefleyen çok sayıda ve farklı alanlarda çalışmalar yer almaktadır. Bu çalışmalar arasında makine öğrenmesi temelli çalışmalar öne çıkmaktadır [(Bayraktar ve Gökçen, 2020), (Zhao vd., 2014), (Ferreira vd., 2007), (Das vd., 2020), (Himeur vd., (2021), (Theumer vd., (2021), (Lu vd., (2017), (Hiruta vd., 2021), (Hızarcı vd., 2021), (Jami ve Heyns, 2018), (Takiddin vd., 2022), (Givnan vd., 2022), (Homayouni vd., 2020), (Kalıpcıoğlu vd., 2019), (Zavrak, 2020), (Seitz vd., 2023), (Bingchen vd., 2023), (Yuying vd., 2023), (Mahsan vd., 2023), (Armin Masoumian vd., 2023), (Brummer ve Vleeschouwer, 2023), (Antonenko ve Read 2023), (Martino vd., 2022), (Jamjoom vd., 2023), (Ye ve Bors, 2023), (Lopes ve Pedronette 2023), (Tavares vd., 2023), (Lambrecht, 2023), (Volya ve Mishra, 2023), (Rahman vd., 2023), (Sauvalle ve Fortelle, 2023), (Hassanaly vd., 2023), (Zhu vd., 2023), (Li vd., 2019), (Maasoum vd., 2020), (Roy vd., 2018), (Ammann vd., 2020), (Ahmad vd., 2020), (Borghesi vd., 2019), (Meire ve Karsmakers, 2019), (Ferreira vd., 2019), (Tolpin, 2019), (Zhang vd., 2023), (Kirchgässner vd., 2020)].

Bayraktar ve Gökçen (2020), enerji optimizasyonu hakkında farklı anomali tespit algoritmaları üzerine çalışmışlardır. Çalışmalarında deney seti olarak yüksek raflı bir konveyör sistemi kullanmışlar ve bu deney setinden elde edilen veriler üzerine çalışmalarını gerçekleştirmişlerdir. Naive Bayes, Yapay Sinir Ağları, C4.5 Karar Ağacı, Rastgele Orman, k En Yakın Komşu, Lojistik Regresyon, Destek Vektör Makineleri yöntemlerinin başarımlarını karşılaştırmak için kullanmışlardır. Başarım değerlendirmesini karmaşıklık matrisi sonuçlarına göre; model doğruluğu, model kesinliği, modelin duyarlılığı, F ölçütü, hata oranı ölçütleri bakımından karşılaştırmışlardır. Yapay Sinir Ağları, C4.5 Karar Ağacı, Rastgele Orman, k En Yakın Komşu yöntemlerinin başarımları diğer yöntemlere göre daha yüksek çıkmıştır. Bu yöntemlerin içerisinden de Rastgele Orman algoritması ile oluşturulan modelin anomalilerin tespitindeki başarımı %98 oranında bulmuşlardır. Yaptıkları çalışmalarının sonucunda sistemdeki anomalileri büyük bir başarımla tespitini gerçekleştirmişler ve sistemdeki enerji optimizasyon başarımını artırdığı sonucuna varmışlardır.

Zhao ve diğerleri (2014), çelik endüstrisinde enerji sistemlerindeki güvenilirliği ve verimliliği artırmak için farklı anomali tespiti yöntemleri üzerine çalışmışlardır. Çelik endüstrisinde bulunan enerji sisteminin sahip olduğu verilerin özelliklerine göre anomalileri sözde periyodik veriler için trend anomalileri ve genel veriler için sapmalar olmak üzere ikiye ayırmışlardır. Anomali tespiti için kullandıkları yöntemler trend anomaliler için; AFCM (Adaptive Fuzzy C Means) tabanlı yöntem, DTW tabanlı AFCM, SWM tabanlı yöntemler, sapmalar için ise KNN-AFCM tabanlı yöntem SWM tabanlı yöntem, AFCM tabanlı yöntemleri kullanmışlardır. Anomali tespiti hesaplama süreleri, anomali index değerleri, boru hattı basıncı gibi başarım karşılaştırma parametreleri kullanmışlardır. AFCM tabanlı yöntemler önerilmiş olup SWM tabanlı yöntemlerden iyi sonuçlar vermiştir. Başarım performansları için en iyi sonuç veren yöntemler trend anomali tespiti için DTW tabanlı AFCM yöntemi önerilmiş olup, sapmalar için ise KNN-AFCM yöntemi önerilmiştir.

Ferreira ve diğerleri (2007), kararlı durum besleme anomalilerinin 3 fazlı sincap kafesli asenkron motorlara etkisini incelemişler ve çalışmalarını bunun üzerine gerçekleştirmişlerdir. Motorlar üzerinde güç düşüşüne etkisi olan bölümler olan, gerilim dengesizliği, Gerilim bozulması, gerilim büyüklüğü sapması, kombine efektler, kutup numarasının motor güç azalmasındaki etkisi, rotor-bar tipinin motor güç azalmasındaki etkisi, çerçeve tipinin motor güç azalmasındaki etkisi üzerine başarım sonuçlarını çalışmalarında karşılaştırmışlardır. Karşılaştırma sonuçlarına göre çalışmalarında motorlarda oluşan güç anomalilerinin motorların üzerindeki etkisini değerlendirmek için çok girişli, ileri beslemeli, yapay sinir ağları tabanlı bir yaklaşım önermişlerdir.

Endüstriyel kontrol sistemlerinde yapılan siber saldırılar sistemin güvenliği ve verimi için çok önemlidir. Siber saldırıların önüne geçebilmek ve anomalileri gerçek zamanlı tespit edebilmek amacıyla Das ve diğerleri (2020), endüstriyel kontrol sistemlerinde oluşan anomaliler üzerine çalışmalarını yapmışlardır. Çalışmalarında Güvenli Su Arıtma Sistemi'nin test ortamında verilerini almışlar ve bu veriler üzerine çalışmalarını gerçekleştirmişlerdir. Kullandıkları verileri anomali olmayan kısmına kararlı, anomali olan kısmına kararsız yapı olarak ikiye sınıflandırmışlardır. Sınıflandırıcının performansları sınıflarına (kararlı ve kararsız) göre kesinlik, geri çağırma, F1 puanına

göre karşılaştırılmıştır. Çalışmaları sonucunda LAD (Logical Analysis of Data) tabanlı ADS (Anomaly Detection System) yaklaşımı ile anomalileri gerçek zamanlı tespit etmişlerdir. Önerdikleri bu çalışmanın LAD tabanlı olması sebebiyle ayrıca anomalilerin lokalize tespitini yapmışlardır.

Himeur ve diğerleri (2021), binaların enerji tüketimindeki anomali değerlerini yapay zekâ yöntemleri ile tespiti üzerine çalışmışlardır ve önerilen yaklaşımın faydalarından bahsetmişlerdir. Çalışmalarında; denetimli öğrenme, denetimsiz öğrenme, toplu öğrenme, özellik çıkarma, hibrit öğrenme ve görselleştirme, sıkıştırma algılama gibi diğer uygulama teknikleri kullanmışlardır. Enerji sistemlerindeki anomali tespitlerini yaptıkları çalışmalarında; yapay zekâ tabanlı tahmin, enerji mikro anlarının ve doluluk verilerinin yapay zekâ ile sınıflandırılması, enerji verilerinin tek sınıf sınıflandırılması olmak üzere üç aşamalı incelemişlerdir. Çalışmalarının sonucunda enerji sistemlerine uygulanan anomali tabanlı yaklaşımların enerji sistemlerinin üzerindeki yaklaşım sonuçları başarımlarının yüksek olduğunu sunmuşlardır.

Theumer ve diğerleri (2021), endüstriyel tesislerde enerji verimliliğini kontrol altında tutmak ve enerji verimliliğini artırmak için çalışmışlardır. Endüstriyel tesislerde sürekli ve otonom olarak izlenmesi ile farklı zaman serilerindeki anomali tespitini gerçekleştirmişlerdir. Nokta anomalileri ve toplu anomalileri önerdikleri yöntemle tespit edebilmişlerdir. Anomali türlerine göre yaklaşımları farklı seçilmiştir. Başarımları karşılaştırılırken anomali puanları hesaplanarak gerçekleştirilmiştir. Anomali puan hesaplaması anomali türlerine göre hesaplanmıştır. Tek değişkenli zaman serilerindeki anomali değerlerini tespit ettikleri Otokodlayıcı önermişlerdir. Bu önerilen yöntemin avantajları ve dezavantajlarını değerlendirmişlerdir.

Lu ve diğerleri (2017), insansız hava aracı motorlarında oluşan anomalilerini tespit etmek için makine öğrenmesi yöntemi olan takviyeli öğrenme üzerine çalışmışlardır. Çalışmalarının amacı İHA motorunun anormal sıcaklıklarda çalışmasını önleyecek bir anomali tespit sisteminin gerçekleştirilmesidir. Motor sıcaklığı otomatik oluşturulan bir eşiği aştığında drone'un güvenli bir iniş yapabilmesini sağlamışlardır. Anormal ses tespiti, kızıl ötesi görüntüleme tespiti, anormal rotasyon tespiti gibi farklı anomali tespit

yöntemleri kullanmışlardır. Yaptıkları çalışmada deney düzeneği için sıcaklık sensörleri ve Raspberry Pi 2B CPU ile donatılmış bir drone ve bir tablet kullanmışlardır. Çalışmalarında; drone'un düşme ve tırmanma durumları hız ve sıcaklık değişimleri ve yavaşlama oranlarını hesaplayarak başarım değerlendirmesini yapmışlardır. Önerdikleri Takviyeli öğrenmeli yöntem Drone'ların güvenli kullanımında yüksek başarım göstermiştir.

Hiruta ve diğerleri (2021), motor rulmanlarındaki anomalileri tespiti için makine öğrenmesi yöntemlerinden denetimsiz öğrenme tabanlı yöntem kullanmışlardır. Deney seti düzeneklerini kendileri oluşturmuşlardır. Veri seti olarak oluşturdukları düzenekten aldıkları verileri kullanmışlardır. Başarım karşılaştırmaları için önerdikleri denetimsiz öğrenme modeli ile geleneksel modeller (denetimli öğrenme modelleri) ile gerçekleştirmişlerdir. Çalışmalarının sonucunda yetersiz gresten kaynaklı oluşan motor yağı arıza modunu yakalamışlar ve yatak durumunda artan anomalileri göstermişlerdir.

Hızarcı ve diğerleri (2021), sonsuz dişli kutusunun çalışması sırasında oluşan anomali tespitinin üzerine çalışmışlardır. Çalışmalarında bir AC Motor, bir sonsuz dişli kutusu ve bir elektromanyetik frenden oluşturdukları bir deney düzeneğinin verilerini kullanmışlardır. Anomali başarımlarını ölçmek için; geleneksel ve multidisipliner yöntemleri uygulamış ve karşılaştırmışlardır. Çalışmalarının sonucunda Karar Ağacı tabanlı bir yöntem önermişlerdir. Önerdikleri bu yöntem sonucunda dişli kutularında meydana gelen farklı tipteki anomalileri büyük bir başarım ile tespit etmişlerdir.

Jami ve Heyns (2018), değişken akış koşularında pervane arıza tespitini gerçekleştirmişlerdir. Deneysel çalışmalarında pervane titreşimlerini ölçmek için kullanabilecekleri deney düzeneği geliştirmişlerdir. Deneysel çalışmalarında yedi pervane koşulunu incelemişlerdir. YSA yapılarını eğitirken; TDA, FA, TFA özellik çıkarımları ile üç ayrı yapay sinir ağı yapısı geliştirmişlerdir. Başarım değerlendirmesi için hata matrisi tablolarını karşılaştırma olarak kullanmışlardır. Hata matrisi tabloları ise TDA özellikleri kullanılarak eğitilmiş YSA yapısı, FA özellikleri kullanılarak eğitilmiş YSA yapısı, WPT özellikleri kullanılarak eğitilmiş bir YSA yapısında oluşturmuşlardır. Ayrıca başarım değerlendirmelerinde eğitim algoritmaları, farklı transfer

fonksiyonlarının etkinliğini ve ağdaki gizli düğümleri karşılaştırmışlardır. Çalışmalarının sonucunda ayrıştırılmış titreşim sinyalleri ve enerji tabanlı özellikler kullanılarak geliştirilen yapay sinir ağı modelinin tahminini yüksek doğrulukta olduğunu göstermişlerdir. WPT-ANN tabanlı önerilen modelin WPT' den ayrılan titreşiminin enerji kullanılarak önemli ölçüde artırıldığını kanıtlamışlardır. Ayrıca yapay sinir ağları ile kararsız akış hızları altında pervane arıza tespit tekniklerini karşılaştırmış ve pervane kusurlarını sınıflandırmak için etkin bir şekilde kullanılabileceğini belirtmişlerdir.

Takiddin ve diğerleri (2022), çalışmasında, akıllı şebekelerdeki elektrik hırsızlığı siber saldırıları için otokodlayıcı tabanlı anomali tespiti üzerine çalışmışlardır. Çalışmalarında akıllı şebekelerde oluşan anomali verileri tespit edebilmek için siber saldırı dedektörü tasarlamayı amaçlamışlardır. İlk olarak seq2seq (sequence-to-sequence) temelli bir LSTM (long-short-term-memory) yapısı oluşturmuşlardır. Oluşturdukları seq2seq yapısı temelinde; basit otokodlayıcı modeli, varyasyonel otokodlayıcı modeli ve AEA (autoencoder with attention) otokodlayıcı modellerine ait performansları karşılaştırmışlardır. Çalışmaların da kullandıkları veri setlerinden ilki; Çin Devlet Şebeke Şirketi (SGCC)'den gerçek müşterilerin enerji tüketimindeki etiketlenen veri setlerinden faydalanmışlardır, ikinci veri seti olarak ise İrlanda Akıllı Enerji Denemesi (ISET) tarafından sağlanan gerçek enerji tüketim verileri ve yapay oluşturulmuş sistem siber saldırı verileri kullanılmıştır. Çalışmalarının simülasyon başarım değerlendirmelerine göre AEA dedektörü ile algılama oranında %17'lik bir artış ve yanlış alarm oranı %9'luk artarak iyileştirme gerçekleştirmişlerdir. Çalışmalarının sonucunda oluşturdukları farklı yapılardaki otokodlayıcı modellerinin içerisinden; AEA (autoencoder with attention) dedektörünü mevcut son teknoloji dedektörlere kıyasla daha fazla başarım gösterdikleri için önermişlerdir.

Givnan ve diğerleri (2022), endüstriyel motorlarda makine öğrenmesi yaklaşımı ile otomatik kodlayıcı kullanarak, anomalileri tespit etmişler ve makine arızası öncesi uyarı vermesini sağlamışlardır.

Homayouni ve diğerleri (2020), çok değişkenli zaman serisi verilerindeki anomalilikleri tespit etmek ve bu anomalilikleri raporlamak için LSTM-Otokodlayıcı tabanlı bir yaklaşım üzerine model geliştirmişlerdir. Oluşturdukları model için Yahoo sunucularının, NASA Shuttle ve Colorado Eyalet Üniversitesi Enerji Enstitüsü'nden alınan veri setlerini kullanmışlardır. Başarım değerlendirmesi için mevcut stokastik ve ML tabanlı anomali tespit teknikleri ile karşılaştırmışlardır. Çok değişkenli zaman serilerinde oluşan anomalileri tespit edebilen LSTM tabanlı Otokodlayıcı modeli olan IDEAL'i (Interactive Detection and Explanation of Anomalies) önermişlerdir. Ayrıca önerdikleri model yapay sinir ağının girişi için otomatikleştirilmiş korelasyona dayalı bir pencereleme yaklaşımını da içermektedir.

Kalıpcıoğlu ve diğerleri (2019), çalışmalarında son kullanıcıların sistemlerine gelen siber saldırılar (anomaliler) için geliştirilmiş farklı anomali saldırı tespit sistemlerinin karşılaştırmaları üzerine çalışmışlardır. Siber saldırı tespit sistemlerinde kullanılan enformasyon teorisi yöntemi, en yakın komşu yöntemi, sınıflandırma yöntemi, kümeleme yöntemi ve istatistiksel anomali tespit yöntemlerini karşılaştırmışlardır.

Zavrak (2020), yazılım tanımlı ağlardaki akış özellikleri kullanarak akıştaki bilinmeyen saldırıları tanımlamak için birbirinden farklı otomatik kodlayıcı yöntemleri kullanmıştır. Yarı denetim tabanlı öğrenme ile AE (otokodlayıcı), VAE (Varyasyonel Otokodlayıcı) ve OCSVM (Tek Sınıflı Destek Vektör Makinesi) yöntemlerinin başarım sonuçlarını AUC ölçütleri, ROC eğrileri bakımından karşılaştırmıştır. Sonuçlarda VAE'nin anomalileri tespit etme oranı AE ve OCSVM'den daha yüksek olduğunu sunmuştur. Deneysel sonuçlarda LSTM ağlarının özel bir çeşidi olan EncDecAD, literatürdeki yöntemlerden daha iyi başarım sergilediği için önermiştir. Ayrıca çalışmasında SAnDet (Yazılım Tanımlı Ağlar için Anomali Tespit Sistemi) uygulaması sistemlerdeki saldırı olan anomalileri tespit edip ve olacak saldırıları engelleyebildiğini göstermiştir.

Otokodlayıcılar, sistemlerde bulunan anomalileri ve farklılıkları tespit ettiği gibi benzer şekilde malzemelerin yapısal özelliklerine uygun seçim işlemlerini de yapabilmektedir. Anomalileri ayırt ederken kullanılan çok aşamalı yapı kullanılarak yapısal özellikler tespit edilip malzeme özelliklerini iyileştirme ve geliştirmede de kullanılabilmektedir. Bu

konu üzerine Seitz ve diğerleri (2021), çalışma gerçekleştirmişlerdir. Yaptıkları çalışmada pil anotları için SiO_2 matris malzemelerinin yapısal özellik analizlerini incelemişlerdir. Çalışmalarında geliştirdikleri otokodlayıcı karmaşık sentez sürecini standart makine öğrenimi yöntemleriyle uyarlamak için malzeme geliştirmede yüksek başarımla kullanılabileceğini göstermişlerdir.

Güç sistemlerinde yığılmış gürültü güç sistemlerinin güvenilirliği açısından önemlidir. Bingchen ve diğerleri (2023), yığılmış gürültü giderici üzerine (anomalileri tespit edip azaltan), güç sistemlerindeki optimal güç akışı (OPF) algoritmasının yerini alacak, Yığılmış Gürültü Giderici Otokodlayıcılı (SDAE) bir yöntem üzerine çalışmışlardır. Güç sistemlerinde güvenilirlik değerlendirmesi durum seçimi, durum değerlendirmesi ve endeks hesaplaması olmak üzere farklı şekillerde yapmışlardır. Önerilen yöntemin güç sistemi işletimi için güvenilirlik değerlendirmesinin etkinliğini büyük ölçüde artırabileceğini göstermişlerdir.

Yuying ve diğerleri (2023), yüksek başarım kazanımları olan çok ölçekli evrişimli bir Otokodlayıcı MrCAE (multi-resolution convolutional autoencoder) yöntemi üzerine çalışmışlardır. Çalışmalarında çoklu ızgara yöntemleri, evrişimli otomatik kodlayıcılar ve transfer öğrenimi olmak üzere üç aşamalı matematiksel yapıyı kullanmışlardır.

Hailan ve diğerleri (2023), çalışmalarında sıkıştırma oranının üst sınırını belirleyip maksimum sıkıştırma oranı için belirli bir kuantum kodlayıcıyı öğrenme kontrol yaklaşımı ile eğitmişlerdir. Ulaştıkları optimum sıkıştırma oranı sayısal sonuçlar ile tutarlı çıkmıştır.

Mahsan ve diğerleri (2023), geliştirdikleri kredi puanlama modeli için torbalama denetimli Otokodlayıcı sınıflandırıcı (BSAC) öğrenme modelini sunmuşlardır. Kredi başvurularının sonucunu tahmin etmede önerilen yöntemin üstünlüğünü ve sağlamlığını deneylerle doğrulamışlardır.

Armin Masoumian ve diğerleri (2023), çalışmalarında monoküler derinlik tahmini için çok ölçekli bir grafik evrişim ağına (GCN) dayalı kendi kendini denetleyen bir derin

öğrenme modeli geliştirmişlerdir. Önerdikleri yöntem ile kullanılan veri kümeleri için monoküler video tekli görüntüden mevcut yöntemlerle karşılaştırılabilir bir derinlik tahmini elde etmişlerdir.

Brummer ve Vleeschouwer (2023), fotoğrafçılıkta kullanılan görüntü gürültüsünü otokodlayıcıyla oluşturdukları model ile başarımı yüksek bir şekilde gidermişlerdir. Çalışmalarının sonucunda ise görüntü gürültüsü için oluşturulan otokodlayıcı modeli sunulmuştur.

Antonenko ve Read (2023), kayıp değer atamasını çok etiketli bir sınıflandırma problemi olarak ele almışlar ve otokodlayıcıları gürültüden arındırmak için daha basit bir alternatif olan otoreplikatif (kendi kendini kopyalayan) rastgele ormanları önermişlerdir. Deneylerde, algoritma yalnızca veri kümesinden gelen bilgilere dayanarak eksik değerleri etkili bir şekilde yüklemiş ve herhangi bir ek bilgi gerektirmeyen standart algoritmalardan daha iyi başarım göstermiştir.

Martino ve diğerleri (2022), tarımsal ürün türlerinde meydana gelen hataları ölçüp tahmin ederek veri kümesindeki olası hata oranı ve düzeltme önerileri sunabilen bir otokodlayıcı metodu önermişlerdir. Bu metodu denetimli algoritmalarla karşılaştırdıklarında yeniden etiketleme kalitesinin iyileştiğini görmüşlerdir.

Jamjoom ve diğerleri (2023), çalışmalarında göğüs tomografisi taramalarına dayalı pnömoni enfeksiyonunun teşhisi için yeni bir hibrit derin öğrenme modeli kullanmışlardır. Önerilen gürültü giderme hibriti otomatik kodlayıcı modelde aynı veri kümesine uygulanan diğer modellere göre en yüksek doğruluk değerini elde etmişlerdir.

Ye ve Bors (2023), sürekli öğrenmede kullanıldığında unutma sürecini yorumlayan bir değişken otokodlayıcı (VAE) modelini etkinleştirmek üzere genişletme stratejisi önermişlerdir. Önerdikleri dinamik genişletme modeli (DEM), sürekli öğrenme uygulamalarında diğer modellerden daha iyi başarım göstermiştir.

Lopes ve Pedronette (2023), çalışmalarında farklı öğrenme yöntemlerinin (kendinden denetimli, yarı denetimli, denetimsiz gibi) birleşiminden kendinden denetimli Grafik Evrişimli Sınıflandırma (SGCC) yöntemi sunmuşlardır. Önerdikleri SGCC yöntemini görüntü ve alıntı ağları veri kümelerinde değerlendirip klasik ve yeni kümeleme yöntemleriyle karşılaştırmış tüm senaryolarda yüksek etkili sonuçlar elde etmişlerdir.

Enerji endüstrisindeki Rüzgâr türbinlerindeki oluşan arızalar çok önemlidir. Bu nedenle Rüzgâr türbinlerindeki anomalileri algılamak ve tespitini doğru bir şekilde gerçekleştirmek için Tavares ve diğerleri (2023), frekans yanıt fonksiyonlarını ya da çapraz güç spektral yoğunluklarını sınıflandıran otokodlayıcılı bir algoritma sunmuşlardır. Anomali Algılayan Otokodlayıcı (ADAE) büyüklüğü ve konumu değişen her tür hasar senaryosunu doğru bir şekilde sınıflandırmayı başarması nedeniyle, Çok Değişkenli Gauss Anomali Tespiti (MGAD) algoritmasından daha yüksek hasar tespit doğruluğu elde etmişlerdir.

Lambrecht (2023), piksel izleyicideki bir örnek olay incelemesi üzerine CMS detektörün durumunu ve başarımını özetlemek için üst düzey fizik analizlerinde sonuçların doğruluğunu sağlamada kilit bileşen olarak kullanılan bilgisayar destekli Veri Kalitesi İzleme (DQM) ve Veri Sertifikalandırma (DC) yöntemlerini araştırmışlardır. Otomatik kodlayıcı tekniklerinin, daha önce fazla insan gücü gerektiren manuel sertifika prosedürü için anomali dedektör davranışını doğru bir şekilde tespit edebildiğini göstermişlerdir.

Volya ve Mishra (2023), kuantum verilerinin üzerine yaptığı çalışmada verimliliği artırmak, kuantum kontrolünün hesap sürelerini kısaltmaya yönelik kuantum kayıtlarındaki verileri sıkıştırmak için bir otokodlayıcı modeli önermişlerdir.

Rahman ve diğerleri (2023), eş zamanlı öğrenme (aynı anda birden fazla alanı eğitme) ile ilgili yarı denetimli etki alanı uyarlaması üzerine bir otokodlayıcı yöntemi üzerine çalışmışlardır. Çalışmalarında, kaynak alan ve hedef alan için bir otomatik kodlayıcı ve sınıflandırıcı olmak üzere üç bileşen kullanmışlardır. Otokodlayıcıların eğitimi ve yarı denetimli sınıflandırıcının eğitimi olmak üzere iki aşamalı eğitim modeli kullanmışlardır. Eşzamanlı öğrenmeli otokodlayıcı modeli için deneysel çalışmalarında DomainNet veri

seti, Office-31 veri seti ve Digit Recognition veri seti olmak üzere üç farklı veri seti kullanmışlardır. Çalışmalarının başarım değerlendirmesinde dört alan altında (R: Real, C: Clipart, P: Painting, S: Sketch) hedeflenen doğruluk oranına göre önerdikleri yöntemi, diğer son teknoloji yöntemler olarak karşılaştırmışlardır. Yaptıkları çalışma sonucunda önerdikleri yöntemin başarım değerinin daha iyi olduğu sonucuna ulaşmışlardır.

Hareketli kameralar ile yapılan video kayıtlarında arka planda istenmeyen gürültülerin (anomalilerin) oluşması videoların kalitesini düşürmektedir. Sauvalle ve Fortelle (2023), hareketli video çekimlerinde arka plandaki gürültüleri yakalayıp tekrar yapılandırabilen denetimsiz öğrenme ile modellenen bir otokodlayıcı üzerine çalışmışlardır. Uygulamalarında video görüntüleri için CDnet 2014, LASIESTA ve BMC 2012 veri setlerini kullanmışlar video olmayan görüntüler için ise Clevrtex, ObjectsRoom, ShapeStacks veri setlerini kullanmışlardır. Başarım değerlendirmesinde kayıp fonksiyon tasarımı ve arka plan gürültüsü tahmin katmanının üzerindeki etkilerini değerlendirmişlerdir. Çalışmalarında başarım sonuçları yüksek çıkan arka plan gürültülerini tahmin edebilen bir otokodlayıcı modeli önermişlerdir.

Hassanaly ve diğerleri (2023), yaptıkları çalışmalarında Alzheimer hastaları için FDG PET görüntülerindeki anomalilerin tespiti için 3D değişken otokodlayıcıya (3D VAE) dayalı bir yaklaşım üzerine çalışmışlardır. Çalışmaların da kullanılan veri seti ADNI veri tabanından FDG PET taramalarından elde etmişlerdir. Çalışmalarında öncelikle anomali içermeyen görüntülerin eğitilmesinden sonra, anomali içeren görüntülerle model eğitimlerini yapmışlardır. Bu çalışmalarında farklı derecelerde anomalileri ve farklı demans türlerini farklı koşullarda bir UAD yaklaşımını değerlendirmek için simüle etmişlerdir. Başarım karşılaştırması görüntülerini yeniden yapılandırma ölçümlerine (MSE, MAE, PSNR, SSIM) göre yapmışlardır. Başarım değerlendirmesi bakımından 3D VAE yönteminin FDG PET görüntülerinde anomalileri tespit etme yöntemi olarak önermişlerdir.

Zhu ve diğerleri (2023), yapmış oldukları çalışmalarında denetimsiz alanlarda uyarlamayı (denetimli alandan denetimsiz alana bilgi aktarılması) gerçekleştiren bir yığılmış evrişimli otokodlayıcı modeli geliştirmişlerdir.

Mekanik ekipmanlarda çalışma şartlarına göre farklı anomaliler meydana gelebilmektedir. Li ve diğerleri (2019), derin öğrenme tabanlı otokodlayıcı temelli bir yaklaşım üzerine yapmış oldukları çalışmalarını, mekanik ekipmanlardaki anomalileri tespit etmek için önermişlerdir.

Maasoum ve diğerleri (2020), döner makinalarda arıza tespitini gerçekleştiren, bir otokodlayıcı modeli üzerine çalışmışlardır. Çalışmalarında kullandıkları veri seti için elektromekanik bir deney düzeneği kurmuşlardır. Başarım değerlendirmesi için hata matrisi tablosu kullanmışlardır. Çalışmalarında döner makinelerin çevrimiçi anomali tespitini gerçekleştiren otokodlayıcı modeli önermişlerdir.

Roy ve diğerleri (2018), yapmış oldukları çalışmada endüstriyel sistemlerin proseslerinde bulunan ana parçalarında meydana gelebilecek arızaları önceden tahmin edebilecek bir otokodlayıcı modeli önermişlerdir.

Ammann ve diğerleri (2020), yapmış oldukları çalışmada helikopterlerdeki ivme ölçerler tarafından kaydedilen zaman serisi verilerinde bulunan anomalilikleri tespit etmek için karma mimari ile oluşturulan bir otokodlayıcı modeli önermişlerdir.

Elektrik motorlarının rotorları ve yataklarındaki istenmeyen anomali durumlar (arızalar) sistemde tamamen sorun meydana getirebilmekte ve sistemi tamamen kapatıp sorun yaratabilmektedir. Bu soruna yönelik Ahmad ve diğerleri (2020), çalışmışlardır. Çalışmalarında iki veri seti kullanmışlardır. Başarım karşılaştırmasında izolasyon ormanı yöntemini ve LSTM yapısını da çalışmalarında kullanmışlardır. Çok aşamalı bir anomali tespiti gerçekleştirmişlerdir. İlk aşamada Uzun Kısa Süreli Bellek Otokodlayıcı (LSTM-AE) modeli kullanılarak otomatik özellik çıkarımı yapmışlardır. Sonraki aşamada İzolasyon Ormanı (IF) anomali tespit modelini eğitmişlerdir. Önerilen yöntemin başarım sonucu olarak ortalama F1 puanını %99.6 bulmuşlardır. Sonuçlarda LSTM-Otokodlayıcı modelinin eğitiminde sadece normal sinyaller olmasına rağmen döner bir makinede anomali olup olmadığını tespit etmesini ve ayrıca model tarafından öğrenilen bilginin farklı döner makinelerdeki benzer anomalileri de tespit edebildiğini sunmuşlardır.

Çalışmalarında dönen makinalarda oluşan anomalilerin durumlarını izleyen LSTM tabanlı otokodlayıcı modelini önermişlerdir.

Yüksek başarımlı bilgi işlem (HPC) sistemlerinde anomali tespiti önemlidir. Borghesi ve diğerleri (2019), yüksek başarımlı bilgi işlem (HPC) sistemlerinde anomali tespitini gerçekleştirecek yarı denetimli bir otokodlayıcı modeli sunmuşlardır.

Meire ve Karsmakers (2019), akustik sinyallerde denetimsiz anomali tespitini gerçekleştiren birbirinden farklı otokodlayıcı modellerini karşılaştıran bir çalışma sonucu başarımı yüksek çıkan otokodlayıcı modeli önermişlerdir.

Ferreira ve diğerleri (2019), yapmış oldukları çalışmada demir yolundan alınan gerçek verilerle raylı sistemlerde arıza tespiti ve tahminini yapabilmek için bir otokodlayıcı modeli tasarlamışlardır.

Tolpin (2019), rasgele dağıtılmış yüksek boyutlu verilerdeki popülasyon anomalisini tespit etmek için bir otokodlayıcı yaklaşımı önermiş ve çeşitli uygulama alanlarında değerlendirmiştir. Diğer araştırmalardan farklı olarak önerilen yöntemde popülasyon anomaliliklerinin tespitine yönelik bir kara kutu yaklaşımı kullanılmış ve veri seti boyutlarına, veri türlerine ve dağılımlarına karşı dayanıklı olduğu görülmüştür.

Zhang ve diğerleri (2023), rulman arızalarının teşhisi için yarı denetimli bir varyasyonel otokodlayıcı (VAE) yaklaşımı önermişlerdir. Deneysel sonuçlarda, önerilen yarı denetimli öğrenme yaklaşımını, etiketli veri örneklerinin aynı yüzdesine sahip bazı genel denetimli ve yarı denetimli kıyaslamalardan daha iyi başarım gösterdiğini sunmuşlardır.

Endüstriyel sistemlerin bileşenlerinden bazıları sıcaklığa duyarlıdır ve bu bileşenlerin sistemin çalışma başarımı açısından izlenmesi oldukça önemlidir. İzleme görevi için kullanılabilecek farklı çalışma teknikleri (Kirchgässner vd., 2020)'de sunulmuştur.

Literatürden farklı olarak bu tez çalışmasında, DC motor sıcaklığındaki anomali değerler oluşturulan bir Otokodlayıcı modeli ile tespit edilmiştir.

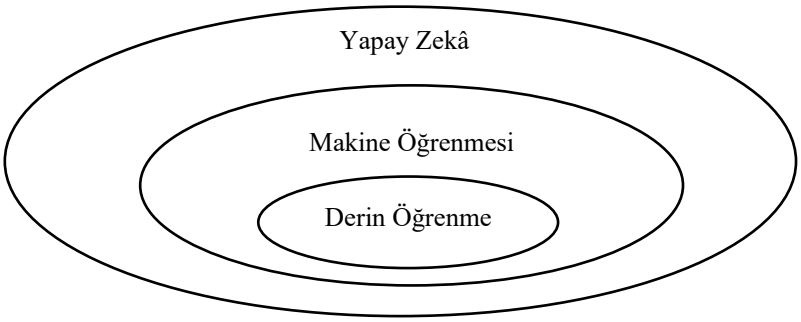
2.1. Yapay Zekâ

Teknolojinin gelişmesine paralel olarak yaşam standartlarımızın kalitesi de artmıştır. Bu artış ile teknolojiye olan bağımlılık ve bilginin önemi de artmıştır. Bilginin önem kazanması işleme süreçlerini ve araçlarını geliştirmiştir. Bilgi işleme süreçlerindeki en büyük en önemli adım sayısal sistemlerdir (bilgisayarlardır). Bilgisayarlar bilgiyi işleyip depolayabildiği gibi aynı zamanda zeki olma potansiyeline de sahiplerdir. Bilgisayar Bilimlerinin bir dalı olarak kabul edilen Yapay Zekâ (Artificial Intelligence) zeki sayısal sistemler çerçevesinde incelenmektedir (Köse, 2017).

Yapay zekâyı tek başına değerlendirmek yerine makine öğrenmesi ve derin öğrenme gibi alt terimlerle birlikte değerlendirmek daha doğru olacaktır. Yapay zekâ çalışmaları 1950'li yıllarda başlamaktadır, makine öğrenmesi ise 1959 yılında Arthur Samuel tarafından "bilgisayarın açıkça programlanmadan öğrenme yeteneği" olarak tanımlanmıştır (Samuel, 1959).

Yapay zekanın temelinin Alan Mathison Turing tarafından atıldığı kabul edilmektedir. Alan Mathison Turing; geliştirdiği Turing testi ile makinelerin düşünebileceği üzerinde çalışmalar gerçekleştirmiştir. Bu çalışmalardan dolayı Turing, Bilgisayar Bilimlerinin babası olarak kabul edilmektedir (Daylight ve Logic, 2015).

Yapay zekâ, bilgiyi işleyebilme yeteneğidir. Girdi bilgilerini alır ve bunları yorumlayarak tahminler yapar ve çıkış verilerini oluşturur. Kısaca, bilginin yorumlanabilme yeteneğidir. Bu yorum gelecekteki çıkış bilgilerine yön vermektedir. Yapay zekâ için literatürde yer alan genel bir venn şeması Şekil 2.1.'de gösterilmiştir.



Şekil 2.1. Yapay Zekâ için Venn Şeması (Jenis, 2023)

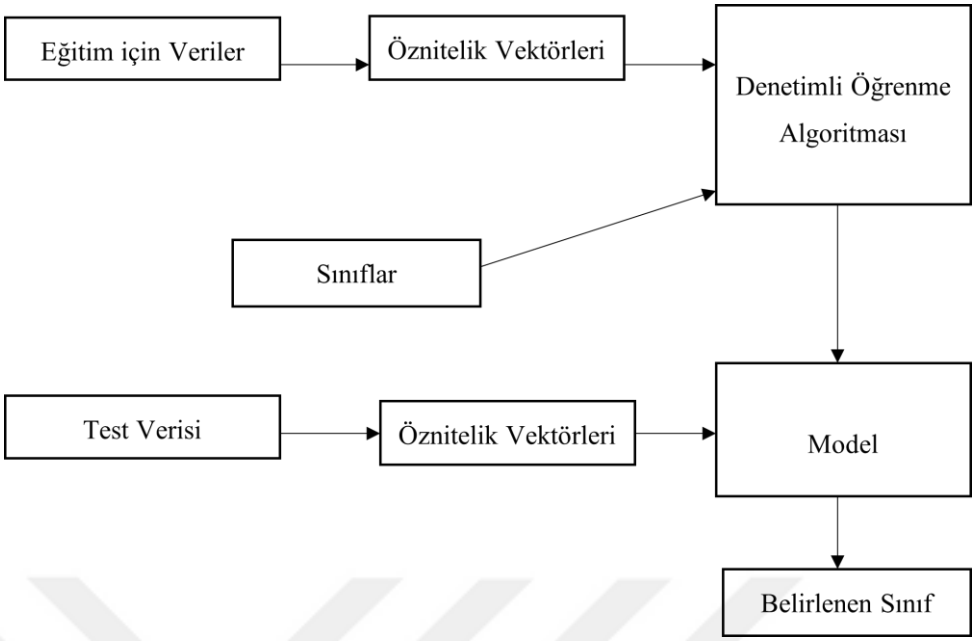
2.2. Makine Öğrenmesi

Makine öğrenmesi, sayısal sistemlerin veri kümeleri aracılığıyla kendi başlarına öğrenebilmeleri üzerinde yoğunlaşan yapay zekânın bir alt bölümü olarak tanımlanmaktadır. Makine öğrenmesi yöntemlerinin başarısı için verilerin kalitesi oldukça önemlidir. Veri kalitesini ise verilerin doğru ve eksiksiz olması belirlemektedir. Makine öğrenmesi yöntemleri öğrenme yaklaşımına göre dört ana başlık altında incelenmektedir. Bu başlıklar denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve takviyeli öğrenmedir.

2.2.1. Denetimli Öğrenme

Denetimli öğrenme, makine öğrenmesi yöntemleri içerisinde en yaygın olarak kullanılan öğrenme yaklaşımıdır. Denetimli öğrenmede, etiketli bir veri kümesi kullanılarak bir model eğitimi gerçekleştirilmektedir. Eğitilen model yeni veriler için bir tahmin üretmektedir.

Denetimli öğrenme diğer öğrenme yaklaşımları ile karşılaştırıldığında daha kolay bir yapıdadır. Etiketlenmiş veriler ile gerçekleştirilen eğitilme, modelin tahmin üretmesini kolaylaştırmaktadır. Denetimli öğrenmede veri kümesinde yer alan tüm verilerin etiketlenmiş olması gerektiğinden, etiketlenmiş verilerin tam ve yüksek doğrulukta olması çok önemlidir. Denetimli öğrenme için temel bir süreç akışı, bir sınıflandırma problemi için, aşağıda Şekil 2.2.'de gösterilmiştir.



Şekil 2.2. Denetimli Öğrenme (Sınıflandırma problemi)

Denetimli öğrenme modellerinin uygulama alanları temel olarak sınıflandırma ve regresyon olarak iki başlıkta toplanabilir.

Etiketlenmiş veri setindeki verilerin etiketlerine uygun biçimde ayrıştırılması işlemine sınıflandırma denmektedir. Başka bir ifadeyle, verilerin özniteliklerine uygun olacak biçimde ayrılması işlemi sınıflandırma olarak adlandırılmaktadır. Sınıflandırma problemlerin çözümünde, literatürde, K-En Yakın Komşu, Naive Bayes, Karar Ağacı ve Destek Vektör Makineleri (DVM) gibi modellerin kullanımı öne çıkmaktadır.

Bağımsız değişkenler ile bağımlı değişkenler arasındaki ilişkinin matematiksel olarak modellenmesi süreci regresyon olarak tanımlanmaktadır. Regresyon problemlerinin çözümünde ise yapay sinir ağları, rastgele orman, temel bileşen regresyonu ve doğrusal regresyon modelleri kullanım sıklıklarıyla öne çıkmaktadır (Metlek ve Kayaalp, 2020).

2.2.2. Denetimsiz Öğrenme

Denetimsiz öğrenmede, denetimli öğrenmeden farklı olarak modelin eğitiminde etiketli verilere ihtiyaç yoktur. Denetimsiz öğrenme modelleri kümeleme, boyut indirgeme, öznitelik ilişkilerinin bulunması gibi amaçlar için kullanılmaktadır. Denetimsiz öğrenme

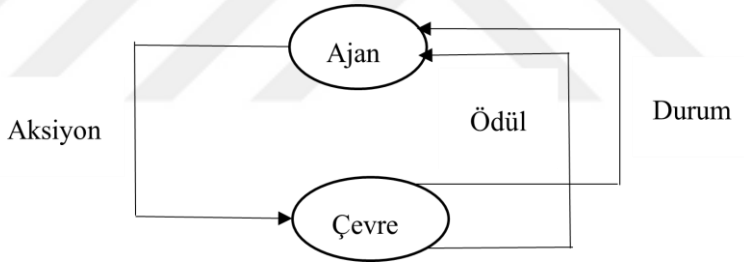
modelinde etiketlenmiş veri olmadığı için kullanıcının veriler hakkında yorum yapabilmesi zaman alıcı olabilmektedir (Bilgin, 2017).

2.2.3. Yarı Denetimli Öğrenme

Yarı denetimli öğrenme; denetimli öğrenme ile denetimsiz öğrenme modellerinin arasında olan bir öğrenme yaklaşımıdır. Bu öğrenme modelinde belirli sayıdaki etiketli veri grubuyla model eğitildikten sonra etiketli olmayan veri grupları için model tahmini gerçekleştirilmektedir. Etiketli olmayan veri grupları model tahmini ile eğitim setine dahil edilerek eğitilme verisi genişletilmektedir.

2.2.4. Takviyeli Öğrenme

Takviyeli öğrenmede de etiketlenmiş veriye ihtiyaç yoktur. Öğrenme süreci çevre ile etkileşime girerek öğrenen bir ajanın ödül-ceza yöntemine dayanmaktadır. Takviyeli, öğrenme sürecine ait algoritma işleyiş şeması Şekil 2.3.'te gösterilmektedir.



Şekil 2.3. Takviyeli Öğrenme Algoritmasının Şeması

Takviyeli öğrenme terminolojisinde ajan, aksiyon, çevre, durum, ödül, politika ve değer kavramları bulunmaktadır. Ajan; oluşturulan çevrede aldığı aksiyonlar ile ödül ve ceza sürecinde öğrenen bir varsayımsal nesne olarak adlandırılmıştır. Aksiyon; ajanın izleyebileceği hareketlerin tamamına denir. Çevre; ajan için oluşturulan senaryolar olarak ifade edilebilir. Durum; ajanın içerisinde bulunduğu mevcut durumu ifade etmektedir. Değer; uzun zamanda azalma oranına bağlı olarak ödül ve cezayı belirlemede oluşturulan bir fonksiyondur. Değer fonksiyonu en büyük ödüle sahip olunabilecek aksiyon politikasına sahiptir. Takviyeli öğrenmede oluşturulan ajan, bulunduğu çevre içerisinde gözlem yapar ve gözlemlerinden çıkardığı sonuca bakarak aksiyonuna karar verir. Karar sonucu aldığı aksiyon sonrasında tekrar gözlem yapar ve seçeneklerini değerlendirerek

bir sonraki aksiyonuna karar verir. Aksiyon sonrası tekrar gözlem yapmaya başlar ve bu gözlem sonrasında da tekrar seçeneklerini değerlendirir ve sonraki aksiyonuna karar verir. Ajan alacağı aksiyonlara karar vermeden önce en yüksek ödül alacağı seçeneği seçerek aksiyon alır. Ajanın takviyeli öğrenme sistemindeki amacı alabileceği en yüksek ödüle ulaşmaktır (Sutton ve Barto, 2018).

2.3. Derin Öğrenme

Derin öğrenme; ham verilerdeki yararlı öznitelikleri ve kalıpları otomatik olarak çıkarmak, aynı zamanda bu kalıpları veya öznitelikleri kullanmak için yapay sinir ağlarının kullanımından oluşan makine öğreniminin bir alt bölümüdür. Yapay sinir ağları kavramı ilk olarak Rosenblatt'ın oluşturmuş olduğu yapay nöronlardan (Perceptron) ortaya çıkmıştır (Rosenblatt, 1958).

Derin öğrenme üç katmandan daha fazla katmana sahip olan derin sinir ağları aracılığı ile giriş verilerinin derin özelliklerini hiyerarşik biçimde öğrenebilmektedir. Derin öğrenme çok kapsamlı matematiksel ve kavramsal bir arka plan barındırmaktadır (Kim, 2016).

Derin öğrenme modellerinin öne çıkan farkı; büyük miktardaki veriyi kolay bir biçimde, bütünleşik olarak ve otomatik bir şekilde işleyip analiz edebilme yeteneğidir. Derin öğrenme çalışmalarından sonra, problemin çözümüne bağlı olarak, makine öğrenmesi yöntemleri ile bütünleşik çözümler aranabilmektedir (Metlek ve Kayaalp, 2020).

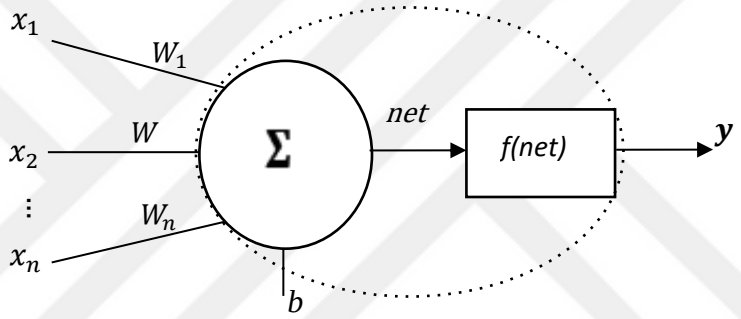
2.3.1. Yapay Sinir Ağları

Biyolojik sinir ağlarından esinlenerek yapay sinir ağları oluşturulmuştur. Yapay sinir ağlarında bilginin işlenmesi ve aktarımı biyolojik sinir ağlarına benzer şekilde gerçekleştirilmektedir. Bu nedenle biyolojik sinir ağlarının yapısı ve temel çalışma prensibinin incelenmesi yapay sinir ağları mimarisinin anlaşılmasına önemli katkılar sağlayacaktır. Biyolojik sinir hücreleri dört bölümden oluşmaktadır. Bu bölümler; dendrit, akson, çekirdek ve bağlantılardır. Dendtrit giriş bilgilerini çekirdeğe iletmekle görevlidir, çekirdek; dendritten gelen bilgileri toplayıp aksona iletmekle görevlidir, akson; çekirdekten gelen verileri bağlantılara iletmekle görevlidir, bağlantılar ise

aksondan gelen verileri diğer nöronlara taşımakla görevlidir [(Öztürk ve Şahin, 2018), (Orhan, 2015)]. Çizelge 2.1.'de yapay sinir ağları ile biyolojik sinir ağlarında kullanılan terimler arasındaki ilişki verilmiştir.

Çizelge 2.1. Biyolojik Sinir Ağları Terimlerine Karşılık Gelen Yapay Sinir Ağları Terimleri (Öztürk ve Şahin, 2018)

Biyolojik Sinir Sistemi	**Yapay Sinir Sistemi**
Nöron	İşlemci Elemanı
Dentrit	Toplama Fonksiyonu
Hücre Gövdesi	Transfer Fonksiyonu
Aksonlar	Yapay Nöron Çıkışı
Sinapslar	Ağırlıklar



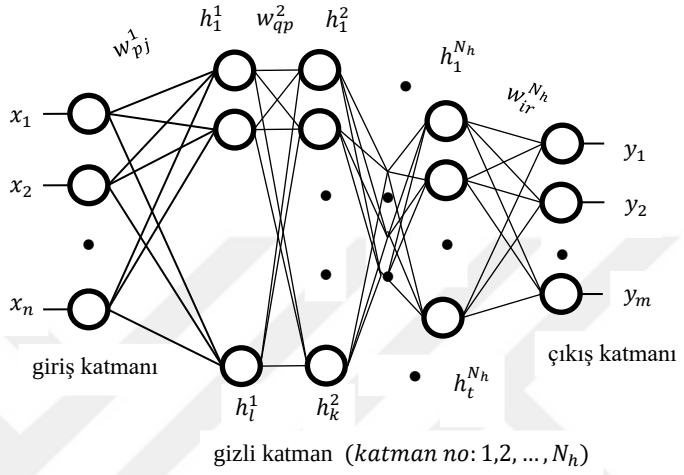
Şekil 2.4. Yapay Sinir Ağı (Algılayıcı) Yapısı (Orhan, 2015)

$$net = b + \sum_{i=1}^{M} x_i W_i \qquad y = f(net) \tag{2.1}$$

Temel bir yapay sinir ağı Şekil 2.4.'de gösterilmiştir. Bu yapı algılayıcı (nöron) olarakta isimlendirilmektedir (Staudemeyer ve Morris, 2019). Eğer bir ağ yapısı iki giriş ve bir çıkıştan oluşuyorsa bu ağ yapısı tek katmanlı (en temel / basit form) yapay sinir ağı olarak adlandırılmaktadır. Giriş verilerinin (*x*); her birisi kendisine ait ağırlık katsayılarına sahiptir. Bu ağırlık katsayıları ile giriş verilerinin çarpılması ve bu çarpımların toplamına *b* (önyargı-bias) eklenmesi ile oluşan sonucun (*net*) bir aktivasyon fonksiyonuna uygulanması sonrasında nöron çıkışı (*y*) elde edilir.

Yapay sinir ağları mimari yapılarına göre literatürde farklı isimlerle anılmaktadırlar. Bu mimariler arasında sıklıkla başvurulanlar arasında tek katmanlı yapay sinir ağları, çok

katmanlı yapay sinir ağları, LVQ (öğrenmeli vektör kuantalama) ağlar, ART (adaptip rezonans teori) ağlar ve SOM (özdüzenleyici haritalar) ağlar sayılabilmektedir [(Öztürk ve Şahin, 2018), (Orhan, 2015)]. Çok katmanlı yapay sinir ağları (ÇKYSA) bir giriş, bir çıkış ile bir veya birden fazla gizli katmana sahip olup her bir nöron bir önceki katmandaki tüm nöronlara bağlıdır (Haykin, 1999). Tipik bir ÇKYSA mimarisi Şekil 2.5'te verilmiştir.



Şekil 2.5. ÇKYSA Mimarisi (Yılmaz, 2016)

İlk gizli katmandaki her bir nöronun çıkışı aşağıdaki şekilde hesaplanır (Yılmaz, 2016).

$$h_p^1 = \alpha_f1\left(\sum_{j=1}^{n} w_{pj}^1 x_j\right) \tag{2.2}$$

α_f1 aktivasyon fonksiyonu olarak ifade edilmektedir. w_{pj}^1 Giriş katmanındaki nöronlar ile ilk gizli katman arasındaki ağırlığı göstermektedir. Çok katmanlı sinir ağlarında daha sonraki çıktılar içinde benzer şekilde; ilk gizli katmandaki çıktılar bir sonraki gizli katmana dağıtılarak h_q^2 aşağıdaki şekilde hesaplanır.

$$h_q^2 = \alpha_f2\left(\sum_{p=1}^{1} w_{qp}^2 h_p^1\right) \tag{2.3}$$

Burada α_f2 aktivasyon fonksiyonu ve w_{qp}^{2} ise ikinci gizli katman nöronları ile birinci gizli katman nöronları arasındaki ağırlık katsayısıdır. Çıkış katmanına gelindiğinde de hesaplamalar gizli katman nöronlarıyla aynı şekilde yapılır ve çıkışlar aşağıdaki gibi elde edilmektedir.

$$y_j = \alpha_f3\left(\sum_{r=1}^{1} w_{ir}^{N_h+1} h_r^{N_h}\right) \tag{2.4}$$

Burada α_f3 bir aktivasyon fonksiyonu ve $w_{ir}^{N_h+1}$ ise çıkış katman nöronları ile son gizli katman arasındaki ağırlık katsayısıdır (Yılmaz, 2016).

ÇKYSA'larda eğitilme algoritması olarak Geriye Doğru Yayılım (GDY) algoritması kullanılmaktadır. GDY algortiması bir hata düzeltme öğrenme kuralı ile çalışmakta ve bir hata yüzeyindeki en küçük hata noktalarını bulmak için Gradient İnişi yöntemini kullanmaktadır [(Haykin, 1999), (Basheer ve Hajmeer, 2000), (Übeyli, 2009)].

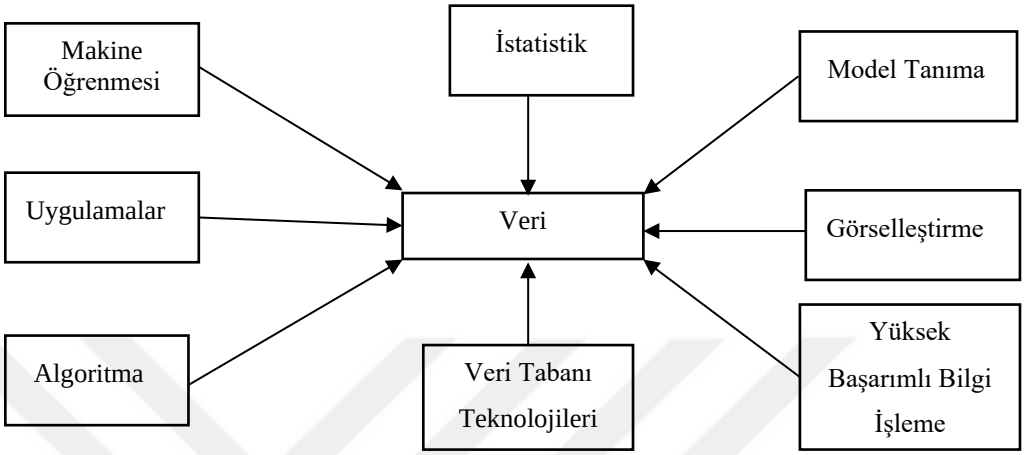
2.4. Anomali Nedir?

Anomali tanımından önce veriyi tanımlamak gerekirse veri; ölçümler, hesaplamalar, gözlemler yolu ile elde edilen bilgidir. Geçmişten günümüze teknolojinin ilerlemesiyle birlikte, verilerde hızlı bir biçimde artmaktadır. Artan veri yapısına bağlı olarak verinin depolanması, işlenmesi ve analizinin yapılarak yorumlanması zamanla daha da zor hale gelmektedir.

Büyük boyutta depolanmış veri yığınları içerisinde bağlantının kurulması, analizlerinin gerçekleştirilmesi ve bu analizlerden anlamlı bir bütün oluşturulması için farklı disiplinlerin (istatistik, bilgisayar teknolojileri, veri tabanı teknolojileri, vb.) bir araya gelmesi ile veri bilimi oluşmuştur (Özbay, 2015).

Veri madenciliğinde oluşturulan modeller iki alt başlık altında sınıflandırılmaktadır. Bu sınıflar tahmin edici ve tanımlayıcı olarak adlandırılmıştır (Özbay, 2015). Veri madenciliği yapısındaki disiplinler Şekil 2.6'da sunulmuştur. Endüstriyel sistemlerde

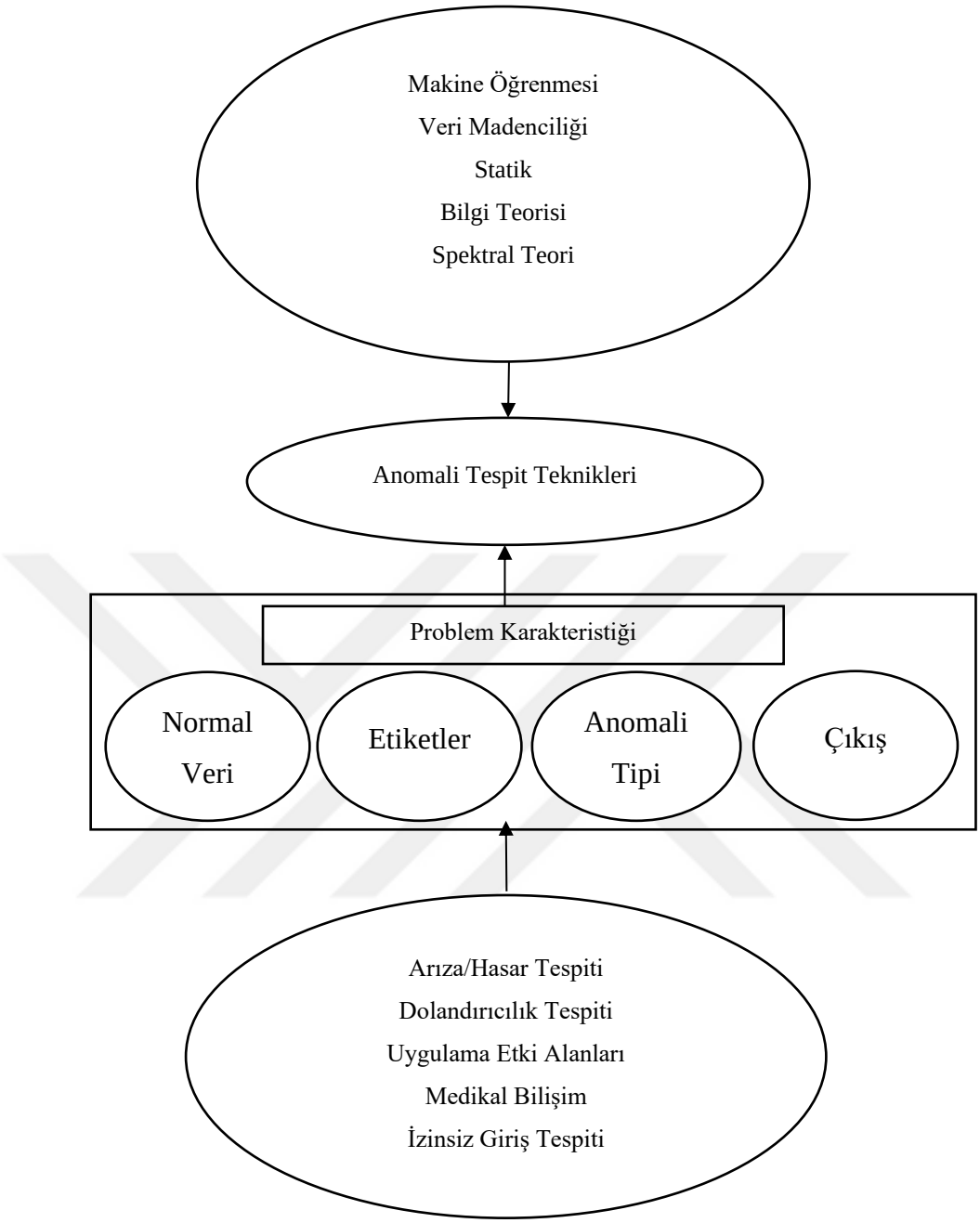
bulunan proseslerin ve ekipmanların takip edilen verilerinin arasından standardın dışında olan verilere anomali veri denmektedir. Yani sistemin verileri içerisinde normal kabul edilen sınırların dışarısına çıkan verilere anomali veri denmektedir.



Şekil 2.6. Veri Madenciliği Yapısındaki Disiplinler (Han vd., 2022)

Anomali, ayrıca sistemdeki farklılık, aykırılık, gürültü, sapma, istisna olarak da ifade edilebilmektedir (Hodge ve Austin, 2004). Anomali tespitinde; bir veri kümesindeki normal davranışlarından sapan veri noktaları, tespit edilmektedir.

Anomali tespitinin kullanıldığı alanlar; ürün kalitesinin artırılması, sistem başarımlarının artırılması, ekipman başarımının artırılması, personel başarımının artırılması, kalite-kestirimci bakım, kullanıcı dostu deneyimin iyileştirilmesi, siber saldırıların tespitinde, arıza tespitinde, dolandırıcılık tespitinde, endüstriyel sistemlerin başarımlarının takibinde, sensör ağlarında sorun tespitinde, görüntülerde bulunan hataların (farklılıkların) tespitinde, kredi kartı kullanım takibinde, cep telefonu takibinde, uydu görüntülerindeki farklılıkların tespitinde, zaman serilerinde izleme, tıbbi durumların takibi, metinlerdeki değişikliklerin takibinde, farmasötik araştırmalarda, yanlış etiketlenen verilerin tespitinde vb. gibi kullanılmaktadır. Anomali tespiti boyut azaltma yöntemleri ile de gerçekleştirilebilmektedir (Chen vd., 2018). Anomali tespiti ile ilgili ana yapı bileşenleri Şekil 2.7'de verilmiştir.



Şekil 2.7. Anomali Tespiti ile İlgili Ana Yapının Bileşenleri (Chandola vd., 2009)

2.4.1 Anomali Tipleri

Anomaliler noktasal anomaliler, bağlamsal anomaliler ve toplu anomaliler olmak üzere üç ana grup altında incelenmektedir. Eğer anomali tek bir değerden oluşuyorsa noktasal anomali olarak ifade edilmektedir. Bağlamsal anomalide, bazı veriler bazı özel durumlara (bağlama) göre bazen anomali bazen de normal olarak değerlendirilmektedirler. Toplu

anomali için ise anomali içeren değerlerin birden fazla olması ve birbiri ile ilişki içinde olması aranmaktadır. Anomali tespitinde literatürde çok sayıda yöntem başarıyla uygulanmaktadır. Bu yöntemler arasında öne çıkanlar arasında K-ortalama, Markov zinciri, izolasyon ormanı, tek sınıflı SVM (destek vektör makineleri) ve otokodlayıcılar gösterilebilmektedir (Acar, 2019).

Sensör verilerinde anomali tespiti yapılırken anomali tipleri çoğunlukla noktasal ve toplu anomali olarak değerlendirilmektedirler. Bu iki anomali tipi arasındaki ayrım problemin çözüm yaklaşımını belirlemektedir.

Toplu anomaliler, seriler biçiminde incelenmektedir. Veri seti içerisinde yer alan bir veri serisi tüm veri setine göre anormal değerler içeriyorsa toplu anomali olarak ifade edilmektedir. Toplu anomalilerde anomali değerleri tek tek incelendiğinde anomali olmayabilirler. Toplu anomaliler, verilerin belirli şekilde birlikte ortaya çıkma durumu olarakta tanımlanabilmektedirler.

Noktasal anomali ile toplu anomali arasındaki temel fark noktasal anomali tüm veri kümelerinde oluşabiliyorken, toplu anomaliler sadece birbirleriyle bağlantılı verilere sahip veri kümelerinde oluşabilmektedir. Toplu anomaliler tespit edilirken veriler arasındaki ilişkinin tespit edilmesi önemlidir. Veriler arasındaki ilişki genellikle sıralı, grafiksel veya uzamsaldır. Noktasal anomaliler en basit anomali formudur. Noktasal anomalinin tespitinde bir dizi ya da bağlantı kurmaya gerek olmadığından daha basit tespit edilebilmektedir. Toplu anomalilerin tespit edilmesi noktasal anomalilere göre daha karmaşık bir çözüm süreci gerektirmektedir [(Chandola vd., 2009), (Shaukat vd., 2021)].

3. MATERYAL ve YÖNTEM

Bu bölümde, anomali veri seti, anomali tespit yöntemleri ve bu yöntemlerin başarımlarının değerlendirilmesinde kullanılan temel yaklaşım ve ilgili başarım ölçütleri sunulmuştur.

3.1. DC Motor Deney Düzeneği ve Anomali Veri Seti

Tez çalışmasında gerçekleştirilen deneylerde kullanılmak üzere bu tez çalışmasına özel bir anomali veri seti oluşturulmuştur. Veri setinin elde edilmesinde kullanılan DC motor deney düzeneği ve bu düzenek ile elde edilen anomali veri setine ait detaylar alt bölümler halinde verilmiştir.

3.1.1. DC Motor Deney Düzeneği

Endüstriyel sistemlerde DC motorlar genellikle değişken yük altında belirli bir konumu veya hızı takip etmek için kullanılmaktadır (Faramarzi ve Sabahi, 2011). Bu çalışmada bir DC motorun hızı, sıcaklığı, akımı izlenmiştir. DC motor dinamik denklemleri 3.1 ve 3.2'de verilmiştir.

$$\frac{dw}{dt} = \frac{1}{J}(K_t i - bw) \tag{3.1}$$

$$\frac{di}{dt} = \frac{1}{L}(V - Ri - K_e w) \tag{3.2}$$

Burada "Kt" motor tork sabiti, "J" rotorun eylemsizlik momentidir, "i" armatür akımıdır, "b" motor viskoz sürtünme sabitidir, "R" elektrik direncidir, "Ke" elektromotor kuvvet sabiti ve "L" elektrik endüktansıdır (Abraham ve Shrivastava, 2018).

DC motorun referans hızı 200 rpm, 255 pwm olarak 8645 veri kaydedilmiştir. Otokodlayıcıda işlenecek veriler için öncelikle bir DC motor deney düzeneği tasarlanmıştır. Bu deney düzeneğinden alınan veriler otokodlayıcıda kullanılmıştır. Deney düzeneği için; 12 Volt 200 rpm özelliğinde bir DC motor kullanılmıştır. Bu DC motor belirli sürelerde çalıştırılıp sıcaklık, hız, akım gibi verileri kaydedilmiştir.

Oluşturulan deney düzeneğinde devre için Arduino Mega 2560 (ATmega2560) kullanılmıştır. Deney düzeneğinin tam turda ürettiği pulse sayısını hesaplamak için 1 adet rotary encoder kullanılmıştır.

DC Motorun gövde sıcaklığının anlık izlenebilmesi için ise motor gövdesine NTC (10K) bağlanmıştır. DC motoru PWM ile sürebilmek için 1adet n-kanal mosfet kullanılmıştır. DC motorun oluşan verilerinin SD karta kaydedilebilmesi için 1 adet SD kart modülü kullanılmıştır.

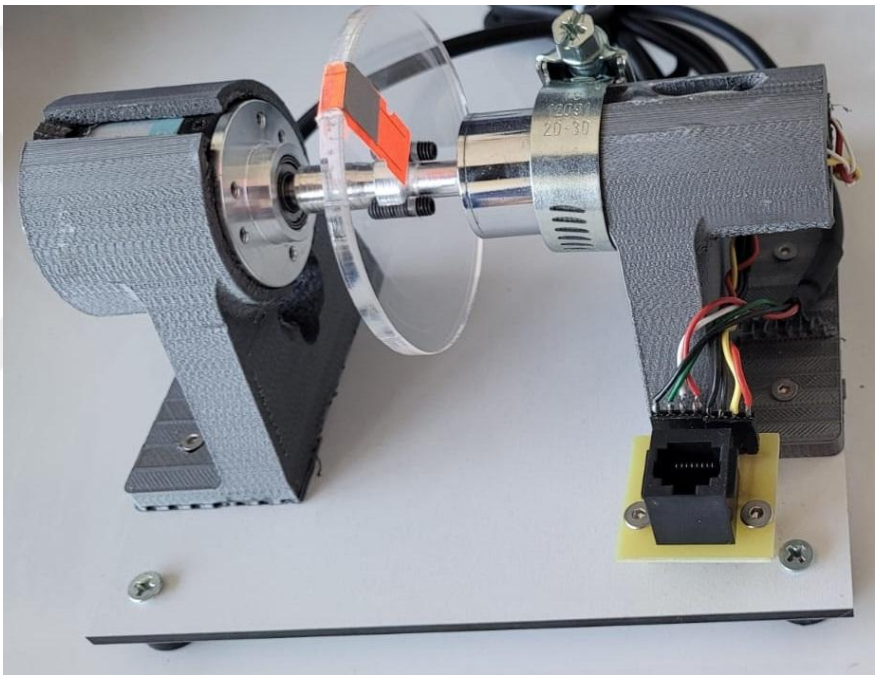
Oluşan DC motor verilerinin ekranda gözlenmesi için 1 adet nextion serial ekran kullanılmıştır. Verilerin gerçek zamanlı kaydedilmesi için 1 adet DS1307 kullanılmıştır. Ekran üzerindeki butonlar ile mikrodenetleyici tarafından 0-255 (%0-%100) aralığında PWM üretilmesi sağlanmıştır. Üretilen bu PWM ile n-kanal mosfet sürücü devresi tasarlanarak, 12 Volt 200 rpm redüktörlü bir dc motorun hız kontrolü sağlanmıştır.

DC Motorun kaplin ile miline bağlanan rotary encoder'ın döndürülmesiyle motor dönüş hızı "rpm" olarak hesaplanarak SD kartta ve ekranda görüntülenmesi sağlanmıştır. DC Motor döndürülürken oluşan sıcaklığın, motorun döndürülmesi engellendiğinde, motorun çektiği zorlama akımının artmasından dolayı sıcaklığın daha hızlı arttığı gözlemlenmiştir.

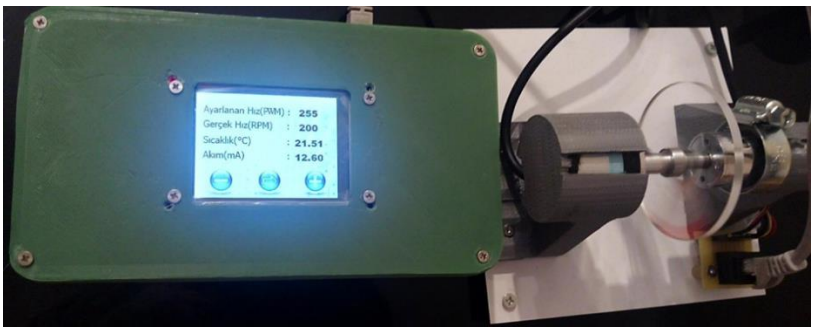
SD kartta bulunan bir switch yardımı ile SD kartın soketine takılı olup olmadığı sürekli kontrol edilmiştir. SD kart soketi üzerinde bulunan anahtar Pull-up devresi ile mikrodenetleyicinin pinine girilmiştir.

Mikrodenetleyicinin pini lojik 1 (+5V) olduğunda soketin takılmadığı, lojik 0 (0V) olduğunda ise soketin takıldığı mikrodenetleyiciye iletilmiştir. Bu soketin kullanılma amacı SPI (Serial Peripheral Interface) haberleşme ile yapılan haberleşme türünde, SD kartın olmaması durumunda SPI haberleşme yavaşlamaya uğrayacak, bu sebeple yazılım döngüsü yavaşlayacaktır. Bu anahtar ile SD kartın olmadığı durumlarda SPI haberleşmenin yapılması engellenmiştir.

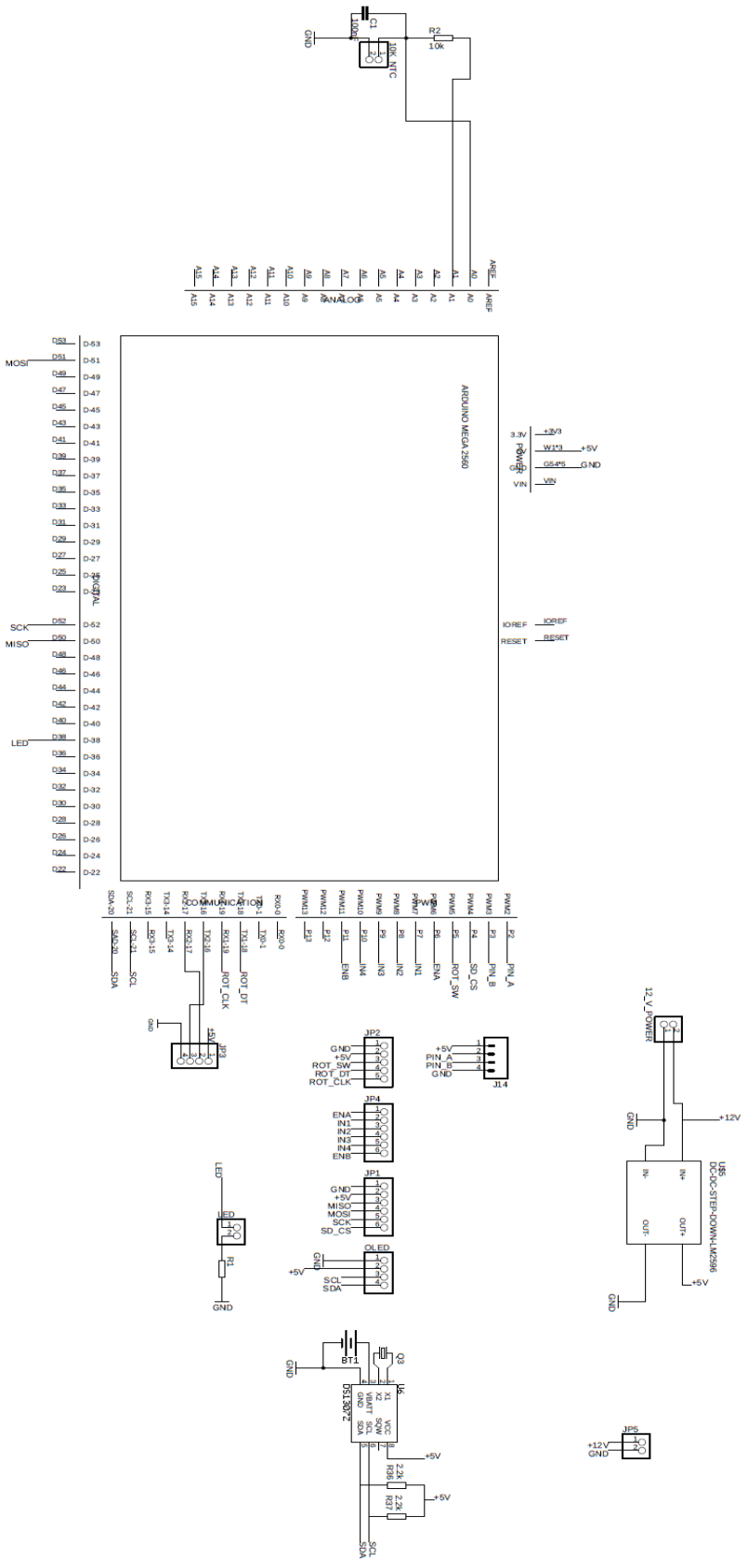
DC motor devremizdeki devir sayısını ve akımı ölçebilmek için INA219 ve motoru sürmek için L298 motor sürücü entegresi kullanılmıştır. DC motor devremizde zaman, pwm, rpm, I²C üzerinden haberleşen DS1307 gerçek zamanlı saat entegresi ile verilerin anlık olarak ekran ve SD kart modülü ile kullanıcıya aktarılması sağlanmıştır. Ayrıca oluşturulan panelin dış aksamı tasarlanarak 3 boyutlu yazıcı tarafından basılmıştır. SD kart modülüne veri kaydı gerçekleştirilmiş olup bu veriler anomali tespiti için kullanılmıştır. Şekil 3.1'de veri seti için oluşturulan DC motor deney düzeneği görüntüsü ve Şekil 3.2.'de panelde izlenebilen değerlerin görüntüsü verilmiştir. Şekil 3.3'te ise DC motor deney düzeneğinin ayrıntılı çizimi sunulmuştur.



Şekil 3.1. Oluşturulan DC Motor Deney Düzeneği



Şekil 3.2. Oluşturulan DC Motor Deney Düzeneği Akım, Hız, Sıcaklık Verilerinin İzlenebildiği Panelin Görüntüsü



Şekil 3.3. Oluşturulan DC Motor Deney Düzeneği Ayrıntılı Çizimi

3.1.2. Anomali Veri Seti

Anomali veri seti, bu çalışma için özel olarak tasarlanmış olan 3.1.1. bölümde anlatılan DC motor deney düzeneği kullanılarak oluşturulmuştur. Düzenekte 12 Volt 200 rpm DC motor kullanılmıştır. DC motor belirli periyotlarla çalıştırılarak sıcaklık, hız ve akım değerlerinden oluşan veriler panelde izlenerek eş zamanlı olarak SD karta kaydedilmiştir.



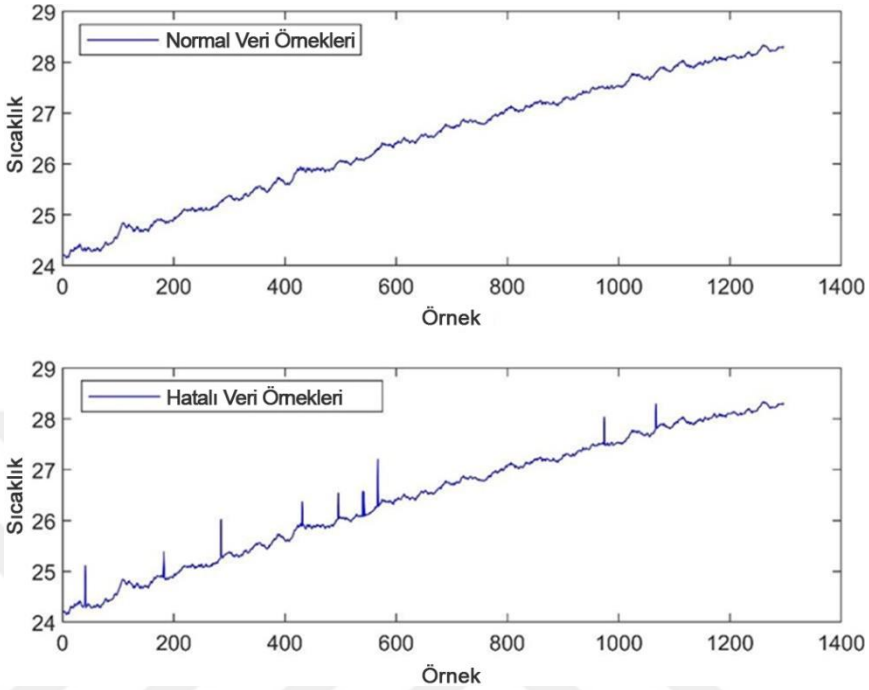
Şekil 3.4. Veri Değerlerinin (Sıcaklık, Hız, Akım) İzlenmesi.

Sıcaklık ölçümünün elektrik motorlarının durumunu izlemek için en popüler tekniklerden biri olması nedeniyle [(Nandi vd., 2005), (Saad, 2015)] bu çalışma, DC motorun vücut sıcaklığındaki anomaliliklerin tespitine odaklanmaktadır. DC motor gövde sıcaklığına yönelik anomalilik veri seti, normal ve anomali (hatalı) olmak üzere iki etiketli zaman serisi verileri içermektedir. Normal veriler stabil çalışma rejimi ile normal oda koşullarında, hatalı veriler ise aynı oda koşullarında deney ekipmanının frenleme kısmında oluşturulan harici arıza altında toplanmıştır. Normal ve anomali veri örneklerinin görselleştirilmesi Şekil 3.5'te verilmiştir.

3.2. Anomali Tespit Yöntemleri

Literatürde çok çeşitli anomali tespiti yöntemleri bulunmaktadır. Bunlar arasında 3-Sigma Aykırı Değeri, K-ortalama, Markov Zinciri, İzolasyon Ormanı, Eliptik Zarf modeli, Tek Sınıflı Destek Vektör Makineleri, Tekrarlı Sinir Ağları ve Otokodlayıcılar en çok tercih edilen yöntemlerdir (Acar, 2019).

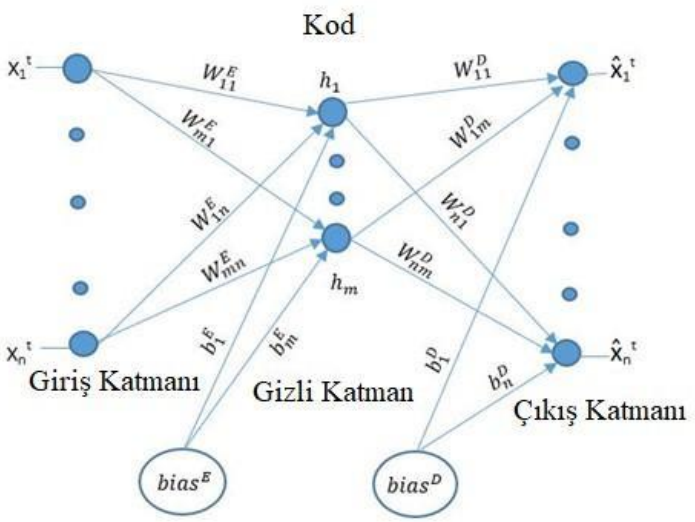
Bunlardan tez çalışmasında Otokodlayıcı, 3-Sigma Aykırı Değeri, İzolasyon Ormanı yöntemleri kullanılmıştır.



Şekil 3.5. Normal ve Hatalı Veri Örneklerinin Görselleştirilmesi.

3.2.1. Otokodlayıcı

Otokodlayıcı (AE), kodlayıcı ve kod çözücü olmak üzere iki alt ağ parçası içeren bir sinir ağıdır. Otokodlayıcılar denetimsiz öğrenmeye aittir ve eğitimi etiketli bir veri kümesi gerektirmez. Kodlayıcı kısım, kodlar olarak adlandırılan sıkıştırılmış (gizli) temsillere anlamlı özellikleri çıkararak girdi vektörlerini eşlerken, kod çözücü kısım kodları orijinal girdi alanına geri eşler ve yeniden yapılandırılmış vektörler üretir. Otokodlayıcılı öğrenme, giriş vektörleri ile yeniden yapılandırılmış vektörler [(Chow vd., 2020), (An ve Cho, 2015), (Chen vd., 2018), (Siddalingappa ve Kanagaraj, 2021)] arasındaki yeniden yapılandırma hatasını en aza indirmeyi amaçlar. Tipik bir Otokodlayıcının mimarisi, Şekil 3.6.'da verilmiştir.



Şekil 3.6. Tipik Bir Otokodlayıcı Mimarisi

Giriş vektörleri, kodlar ve yeniden oluşturulmuş vektörler sırasıyla x, h and $\hat{x}$ olarak temsil edilir ve veri kümesi N giriş vektörüne sahiptir.

Giriş vektörleri $x_t \in R^n$, $x_t = [x_1^t, x_{2,}^t \dots, x_n^t]$, $t = 1,2, \dots, N$ gizli vektörlere (kodlara) eşlenir $h \in R^m$, $h = [h_{1,} h_{2,\dots,} h_m]$, ve $n < m$.

Gizli katmanın p_{th} nöronunun çıktısı Denklem (3.3) kullanılarak hesaplanır.

$$h_p = a_f\ (\sum_{j=1}^{n} w_{pj}^E x_j^t + b_p^E) \tag{3.3}$$

Bu denklemde a_f bir aktivasyon fonksiyonudur, w_{pj}^E , j_{th} girdi katmanı nöronu ile p_{th} gizli katman nöronu arasındaki kodlayıcı ağırlık katsayısını temsil eder ve b_p^E, p_{th} gizli katman nöronu için bir bias terimidir.

Gizli vektörler, yeniden yapılandırılmış vektörlere eşlenir $\hat{x}_t \in R^n$, $\hat{x}_t = [\hat{x}_1^t, \hat{x}_2^t, \dots, \hat{x}_n^t]$, Boyutsal olarak giriş alanları ile eşleştirilir. Çıkış katmanının i_{th} nöronunun çıktısı Denklem (3.4) kullanılarak hesaplanır.

$$\hat{x}_i = a_f\ (\sum_{p=1}^{m} w_{ip}^D h_p + b_i^D) \tag{3.4}$$

Burada a_f, aktivasyon fonksiyonudur, w_{ip}^{D}, p_{th} gizli katman nöronu ile i_{th} çıkış katmanı nöronu arasındaki kod çözücü ağırlık katsayısını temsil eder ve b_i^D , i_{th} çıkış katmanı nöronu için bir bias terimidir.

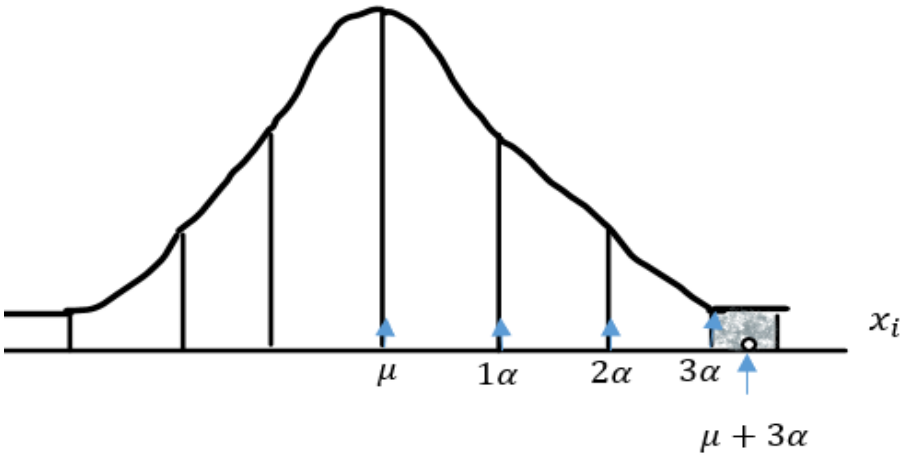
Girdi vektörleri ile yeniden yapılandırılmış vektörler arasındaki yeniden yapılandırma hatası Denklem (3.5) kullanılarak hesaplanır.

$$Error = \sum_{t=1}^{N} \sum_{i=1}^{n} (x_i^t - \hat{x}_i^t)^2 \tag{3.5}$$

3.2.2. 3-Sigma Aykırı Değeri (3-SgOut)

3-Sigma aykırı değeri, anomali tespiti için kontrol grafiği ile ilgili bir tekniktir [(Kasunic vd., 2011)]. Verilerin iki temel istatistiksel özelliğine, ortalama (m) ve standart sapmaya (std) dayanır. Üç sigma aralığının dışında kalan veri noktaları, Denklem (3.6)'daki gibi aykırı değerler olarak adlandırılır [(Kasunic vd., 2011), (Mirzaee vd., 2014)].

$$outliers = data_{points}(out\ of\ [m - 3std, m + 3std]) \tag{3.6}$$



Şekil 3.7. 3-Sigma Aykırı Değeri için Grafiksel Gösterim (Zhao vd., 2013)**.**

3-Sigma aykırı değerinde şekilde gösterildiği gibi ortalamanın 3 standart sapmanın dışına çıkması durumlarını anomali olarak yakalamaktadır (Zhao vd., 2013).

$$|x_i - \hat{\mu}| > 3\sigma \tag{3.7}$$

3-Sigma aykırı değeri anomalilerin tespiti için sık kullanılan yöntemlerden birisidir. 3-Sigma aykırı örnek boyutu yüksek olan anomali değerleri, örnek sayısı düşük olan anomali değerlerine göre daha iyi yakalamaktadır (Zhao vd., 2013).

3.2.4. Karar Ağacı (KA)

Karar ağaçları tahmin ve sınıflandırma problemlerinin çözümünde kullanılmaktadır. Karar ağaçları kolay anlaşılabilir ve yorumlanabilir olmaları nedeniyle literatürde oldukça yoğun bir ilgi görmektedir. Bununla birlikte, maliyetlerinin düşük olması ve güvenilirlik oranlarının yüksek olması nedeniyle en çok tercih edilen yöntemlerin arasında yer almaktadır. Karar ağaçlarında genellikle QUEST, CART, C4.5, C5.0, CHAID ve ID3 gibi algoritmalar kullanılmaktadır. Karar ağacının oluşturulmasında temel ölçüt Entropi ölçümüdür. Entropi ölçümünün olabildiğince düşük olması amaçlanmaktadır. Entropi ölçümünün yüksek olması sonuçların kararsızlığını artırmaktadır (Çalış vd., 2014).

3.2.4.1. İzolasyon Ormanı

İzolasyon ormanı algoritması (Isolation Forest), denetimsiz bir anomali tespit algoritması olup, veri noktalarını izole etmeye dayanmaktadır. Bu toplu anomaliler için genellikle tercih edilen bir algoritmadır.

İzolasyon ormanı algoritması Liu ve arkadaşları tarafından geliştirilmiş olup; büyük veri kümelerindeki anomalileri hızlı ve etkili bir şekilde tespit etmeyi sağlayan bir yapıdadır. İzolasyon ormanı algoritması özellikle toplu anomalilerin tespiti için genellikle kullanımı tercih edilmektedir (Liu vd., 2008).

Bu algoritmada özellikle penceresel özellikler kullanıldığında, zaman serisi verileri ve ardışık verilerde bulunan toplu anomaliler etkili bir şekilde tespit edilebilmektedir.

Finansal dolandırıcılık tespiti, siber güvenlik, sağlık verileri analizleri ve üretim süreçlerindeki arıza tespitleri gibi çeşitli alanlardaki anomalileri hızlı ve etkili bir şekilde tespit eden en popüler yöntemlerden bir tanesidir.

İzolasyon algoritmasının mantığı veri noktalarını rastgele bölmeye dayanır ve bir veri noktasının ne kadar kolay izole edilip edilemediğine bakılarak anomali olup olmadığını değerlendirmektedir. (Liu vd., 2008).

İzolasyon algoritması öncelikle anomali tespitinde kullanılan karar ağacı temelli bir algoritmadır. Bu algoritmanın temel çalışma prensibi ise anomali olarak değerlendirilen veri noktalarının izole bir şekilde olarak ve az sayıda bulunmasıdır. Algoritmanın çalışma mantığında; Rastgele bölme işlemi, izole etme, orman yapısı sonuçları analiz etme bulunmaktadır.

Karar ağacı tabanlı yöntemlerden birisi olan izolasyon ormanında veri kümesindeki örneklerin birbirinden izole bir şekilde ayrılması amaçlanmaktadır. İzolasyon ormanı birden fazla karar ağacı yapısına sahiptir. Rastgele seçtiği gözlemler ve öznitelik yapılarını kullanarak belirlenen en küçük ve en büyük değerlere uygun biçimde örnekleri izole edip dallara ayırır. İzolasyon ormanında yapılan her eğitim vektörünü izole edebilmek için gereken düğüm sayısı hesaplanmaktadır.

İzolasyon ormanı ile yapılan anomali tespitinde dallara ayırma yeteneği ile anomali değerleri diğer dallardan izole edilmektedir. Modelin ayrıştırma başarımının değerlendirilmesinde Silhoutte Skoru kullanılmaktadır [(Toklu vd., 2021) (Lesouple vd., 2021)].

Bir adet izolasyon ağacı oluşturmak için, n adet eğitim verisine sahip olduğumuzu varsayalım. Burada $\{x_1, .., x_n\}$ eğitilme verisi olmak üzere, her bir x_i aşağıdaki biçimde tanımlansın;

$$x_i = \left[x_{i,1} \dots x_{i,d}\right]^T \in R^d \tag{3.8}$$

Bu durumda eğitilme veri kümesi X ile gösterilmek üzere aşağıdaki gibi ifade edilir.

$$X = [x_1 \dots x_n]^T \in R^{nxd} \tag{3.9}$$

Veri kümesini iki alt kümeye bölmek ve rastgele bir düğüm oluşturmak için R^d'nin bir bileşeni (q) rastgele seçilir ve bölüm değeri p aşağıda verilen aralıkta düzgün dağılımlı olacak biçimde örneklenir.

$$\left[min_{i=1,\dots,n}\, x_{i,q}; max_{i=1,\dots n} x_{i,q}\right] \tag{3.10}$$

Böylece veri kümesi sol dal ve sağ dal olmak üzere ikiye bölünür. Sol dalda yer alan örnekler kümesi aşağıdaki gibi oluşturulur.

$$\left\{x_i, x_{i,q} \leq p\right\} \tag{3.11}$$

Sağ dal kısmında yer alacak örnekler kümesi ise aşağıdaki gibi oluşturulur.

$$\left\{x_i, x_{i,q} > p\right\} \tag{3.12}$$

Bu prosedür iteratif bir biçimde her bir dalda tek bir örnek kalana kadar ve belirlenen bir derinliğe ulaşılana kadar devam ettirilerek ağaç oluşturulur.

İzolasyon orman *t* sayıda rastgele izolasyon ağacı kullanılarak oluşturulduktan sonra bir x noktasını izole edebilmek için beklenen yol uzunluğu $h(x)$, her bir ağacın noktayı izole etmek için ihtiyaç duyduğu yol uzunluklarının ortalaması olarak hesaplanır. Sonunda anomali skoru hesabı aşağıdaki gibi gerçekleştirilir.

$$s(x) = 2^{-\frac{E[h(x)]}{c(\Psi)}} \tag{3.13}$$

Burada $c(n)$, n boyutundaki bir veri kümesi için $h(x)$'in ortalama değerini ifade etmektedir ve aşağıdaki biçimde hesaplanmaktadır.

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n} \tag{3.14}$$

Burada $H(n)$, n. harmonik sayıdır ve yaklaşık olarak $\ln(n) + 0.577$ biçiminde hesaplanmaktadır. Böylece, $E[h(x)] = c(n)$ olduğu durumda x'in anomali skoru $s(x,n) = 0.5$ olur. Sonuç olarak $h(x)$, $c(n)$ ile karşılaştırıldığında küçük ise x'in izole bir nokta olacağı kararı verilir ve ilişkili anomali skoruda 1'e yakın bir değer olarak gerçekleşir.

Pratik uygulamalarda $s_0 \epsilon\, [0,1]$ olacak biçimde bir eşik değeri belirlenir. $s(x) > s_0$ olduğu durumda x noktası anomali olarak değerlendirilirken, $s(x) \leq s_0$ olarak belirlendiğinde ise x noktası için normal değer kararı verilecektir (Lesouple vd., 2021).

3.2.5. Başarım Değerlendirmesi

Uygulanan anomalilik tespit yöntemlerinin başarımları; Doğruluk, Geri Çağırma ve Kesinlik ölçümleri açısından değerlendirilmektedir. Denklem (3.15) - (3.17), Çizelge 3.1.'de verilen hata matrisinden hesaplanmaktadır [(Kohavi ve Provost, 1998), (Abdoli vd., 2023)].

Hata matrisi içerisinde tahmin edilen anomali ve normal değerlerin karşılığında gerçekleşen anomali ve hatalı değerler incelenmiştir. Gerçekleşen anomali ve hatalı değerler olarak TP (gerçek pozitif), FN (yanlış negatif) sınıflandırılmışken, gerçekleşen normal değerler FP (yanlış pozitif), TN (gerçek negatif) olarak sınıflandırılmıştır. Sınıflandırmanın başarımı TP, FN, FP, TN hesaplanarak değerlendirmesi yapılmaktadır.

Çizelge 3.1. Hata Matrisi

Tahmin Edilen	Gerçekleşen	
	Anomali (Hatalı)	Normal
Anomali (Hatalı)	TP (Gerçek Pozitif)	FP (Yanlış Pozitif)
Normal	FN (Yanlış Negatif)	TN (Gerçek Negatif)

$$Doğruluk\ (accuracy) = (TP + TN) / (TP + FP + TN + FN) \quad (3.15)$$

$$Geri\ Çağırma\ (recall) = TP / (TP + FN) \quad (3.16)$$

$$Kesinlik\ (precision) = TP / (TP + FP) \quad (3.17)$$

4. BULGULAR ve TARTIŞMA

4.1. Bulgular

Deneylerde DC motor gövde sıcaklığındaki noktasal ve toplu anomalilerin tespiti gerçekleştirilmiştir. Noktasal anomalilerin tespiti için Otokodlayıcı ve 3-SgOut Aykırı Değer yöntemleri kullanılırken toplu anomalilerin tespitinde ise Otokodlayıcı ile İzolasyon Ormanı yöntemi kullanılmıştır. Deneyler MATLAB ortamında gerçekleştirilmiştir (MATLAB Toolbox, 1993). Zaman serisi sinyaline ait örnekler kayan pencereler yoluyla uygulanmıştır. Yöntemlerin başarımları pencere boyutundan etkilendiğinden, en uygun pencere boyutunun seçimi ızgara arama algoritması kullanılarak belirlenmiştir.

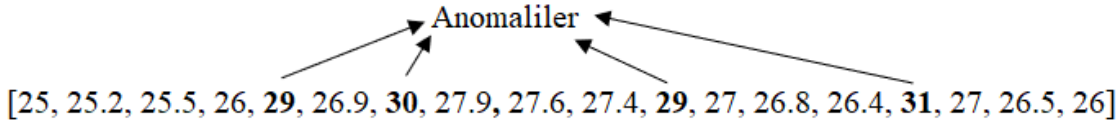
Veri seti noktasal ve toplu olarak sınıflandırılabilecek anomali değerlerine sahip 223 zaman serisi sinyali içermektedir. Her sinyal 1297 veri örneğinden oluşmaktadır. Veri setinde 2109 noktasal anomali örneği yer alırken, toplam örnek sayısı ise 289231'dir. Yöntemlerin başarımları; Doğruluk, Geri Çağırma ve Kesinlik ölçütleri kullanılarak belirlenmiştir.

Her iki anomali tipine ait bulgular alt bölümler içerisinde detaylı bir biçimde sunulmuştur.

4.1.1. Noktasal Anomali Tespiti

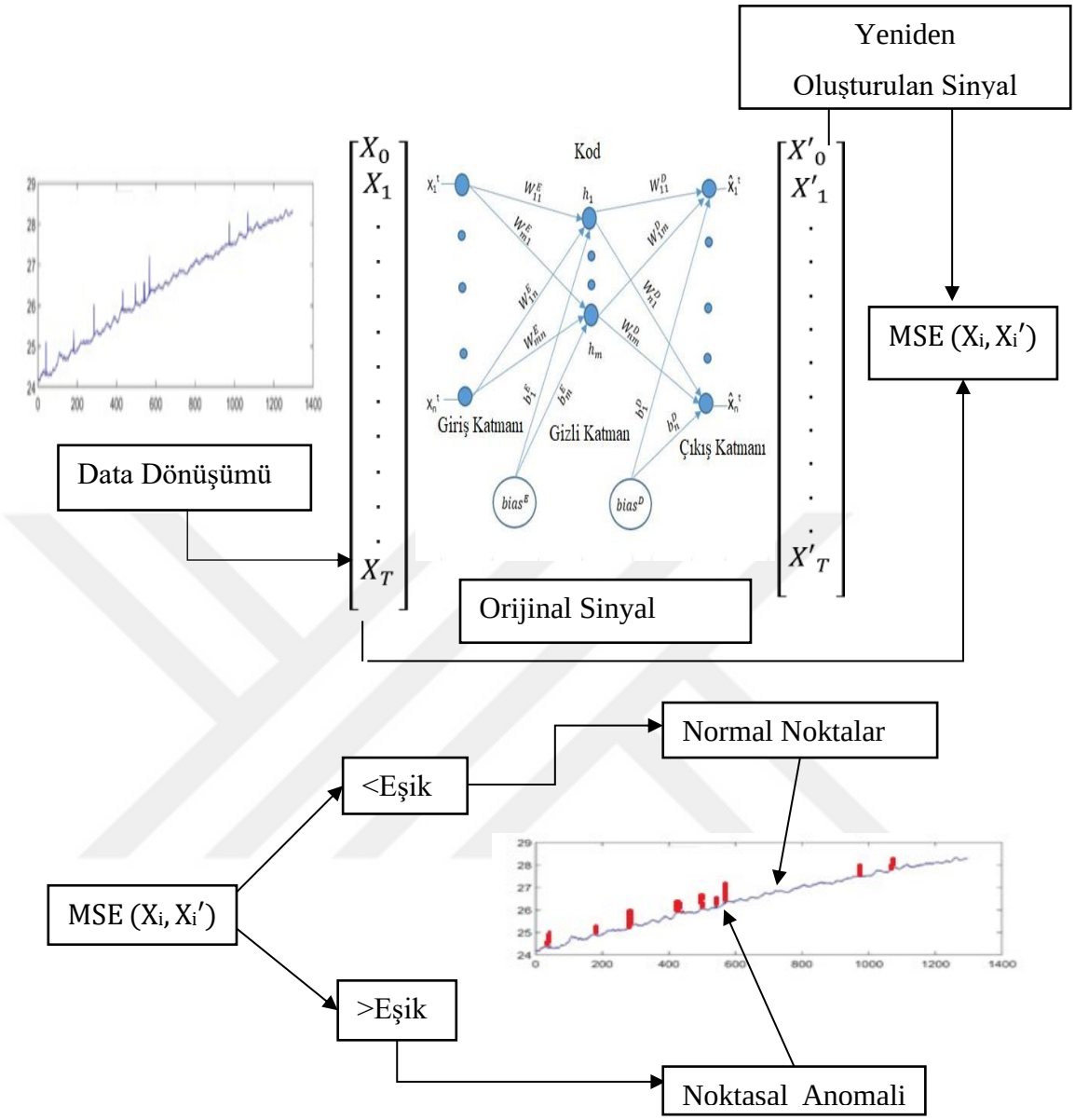
Bu yaklaşımda sıcaklık sensöründen gelen zaman serisi içerisinde yer alan noktasal anomalilerin tespiti gerçekleştirilmiştir. İlgili zaman serisi içerisinde görülen noktasal anomali değerlerine ait bir örnek Şekil 4.1.'de verilmiştir.

Noktasal Anomaliler:



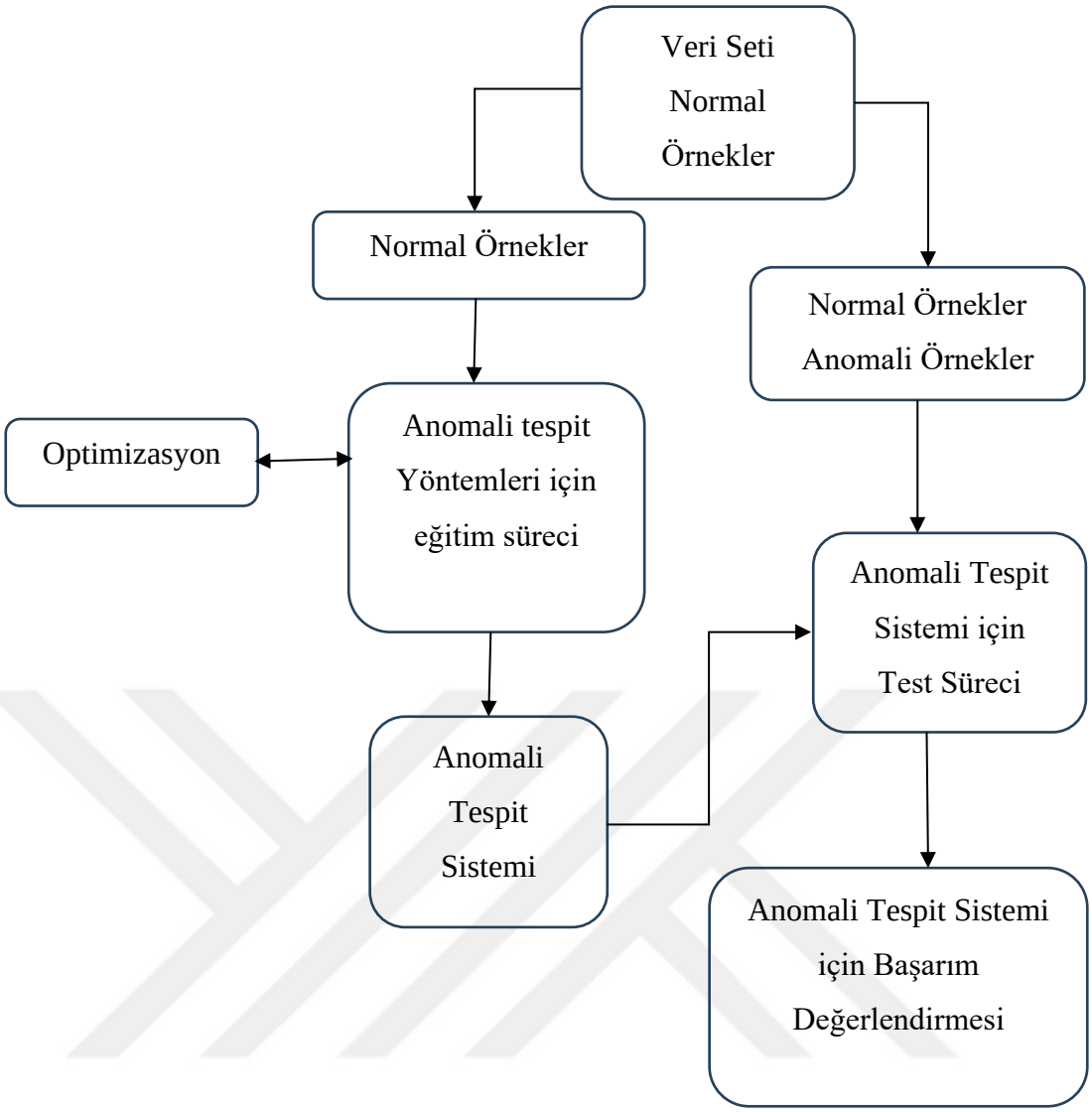
Şekil 4.1. Sıcaklık sensörü zaman serisinde yer alan noktasal anomaliler

Sıcaklık sensörü zaman serisinde yer alan noktasal anomaliler tespit edilirken izlenen yaklaşıma ait bir görsel anlatım Şekil 4.2.'de sunulmuştur.



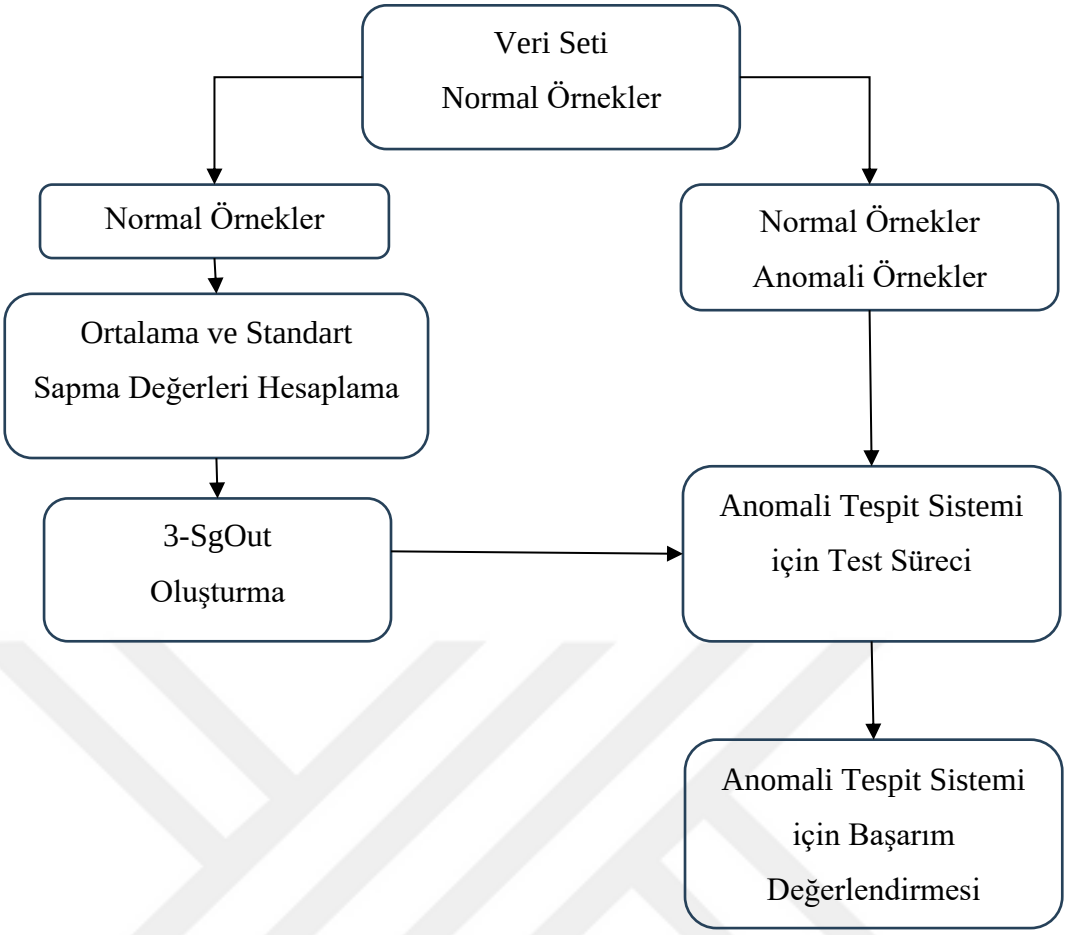
Şekil 4.2. Noktasal Anomali Tespiti Görsel Anlatımı (Nicholaus vd., 2021)

Noktasal anomalilerin tespitinde Otokodlayıcı ve 3-Sigma Aykırı Değer yöntemleri kullanılmıştır. Her iki yöntem ile noktasal anomalilerin tespitinin gerçekleştirilmesine ait akış diyagramları sırasıyla Şekil 4.3. ve Şekil 4.4.'de sunulmuştur.



Şekil 4.3. Otokodlayıcı ile Noktasal Anomali Tespiti için Akış Diyagramı

Veri setinde yer alan noktasal anomalilerin tespiti Otokodlayıcı ve 3-SgOut Aykırı Değer yöntemleri kullanılarak Şekil 4.3. ve Şekil 4.4. ile verilen akış diyagramlarında belirtildiği biçimde gerçekleştirilmiştir. Her iki yönteme ait bulgular alt bölümlerde verilmiştir.



Şekil 4.4. 3-Sigma Aykırı Değeri ile Noktasal Anomali Tespiti için Akış Diyagramı

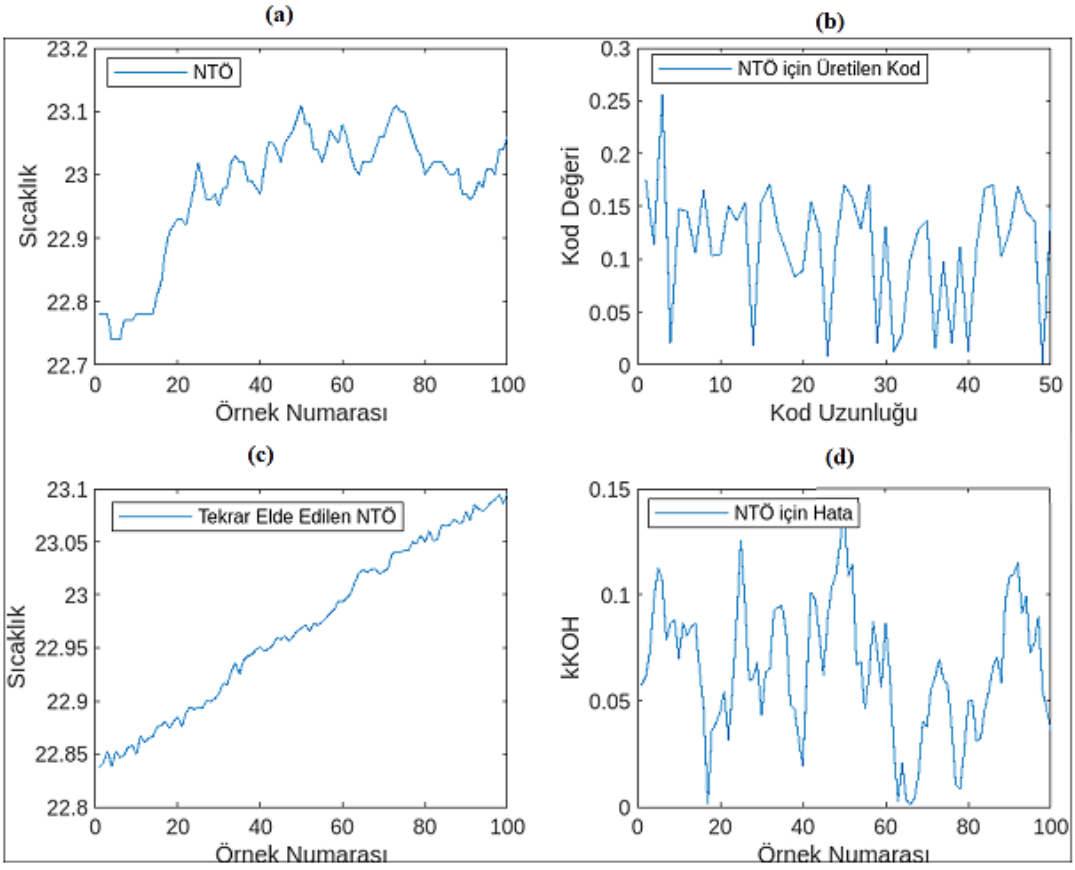
4.1.1.1 Otokodlayıcı ile Anomali Tespiti

Deneylerde, DC motor sıcaklığındaki anomalilerin tespiti için Şekil 3.6.'da verilen tipik bir Otokodlayıcı mimarisi kullanılmıştır. Giriş ve çıkış katmanlarındaki nöron sayısı kayan pencerenin boyutuna, gizli katmandaki nöron sayısı ise kod uzunluğuna eşit olacak şekilde seçilmiştir. Her iki nöron sayısı için en uygun değerler ızgara arama algoritması kullanılarak belirlenmiştir. Mimaride, gizli katman ve çıkış katmanı nöronları için aktivasyon fonksiyonları olarak sırasıyla log-sigmoid (logsig) ve doğrusal transfer fonksiyonu (pureline) kullanılmıştır.

Eğitim aşamasında, Otokodlayıcı 1000 iterasyon boyunca eğitilmiş ve ağırlık katsayıları Ölçeklenmiş Eşlenik Gradyan İnişi Algoritması ile güncellenmiştir. Deneysel sonuçlar, kayan pencere boyutunun 100 ve kod uzunluğunun ise 50 olduğu durumda en iyi başarımın elde edildiğini göstermiştir.

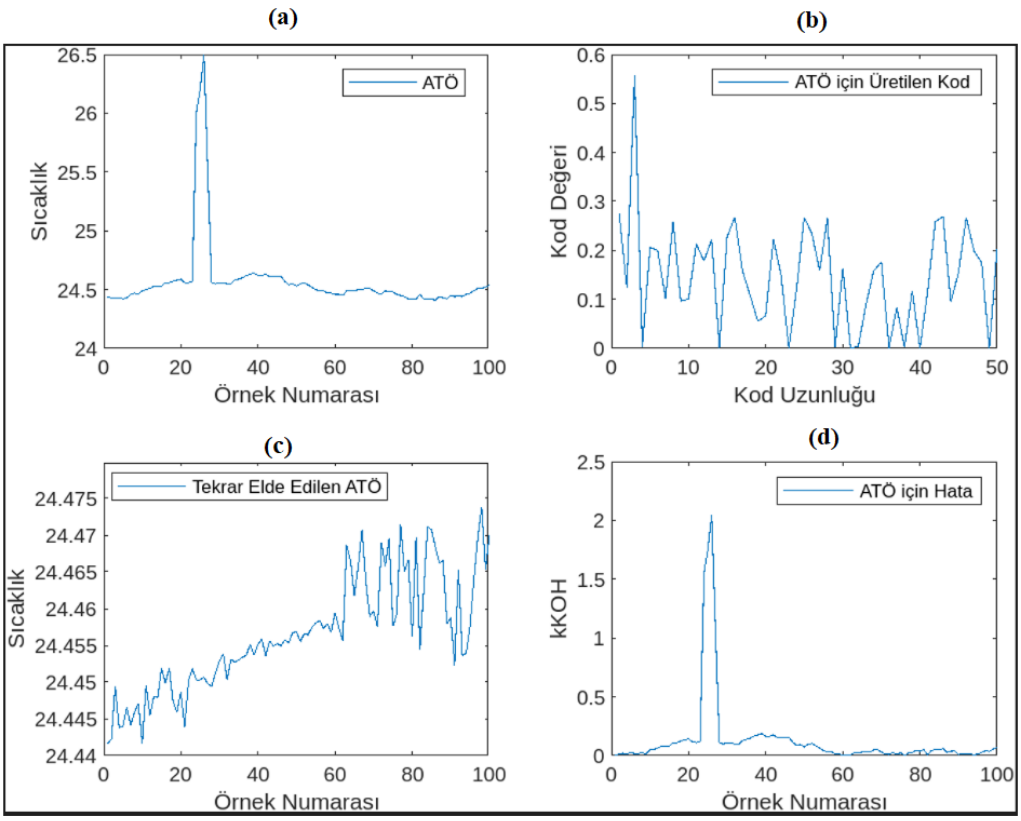
Otokodlayıcı ile noktasal anomali tespitine ait görsel bir yorum Şekil 4.5 ve Şekil 4.6.'da sunulmuştur. Şekil 4.5.'de noktasal anomalilerin yer almadığı bir kayan pencere için sonuçlar görselleştirilirken, Şekil 4.6.'da ise noktasal anomalilerin yer aldığı bir kayan pencereye ait sonuçlar sunulmuştur. Şekillerde kayan pencerede yer alan örnek değerler, kayan pencere için oluşturulan kodlar, kayan pencere için yeniden oluşturulmuş örnek değerler ve yeniden oluşturulma hataları gösterilmektedir.

Bir kayan pencere uzunluğundaki normal bir test örneği (NTÖ) Şekil 4.5.(a)'da gösterilmektedir. Otokodlayıcının NTÖ için ürettiği kod Şekil 4.5.(b)'de verilmiştir. NTÖ'nün yeniden oluşturulmuş hali Şekil 4.5.(c)'de gösterilmektedir. NTÖ ile yeniden oluşturulmuş NTÖ arasındaki ortalama karesel hata ise Şekil 4.5.(d)'de verilmiştir.



Şekil 4.5. Otokodlayıcı ile Noktasal Anomali Tespitinde; Normal Test Örneği için (a) NTÖ (b) NTÖ Kodu (c) Yeniden Oluşturulan NTÖ (d) NTÖ Hatası

Bir kayan pencere uzunluğundaki anomali bir test örneği (ATÖ) Şekil 4.6.(a)'da gösterilmektedir. Otokodlayıcının ATÖ için ürettiği kod Şekil 4.6.(b)'de verilmiştir. ATÖ'nün yeniden oluşturulmuş hali Şekil 4.6.(c)'de gösterilmektedir. ATÖ ile yeniden oluşturulmuş ATÖ arasındaki ortalama karesel hata ise Şekil 4.6.(d)'de verilmiştir.



Şekil 4.6. Otokodlayıcı ile Noktasal Anomali Tespitinde; Anomali Test Örneği için (a) ATÖ (b) ATÖ Kodu (c) Yeniden Oluşturulan ATÖ (d) ATÖ Hatası

Otokodlayıcı ile noktasal anomalilerin tespiti için gerçekleştirilen deney sonuçları bir hata matrisi kullanılarak Çizelge 4.1.'de sunulmuştur.

Çizelge 4.1. Otokodlayıcı için Hata Matrisi.

Tahmin edilen	Gerçekleşen	
	Anomali (Hatalı)	Normal
Anomali (Hatalı)	2089	0
Normal	73	287069

Otokodlayıcının noktasal anomali tespitindeki başarımına ait Doğruluk, Geri Çağırma ve Kesinlik ölçütlerine ait değerler Çizelge 4.1.'de verilen Hata Matrisi kullanılarak Çizelge 4.2.'de yer alan biçimde hesaplanmıştır.

Çizelge 4.2. Otokodlayıcı Başarım Sonuçları

	Doğruluk (Accuracy)	Duyarlılık (Recall)	Kesinlik (Precision)
Otokodlayıcı	%99.97	%99.05	%100

4.1.1.2. 3-SgOut ile Anomali Tespiti

Deneylerde, DC motor gövde sıcaklığındaki anomalilerin tespitinde Otokodlayıcının başarımını analiz etmek için 3-SgOut Aykırı Değer yöntemi bir kıyaslama yöntemi olarak kullanılmıştır. Bu nedenle Otokodlayıcının en iyi başarımı verdiği kayan pencere boyutu olan 100 örnekli pencere boyutu 3-SgOut içinde kullanılmıştır. 3-SgOut ile elde edilen deney sonuçları Çizelge 4.3.'de yer alan Hata Matrisi ile verilmiştir.

Çizelge 4.3. 3-SgOut için Hata Matrisi.

Tahmin edilen	Gerçekleşen	
	Anomali (Hatalı)	Normal
Anomali (Hatalı)	389	0
Normal	1773	287069

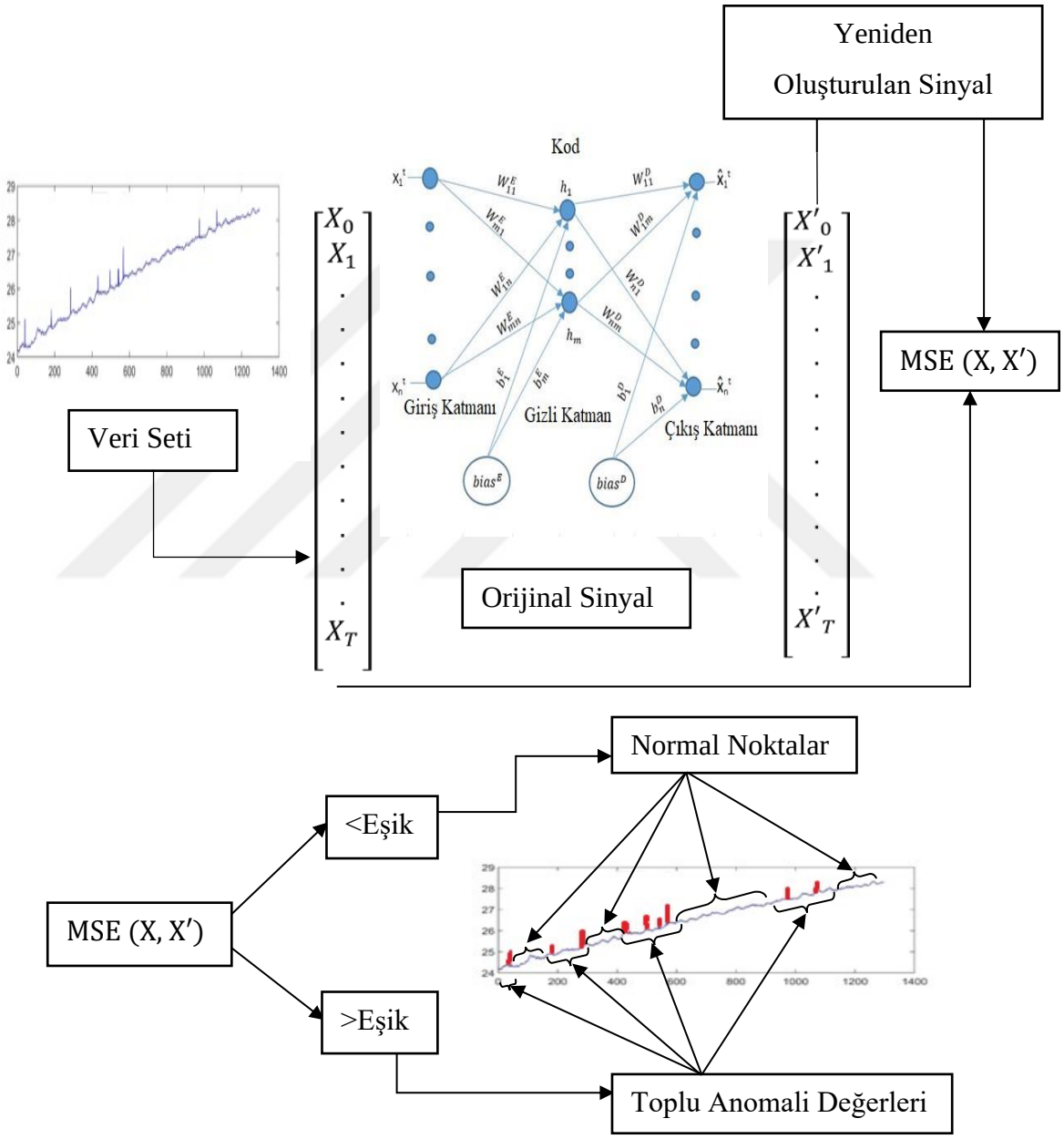
3-SgOut Aykırı Değer yönteminin noktasal anomali tespitindeki başarımına ait Doğruluk, Geri Çağırma ve Kesinlik ölçütlerine ait değerler Çizelge 4.4.'de verilen Hata Matrisi kullanılarak aşağıdaki biçimde hesaplanmıştır.

Çizelge 4.4. 3-SgOut Başarım Sonuçları

	Doğruluk (Accuracy)	Duyarlılık (Recall)	Kesinlik (Precision)
3-SgOut	%99.38	%17.99	%100

4.1.2. Toplu Anomali Tespiti

Bu yaklaşımda sıcaklık sensöründen gelen zaman serisi içerisinde yer alan toplu anomalilerin tespiti gerçekleştirilmiştir. Toplu anomali tespiti için bir görsel anlatım Şekil 4.7'de sunulmuştur.



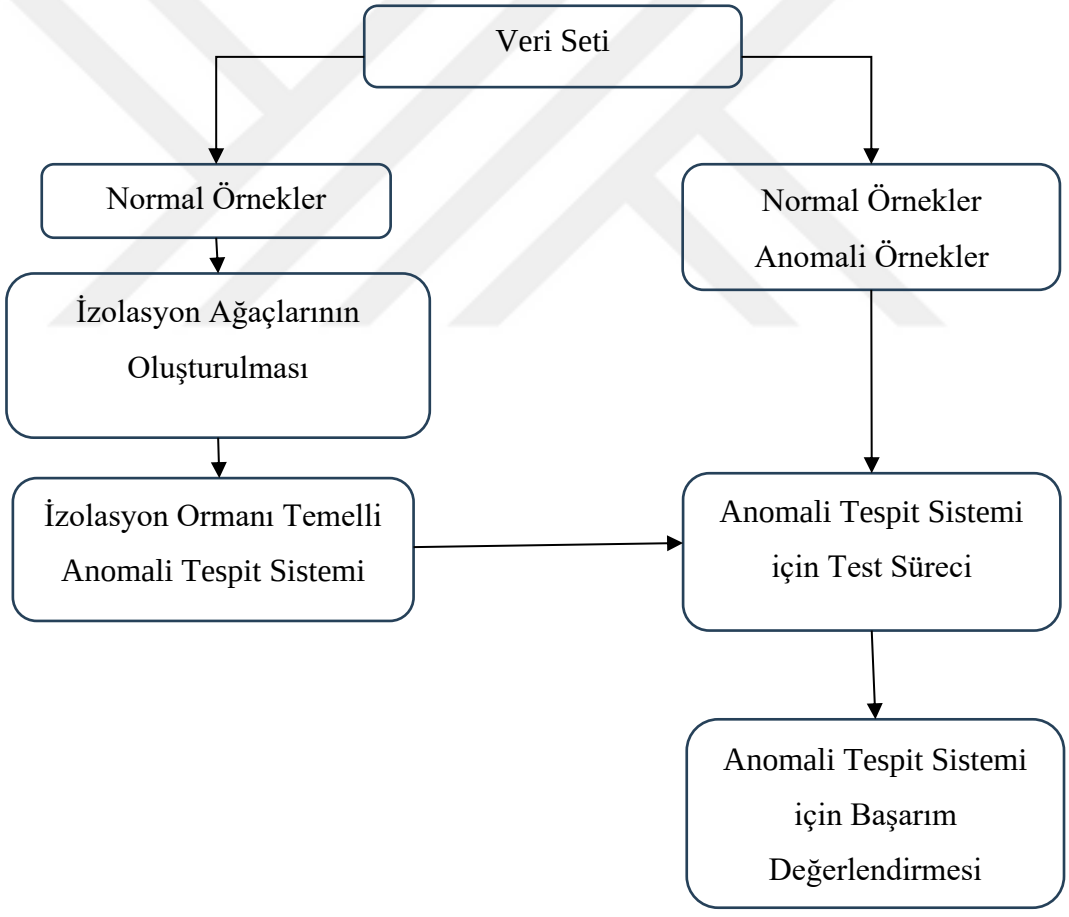
Şekil 4.7. Toplu anomali tespiti görsel anlatımı (Nicholaus vd., 2021)

Sıcaklık sensörüne ait zaman serileri içerisinde görülen toplu anomali değerlerine için örnekler, Çizelge 4.5'te verilmiştir.

Çizelge 4.5. Sıcaklık Sensörü Zaman Serilerinde Yer Alan Toplu Anomaliler

[24.8 25.1 25.5 25.4 26.2 25.7 25.3 25.4 23.9 24.2]	Normal Veri
[29.1 29.2 29.3 **30.5 28 28.1 30.6** 29.1 29.1 29.2]	Anomali Veri
[26 26.2 26.8 27.3 27.5 **29.1 28.7 30.4 31.2 32.3**]	Anomali Veri

Toplu anomalilerin tespitinde; Otokodlayıcı ve İzolasyon Ormanı yöntemleri kullanılmıştır. Otokodlayıcı ile toplu anomalilerin tespitinin gerçekleştirilmesine ait akış diyagramı noktasal anomali tespitinde kullanılan ve Şekil 4.3. ile verilen akış diyagramı ile aynıdır. İzolasyon ormanına ait akış diyagramı Şekil 4.8.'de verilmiştir.



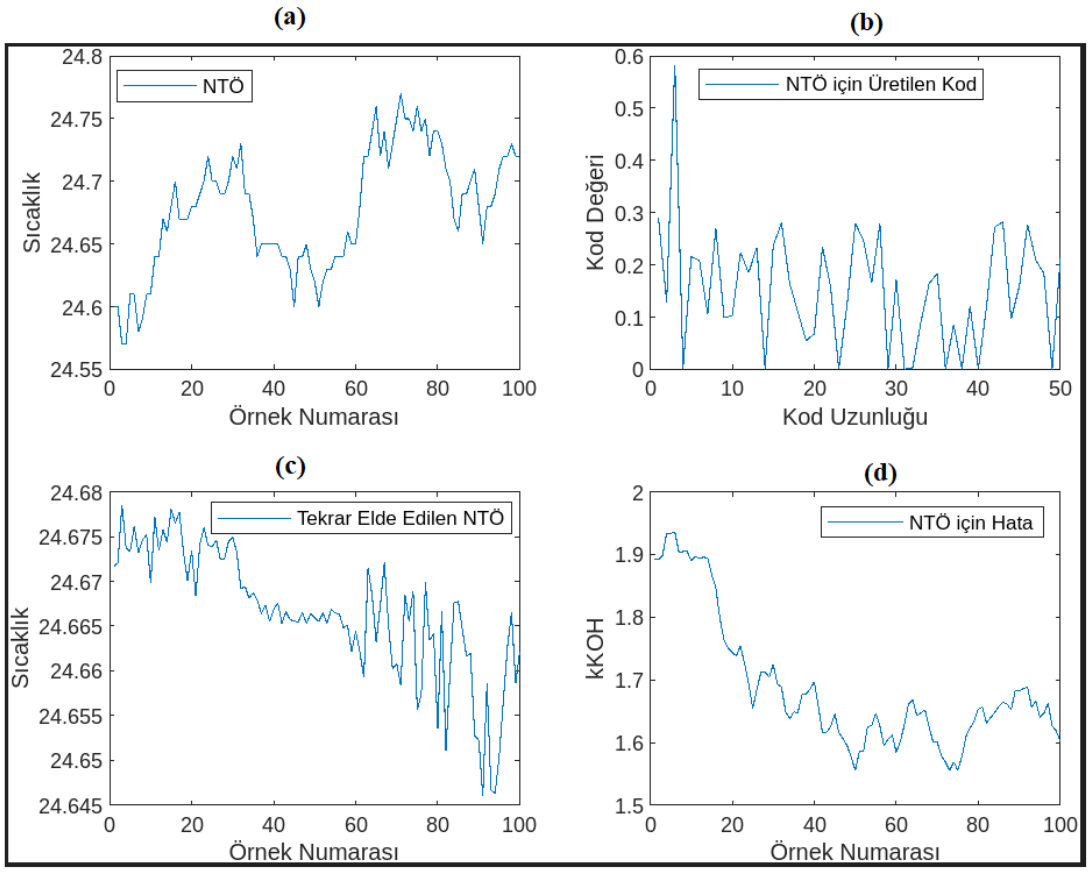
Şekil 4.8. İzolasyon Ormanı ile toplu anomalilerin tespiti için akış diyagramı

Veri setinde yer alan toplu anomalilerin tespiti Otokodlayıcı ve İzolasyon Ormanı yöntemleri kullanılarak Şekil 4.3. ve Şekil 4.8. ile verilen akış diyagramlarında belirtildiği biçimde gerçekleştirilmiştir. Her iki yönteme ait bulgular alt bölümlerde verilmiştir.

4.1.2.1. Otokodlayıcı ile Anomali Tespiti

Deneylerde, Şekil 3.6.'da verilen Otokodlayıcı mimarisi kullanılmıştır. Giriş ve çıkış katmanlarındaki nöron sayıları, kod uzunluğu, gizli katman ve çıkış katmanında yer alan aktivasyon fonksiyonları ile ağırlık katsayılarının güncellenmesinde kullanılan eğitilme algoritması noktasal anomali tespitindeki Otokodlayıcı yapısıyla aynı olarak seçilmiştir.

Otokodlayıcı ile toplu anomalilerin tespitine ait görsel bir yorum Şekil 4.9. ve Şekil 4.10.'da sunulmuştur. Şekil 4.9.'da toplu anomalilerin yer almadığı bir kayan pencere için sonuçlar görselleştirilirken, Şekil 4.10.'da ise toplu anomalilerin yer aldığı bir kayan pencereye ait sonuçlar sunulmuştur.



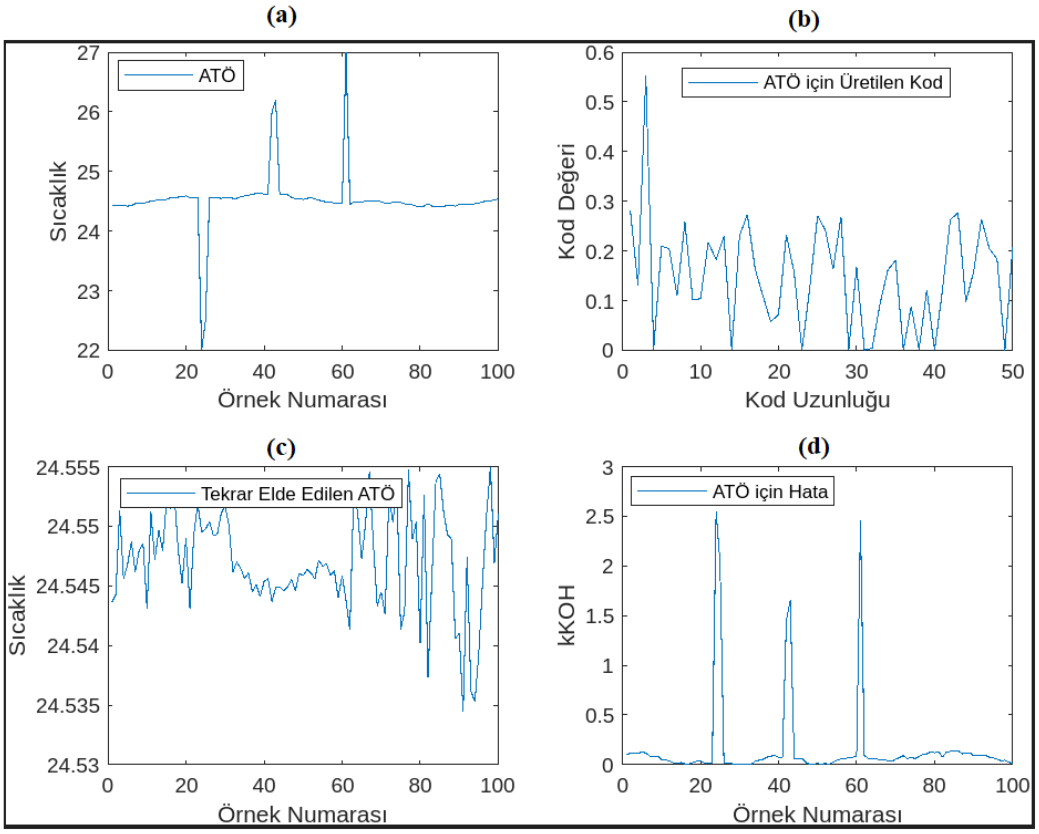
Şekil 4.9. Otokodlayıcı ile Toplu Anomali Tespitinde; Normal Test Örneği için (a) NTÖ (b) NTÖ Kodu (c) Yeniden Oluşturulan NTÖ (d) NTÖ Hatası

Bir kayan pencere uzunluğundaki normal bir test örneği (NTÖ), Şekil 4.9.(a)'da gösterilmektedir. Otokodlayıcının NTÖ için ürettiği kod Şekil 4.9.(b)'de verilmiştir. NTÖ'nün yeniden oluşturulmuş durumu Şekil 4.9.(c)'de gösterilmektedir. NTÖ ile yeniden oluşturulmuş NTÖ arasındaki ortalama karesel hata ise Şekil 4.9.(d)'de verilmiştir.

Şekillerde kayan pencerede yer alan örnek değerler, kayan pencere için oluşturulan kodlar, kayan pencere için yeniden oluşturulmuş örnek değerler ve yeniden oluşturulma hataları gösterilmektedir.

Bir kayan pencere uzunluğundaki anomali bir test örneği (ATÖ) Şekil 4.10.(a)'da gösterilmektedir. Otokodlayıcının ATÖ için ürettiği kod Şekil 4.10.(b)'de verilmiştir. ATÖ'nün yeniden oluşturulmuş durumu Şekil 4.10.(c)'de gösterilmektedir. ATÖ ile

yeniden oluşturulmuş, ATÖ arasındaki ortalama karesel hata ise Şekil 4.10.(d)'de verilmiştir.



Şekil 4.10. Otokodlayıcı ile Toplu Anomali Tespitinde; Anomali Test Örneği için (a) ATÖ (b) ATÖ Kodu (c) Yeniden Oluşturulan ATÖ (d) ATÖ Hatası

Otokodlayıcı ile toplu anomalilerin tespiti için gerçekleştirilen deney sonuçları bir hata matrisi kullanılarak Çizelge 4.6.'da sunulmuştur.

Çizelge 4.6. Otokodlayıcı için Hata Matrisi Çizelgesi

Tahmin edilen	Gerçekleşen	
	Anomali (Hatalı)	Normal
Anomali (Hatalı)	1899	291
Normal	0	1648

Otokodlayıcının toplu anomali tespitindeki başarımına ait Doğruluk, Geri Çağırma ve Kesinlik ölçütlerine ait değerler Çizelge 4.6.'da verilen Hata Matrisi kullanılarak Çizelge 4.7.'de yer alan biçimde hesaplanmıştır.

Çizelge 4.7. Otokodlayıcı için Başarım Sonuçları

	Doğruluk (Accuracy)	Geri Çağırma (Recall)	Kesinlik (Precision)
Otokodlayıcı	%93.39	%100	%86.71

4.1.2.2. İzolasyon Ormanı ile Anomali Tespiti

Deneylerde, DC motor gövde sıcaklığındaki toplu anomalilerin tespitinde, Otokodlayıcının başarımını analiz etmek için İzolasyon Ormanı yöntemi bir kıyaslama yöntemi olarak kullanılmıştır. Bu nedenle Otokodlayıcının en iyi başarımı verdiği kayan pencere boyutu olan 100 örnekli pencere boyutu İzolasyon Ormanı içinde kullanılmıştır. İzolasyon Ormanı ile elde edilen deney sonuçları Çizelge 4.8.'de yer alan Hata Matrisi ile verilmiştir.

Çizelge 4.8. İzolasyon Ormanı için Hata Matrisi Çizelgesi

Tahmin edilen	Gerçekleşen	
	Anomali (Hatalı)	Normal
Anomali (Hatalı)	1872	1060
Normal	27	879

İzolasyon Ormanı yönteminin toplu anomali tespitindeki başarımına ait Doğruluk, Geri Çağırma ve Kesinlik ölçütlerine ait değerler Çizelge 4.8.'de verilen Hata Matrisi kullanılarak Çizelge 4.9.'de yer alan biçimde hesaplanmıştır.

Çizelge 4.9. İzolasyon Ormanı için Başarım Sonuçları

	Doğruluk (Accuracy)	Geri Çağırma (Recall)	Kesinlik (Precision)
İzolasyon Ormanı	%72.43	%98.57	%63.85

4.2. Tartışma

Bu bölümde noktasal anomali ve toplu anomali olmak üzere anomali tespit yöntemlerinin başarım sonuçları iki ayrı alt bölümde karşılaştırılmıştır.

4.2.1. Noktasal Anomali Tespiti

Otokodlayıcı ve 3-SgOut Yöntemlerinin başarım sonuçları karşılaştırılması Doğruluk, Geri Çağırma ve Kesinlik olarak Çizelge 4.10.'da verilmiştir.

Çizelge 4.10. Otokodlayıcı ve 3-SgOut için Başarım Sonuçları

	Doğruluk	Geri Çağırma	Kesinlik
Otokodlayıcı	%99.97	%99.05	%100
3-SgOut	%99.38	%17.99	%100

Her iki yöntemin Doğruluklarının %99'un üzerinde ve Kesinliklerinin %100'e eşit olduğu, Geri Çağırmaların ise Otokodlayıcı ve 3-SgOut için sırasıyla %99.05 ve %17.99 olduğu görülmektedir.

Doğruluk, doğru şekilde tespit edilen (normal veya anomali) numunelerin toplam numunelere oranıdır. Geri çağırma, yöntemin anomalilik içeren örnekleri doğru şekilde tespit etme yeteneğini gösterir. Kesinlik, anomaliliğe sahip tüm tespit edilen numuneler arasında gerçekten tespit edilen anomali numunelerin yüzdesini verir.

Veri setindeki anomalili örnek sayısının %1'den az olması nedeniyle doğruluk oranı, yöntemlerin başarım karşılaştırması açısından anlamlı değildir. Diğer iki ölçüm olan Kesinlik ve Geri Çağırma, başarımı hakkında önemli bilgiler verir. Her iki yöntem de %100 Kesinlik oranlarına sahiptir. Bu, yöntemlerin yanlış pozitif hatası olmadığı anlamına gelir.

3-SgOut'un geri çağırma oranı %17.99 iken Otokodlayıcı için bu oran %99.05'tir. Yöntemlerin Geri Çağırma oranlarına göre 3-SgOut'un toplam anomali örneklerinin

yalnızca %17.99'unu anomali olarak tespit ettiği, Otokodlayıcının ise %99.05 oranında doğru tespit ettiği sonucuna varılmıştır.

3-SgOut yönteminin bir eğitim aşaması yoktur ve yalnızca verilerin ortalamasına ve standart sapmasına dayanarak aykırı değerleri tespit eder. 3-SgOut'un sonuçları, sıcaklık verilerindeki anomaliliklerin tespitinin yalnızca aykırı değerlerin bulunmasıyla etkili bir şekilde yapılamayacağını göstermektedir.

4.2.2. Toplu Anomali Tespiti

Otokodlayıcı ve İzolasyon Ormanı Yöntemlerinin başarım sonuçlarının karşılaştırılması Doğruluk, Geri Çağırma ve Kesinlik olarak Çizelge 4.11.'de verilmiştir.

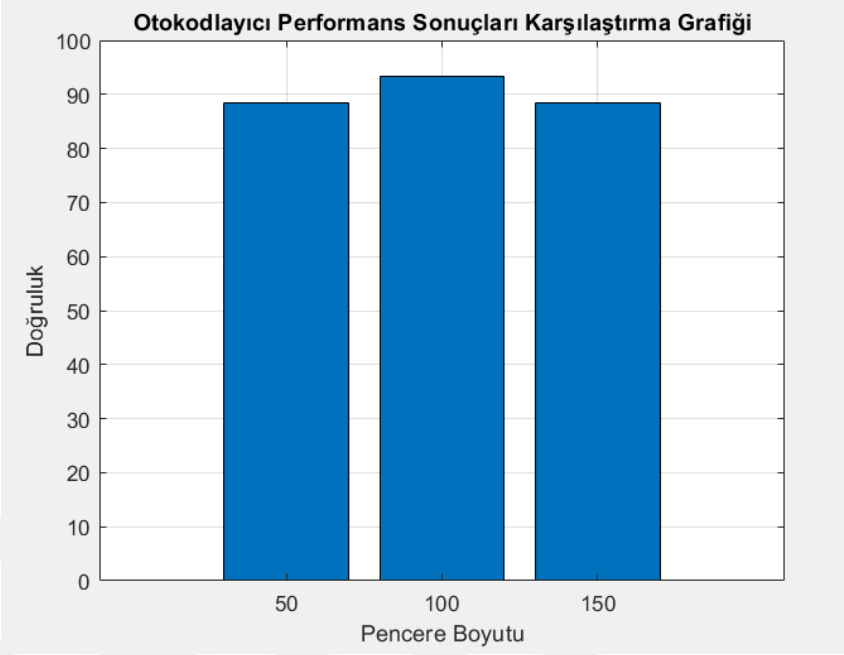
Çizelge 4.11. Otokodlayıcı ve İzolasyon Ormanı için Başarım Sonuçları

	Doğruluk	Geri Çağırma	Kesinlik
Otokodlayıcı	%93.39	%100	%86.71
İzolasyon Ormanı	%72.43	%98.57	%63.85

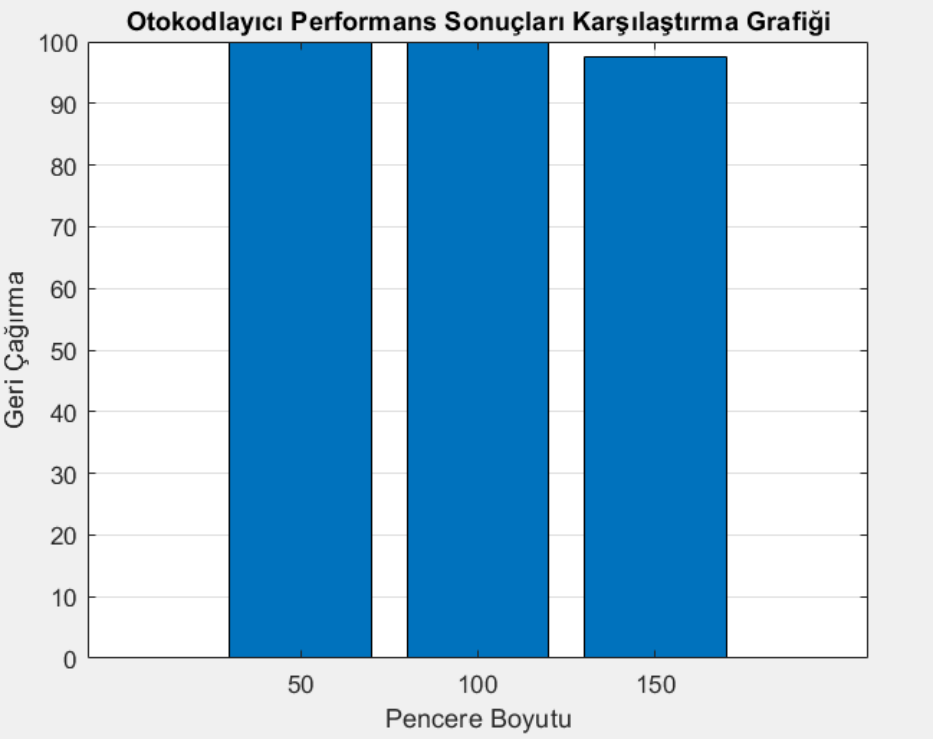
Toplu anomali tespitinde Otokodlayıcı yönteminin Doğruluk oranı %93.39 iken İzolasyon Ormanı Yönteminin doğruluk oranı %72.43 çıkmıştır. Geri Çağırma yüzdeleri ise Otokodlayıcı için %100 ve İzolasyon Ormanı için %98.57 çıkmıştır. Kesinlik olarak yüzdeleri Otokodlayıcı için %86.71 ve izolasyon ormanı için %63.85 olduğu gözükmektedir.

4.2.3. Pencere Boyutu Seçimi ve Başarım Sonuçlarına Etkisi

Anomali tespit denemelerinden farklı pencere boyutları ile çalışılmıştır. Izgara arama yöntemi ile başarımın en yüksek olduğu pencere boyutu belirlenmiştir. Aşağıda denemelerde kullanılan üç farklı pencere boyutu (50,100,150) için başarım değerlerinin değişimi grafiksel olarak sunulmuştur.

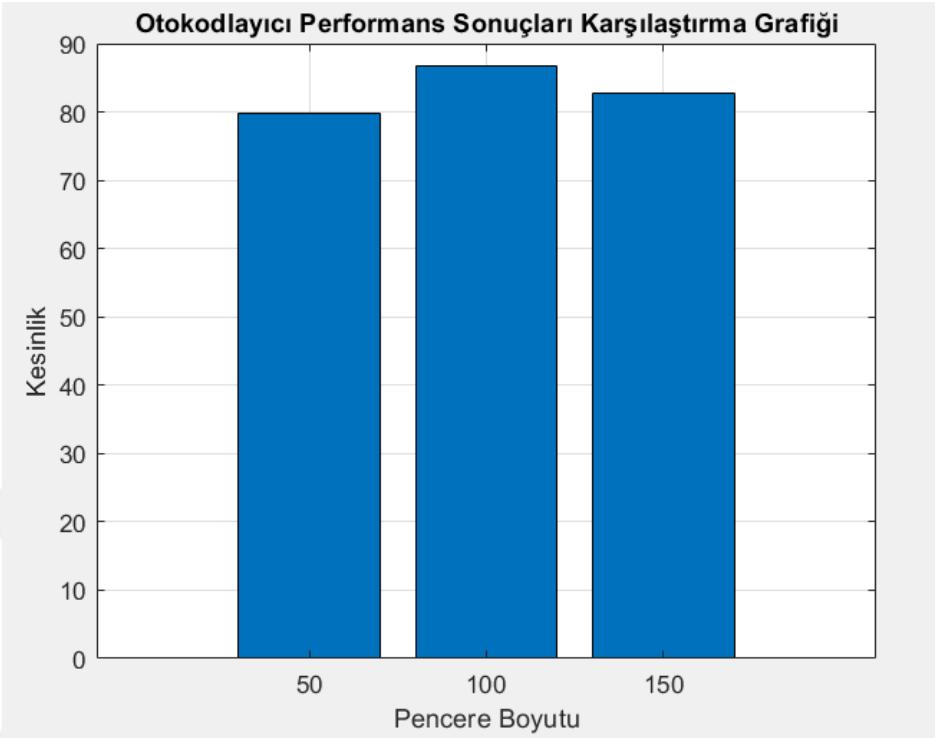


Şekil 4.11. Otokodlayıcının Doğruluk Yüzdesinin Pencere Boyutu ile Değişimi



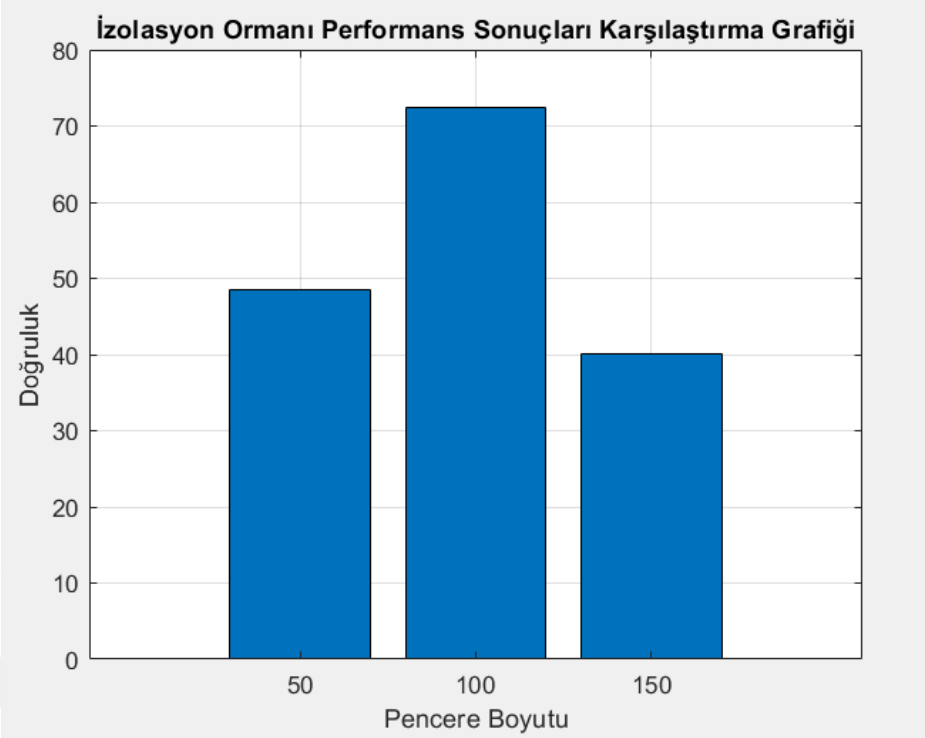
Şekil 4.12. Otokodlayıcının Geri Çağırma Yüzdesinin Pencere Boyutu ile Değişimi

Otokodlayıcı için başarım sonuçlarının pencere boyutu ile değişimine ait grafikler Şekil 4.11-Şekil 4.13'te sunulmuştur.

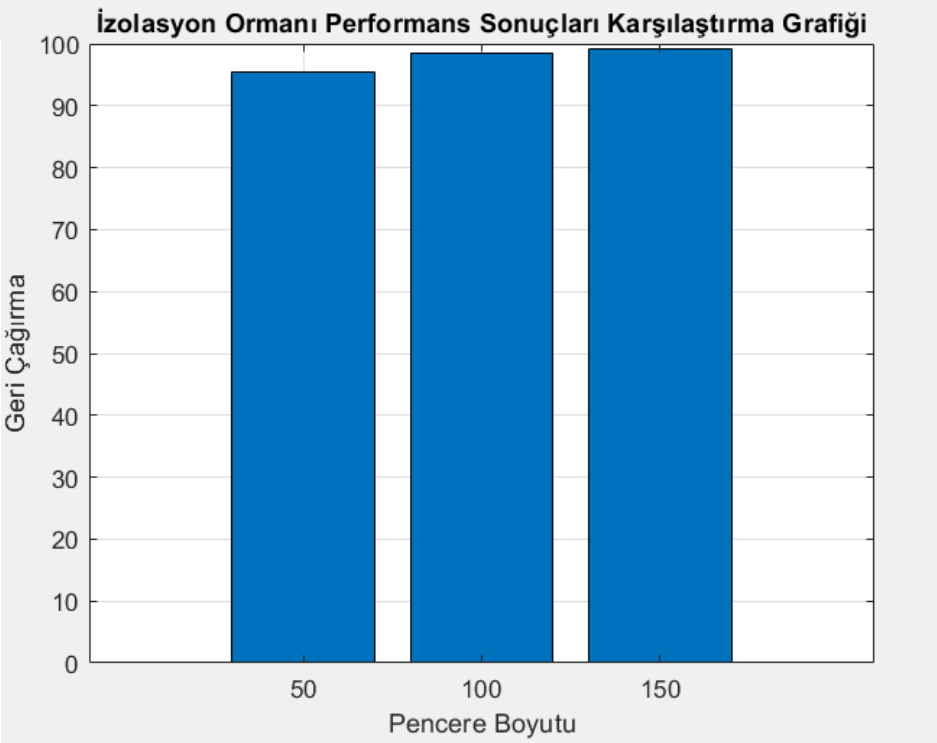


Şekil 4.13. Otokodlayıcının Kesinlik Yüzdesinin Pencere Boyutu ile Değişimi

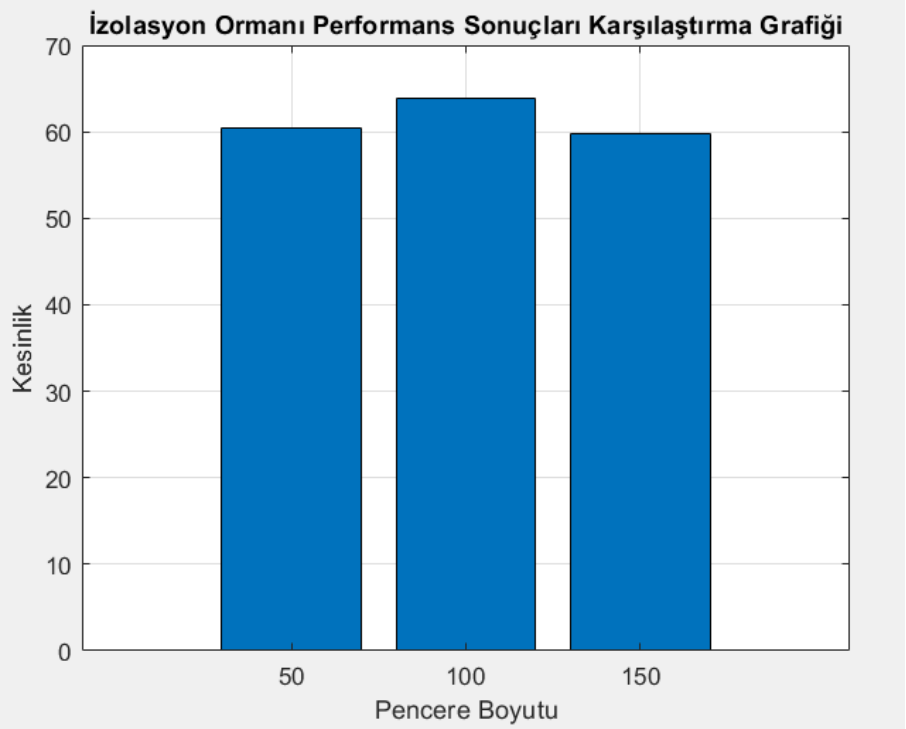
İzolasyon ormanı için başarım sonuçlarının pencere boyutu ile değişimine ait grafikler ise Şekil 4.14-Şekil 4.16'da sunulmuştur.



Şekil 4.14. İzolasyon Ormanının Doğruluk Yüzdesinin Pencere Boyutu ile Değişimi



Şekil 4.15. İzolasyon Ormanının Geri Çağırma Yüzdesinin Pencere Boyutu ile Değişimi



Şekil 4.16. İzolasyon Ormanının Kesinlik Yüzdesinin Pencere Boyutu ile Değişimi

Başarım sonuçlarından da görüleceği üzere tüm değerlendirmede en iyi sonucun pencere boyutunun 100 örnek olduğu durum için alındığı gözlemlenmiştir. Bu nedenle, tez çalışmasında gerçekleştirilen anomali tespiti sonuçları ve karşılaştırmaları pencere boyutunun 100 örnek olduğu durum için sunulmuştur.

5. SONUÇ

Bu çalışmada DC motor sağlık izlemesinde temel gösterge olan gövde sıcaklığındaki anomalilerin tespiti Otokodlayıcılar ile gerçekleştirilmiştir. Gövde sıcaklığı verisinde yer alan noktasal ve toplu anomalilerin tespitinde otokodlayıcıların başarım analizleri yapılmıştır. Çalışmada kullanılan veri seti, bu çalışma için özel olarak tasarlanmış bir DC motor deney düzeneği kullanılarak oluşturulmuştur. Bu veri seti üzerinde başarım incelemesi; Doğruluk, Geri Çağırma ve Kesinlik ölçütleri kullanılarak yapılmıştır. Başarım karşılaştırmasında, noktasal anomaliler için 3-SgOut Aykırı Değer yöntemi ve toplu anomaliler için İzolasyon Ormanı yöntemi kullanılmıştır.

Noktasal anomalilerin tespiti için yapılan deneyler ile Otokodlayıcı ve 3-SgOut Aykırı Değer yöntemlerinin başarım karşılaştırması yapılmıştır. Otokodlayıcı'nın Doğruluk, Geri Çağırma ve Kesinlik için sırasıyla %99.97, %99.05 ve %100 gibi dikkat çekici başarım sonuçları elde ettiği ve DC motor gövde sıcaklığındaki noktasal anomali tespitinde başarıyla kullanılabileceği görülmüştür.

Toplu anomalilerin tespiti için yapılan deneyler ile Otokodlayıcı ve İzolasyon Ormanı yöntemlerinin başarım karşılaştırması yapılmıştır. Otokodlayıcı'nın Doğruluk, Geri Çağırma ve Kesinlik için sırasıyla %93.39, %100, %86.71 gibi yüksek başarım sonuçları elde ettiği ve DC motor gövde sıcaklığındaki toplu anomali tespitinde başarıyla kullanılabileceği görülmüştür.

Her iki anomali türünde göstermiş olduğu yüksek tespit başarımı nedeniyle Otokodlayıcı temelli anomali tespit yaklaşımlarının endüstriyel sistemlerde farklı sensör tiplerinden gelen zaman serisi verilerindeki anomalilerin tespitinde başarıyla uygulanabileceği gösterilmiştir.

Bu çalışmanın gelecekteki araştırma yönleri iki ana başlıkta toplanabilir. Bunlardan ilki, önerilen Otokodlayıcı tabanlı anomali tespit yönteminin çok sensörlü verilerin üzerindeki başarımlarının incelenmesidir. İkinci konu ise çok sensörlü verilerle çalışırken Otokodlayıcı mimarisinde derin öğrenme yöntemlerinin kullanılmasıdır.

KAYNAKLAR

Abdoli, M., Akbari, M., & Shahrabi, J. (2023). Bagging Supervised Autoencoder Classifier for Credit Scoring. *Expert Systems with Applications*, *213*, 118991.

Abraham, J. A., & Shrivastava, S. (2018). DC Motor Speed Control Using Machine Learning Algorithm. *International Journal of Engineering Research Technology*, *7*(04), 33-37.

Acar, O. C. (2019). *Akış verisinde anormallik tespiti* (Master's thesis, Fen Bilimleri Enstitüsü).

Ahmad, S., Styp-Rekowski, K., Nedelkoski, S., & Kao, O. (2020, December). Autoencoder-based condition monitoring and anomaly detection method for rotating machines. In *2020 IEEE International Conference on Big Data (Big Data)* (pp. 4093-4102). IEEE.

Ammann, O., Michau, G., & Fink, O. (2020). Anomaly Detection And Classification In Time Series With Kervolutional Neural Networks. *arXiv preprint arXiv:2005.07078*.

Antonenko, E., & Read, J. (2023). Chains of Autoreplicative Random Forests for missing value imputation in high-dimensional datasets. *arXiv preprint arXiv:2301.00595*.

Bayraktar, C., & Gökçen, H. (2020). Yüksek Raflı Depolama Sistemlerinin Enerji Optimizasyonunda Anomali Tespiti İçin Sınıflama Algoritmalarının Karşılaştırılması. *Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi*, *4*(2), 89-109.

Bilgin, M. (2017). Gerçek veri setlerinde klasik makine öğrenmesi yöntemlerinin performans analizi. *Breast*, *2*(9), 683.

Borghesi, A., Bartolini, A., Lombardi, M., Milano, M., & Benini, L. (2019). A semisupervised autoencoder-based approach for anomaly detection in high performance computing systems. *Engineering Applications of Artificial Intelligence*, *85*, 634-644.

Brummer, B., & De Vleeschouwer, C. (2023). On the Importance of Denoising when Learning to Compress Images. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2440-2448).

Basheer, I.A., Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of Microbiological Methods*, 43(1), 3–31.

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys*, *41*(3), 1-58.

Chen, Z., Yeo, C. K., Lee, B. S., & Lau, C. T. (2018, April). Autoencoder-based network anomaly detection. In *2018 Wireless telecommunications symposium (WTS)* (pp. 1-5). IEEE.

Çalış, A., Kayapınar, S., & Çetinyokuş, T. (2014). Veri Madenciliğinde Karar Ağacı Algoritmaları ile Bilgisayar ve İnternet Güvenliği Üzerine Bir Uygulama. *Endüstri Mühendisliği*, *25*(3), 2-19.

da Silva Ferreira, M., Vismari, L. F., Cugnasca, P. S., de Almeida, J. R., Camargo, J. B., & Kallemback, G. (2019, December). A comparative analysis of unsupervised learning techniques for anomaly detection in railway systems. In *2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA)* (pp. 444-449). IEEE.

Das, T. K., Adepu, S., & Zhou, J. (2020). Anomaly detection in industrial control systems using logical analysis of data. *Computers & Security*, *96*, 101935.

Daylight, E. G. (2015). Towards a historical notion of 'Turing—the father of Computer Science'. *History and Philosophy of Logic*, *36*(3), 205-228.

Demircioğlu, E. H., & Yılmaz, E. (2023). A Method Based on an Autoencoder for Anomaly Detection in DC Motor Body Temperature. Applied Sciences, 13(15), 8701.

Di Martino, T., Guinvarc'h, R., Thirion-Lefevre, L., & Colin, E. (2022). FARMSAR: Fixing AgRicultural Mislabels Using Sentinel-1 Time Series and AutoencodeRs. *Remote Sensing*, *15*(1), 35.

Faramarzi, A., & Sabahi, K. (2011, August). Recurrent fuzzy neural network for DC-motor control. In *2011 Fifth International Conference on Genetic and Evolutionary Computing* (pp. 93-96). IEEE.

Ferreira, F. J., de Almeida, A. T., Deprez, W., Belmans, R., & Baoming, G. (2007). Impact of steady-state voltage supply anomalies on three-phase squirrel-cage induction motors. In *2007 International Aegean Conference on Electrical Machines and Power Electronics* (pp. 607-615). IEEE.

Givnan, S., Chalmers, C., Fergus, P., Ortega-Martorell, S., & Whalley, T. (2022). Anomaly detection using autoencoder reconstruction upon industrial motors. *Sensors*, *22*(9), 3166.

Han, J., Pei, J., & Tong, H. (2022). *Data mining: concepts and techniques*. Morgan kaufmann.

Hassanaly, R., Bottani, S., Sauty, B., Colliot, O., & Burgos, N. (2023, April). Simulation based evaluation framework for deep learning unsupervised anomaly detection on brain FDG-PET. In *Medical Imaging 2023: Image Processing* (Vol. 12464, pp. 511-518). SPIE.

Haykin, S. (1999). Neural networks-A comprehensive foundation. New Jersey: Prentice Hall.

Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A. (2021). Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy*, *287*, 116601.

Hiruta, T., Maki, K., Kato, T., & Umeda, Y. (2021). Unsupervised Learning Based Diagnosis Model for Anomaly Detection of Motor Bearing with Current Data. *Procedia CIRP*, *98*, 336-341.

Hızarcı, B., Ümütlü, R. C., Kıral, Z., & Öztürk, H. J. S. A. S. (2021). Fault severity detection of a worm gearbox based on several feature extraction methods through a developed condition monitoring system. *SN Applied Sciences*, *3*(1), 1-12.

Homayouni, H., Ghosh, S., Ray, I., Gondalia, S., Duggan, J., & Kahn, M. G. (2020). An autocorrelation-based LSTM-Autoencoder for anomaly detection on time-series data. In *2020 IEEE international conference on big data (big data)* (pp. 5068-5077). IEEE.

Jami, A., & Heyns, P. S. (2018). Impeller fault detection under variable flow conditions based on three feature extraction methods and artificial neural networks. *Journal of Mechanical Science and Technology*, *32*, 4079-4087.

Jamjoom, M., Mahmoud, A. M., Abbas, S., & Hodhod, R. (2023). Gaussian Mixture with Max Expectation Guide for Stacked Architecture of Denoising Autoencoder and DRBM for Medical Chest Scans and Disease Identification. *Electronics*, 105-105.

Kagermann, H., Lukas, W. D., & Wahlster, W. (2011). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. VDI nachrichten, 13(1), 2-3

Kalıpcıoğlu, K. C., Toğay, C., & Yolaçan, E. N. (2019). Son Kullanıcılar için Anomali Saldırı Tespit Sistemleri. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, *27*(3), 199-212.

Kasunic, M., McCurley, J., Goldenson, D., & Zubrow, D. (2011). An investigation of techniques for detecting data anomalies in earned value management data. *Software Engineering Institute: Pittsburgh, PA, USA*.

Kılıç, S. (2023). Yapay Zekâ & Teori ve Uygulamalar.

Kim, K. G. (2016). Book review: Deep learning. *Healthcare informatics research*, *22*(4), 351-354.

Kirchgässner, W., Wallscheid, O., & Böcker, J. (2020). Estimating electric motor temperatures with deep residual machine learning. *IEEE Transactions on Power Electronics*, *36*(7), 7480-7488.

Köse, U. (2017). *Yapay zekâ tabanlı optimizasyon algoritmaları geliştirilmesi* (Doctoral dissertation, Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği ABD).

Lambrecht, L., & CMS Collaboration. (2023). Prospects for computer-assisted data quality monitoring at the CMS pixel detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *1045*, 167495.

Lesouple, J., Baudoin, C., Spigai, M., & Tourneret, J. Y. (2021). Generalized isolation forest for anomaly detection. *Pattern Recognition Letters*, 149, 109-119

Li, Z., Li, J., Wang, Y., & Wang, K. (2019). A deep learning approach for anomaly detection based on SAE and LSTM in mechanical equipment. *The International Journal of Advanced Manufacturing Technology*, *103*, 499-510.

Liu, Y., Ponce, C., Brunton, S. L., & Kutz, J. N. (2023). Multiresolution convolutional autoencoders. *Journal of Computational Physics*, *474*, 111801.

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008, December). Isolation forest. In *2008 eighth ieee international conference on data mining* (pp. 413-422). IEEE.

Lopes, L. T., & Pedronette, D. C. G. (2023). Self-Supervised Clustering based on Manifold Learning and Graph Convolutional Networks. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 5634-5643).

Lu, H., Li, Y., Mu, S., Wang, D., Kim, H., & Serikawa, S. (2017). Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. *IEEE internet of things journal*, *5*(4), 2315-2322.

Ma, H., Huang, C. J., Chen, C., Dong, D., Wang, Y., Wu, R. B., & Xiang, G. Y. (2023). On compression rate of quantum autoencoders: Control design, numerical and experimental realization. *Automatica*, *147*, 110659.

Maasoum, S. M. H., Mostafavi, A., & Sadighi, A. (2020). An autoencoder-based algorithm for fault detection of rotating machines, suitable for online learning and standalone applications. In *2020 6th iranian conference on signal processing and intelligent systems (ICSPIS)* (pp. 1-6). IEEE.

Masoumian, A., Rashwan, H. A., Abdulwahab, S., Cristiano, J., Asif, M. S., & Puig, D. (2023). Gcndepth: Self-supervised monocular dest estimation based on graph convolutional network. *Neurocomputing*, *517*, 81-92.

MATLAB Toolbox, S. M. (1993). Matlab. *Mathworks Inc*.

Meire, M., & Karsmakers, P. (2019). Comparison of deep autoencoder architectures for real-time acoustic based anomaly detection in assets. In *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)* (Vol. 2, pp. 786-790). IEEE.

Metlek, S., & Kayaalp, K. (2020). Makine Öğrenmesinde, Teoriden Örnek MATLAB Uygulamalarına Kadar Destek Vektör Makineleri. *İksad Yayınevi*.

Mirzaee, S., Motagh, M., Arefi, H., & Nooryazdan, M. (2014). Classification of agricultural fields using time series of dual polarimetry TerraSAR-X images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *40*, 191-196.

Nandi, S., Toliyat, H. A., & Li, X. (2005). Condition monitoring and fault diagnosis of electrical motors—A review. *IEEE transactions on energy conversion*, *20*(4), 719-729.

Nicholaus, I. T., Park, J. R., Jung, K., Lee, J. S., & Kang, D. K. (2021). Anomaly detection of water level using deep autoencoder. *Sensors*, *21*(19), 6679.

Orhan, U. (2015). Makine Öğrenmesi ders notları. *Çukurova Üniversitesi http://bmb. cu. edu. tr/uorhan/MLearn. htm*, *2*.

Ötleş, S., & Özyurt, V. H. (2016). Endüstri 4.0: Büyüme ve Verimlilik İçin Dijitalleşme. Plastik&Ambalaj Teknolojisi Dergisi, 4, 54-57.

Özbay, Ö. (2015). Veri madenciliği kavramı ve eğitimde veri Madenciliği uygulamaları. *Uluslararası Eğitim Bilimleri Dergisi* (5), 262-272.

Öztürk, K., & Şahin, M. E. (2018). Yapay sinir ağları ve yapay zekâ'ya genel bir bakış. *Takvim-i Vekayi*, *6*(2), 25-36.

Rahman, M. M., Panda, R., & Alam, M. A. U. (2023). Semi-Supervised Domain Adaptation with Auto-Encoder via Simultaneous Learning. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 402-411).

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

Roy, M., Bose, S. K., Kar, B., Gopalakrishnan, P. K., & Basu, A. (2018, November). A stacked autoencoder neural network based automated feature extraction method for anomaly detection in on-line condition monitoring. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1501-1507). IEEE.

Saad, S. A. (2015). *The Utilisation of Information Available in a Sensorless Control System of an AC Induction Motor for Condition Monitoring* (Doctoral dissertation, University of Huddersfield).

Samuel, A. L. (1959). Machine learning. *The Technology Review*, *62*(1), 42-45.

Sauvalle, B., & de La Fortelle, A. (2023). Autoencoder-based background reconstruction and foreground segmentation with background noise estimation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3244-3255).

Seitz, P., Scherdel, C., Reichenauer, G., & Schmitt, J. (2023). Machine Learning in the development of Si-based anodes using Small-Angle X-ray Scattering for structural property analysis. *Computational Materials Science*, *218*, 111984.

Siddalingappa, R., & Kanagaraj, S. (2021). Anomaly detection on medical images using autoencoder and convolutional neural network. *International Journal of Advanced Computer Science and Applications*, (7).

Shaukat, K., Alam, T. M., Luo, S., Shabbir, S., Hameed, I. A., Li, J., Abbas, S. N. & Javed, U. (2021). A review of time-series anomaly detection techniques: A step to future perspectives. In Advances in Information and Communication: Proceedings of the 2021 Future of Information and Communication Conference (FICC), Volume 1 (pp. 865-877). Springer International Publishing.

Staudemeyer, R. C., & Morris, E. R. (2019). Understanding LSTM--a tutorial into long short-term memory recurrent neural networks. *arXiv preprint arXiv:1909.09586*.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Takiddin, A., Ismail, M., Zafar, U., & Serpedin, E. (2022). Deep autoencoder-based anomaly detection of electricity theft cyberattacks in smart grids. *IEEE Systems Journal*, *16*(3), 4106-4117.

Tavares, A., Lopes, B., Di Lorenzo, E., Cornelis, B., Peeters, B., Desmet, W., & Gryllias, K. (2022, June). Machine Learning Techniques for Damage Detection in Wind Turbine Blades. In *European Workshop on Structural Health Monitoring* (pp. 176-189). Cham: Springer International Publishing.

Theumer, P., Zeiser, R., Trauner, L., & Reinhart, G. (2021). Anomaly detection on industrial time series for retaining energy efficiency. *Procedia CIRP*, *99*, 33-38.

Toklu, Y. C., Bekdaş, G., & Nigdeli, S. M., (2021). S-20 Sezgisel Algoritmalarla Toplam Potansiyel Enerji Optimizasyonu Yönteminin Yapısal Çözümlemede Üstünlükleri ve Gelişmeler. *In 1st International Congress On Artificial Intelligence And Data Science Proceeding Book* (P. 144).

Tolpin, D. (2019). Population anomaly detection through deep gaussianization. Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing,

TÜSİAD. (2016). Türkiye'nin Küresel Rekabetçiliği İçin Bir Gereklilik Olarak Sanayi 4.0 Gelişmekte Olan Ekonomi Perspektifi. The Boston Consulting Group. Yayın No: TÜSİADT/2016-03/576. İstanbul.

Übeyli, E. D. (2009). Combined neural networks for diagnosis of erythemato-squamous diseases. *Expert Systems with Applications*, 36 (3), 5107–5112.

Volya, D., & Mishra, P. (2023). Quantum Data Compression for Efficient Generation of Control Pulses.

Ye, F., & Bors, A. G. (2023). Lifelong Variational Autoencoder via Online Adversarial Expansion Strategy.

Yılmaz, E. (2016). Fetal state assessment from cardiotocogram data using artificial neural networks. Journal of Medical and Biological Engineering, 36, 820-832.

Zavrak, S. (2020). Yazılım-tanımlı ağlar için uyarlanabilir bir saldırı tespit ve önleme sistemi tasarımı.

Zhang, B., Hou, K., Dong, Z., Liu, Z., & Wang, X. (2023). Stacked-denoising-auto-encoder-based reliability assessment method for power system operation. *Energy Reports*, *9*, 505-511.

Zhao, J., Liu, K., Wang, W., & Liu, Y. (2014). Adaptive fuzzy clustering based anomaly data detection in energy system of steel industry. *Information Sciences*, *259*, 335-345.

Zhao, Y., Lehman, B., Ball, R., Mosesian, J., & de Palma, J.-F. (2013). Outlier detection rules for fault detection in solar photovoltaic arrays. In 2013 *twenty-eighth annual IEEE applied power electronics conference and exposition (APEC)*, (pp. 2913-2920). IEEE.

Zhu, Y., Zhou, X., & Wu, X. (2023). Unsupervised Domain Adaptation via Stacked Convolutional Autoencoder. *Applied Sciences*, *13*(1), 481.