MARMARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ İLE GEMİ MAKİNELERİNİN HATA ANALİZİ

FUNDA KAYA İNCEİŞÇİ

YÜKSEK LİSANS TEZİ
Bilgisayar Mühendisliği
Anabilim Dalı
Bilgisayar Mühendisliği Programı

DANIŞMAN
Dr. Öğr. Ayça AK

EŞ-DANIŞMAN
Prof. Dr. Vedat TOPUZ

İSTANBUL, 2021

ÖNSÖZ

Tez çalışmamın gerçekleştirilme sürecindeki destek ve katkılarından dolayı değerli hocalarım Dr. Öğr. Üyesi Ayça Ak'a ve Prof. Dr. Vedat Topuz'a,

Desteği ve sevgisi ile her an yanımdaymış hissi veren ablama,

Yardımları ve destekleriyle yanımda olan sevgili anneme ve babama,

Yüksek lisans eğitimim boyunca her aşamada yanımda olan sevgili eşim BARIŞ İNCEİŞÇİ'ye,

Teşekkür ederim.

İÇİNDEKİLER

ÖNSÖZ ..... **ii**

İÇİNDEKİLER ..... **iii**

ÖZET ..... **v**

ABSTRACT ..... **vii**

SEMBOLLER ..... **viii**

KISALTMALAR ..... **ix**

ŞEKİL LİSTESİ ..... **xi**

TABLO LİSTESİ ..... **xiii**

BÖLÜM 1 – GİRİŞ **1**

1.1. Amaç 4

1.2. Gemi Makinelerinde Hata Analizi ve Makine Öğrenmesi 4

BÖLÜM 2 – MATERYAL VE YÖNTEM **9**

2.1 Gemi Makineleri 9

2.1.1 Gemi Makineleri Sistem Modeli 10

2.1.2 Gemi Makineleri Gürültü Modeli 18

2.2. Makine Öğrenmesi Kavramı 19

2.2.1. Makine Öğrenmesinin Çeşitleri 20

2.2.2 Regresyon Yöntemleri 20

2.2.2.1 Doğrusal Regresyon 21

2.2.2.2 Karar Ağaçları Regresyon 21

2.2.2.3 K-En Yakın Komşular Regresyon 22

2.2.2.4 Rastgele Orman Regresyon 23

2.2.2.5 Bayesian Ridge Regresyon 24

2.2.2.6 Ekstra Ağaç Regresyon 25

2.2.2.7 Doğrusal Destek Vektör Makineleri Regresyon.....26

2.2.3. Yapay Sinir Ağları 27

2.3. Derin Öğrenme Modelleri 30

2.3.1 ESA Modeli 30

BÖLÜM 3 – BULGULAR VE TARTIŞMA 33

3.1 Modelin Değerlendirilmesinde Kullanılan Metrikler 33

3.1.1 Hata Metrikleri.....33

3.1.1.1 Korelasyon Katsayısı 34

3.1.1.2 Ortalama Mutlak Hata.....34

3.1.1.3 Ortalama Kare Hata 34

3.1.1.4 Ortalama Mutlak Yüzde Hata 35

3.1.1.4 Doğruluk 35

3.2 Makine Öğrenmesi Yöntemleri ile Sonuçların Değerlendirilmesi 36

3.3 ESA Yöntemi ile Sonucun Değerlendirilmesi.....46

BÖLÜM 4 – SONUÇLAR.....54

KAYNAKLAR 55

ÖZGEÇMİŞ.....61

ÖZET

MAKİNE ÖĞRENMESİ İLE GEMİ MAKİNELERİNİN HATA ANALİZİ

Gemi sistemlerinin bakım onarımının yapılması ve hata olasılıklarının tahmin edilmesi dinamik gemi sistemleri için önemli bir faktördür. Sistemlerin kapalı kalma süresini en aza indirmek ve maliyeti minimize etmek güncel çalışma alanlarından biridir. Son yıllarda bilgi teknolojilerinin hızlı gelişimi ve buna bağlı olarak gelişen makine öğrenimi yaklaşımları sistemlerin bakım onarımlarına makine öğrenimini entegre etmeyi mümkün kılmıştır. Yapılan çalışmalarda makine öğrenmesi kullanılması, bu yöntemin bakım onarım ve hata analizi alanlarında denenmesine olanak sağlamıştır. Sistemlerde oluşabilecek hatayı tahmin etmek ve buna bağlı olarak arızayı önlüyor olabilmek yeni tekniklerin ortaya çıkmasını sağlamıştır. Bu sistemlerin en önemlilerinden biri olan akıllı gemi makine sistemleri açık denizlerdeki gemilerin yaşam ömrünü artırmaktadır. Gerçekleştirilen bu çalışmada gemi makinelerinin makine öğrenmesi yöntemleri kullanılarak hata durumu araştırılmıştır. Geleneksel yöntemlerin dışında bu çalışmada iki farklı analiz yapılarak makine öğrenmesi yöntemleri kullanılmıştır. Yapılan ilk çalışmada LM-2500 tipi bir gemi makinesinden alınan; gemi hızı, gaz türbini (GT) şaft torku, gaz türbini devir hızı, gaz jeneratörü devir hızı, sancak pervane torku, iskele pervane toku, yüksek güçlü (YG) türbin çıkış sıcaklığı, GT kompresör giriş hava sıcaklığı, GT kompresör çıkış hava sıcaklığı, YG türbin çıkış basıncı, GT kompresör giriş hava basıncı, GT kompresör çıkış hava basıncı, gaz türbini egzoz gazı basıncı, türbin enjeksiyon kontrolü ve yakıt akışı verileri gemi makinelerinin hata durumunu tahmin etmek için regresyon ve yapay sinir ağları (YSA) algoritmaları ile analiz edilmiştir. Doğrusal Regresyon, Karar Ağacı Regresyon, K-En Yakın Komşu Regresyon (K-NN), Rastgele Orman Regresyon, Bayesian Ridge Regresyon, Ekstra Ağaç Regresyon, Doğrusal Destek Vektör Makineleri (DVM) Regresyonu ve YSA algoritmalarından elde edilen sonuçlar karşılaştırılmıştır. Analiz sonucunda YSA yönteminin makine hatası tahmininde regresyon yöntemlerine göre daha iyi sonuçlar verdiği ortaya çıkmıştır.

Yapılan diğer çalışmada ise bir makineden alınan akustik sinyallerden normal ve anormal çalışma durumunu gösteren veri kümeleri oluşturulmuş ve spektrogram grafiklerine

çevrilmiştir. Elde edilen bu görüntüler derin öğrenme yöntemi ile analiz edilmiştir. Bu analiz ile akustik veri kümeleri kullanılarak gemi motorlarının hata-arızalarının modellemesi ve sınıflandırılması sağlanmıştır.

ABSTRACT

FAULT ANALYSIS OF SHIP MACHINES WITH MACHINE LEARNING

Maintenance and repair of ship systems and estimation of fault probabilities are important factors for dynamic ship systems. Minimizing the downtime of the systems and the cost is one of the current study areas. Recently, the rapid development of information technologies and machine learning approaches developed accordingly have made it possible to integrate machine learning into maintenance and repair of systems. The use of machine learning in the studies has enabled this method to be tested in the fields of maintenance, repair and fault analysis. To predict the fault that may occur in the systems and prevent the failure accordingly, has led to the emergence of new techniques. Smart ship machine systems, one of the most important of these, increase the lifespan of ships in open seas. In this study, fault prediction of ship machines was investigated by using machine learning methods. Apart from traditional methods, two different analyzes were made and machine learning methods were used in this study. Taken from a ship LM-2500 type machine; ship speed, gas turbine (GT) shaft torque, gas turbine rate of revolutions, gas generator rate of revolutions, starboard propeller torque, port propeller torque, HP Turbine outlet temperature, GT compressor inlet air temperature, GT compressor outlet air temperature, HP Turbine outlet pressure, GT compressor inlet air pressure, GT compressor outlet air pressure, gas turbine exhaust gas pressure, turbine injection control and fuel flow data were analyzed with regression and Artificial Neural Networks (ANN) algorithms to predict the fault prediction of ship machinery. The results obtained from linear regression, decision tree regression, nearest neighbor regression, random forest regression, Bayesian Ridge regression, extra-tree regression, linear SVM regression and ANN algorithms were compared. As a result of the analysis, it was revealed that the ANN method gave better results than the regression methods in estimating the machine fault.

In another study, data sets showing normal and abnormal operating status were created from acoustic signals received from a machine and converted into spectrogram graphics. These images were analyzed by deep learning method. With this analysis, modeling and classification of ship engine faults-failures were provided by using acoustic data sets.

SEMBOLLER

%	**:** Yüzde
°C	**:** Sıcaklık
yi	**:** Tahmin değeri
ϵ_i	**:** Artık değer
kg/s	**:** Kütle
$\sum$	**:** Toplam
y	**:** Gerçek değerler
$\overline{y}$	**:** Gerçek değerlerin ortalaması
$\hat{y}$	**:** Tahmin edilen değerler
N	**:** Veri Sayısı
$\in$	: Elemanıdır
F(f)	**:** Frekans alanı
x(t)	**:** Zaman alanı
π	**:** Pi değeri
ξ	**:** Duyarsız kayıp fonksiyonu
b	**:** Bias

KISALTMALAR

AC : Alternative Current

ANN : Artificial Neural Network

CODLAG : Combined Diesel Electric And Gas

DC : Direct Current

DVM : Destek Vektör Makineleri

DVS : Destek Vektör Sınıflandırması

ESA : Evrişimsel Sinir Ağları

FD : Fourier Dönüşümü

FFT : Fast Fourier Transform

GT : Gaz Türbini

HP : High Power

K-NN : K- En Yakın Komşular

MATLAB : Matrix Laboratory

MÖ : Makine Öğrenmesi

NAR : Nonlinear Autoregressive

NARX : Nonlinear Autoregressive Exogenous

OKH : Ortalama Kare Hata

OMH : Ortalama Mutlak Hata

OMYH : Ortalama Mutlak Yüzde Hata

RF : Rastgele Orman

RPM : Rotations Per Minute

SVM : Support Vector Machines

TSA : Tekrarlayan Sinir Ağları

YSA : Yapay Sinir Ağları

YG **:** Yüksek Güçlü

YZ **:** Yapay Zekâ

ŞEKİL LİSTESİ

Şekil 2.1. Birinci Veri Kümesine Uygulanan İşlemler.....9

Şekil 2.2. İkinci Veri Kümesine Uygulanan İşlemler 10

Şekil 2.3. Bir Geminin Yapısında Bulunan Bölümler.....10

Şekil 2.4. Ana Makine ve Pervane Arasında Bulunan Gemi Elemanları.....11

Şekil 2.5. Araştırmada Kullanılan GT ve Pervane Yapıları.....12

Şekil 2.6. LM-2500 Motorunun Yapısı.....12

Şekil 2.7. Tüm Parametrelere İlişkin Histogram Grafikleri 17

Şekil 2.8. Parametreler Arasındaki Korelasyona İlişkin Isı Matrisi.....17

Şekil 2.9. Havuz Pompası Makinesi.....18

Şekil 2.10. Karar Ağaçları Yapısı 21

Şekil 2.11. K-NN Yapısı 23

Şekil 2.12. Rastgele Orman Yapısı 24

Şekil 2.13. Yapay Sinir Ağı Yapısı 28

Şekil 2.14. YSA'da Bir Nöron Modeli.....29

Şekil 2.15. ESA Mimarisi 31

Şekil 3.1. Makine Öğrenmesi Algoritmaları 37

Şekil 3.2. Doğrusal Regresyon Analizine Göre Gerçek ve Tahmin Edilen Değerler 38

Şekil 3.3. Karar Ağacı Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler.....39

Şekil 3.4. K-NN Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler.....39

Şekil 3.5. Rastgele Orman Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler.....40

Şekil 3.6. Bayesian Ridge Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler 41

Şekil 3.7. Ekstra Ağaç Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler.....41

Şekil 3.8. Doğrusal DVM Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler.....42

Şekil 3.9. YSA Modeline İlişkin Sonuçlar.....43

Şekil 3.10. YSA Modeline Göre Gerçek ve Tahmin Edilen Değerler 44

Şekil 3.11. Normal Akustik Sinyaller 48

Şekil 3.12. Anormal Akustik Sinyaller 49

Şekil 3.13. Normal Ses Spektrogramı 50

Şekil 3.14. Anormal Ses Spektrogramı 51

Şekil 3.15. ESA Modeli 52

Şekil 3.16. ESA Modeli Doğruluk Değeri 53

Şekil 3.17. Anormal Verilere İlişkin Sinyal Görüntülerinin Sınıflandırılması 53

Şekil 3.18. Normal Verilere İlişkin Sinyal Görüntülerinin Sınıflandırılması 53

TABLO LİSTESİ

Tablo 2.1. Araştırmada Kullanılan Parametreler 13

Tablo 2.2. Araştırmada Kullanılan Örnek Parametre Değerleri 14

Tablo 2.3. Sayısal Nitelikler için Ekstra Ağaçların Bölme Prosedürü 26

Tablo 3.1. Regresyon ve YSA Modellerinin Eğitim Parametreleri 36

Tablo 3.2. Regresyon ve YSA Modellerine İlişkin MAE, MSE, MAPE ve R^2 Değerlerinin Karşılaştırması.....45

1. GİRİŞ

Makinelerin çalışma durumunun takibi ve olası arızaların tespiti için geliştirilen yöntemler son yıllarda büyük ilgi görmekte ve bu sayede çeşitli endüstri dallarında büyük ekonomik faydalar elde edilmektedir. Makinelerin çalışmasına ilişkin çeşitli durum bilgilerinin analiz edilmesinden önce arızalara karşı duyarlı bazı faktörler bulunabilir. Bu faktörlere dayanarak meydana gelen hatalar, güvenilir bir şekilde belirlenebilmekte ve makinelerin çalışmasına ilişkin eğilimler tahmin edilebilmektedir.

Makinelerin durumunu takip etmenin endüstride önemli etkileri olduğundan, uzman ve pratik bakım alanında büyük ilgi görmüştür (Widodo ve Yang, 2007). Sistem bakımı ve proses otomasyonu düşünüldüğünde makinelerin çalışma durumunun takip edilmesi önemlidir (Lou ve Loparo, 2004). Makinelerin çalışma durumunun takibi, özellikle dönerek çalışan makinelerin sağlamlığı ve bakım gereksinimleri hakkında önemli bilgiler sağlamakta ve birçok endüstriyel uygulamada kullanılmaktadır (Huang, 2001). Hata tespit sistemleri temel olarak titreşim ve teknolojik parametre ölçümlerine dayanan herhangi bir makinede kullanılabilir (Vasylius, Didžiokas, Mažeika ve Barzdaitis, 2008). Ekipmanın mekanik durumunu izlemek için titreşim, sıcaklık, yağlayıcı kalitesi ve akustik emisyon gibi parametreler uygulanabilir. Hata tespiti, mevcut bir sistemin güvenilirliğini ve kullanılabilirliğini artırmaktadır. Çeşitli arızalar nispeten yavaş bir şekilde azaldığından, erken dönemde arıza teşhisi önemli bir potansiyele sahiptir (Moosavian, Ahmadi, Tabatabaeefar ve Khazaee, 2013).

Denizcilik endüstrisi, dünya çapında üretilen malların toplu taşımacılığından sorumlu olmasına rağmen, son zamanlarda gemilerin güvenilirliğinin, kullanıma uygunluklarının ve karlılığının artırılmasını araştıran yeni yaklaşımlar geliştirilmiştir (Lazakis ve Ölçer, 2015). Denizcilik diğer endüstrilerle karşılaştırıldığında, farklı koşullardaki benzer ekipmanların farklı arıza modelleri ortaya çıkması nedeniyle veri havuzlaması her zaman mümkün olmamaktadır. Ayrıca sürekli yeni ekipmanların ortaya çıkması ve bu gelişmenin geçmişteki kayıtları geçersiz kılması da önemli bir sorundur. Bunun yanı sıra, veriler standartlaştırılmış bir şekilde toplanmamaktadır. Böylece hem daha bilinçli ve etkili karar alma sağlanırken, hem de teknolojik gelişmeler, sürekli çalışan ekipman ve yüksek bakım maliyeti açısından gelişmiş bakım tekniklerine büyük ilgi uyanmasına yol açmıştır (INCASS, 2015). Raza ve Liyanage (2009) gemilerde mevcut durum izleme

programlarının bir parçası olarak akıllı tekniklerin test edilmesi ve uygulanması için artan bir talebin var olduğunu belirtmişlerdir. Veri miktarı, türü ve bunun ne sıklıkta toplanması gerektiği ve nasıl ortaya çıktığı sorusu, her ne kadar bakım şemaları benimsenmiş olsa da kayıtlı operasyonel bilgilerin işlenmesi, analizi ve kullanımında sorunlar olduğunu göstermektedir. Günümüzde gemi makinelerinde kullanılan sensörler tarafından toplanan veriler sıcaklık, basınç, akış, titreşim ve akım sensörleri gibi önemli miktarlarda ölçüm parametresi içermektedir. Uzay-zamansal veri kümelerindeki içsel örüntüleri keşfetmek, ortaya çıkarmak ve genelleştirmek için kümeleme algoritmaları önemli rol üstlenmektedir (Hagenauer ve Helbich, 2013).

Nideerlinski'nin 1971 yılında arızalı bir sistemdeki kontrol problemini çözmek için integral kontrol yöntemi teorisini ortaya atmasından itibaren birçok hata toleranslı kontrol yöntemi geliştirilmiştir. Günümüzde hata analizlerine ilişkin yapılan bu araştırmalar, mühendislik alanındaki önemli araştırma konularından biri haline gelmiştir (Tang ve Yao, 2005; Xie, Wang, He, Diallo ve Claramunt, 2020). Literatürde kestirimci (predictive) bakım olarak adlandırılan uygulamalar, gemilerin fiziksel durumunu kontrol etmek için denizcilik alanındaki güncel eğilimler arasındadır. Burada hata analizi, durum ve süreçler izlenerek gemilerdeki önemli sistemlerde uygulanmaktadır. Durum ve süreç izlemenin amacı, incelenen sistemin durumunu (arızalı, bozulmuş vb.) değerlendirmek ve gelişen hataları tespit etmek için belirli gemi sistemlerinden alınan bilgileri kullanmaktır (Kobbacy, 2008; Mohanty, 2015). Bakım işleminin başarılı bir şekilde uygulanması; verilerin kalitesine, sistem durum değerlendirmesine ve arıza tespiti için kullanılabilen çeşitli algoritmaların uygulanabilirliğine bağlıdır (Cheliotis, Lazakis ve Theotokatos, 2020). Öte yandan gemi bakımı, gemicilik endüstrisinin güvenilirlik ve kullanılabilirlik düzeylerini etkilemekte ve bir geminin operasyonel giderlerinin %20-%30'unu oluşturmaktadır. Ayrıca gemi bakımı arıza süresini en aza indirebilen ve işletme maliyetlerini düşürebilen önemli bir gemi yaşam döngüsü faktörüdür (Lazakis, Raptodimos ve Varelas, 2018).

Özellikle teknolojinin gelişmesine paralel olarak makine öğrenmesine yönelik çalışmaların sayısı artmıştır. Makine öğrenmesi modeli ve modelden elde edilen verilerin dağılımları, çoğunlukla bir hatayı minimum yaparak işlevini en aza indirerek modelleri eğitir. Eğitilen modeller, modeldeki sınıflardan benzer olan sınıflara tahminde bulunmak için uygulanmaktadır (Aydın, 2018).

Son yıllarda gemi makinelerinde makine öğrenmesine yönelik pek çok araştırma göze çarpmaktadır. Gemi makinelerine yönelik literatürdeki çalışmalarda karar ağacı metodu (Saatçioğlu, Göksu, Yüksel ve Gülmez, 2017), dalgacık dönüşümü (Li, Monti ve Ponci, 2014), zaman serisi (Makridis, Kyriazis ve Plitsos, 2020) ve yapay sinir ağları (Pantelelis, Kanarachos ve Gotzias, 2000; Köroğlu, 2007; Guoqiang, 2018; Pal, Datta, Segev ve Yasinyac, 2019) gibi farklı yöntemlere yer verilmiştir.

Bu araştırmada kullanılacak Yapay Sinir Ağları (YSA), beyin gibi biyolojik sinir sistemlerinin bilgiyi işleme şeklinden esinlenen bir bilgi işleme paradigmasıdır. Bu paradigmanın en önemli unsuru, bilgi işleme sisteminin yeni yapısıdır (Zhongling, Ning ve Yan, 2005). Bu sistem, belirli problemleri çözmek için birlikte çalışan birbiriyle bağlantılı çok sayıda işlem elemanlarından (nöronlardan) oluşmaktadır. Sinir ağları, karmaşık veya kesin olmayan verilerden anlam çıkarma konusundaki güçlü yönüyle desenler/örüntüler çıkarmak ve insanlar ya da diğer bilgisayar teknikleri tarafından fark edilemeyecek kadar karmaşık olan eğilimleri tespit etmek için kullanılmaktadır (Zhou ve Xu, 2011). Öğrenme ve adaptasyon, belirsiz ortamlarda sistemlerin performansını iyileştirmek amacıyla geri bildirim kontrol sistemleri için önemli becerilerdir. Uyarlanabilir sinir ağı kontrolü ile ilişkili bir diğer önemli konu ise, geribildirim kontrol sistemlerinin performansını iyileştirmek için "yaparak öğrenme" ve "öğrenilen bilgilerle yapma" anlamına gelen öğrenme becerisidir (Dai, Wang ve Wang, 2016).

Bu tez çalışması, sonuç bölümü de dahil olmak üzere beş bölümden oluşmaktadır. Çalışmanın birinci bölümünde bu tezin amacı anlatılmış ve gemi arızaları hakkında yapılan literatür taramasına yer verilmiştir. İkinci bölümde gemi sistemleri hakkında bilgi verilmekte ve tezde kullanılan veri kümeleri ile kullanılan yöntemler açıklanmaktadır. Üçüncü bölümde önerilen sistemin gerçekleştirilme şekli ve sonuçlar sunulmuştur. Dördüncü bölümde kullanılan yöntemlerin karşılaştırılması verilmiştir. Beşinci bölümde sonuçlar ve gelecekteki çalışmalara yönelik öneriler sunulmuştur.

1.1. Amaç

Denize kıyısı bulunan ülkelerin ekonomik, politik ve çevresel kararları daha tutarlı alabilmesi için mevcut gemiler üzerinde bulunan sistemleri hakkında daha fazla analize sahip olması gerekir. Bu sebeple büyük maliyetler harcadıkları sistemleri daha uzun süre kullanmalarına katkı sağlayabileceği düşünülmektedir. Bu tez çalışmasının amacı; makine öğrenmesi algoritmalarının gemi sistemleri üzerindeki hata olasılığı tahmin yeteneklerini değerlendirmektir. Bu tezde yapılacak çalışma ile gemi makinelerinin makine öğrenmesi algoritmalarını kullanarak bir sonraki bakım onarım zamanını tahmin edebilecek alternatif bir model geliştirmek mümkün kılınacaktır.

Bu çalışmada regresyon ve sınıflandırma algoritmaları kullanılmıştır. Yapılan ilk analizde en doğru algoritmanın seçilmesi zor olacağı için sekiz (8) farklı regresyon algoritması seçilmiştir. İkinci analizde ise akustik sinyallerden oluşturulan spektrogram görüntüleri evrişimsel sinir ağları algoritması ile analiz edilmiştir. Yapılan çalışma ile mevcut gemi sistemlerinin hata oranlarını tahmin ederek bakım onarımı ile ilgili bilgiye ulaşılacaktır. Böylece makine öğrenmesi algoritması kullanım alanlarının genişletilmesi ve yapılacak olan bakım onarım zamanının gerçeğe daha yakın olması hedeflenmektedir. Modelin performansı değerlendirildikten sonra gemiler üzerindeki sistemlere entegre edilmesi hedeflenmektedir. Sonuç olarak da mevcut sistemlerin hata oranları belirlenecektir. Ortaya çıkacak olan sonuçlara bakılarak bakım onarımın zamanları için yorumda bulunulacaktır. Buna bağlı olarak da gemi sistemlerinde, sistemlerin bakım onarım ve hata olasılıklarının bilinmesi kullanılabilirliği arttıracak ve maliyet azalacaktır.

1.2. Gemi Makinelerinde Hata Analizi ve Makine Öğrenmesi

Araştırmanın giriş bölümünde belirtildiği gibi, gemi makinelerinde hata analizi ve makine öğrenmesine ilişkin son yıllarda birçok çalışmanın göze çarptığı tespit edilmiştir. 2000'li yılların başında yapılan çalışmalardan Pantetelis, Kanarachos ve Gotzias (2000) gemilerde rotor, temel ve hidrodinamik yataklardan sinir ağları ve gerçek titreşim verileri kullanarak hata analizi yapmak amacıyla basit sonlu eleman modelleri oluşturmuşlardır. Çalışmada tüm sistemin dinamik simülasyonlarına dayanan basit ve gerçekçi sonlu eleman modellerinin geliştirilmesi doğrultusunda, fiziksel simülasyonların ve ölçümlerin

yapay sinir ağları ile birleştirilmesi sayesinde fiziksel temelli ve daha detaylı hata ölçümü yapılabildiği belirlenmiştir.

Köroğlu (2007) tarafından yapılan araştırmada stifnerli bir panelin tasarım parametreleri kullanılarak yapay sinir ağlarıyla burkulma modellemesi yapılmıştır. MATLAB Neural Networks Toolbox yazılımı kullanılarak yapılan analizde Levenberg-Marquardt yöntemi ve ağ budama yaklaşımı kullanılmıştır. 12 nöron içeren ağda 30 adet verinin ortalama hatası %23 bulunurken, 350 adet veride bu oran %1,5'e düşmüştür. Nöron sayısının performansa etkisi incelendiğinde, 24 nöronun hata oranı %4,3 bulunurken 8 nöron %2,1 hata oranı ile en iyi performansı göstermiştir. Çeyrek 2k, merkezi kompozit ve yapay sinir ağları ortalama performansları karşılaştırıldığında ise, 250 veri içeren yapay sinir ağlarının %1,90 hata oranı ile en iyi sonucu verdiği tespit edilmiştir.

Li, Monti ve Ponci (2014) dalgacık dönüşümü çoklu çözünürlüklü analiz yöntemini yapay sinir ağları ile birleştirerek, orta gerilim DC gemi güç sistemlerinde bir hata algılama ve sınıflandırma tekniği incelemişlerdir. Hata sinyallerinin farklı çözünürlük seviyelerindeki enerji değişimi öznitelik vektörleri olarak seçilmiştir. Analiz ve karşılaştırmalarda Daubechies 10 (db10) dalgacık fonksiyonu ve 9. ayrışma seviyesi kullanılmıştır. Hata türlerini çıkarılan özelliklere göre otomatik olarak sınıflandırmak için yapay sinir ağı kullanılmıştır. Hem DC hem de AC için kısa devre arızaları ve toprak arızası gibi farklı arıza türleri, önerilen yöntemin etkinliğini doğrulamak için test edilmiş, bu hatalar dijital simülatör ile gerçek zamanlı olarak simüle edilerek elde edilen veriler MATLAB ile analiz edilmiştir. Çalışma sonucunda benimsenen yöntemin farklı frekans bantlarında sinyal bileşenlerini çıkarmak için verimli olduğu, sinir ağının gemi makine özelliklerini karşılaştırmada daha akılcı bir yöntem olduğu, önerilen algoritmanın daha az veri girdisi içermesi nedeniyle hesaplama yükünü azalttığı, yüksek doğruluğa sahip olmakla birlikte çeşitli hatalara karşı dayanıklı olduğu, performansın hata direnci ve yükler gibi elektrik parametrelerindeki değişikliklerden etkilenmediği belirlenmiştir.

Coraddu ve ark. (2015) gemilerde gaz türbini tahrik sisteminde koşullara bağlı bakım açısından makine öğrenme modeli sunmuşlardır. Gaz türbini ve kompresörlerde 10, 25, 50, 100, 250 ve 500 öğrenme değerleri için düzenlenmiş en küçük kareler ve destek vektör regresyonu yöntemleri karşılaştırılmıştır. Karşılaştırma sonucunda destek vektör regresyonunun ortalama hatası daha küçük bulunmuş ve varyans değerleri, öğrenme seti

değiştikçe bu yaklaşımın düzenlenmiş en küçük kareler yönteminden daha kararlı olduğunu göstermiştir. İki bozulma faktörünün eşzamanlı varyasyonlarına rağmen, öğrenme sayısı arttıkça hatanın eğiliminin önemli ölçüde azaldığı belirlenmiştir. Bunların yanı sıra düşük hata oranları elde etmek için yaklaşık 100 veriye ihtiyaç duyulduğu sonucuna varılmıştır.

Saatçioğlu, Göksu, Yüksel ve Gülmez (2017) yaptıkları araştırmada farklı ülkelerde yayınlanmış kaza raporlarını inceleyerek, gemilerde makine dairelerinde farklı kaza faktörlerini, kaz oluşum sıklıklarını ve nedenlerini karar ağacı yöntemi kullanarak belirlemişlerdir. Araştırma bulgularına göre, tüm makine dairesi kazalarının %90,43'ünün insan hatasından kaynaklandığı tespit edilmiştir. İnsan kaynaklı hataların ana nedenlerinin temel eğitim ve deneyim eksikliği, yorgunluk ve aşırı iş yükü olduğu belirtilmiştir. Kazalara neden olan diğer faktörün ise tasarım ve üretim hataları olduğu görülmüştür.

Bilgili (2018) yaptığı araştırmada kuru yük gemilerinin raporlarına dayanarak sefer süresi, makine çalışma devri, gemi hızı, hava durumu, deniz durumu ve ortalama draft verilerini girdi olarak kullanmış ve çıktı olarak emisyon verisini içeren yapay sinir ağı modellemesi yapmıştır. En iyi sonucu sağlayan yapay sinir ağı modeli aracılığıyla ulaşılan emisyon değerleri ile gerçek değerler arasındaki ortalama fark %1,57 olarak bulunmuştur. Bu emisyon verisi daha sonra yakıt maliyet analizinde kullanılmış, böylece gemilerin daha yumuşak iklimi şartlarına sahip rotalarda daha az emisyon ortaya çıktığı tespit edilmiştir. Dolayısıyla çevresel ve ekonomik performans artırılmıştır.

Öztanır (2018) yaptığı araştırmada kestirimci bakımla kalan faydalı ömrün hesaplanması ve bakım takviminin planlanmasını amaçlamıştır. Kalan faydalı ömür hesaplamasında regresyon analizi kullanılmıştır. Turbo fan jet motoruna ait açık veri kümesinde makine öğrenmesi ve yapay sinir ağı yöntemleri kullanılmıştır. İki farklı veri kümesinin eğitimi ile modeller oluşturularak, performanslar ortalama değerler ile karşılaştırılmıştır. Regresyon yaklaşımında tekrarlayan sinir ağı türevi olan Uzun Kısa Süreli Bellek kullanılarak oluşturulan model, diğer yöntemlere göre anlamlı şekilde daha başarılı sonuçlar üretmiştir.

Elidolu (2019) gemi personelinin emniyetli çalışma yönelimlerini makine öğrenmesiyle tahminlemeyi amaçlamıştır. 115 kişiye uygulanan çevrimiçi anketler sonrasında ulaşılan

verilere makine öğrenmesi uygulanarak test verisini tahminlemesi hedeflenmiştir. Makine öğrenmesi kapsamında çoklu doğrusal regresyon, karar ağacı regresyonu, destek vektör regresyonu ve rastgele orman regresyonu yöntemleri kullanılarak performansları karşılaştırılmıştır. Emniyetli çalışma davranışı ve çalışma memnuniyeti tahminlemede çoklu doğrusal regresyon yönteminin başarı oranı sırasıyla %95,95 ve %90,83 bulunmuş ve diğer yöntemlere göre daha iyi sonuç vermiştir.

Pal, Datta, Segev ve Yasinyac (2019) bir firkateyn üzerine monte edilmiş gaz türbininin kompresör bozulma katsayısını ve türbin bozulma durumu katsayısını modellemek ve tahmin etmek için yapay sinir ağı ve temel bileşen analizi yapmışlardır. Giriş parametreleri gaz türbinine ait parametreler, çıkış parametreleri ise gaz türbin kompresör ve türbin bozulma durumu katsayılarıdır. Çok sayıda girdinin varlığı nedeniyle daha fazla gizli katmana ihtiyaç duyulmuş ve derin bir sinir ağının uygun olduğu ortaya çıkmıştır. Simülasyon sonucunda önerilen modellerin birçoğu, deniz tahrikinin gaz türbininin bozulma durumu katsayılarını tahmin etmiştir. Çalışma sonucunda giriş ve çıkışla orantılı olarak sürekli azalan gizli katman boyutunun, diğer sinir ağı mimarilerinden daha iyi performans gösterdiği belirlenmiştir. Ayrıca yapay sinir ağı bulgularının birçoğu, yapay sinir ağı ve temel bileşen analizi kombinasyonundan daha iyi performans göstermiştir.

Ateş (2020) gemilerde T şeklindeki birleşim noktalarında meydana gelen gerilme yığılması tahminini kaba ağ yapısı ve makine öğrenmesi ile gerçekleştirmiştir. Kaba ağ yapılı analizden elde edilen gerilmelerin dağılımı, varyansı, ağ boyutu, maksimum gerilmesi gibi parametreler makineye öğretilerek referans gerilme değeri tahmin edilmiştir. 200 veri için yapılan analizlerde yarıçapın değişmesi hata oranını etkilemiştir. En az hata oranı, 0.2 ve 0.4 yarıçap değerlerinde bulunmuştur. En iyi değerleri veren yarıçaplarla 2000 veri modeli çalıştırılmış ve kullanılan yöntemlerden en iyi sonucu Gauss süreci regresyonu vermiştir. Gauss süreci regresyonu ardından en iyi sonuçlar sırasıyla gradyan artırıcı regresyon, rastgele orman regresyonu, karar ağacı regresyonu ve Ridge regresyonu ile elde edilmiştir.

Kimera ve Nangolo (2020) makine öğrenmesi ile bir erken bakım/arıza uyarı sistemine yönelik kestirimci yaklaşım sağlamayı amaçlamışlardır. MATLAB aracılığıyla verilerden çıkarımlar elde etmek için bağlantı pompası çalışma parametrelerini işlemek ve analiz etmek için bir makine öğrenimi metodolojisi kullanılmıştır. Pompanın çalışma

parametreleri için hangi parametrelerin pompa arızasını doğru bir şekilde tahmin edeceğini tanımlamak amacıyla temel bileşen analizi gerçekleştirilmiştir. Geri basınç, akış hızı, amper, dönüş hızı (RPM) ve emme basıncı 40 hafta süreyle izlenmiş, bağlantı pompasının 7. ve 8. haftalar arasında arızalanabileceği veya bakım gerektirebileceği öngörülmüştür. Bu tahmin, bağlantı pompasının arızalanması için gerçek süre olan 9 haftadan farklılık göstermiştir. Gerçek arıza süresinden ortaya çıkan %13,85'lik bu sapmanın, kaydedilen işletim verilerinin kalitesine ve verinin az olmasına bağlanmıştır.

Raptodimos ve Lazakis (2020) bir geminin ana motor silindirinin egzoz gazı çıkış sıcaklığının gelecekteki değerlerini tahmin etmek için Dışsal Girdili Doğrusal Olmayan Otoregresif (NARX) yapay sinir ağı geliştirmişlerdir. NAR ve NARX modelleri kullanılarak tahminlerde yüksek doğruluğu sağlanmıştır. NAR modellerinin doğruluğuna kıyasla, NARX modellerinde dışsal girdi olarak ana motor devri tanıtılarak tahmin doğruluğunun %35'e kadar artırılabileceği görülmüştür. NAR modelinde 23. saat tahmini ile gerçek egzoz gazı çıkış sıcaklığı değeri arasındaki mutlak hata %38,10'a eşitken, NARX modeli için sonuç %4,52'lik mutlak hata ile sonuçlanmış, diğer bir ifadeyle hata oranında %33,58 büyüklüğünde bir azalmaya yol açmıştır. Modelde ortalama mutlak yüzde hatası %4,64'ten %1,02'ye düşürülmüştür.

Uyanık, Karatuğ ve Arslanoğlu (2020) bir konteyner gemisinde yakıt tüketimi için çoklu doğrusal regresyon, Ridge ve LASSO regresyonu, destek vektör regresyonu, karar ağacı algoritması, boosting algoritmaları gibi çeşitli tahmin modelleri kurmuştur. Modellerin doğruluğu, K-katlamalı çapraz doğrulama ile belirlenmiştir. Tahmin modellerinin doğruluğunu değerlendirmek için kök ortalama kare hatası, ortalama mutlak hata, R^2 gibi hata metrikleri kullanılmış ve değişkenler arası korelasyon analizi yapılmıştır. Ana motor devri, ana motor silindir değerleri, şaft göstergeleri gibi parametreler yakıt tüketimi ile yüksek oranda ilişkili bulunmuştur. Çeşitli dış faktörlerin etkisi altında gerçek yakıt tüketimi verilerine en yakın tahmin; 0.0001 ortalama karekök hata, 0.002 ortalama mutlak hata ve R^2=%99,9 katsayısı ile çoklu doğrusal regresyon ve Ridge regresyon ile sağlanmıştır.

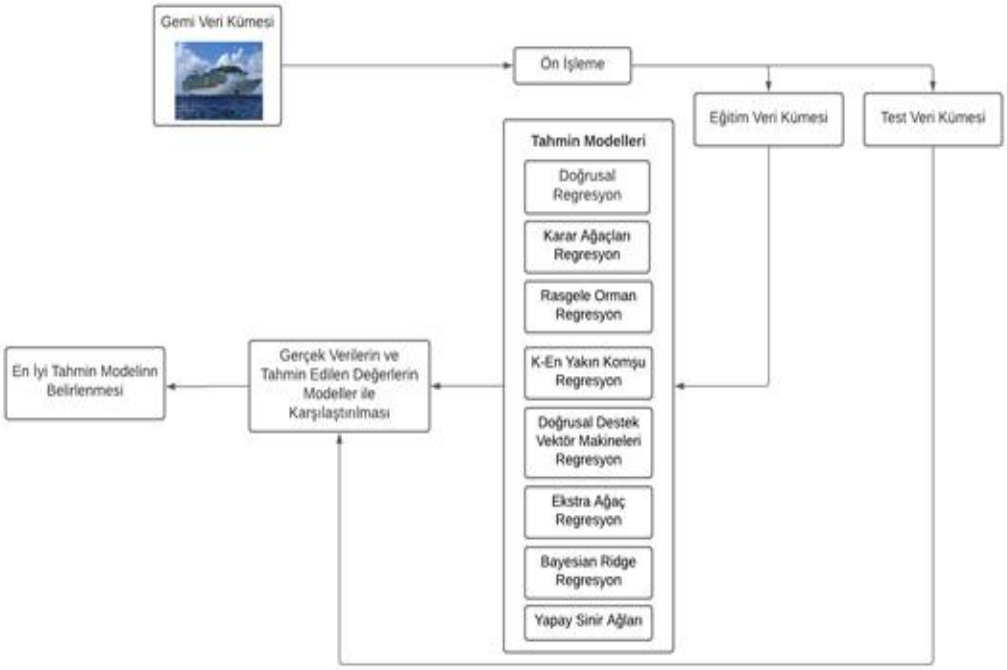
2. MATERYAL VE YÖNTEM

Bu bölümde öncelikle gemi makineleri sistem modeli hakkında bilgi verilecektir. Bölüm 2.1.2'de gürültü modeli anlatılacaktır. Makine öğrenmesi kavramı ve yöntemleri bölüm 2.2'de açıklanacaktır. Araştırmada kullanılan diğer yöntem olan derin öğrenme modeli ise bölüm 2.3'te anlatılacaktır.

2.1 Gemi Makineleri

Bu bölümde çalışmada kullanılan veri kümeleri açıklanacaktır. Çalışmada iki tip veri kümesi kullanılmıştır.

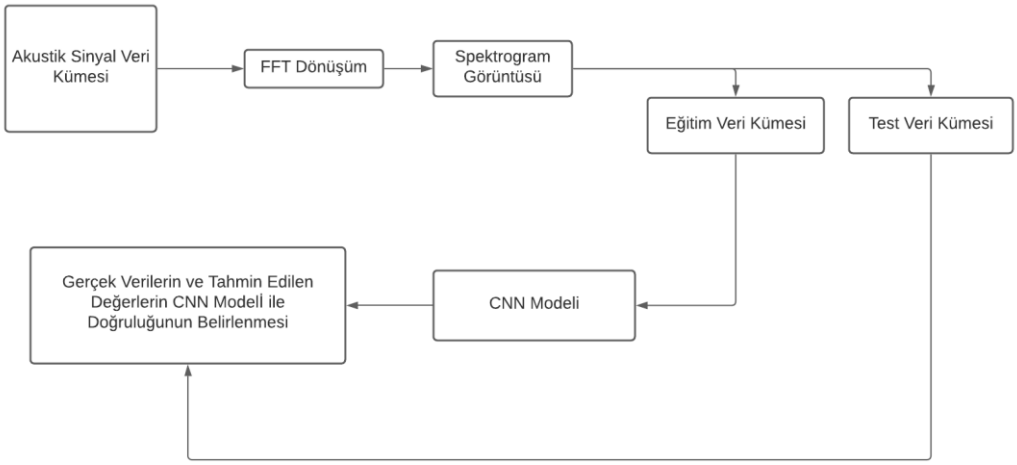
İlk veri kümesi dizel elektrik ve gaz makinesi ile karakterize edilmiş LM-2500 tipi gemi motorundan alınan verilerdir. Bu veriler basınç, hız, sıcaklık, tork, akış ve devir hızı verileridir. Bu veriler ile regresyon ve yapay sinir ağları modelleri kullanılarak gemi motorunun türbin bozulma katsayısı tahmin edilmiştir. Araştırmada kullanılan birinci veri kümesine ait blok şema Şekil 2.1'de gösterilmiştir.



Şekil 2.1. Birinci Veri Kümesine Uygulanan İşlemler

İkinci veri kümesi havuz pompa makinesinden alınan normal ve anormal ses veri kümesidir. Bu veri kümesinde ses verilerinin spektrogramlarının grafikleri çıkarılmış ve

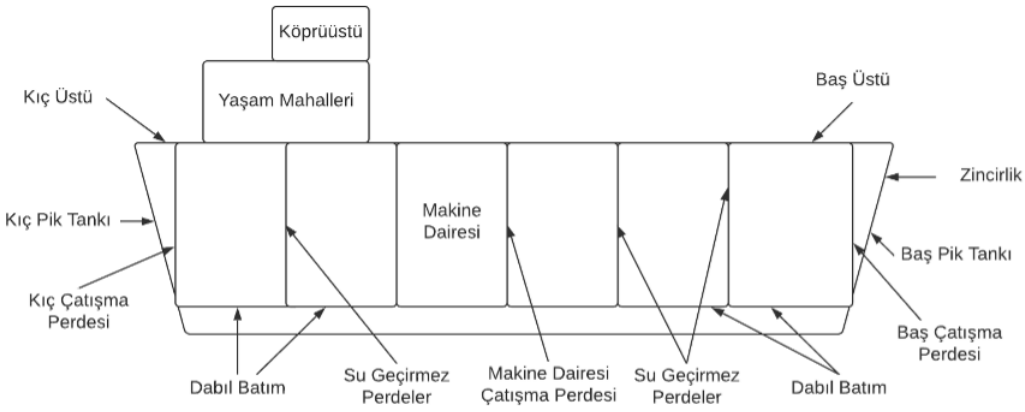
bu grafikler görüntü olarak .png formatında kaydedilmiştir. Daha sonra bu görüntüler ESA algoritması ile eğitilmiştir ve normal ve anormal sesler tahmin edilmiştir. Havuz pompa makine veri kümesi ile elde edilen bu model ile gemi makinelerinin de ses analizinde kullanılması amaçlanmaktadır. Araştırmada kullanılan ikinci veri kümesine ait blok şema Şekil 2.2'de gösterilmiştir.



Şekil 2.2. İkinci Veri Kümesine Uygulanan İşlemler

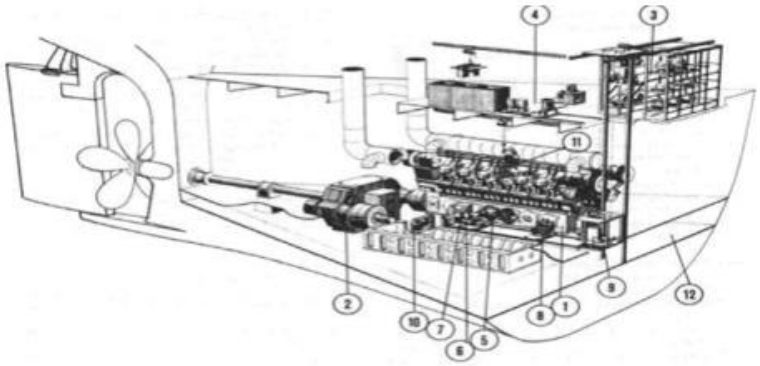
2.1.1 Gemi Makineleri Sistem Modeli

Gemi makine dairesi, tanklar, ambarlar, pik tankları, dabılbatım, yaşam mahalleri, köprüüstü, portuç, zincirlik, kırlangıç ve kasaradan oluşmaktadır.



Şekil 2.3. Bir Geminin Yapısında Bulunan Bölümler

Makine dairesi, geminin makinelerinin bulunduğu ana bölmedir. Makine dairesinde ana makineler, yardımcı makineler, kontrol odası, kontrol paneller ve atölyeler bulunur. Gemi ana makinesi, gemiyi hareket ettiren gücü üreten makinelerdir. Geçmişte pervane veya jetleri tahrik ederek tahrik sağlayan gemi ana makine sistemleri, günümüzde gemi tipine ve boyutuna bağlı olarak buhar motoru, buhar türbini, içten yanmalı motor veya gaz türbini olabilmektedir. Bu amaç dışındaki tüm sistemler yardımcı makineler olarak tanımlanır. Yardımcı makinelere jeneratörler ve kompresörler örnek verilebilir. Jeneratör, geminin elektriğini üreten sistemdir. Bağımsız makinelerde olduğu gibi, genellikle geminin kardan şaftına aktarılan gücü kullanan şaft jeneratörleri kullanılır. Genellikle alternatör ve dizel motordan oluşan bir sistemdir. Kompresör, genellikle hava veya diğer gazları atmosfer basıncından daha yüksek basınçlara sıkıştırmak için kullanılan bir makinedir.

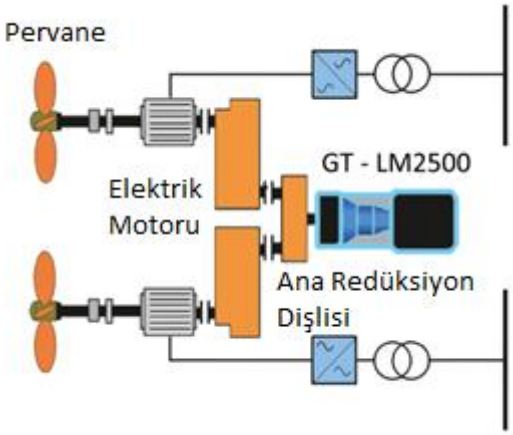


1-Ana makine	5-Piston söküm kaldıracı	9-Depo vinci
2-Dişli kutusu (Redakşınger)	6-Kaldıraç bağlantı ünitesi	10-Ana yatak krikosu
3-Yedek parça ambarı	7-Çalıştırma ünitesi	11-Supap sökme aparatı
4-Atölye	8-Piston taşıma kovanı	12-Makine dairesi dip sacı

Şekil 2.4. Ana Makine ve Pervane Arasında Bulunan Gemi Elemanları

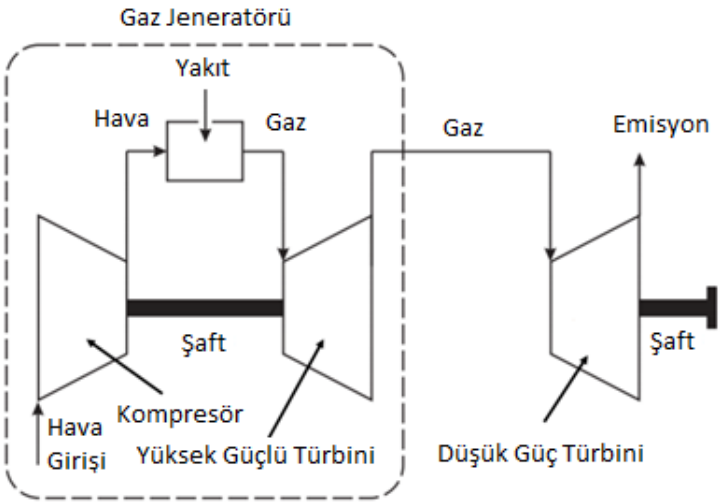
Şekil 2.4'te pervane ve devre elemanlarını bulunmaktadır. Bu görselde bir geminin ana makinadan pervaneye kadar olan bölümleri vardır. Gemide gücün ana makinelerden pervaneye kadar aktarılmasını sağlayan sistemlere itici sevk sistemleri denir. İtici sistemlerinin çalışması şu şekilde gerçekleşir; Ana makinenin çalışması ile elde edilen dairesel hareket krank şaft ve dişliler (reduction gear ya da şanzıman) aracılığı ile hareketi srast yatağına ve ara şafta, oradan da stern tüpten pervane şaftı kanalıyla pervaneye iletir. Pervanenin hareket etmesiyle de geminin hareket etmesi sağlanır.

Bu çalışmada kullanılan veriler bir Gaz Türbini (GT) sevk tesisi ile karakterize edilen bir deniz gemisinin (Fırkateyn) sayısal bir simülatörü vasıtasıyla gerçekleştirilmiştir. Bu gözlemler ışığında, mevcut veriler olası gerçek bir gemi ile uyum içindedir. Kullanılan veritabanı, MATLAB Simulink'te geliştirilmiş, test edilmiş ve doğrulanmıştır. Simülatörde oluşturulan farklı bloklar (Pervane, Muhafaza, GT, Dişli Kutusu ve Denetleyici), yıl boyunca birkaç benzer gerçek sürücü tesisinde geliştirilmiş, ince ayar yapılmış ve test edilmiştir. Kullanılan veri tabanının oluşturulduğu gemi makine ve pervane örneği Şekil 2.5'te gösterilmiştir. Bu sistemde LM-2500 tipi ana motor kullanılmıştır. Burada 2 adet pervane ve 1 adet gaz türbini bulunmaktadır.



Şekil 2.5. Araştırmada Kullanılan GT ve Pervane Yapıları (Coraddu ve ark. (2015))

Araştırmada kullanılan LM-2500 gaz türbini motor görseli Şekil 2.6'da gösterilmiştir.



Şekil 2.6. LM-2500 Motorunun Yapısı (Coraddu ve ark. (2015))

LM-2500 motorda, hava ilk önce havanın basınçlı olduğu kompresöre girer (parametre 12) ve ısınır (parametre 9). Basınçlı havaya (parametre 16) yakıt eklenir ve yanma, yüksek güçlü (YG) türbin içinden akan daha da yüksek basınç oluşturur. Burada basınç ve sıcaklık (parametre 10 ve parametre 11) kaydedilir. YG türbini dakika başına dönüş (RPM) olarak ölçülen gaz üreteci şaftına güç verir ve döndürür (parametre 5). YG türbinini terk eden gazlar, torkun olduğu düşük güç (DG) türbine akar (parametre 3) ve hız (parametre 4) izlenir. Parametre 14 itici motor için yakıt akış yüzdesi anlamında yakıtla ilgilidir. Kol konumu (parametre 1) geminin hızını (parametre 2) kontrol ve bu sisteme giren tek kullanıcıdır ve geminin hızını kontrol eder. Yaklaşık olarak 3 kol konumu 1 gemi hızına eşittir. Parametre 17 çıkış parametresidir. Türbin bozulma durumu katsayısı, 16 öznitelik vektörü analiz edilerek önceden belirlenir ve türbin bozulma durumu katsayısı hesaplanır. Türbin bozulma durumu katsayısı 0,975 ile 1 arasında optimize edilmiştir.

Araştırmada kullanılan veri kümesindeki model, kombine dizel elektrik ve gaz (CODLAG) makinesi ile karakterize edilmiştir. Bu modelde, bir geminin birden çok bileşeni ve alt sistemi modellenmiştir. Oluşturulan veri kümesi 589.223 gözlemden ve 17 özellikten oluşmaktadır. Kullanılan parametreler Tablo 2.1'de verilmiştir.

Tablo 2.1. Araştırmada Kullanılan Parametreler

No.	Parametreler	Birim
1	Kol konumu	[lp]
2	Gemi hızı	[knots]
3	Gaz türbini şaft torku (GGT)	[kN]
4	Gaz türbini devir hızı (GTn)	[rpm]
5	Gaz jeneratörü devir hızı (GGn)	[rpm]
6	Sancak pervane torku (Ts)	[kN]
7	İskele pervane toku (Tp)	[kN]
8	YG Türbin çıkış sıcaklığı (T48)	[C]
9	GT kompresör giriş hava sıcaklığı	[C]
10	GT kompresör çıkış hava sıcaklığı	[C]
11	YG Türbin çıkış basıncı (P48)	[bar]
12	GT kompresör giriş hava basıncı	[bar]
13	GT kompresör çıkış hava basıncı	[bar]
14	Gaz türbini egzoz gazı basıncı	[bar]
15	Türbin enjeksiyon kontrolü (TIC)	[%]
16	Yakıt akışı (mf)	[kg/s]
17	GT türbin bozulma durumu katsayısı	-

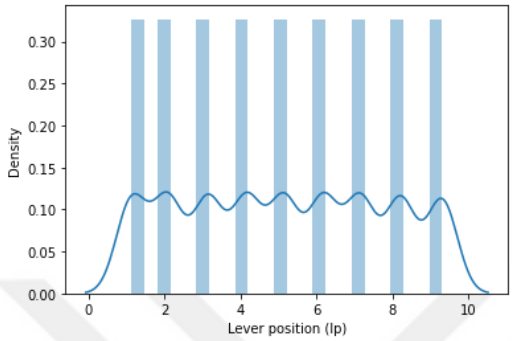
Kullanılan parametrelere ait örnek beş (5) değer Tablo 2.2'de verilmiştir.

Tablo 2.2. Araştırmada Kullanılan Örnek Parametre Değerleri

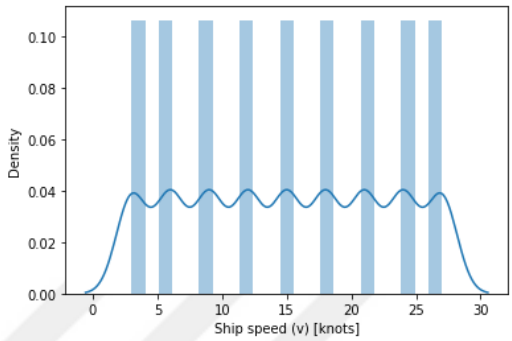
Parametreler	Değer 1	Değer 2	Değer 3	Değer 4	Değer 5
Kol konumu	6.175	2.088	6.175	1.138	9.3
Gemi hızı	18	6	18	3	27
Gaz türbini şaft torku (GGT)	29791.09	6441.58	29761.84	2065.65	72769.61
Gaz türbini devir hızı (GTn)	2307.427	1355.95	2306.842	1401.12	3560.4
Gaz jeneratörü devir hızı (GGn)	8819.747	6771.17	8792.163	6678.14	9743.582
Sancak pervane torku (Ts)	246.271	24.434	246.018	10.457	645.07
İskele pervane toku (Tp)	246.271	24.434	246.018	10.457	645.07
YG Türbin çıkış sıcaklığı (T48)	761.14	589.657	770.037	496.166	1052.557
GT kompresör giriş hava sıcaklığı	288	288	288	288	288
GT kompresör çıkış hava sıcaklığı	664.474	567.062	664.637	553.649	771.195
YG Türbin çıkış basıncı (P48)	2.505	1.312	2.514	1.172	4.531
GT kompresör giriş hava basıncı	0.998	0.998	0.998	0.998	0.998
GT kompresör çıkış hava basıncı	13.113	7.045	13.319	6.305	22.631
Gaz türbini egzoz gazı basıncı	1.031	1.02	1.031	1.019	1.051
Türbin enjeksiyon kontrolü (TIC)	33.433	18.963	34.108	0	87.417
Yakıt akışı (mf)	0.662	0.256	0.675	0.121	1.73
GT türbin bozulma durumu katsayısı	0.997	0.996	0.983	0.986	0.992

Kullanılan veri tabanında girdi ve çıktı değişkenlerinin uygun dağılım gösterdiği görülmektedir. Bu durum, çalışmada sunulan hesaplama yöntemlerinin uygulanmasını kolaylaştırmaktadır. Kullanılan veri tabanından elde edilen histogram grafikleri Şekil

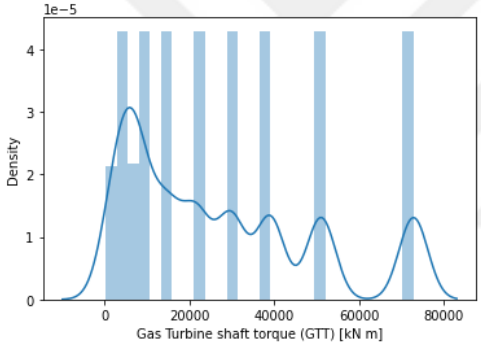
2.7’de verilmiştir. Grafiklerdeki x eksenindeki her sütun verileri temsil ederken, y ekseni ise çekirdek yoğunluğu tahmini için olasılık yoğunluk işlevini göstermektedir. Olasılık yoğunluğu, x eksenindeki birim başına olasılıktır. Parametrelerin histogram grafikleri, verilerin makine öğrenimi modellerine daha iyi uyum sağlaması için hangi yöntemin en uygun olduğu konusunda daha iyi bir fikir vermiştir.



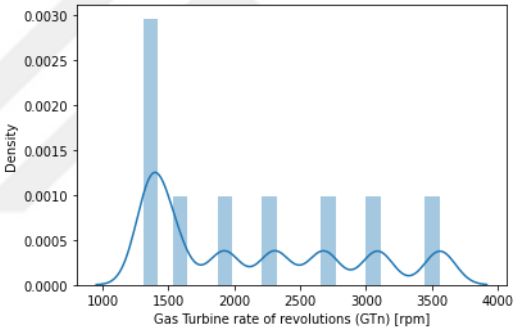
(a) Kol Konumu HG



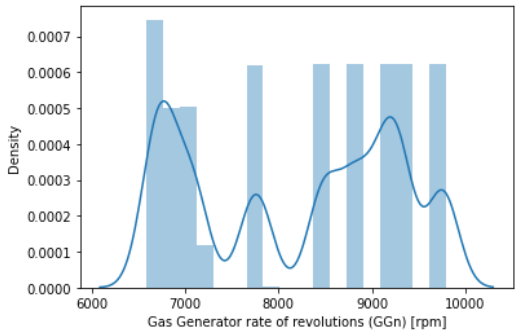
(b) Gemi Hızı HG



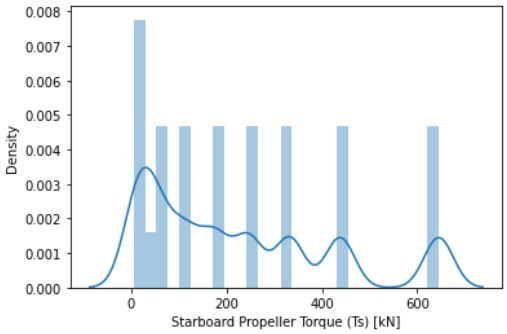
(c) Gaz türbini şaft torku (GGT) HG



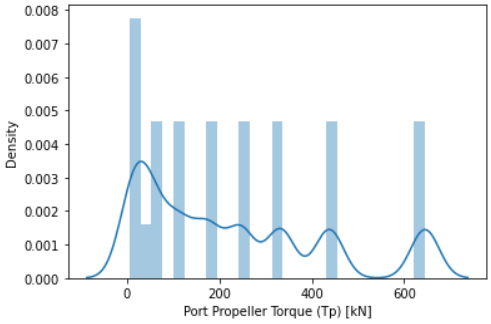
(d) Gaz türbini devir hızı (GTn) HG

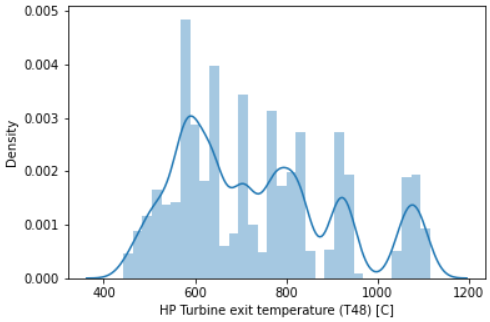


(e) Gaz türbini devir hızı (GTn) HG

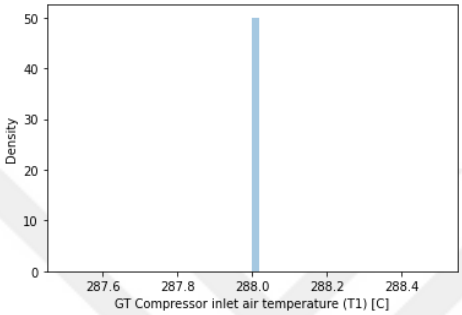


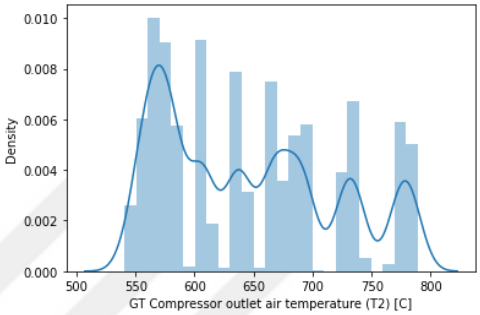
(f) Sancak pervane toku (Tp) HG



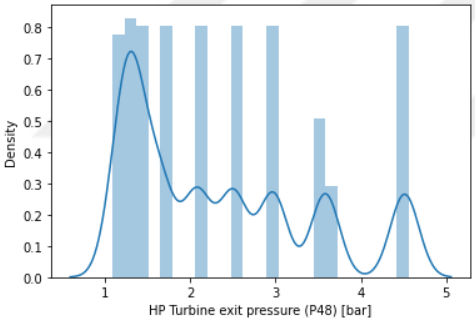


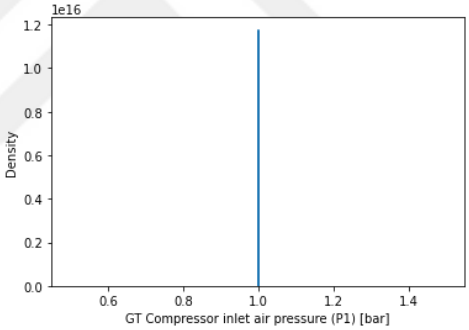
(g) İskele pervane toku (Tp) HG (h) HP Türbin çıkış sıcaklığı (T48) HG



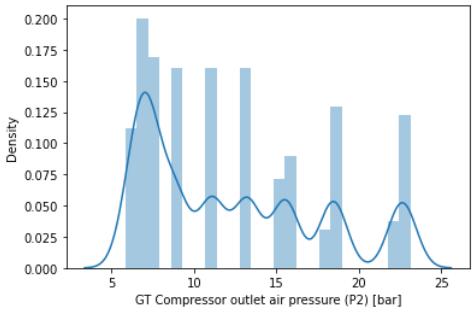


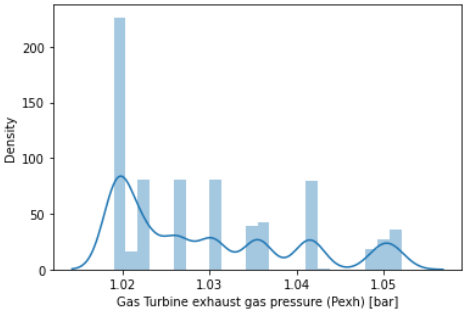
(ı) GT kompresör giriş hava sıcaklığı HG (i) GT kompresör çıkış hava sıcaklığı HG



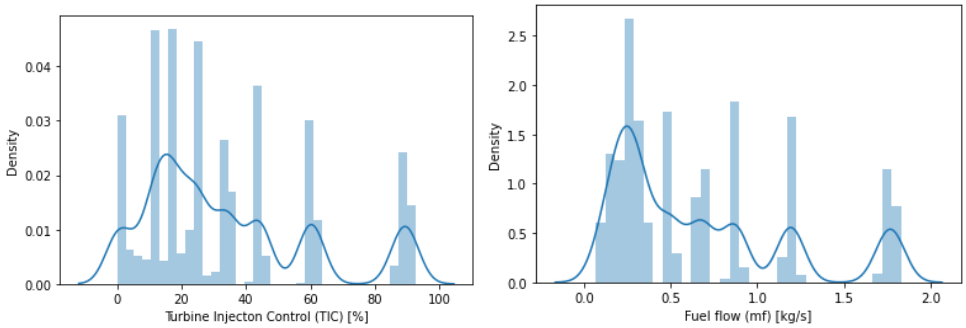


(j) HP Türbin çıkış basıncı (P48) HG (k) GT kompresör giriş hava basıncı HG





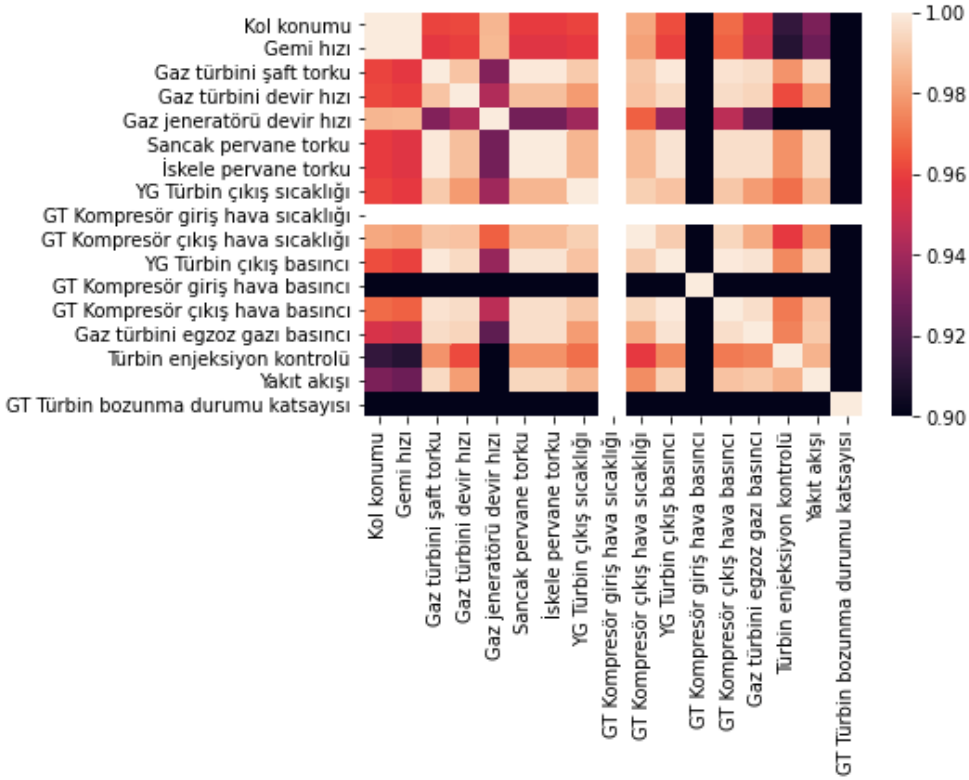
(l) GT kompresör çıkış hava basıncı HG (m) Gaz türbini egzoz gazı basıncı HG



(n) Türbin enjeksiyon kontrolü HG (o) Yakıt akışı (mf) HG

Şekil 2.7. Tüm Parametrelere İlişkin Histogram Grafikleri

Bu çalışmada toplanan veriler eğitim seti ve test seti olmak üzere ikiye bölünerek modeller oluşturulmuştur. Nihai sonucu üretmek için eğitim seti ve hata oranını tahmin etmede modelin doğruluğunu göstermek için test seti kullanılmıştır. Veri kümesindeki özelliklerin ısı matrisi Şekil 2.8'de verilmiştir. Isı matrisindeki koyu siyah renk, değişkenler arasındaki güçlü negatif ilişkiyi gösterirken, beyaz renk ise güçlü pozitif ilişki olduğunu göstermektedir. Değişkenler arasındaki ilişkinin gücü arttıkça renk beyazlaşmaktadır.

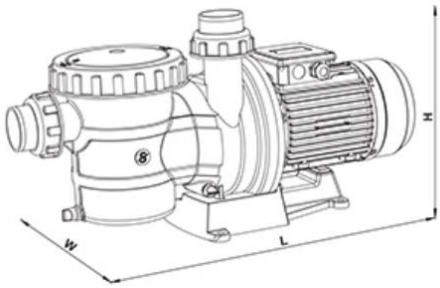


Şekil 2.8. Parametreler Arasındaki Korelasyona İlişkin Isı Matrisi

2.1.2 Gemi Makineleri Gürültü Modeli

Endüstride hemen hemen her alanda ve her tür makine üzerinde endüstriyel makineleri görmek mümkündür. Kullanılan makineler, uygulama alanına göre farklılık göstermektedir. Bu makinelere pompalar valfler örnek verilebilir. Pompa makinelerinin en çok kullanıldığı alanlar havuzlardır. Havuz pompaları bir havuzdan suyu tahliye eden ve havuza su tahliye eden pompalardır. Bu pompalar havuz sisteminin kalbi olarak tarif edilir.

Havuz pompaları havuz içerisindeki sirkülasyonu sağlayan makinelerdir. Bu makineler havuzdan suyu çekerek pompanın ağzında bulunan filtreler aracılığıyla suyu böcek, yaprak, dal gibi istenmeyen maddelerden arındırır. Aynı zamanda suyu sürekli devir daim yaptırdığı için bakterilere karşı kullanılan havuz kimyasalları ve klorun suya eşit miktarlarda dağıtılmasını sağlar. Örnek bir havuz pompası makinesi Şekil 2.9'da gösterilmiştir.



Şekil 2.9. Havuz Pompası Makinesi

Bu çalışmada makine gürültü modeli için MIMII veri kümesi kullanılmıştır. Bu veri kümesi dört farklı tür makinenin sesini içerir: valfler, pompalar, fanlar ve sürgülü raylar. Çalışmada havuz pompaları veri kümesi kullanılmıştır. Veri kümesinde veriler normal sesler (5000 saniyeden 10000 saniyeye kadar) ve anormal seslerden (yaklaşık 1000 saniye) oluşmaktadır. Anormal seslerin oluşumunda gerçek hayattaki bir senaryoyu andırmak için, çeşitli sesler kaydedilmiştir (örneğin, kirlenme, sızıntı, dönen dengesizlik ve ray hasarı). Anormal seslere ek olarak gerçek fabrikada kaydedilen arka plan gürültüsü, makine sesleri de karıştırılmıştır. Sesler, 16 kHz örnekleme hızı ve örnek başına 16 bit ile sekiz kanallı mikrofon dizisi ile kaydedilmiştir. Pompa sesleri 1006 normal ve 143 anormal ses içermektedir.

MIMII veri kümesi, gemi motorlarında elde edilecek akustik verilerin Yapay Zekâ (YZ) ve Makine Öğrenmesi (MÖ) algoritmaları ile analiz edilmesi için örnek oluşturmaktadır. Bu çalışma ile motorun arızalı/arızasız olma durumu kestirilebilecek ve arıza cinsine göre sınıflandırabilecektir. Gemi makinelerinin arızalarının tespit edilmesi maliyeti, zamanı azaltacaktır. Oluşan arızalarının kısa süre içerisinde tespit edilmesiyle kükürt ve kurum miktarının azaltılması sağlanacaktır. Böylece yeşil dönüşüm kapsamında ülkemize önemli katkı sağlanmış olacaktır. Bu durum Avrupa Birliğini'nin hedeflerinden biri olan 2050'ye kadar net sera gazı emisyonlarının sıfırlanması ve ekonomik büyümenin kaynak kullanımına bağlılığının sona ermesi" stratejisini desteklemektedir. Böyle bir çalışmanın temelini oluşturacak olan arıza tespiti ülkemize yenilik getirecektir.

2.2. Makine Öğrenmesi Kavramı

Yapay zekanın bir dalı olan makine öğrenimi, verilerden karmaşık kalıpları tanımlamak ve rasyonel kararlar vermek için istatistikleri ve bilgisayarların bilgi işlem gücünü kullanır. Makine öğrenimi teknikleri, sınıflandırma ve regresyon problemlerinde başarıyla kullanılmaktadır (Nguyen, Armitage, 2008). Bazı makine öğrenme sistemleri insan sezgisine olan ihtiyacı tamamen ortadan kaldırmaya çalışırken, bazıları insan ve makine arasında işbirlikçi bir yaklaşım benimser. Ancak sistemi tasarlayan kişinin verilerin kodlanma şekline tamamen hâkim olması, insan sezgisinin tamamen ortadan kaldırılmasını imkânsız hale getirir. Makine öğrenimi deneysel algoritmanın otomatikleştirilmesi olarak görülür.

Yapay zekâ çalışmalarında makine öğrenimi önemli bir yere sahiptir. Bir sistemin öğrenme yeteneği yoksa gerçekten zeki olduğunu iddia etmek zordur, geçmişteki akıllı sistemler genellikle öğrenme yeteneği yoktur. Bu sistemler bir hatayı kendi kendine düzeltemez, deneyim yoluyla performansını iyileştiremez, gerekli bilgiyi otomatik olarak alamaz ve keşfedemez. Bu nedenle, sadece mevcut olgu ve teoremleri ispatlayabilir, ancak yeni teorem, kanun ve kurallar keşfedemezler. Yapay zekânın gelişmesiyle birlikte, makine öğrenimi yapay zekâ araştırmalarının merkezi haline gelmiştir. En popüler makine öğrenmesi uygulamaları doğal dil anlama, örüntü tanıma, bilgisayarla görme, akıllı robotik ve otomatik akıl yürütmedir (Xiao, 2015).

Bu uygulamalara ek olarak son zamanlarda yapılan çalışmalar, gemi sistemlerinin tahminlemesi için makine öğrenmesinin en elverişli yöntem olduğunu göstermiştir.

Yapılan çalışmalara bakıldığında hata analizinin yapay sinir ağları modelinin kullanılması ile tespit edilmesini mümkün kılmıştır.

2.2.1. Makine Öğrenmesinin Çeşitleri

Makine öğrenmesi algoritmaları öğrenim özelliklerine göre Denetimli Öğrenme ve Denetimsiz Öğrenme olmak üzere ikiye ayrılır. Bu iki öğrenmeye ek olarak literatürde kullanılan Yarı Denetimli Öğrenme, Takviyeli Öğrenme ve Yoğun Öğrenme de vardır.

Bu öğrenme tekniklerinden bazıları regresyon ve sınıflandırma yaparken bazıları kümeleme yapar.

Denetimli öğrenmede, veri kümesindeki her örnek çifti bir giriş değeri ve tahmin etmeye çalıştığımız bir çıktı değeri içerir. Denetimli bir öğrenme algoritması altında eğitim seti analiz edilerek bir tahmin fonksiyonu üretilir. Bu fonksiyon, örnekleri haritalamak veya tahmin etmek için kullanılır. Denetimli öğrenme hem sınıflandırma hem de regresyon problemlerinde kullanılır.

Denetimsiz öğrenmede, denetimli öğrenmeden farklı olarak, bir çıkış değeri yoktur. Denetimsiz öğrenmede yalnızca girdi değeri vardır ve amaç etiketlenmemiş veri kümesindeki örnekler arasındaki benzerlikleri bulmaktır. Denetimsiz öğrenmeyi gerçekleştirmenin iki yöntemi vardır. Bunlardan biri, boyut küçültmedir. Boyut küçültme, çıkış değerini bulmak için değişken sayısını azaltır. Diğer yöntem ise kümelemedir. Kümeleme, benzer verileri kümeleyerek tahminde bulunur. Denetimsiz öğrenme, daha çok veri madenciliği durumudur. Verilerin yapısını tanımlamak için kullanılır.

Bu çalışmada denetimsiz makine öğrenmesi teknikleri kullanılmıştır. Kullanılan teknikler aşağıdaki bölümlerde açıklanmıştır.

2.2.2 Regresyon Yöntemleri

Regresyon analizi, iki veya daha fazla değişken arasındaki bağlantıyı değerlendirmek için tanımlanan bir analiz yöntemidir. Sadece bir değişken ile analiz yapıldığında tek değişkenli regresyon, birden çok değişken ile analiz yapıldığında çok değişkenli regresyon olarak isimlendirilir. Regresyon analizi ile değişkenler arasındaki ilişkinin varlığı ve eğer ilişki varsa bu ilişkinin gücü hakkında bilgi elde edilir.

2.2.2.1 Doğrusal Regresyon

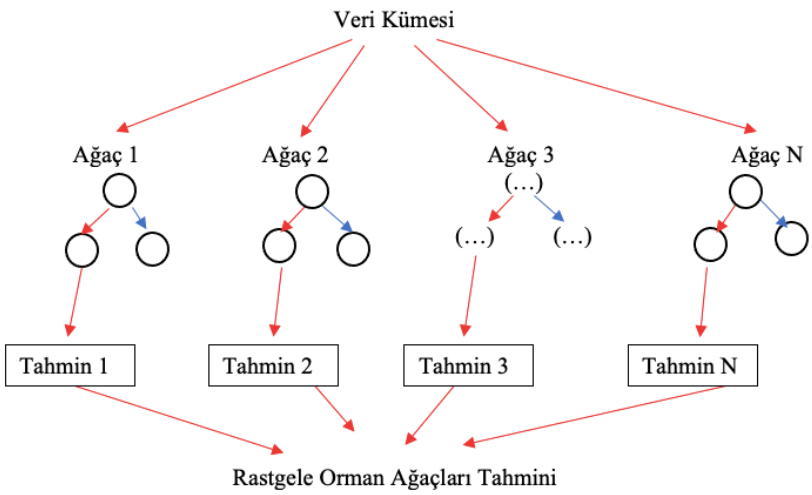
Doğrusal Regresyon algoritması, bir veya daha fazla bağımsız değişken ile başka bir bağımlı değişken arasındaki ilişkiyi modellemek için kullanılan bir yöntemdir. Doğrusal regresyon aşağıdaki formül ile ifade edilir;

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \epsilon_i \tag{2.1}$$

Formülde her tahmin için i = 1, ..., n kadar, y_i tahmin değeridir. β_0 tahmin değerinin kesme noktasının tahmin katsayısı, $\beta_1 X_{i1}$ eğim tahmincisi katsayısı ve ϵ_i artık değerdir. Satır vektörleri *n* boyutludur ve sütun vektörleri *X* matrisi ile ifade edilir.

2.2.2.2 Karar Ağaçları Regresyon

Karar ağaçları algoritması, en çok kullanılan makine öğrenme yöntemlerinden biridir. Karar ağaçları, incelenen veri kümesinin yapısına göre sınıflandırma ve regresyon algoritması ile bir ağaç yapısı oluştururlar (Kavzoğlu, Şahin ve Çölkesen, 2012). Karar ağaçları, bir sorunun çözümünde çok katmanlı ve sıralı bir yaklaşımla karmaşık verileri fazlara dönüştürerek basit bir karar verme süreci gerçekleştirir (Safavian ve Landgrebe, 1991). Karar ağaçları, Şekil 2.10'da görüldüğü gibi kök düğüm (dataset) ara düğüm (Ağaç 1, Ağaç 2, ...) ve son düğüm (Tahmin 1, Tahmin 2, ...) olmak üzere 3 bölümden oluşur. Kök düğüm, ana düğümdür. Ara düğüm karar düğümüdür. Alınan kararlara göre ağaçlar oluşturulur. Ara düğüm, ilgili değişkenin sonucuna göre bölünür. Son düğüm yaprak düğümdür.



Şekil 2.10. Karar Ağaçları Yapısı

Karar ağacı yapısı oluşturmadaki temel ilke, verilerle ilgili bir dizi soru sorarak karar kurallarını oluşturmaktır. Bu işlem için ağaç yapısının temel öğesi olan kök düğümde sorular sorulmaya başlanır ve ağacın dallanması ile ağaç yapısının son öğesi olan yapraklara ulaşılıncaya kadar devam eder (Pal ve Mather, 2003).

2.2.2.3 K-En Yakın Komşular Regresyon

K-En Yakın Komşular Regresyon (K-NN) algoritması 1967'de T. M. Cover ve P. E. Hart tarafından önerilmiştir. Algoritma, sınıfları bilinen bir örnek kümeden elde edilen verilerden yararlanılarak kullanılır. K-NN algoritmasında eğitim kümesindeki örnekler n boyutlu sayısal özelliklerle belirtilir. Tüm eğitim örnekleri n-boyutlu bir örnek uzayda tutulur ve her örnek n-boyutlu uzayda bir noktayı temsil eder. Bilinmeyen bir örnekle karşılaşıldığında, eğitim kümesinden en yakın k örnek belirlenir ve yeni örneğin sınıf etiketi, en yakın k komşunun sınıf etiketlerinin çoğunluk oyu ile atanır (Kamber, Pei, 2006).

K-NN algoritması, sınıflandırma ve regresyon için basit bir yöntemdir. Algoritmanın basitliğine rağmen çok iyi performans gösterir ve önemli bir kıyaslama yöntemidir. K-NN algoritması, eğitim verilerini ve bu verilerin sınıflarının konumunu tutar. Yeni gelen verinin hangi sınıfta olacağına karar vermek gerektiğinde sorgu verisi ile eğitim örnekleri arasındaki mesafe hesaplanır. En az mesafeli örnekler seçilir ve daha fazla örnek gelen sınıf sonucudur (Moosavian, Ahmadi, Tabatabaeefar ve Khazaee, 2013). Örnek veri kümesine dahil edilecek yeni verinin mesafesi mevcut verilere göre tek tek hesaplanır. İlgili mesafelerden en yakın komşular dikkate alınır. Öznitelik değerlerine göre k komşu veya komşu sınıfına atanır. Seçilen sınıf, tahmin edilmesi beklenen gözlem değerinin sınıfı olarak kabul edilir. Başka bir deyişle, yeni veriler etiketlenir. Şekil 2.11 bu yöntemi göstermektedir. Şekilde, 2 sınıf (çember) ve 2 özellik (üçgen ve daire) içeren bir durumu göstermektedir. Verinin ait olduğu sınıf, kürenin merkezinden olan uzaklığa bakılarak bulunur. Örneğin mesafe = 3 için bakıldığında 3 birim uzaklık hesaplanır. Bu uzaklık içteki küçük çemberi ifade eder ve küçük çember 2 üçgen ve 1 daire içerir. Diğer bir deyişle küçük çember üçgen sınıfıdır. Mesafe = 5 için, büyük çembere bakılır ve bu çemberin içinde 11 üçgen ve 9 daire vardır, dolayısıyla sonuç yine üçgen sınıfıdır.

Şekil 2.11. K-NN Yapısı

K-NN algoritması, çok boyutlu bir girdi uzayında mesafeleri hesaplamak için basit ve uygulaması kolay bir yöntemdir ve en karmaşık makine öğrenimi yöntemleriyle karşılaştırıldığında bile rekabetçi sonuçlar verebilir (Moosavian, Ahmadi, Tabatabaeefar ve Khazaee, 2013). K-NN algoritmasında mesafe değeri aşağıdaki formülle hesaplanır:

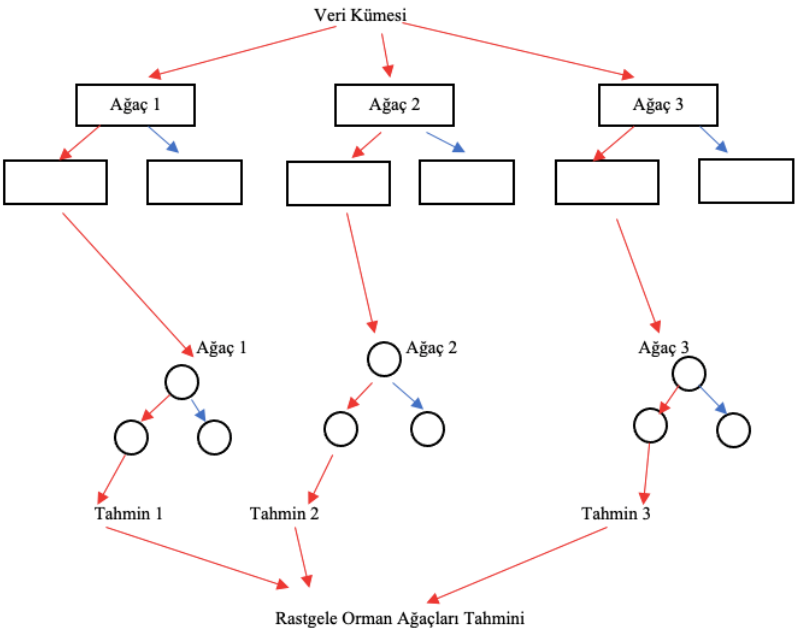
$$d(A,B) = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{2.2}$$

Burada A ve B mesafeyi ölçecek noktalardır. A ve B vektör ile ifade edilir ve vektörleri, $A = x_1, x_2, \ldots, x_k$ ve $B = y_1, y_2, \ldots, y_k$ ile ifade edilebilir. Formülde k vektörün boyutudur yani özellik alanıdır.

2.2.2.4 Rastgele Orman Regresyon

Rastgele Orman (RF) algoritması, karar ağacı sınıflandırıcılarından türetilmekte, sınıflandırma ve regresyon ağaçlarını içermektedir. Rastgele ağaçların bir araya getirilmesiyle orman yapıları elde edilmektedir (Yang, Di ve Han, 2008). Bu yöntemde ilk adım, özyinelemeli bölme prosedürü ile düğümler tarafından oluşturulan bir ağacı kapsamaktadır. Burada bir çıktı değişkenini hesaplamak için birçok farklı karar ağacı oluşturulmaktadır. Tüm terminal düğümler belirli sınıflara atanmakta ve tüm çıktılar her bir ağaçtan değerlendirilmektedir. Nesne için en çok tahmin edilen sınıf atanmaktadır. Rastgele bir ormanın başarısı, her ağacın eğitim domeninden rastgele seçilen verilerle oluşturulduğu, kalanların model doğruluğunu test etmek için kullanıldığı ve ardından ikinci rastgele örneklemenin ağaçtaki her bir düğümün bölünmesini belirlediği iki adıma bağlıdır. Kullanıcı veya RF algoritmasının kendisi tarafından seçilebilen tahmin değişkenleri, her düğümde ikili kural oluşturmak için rastgele seçilmektedir. Bu nedenle rastgeleleştirme, ağaçlar arasında daha az korelasyon ve daha düşük hata oranları

sağlamaktadır (Breiman, 2001; Horning, 2010). RF'nin yapısı Şekil 2.12'de gösterilmektedir.



Şekil 2.12. Rastgele Orman Yapısı

RF algoritmasının avantajı, büyümenin çok geniş olduğu ve çok küçük eğitim verisi alt kümeleri ile ağacı basitleştirmek için muhtemelen gürültü içeren terminal düğümlerin kaldırılmasını ifade eden "budama" işlemidir. Kurallar oluşturulduğunda yorumlama da kolaydır ve sınıflandırma hızlıdır. Liaw & Wiener'e (2002) göre RF uygulamalarında en yaygın parametreler girdi eğitim verileri, ağaç sayısı, tahmin değişkenlerinin sayısı, hata ve değişken anlamlılığını hesaplamak için kullanılan parametrelerdir.

2.2.2.5 Bayesian Ridge Regresyon

Ridge regresyonu, alt küme seçimine karşı bir alternatiftir. Bazı durumlarda en küçük kareler yöntemine göre, katsayıları daha doğru tahmin etmek için kullanılmaktadır. Ridge regresyon, regresyon modelindeki tahminci sayısının azaltılmasına olanak sağlamaktadır (Uyanık, Karatuğ ve Arslanoğlu, 2020). Bayesian Ridge regresyonu, regresyon probleminin olasılıksal bir modelini tahmin etmektedir. Katsayı tahminleri sıfıra daha yakın (veya eşit) ötelenmektedir (Whittaker, Thompson ve Denham, 2000). Ridge regresyon tahmin edicisinin Bayesian yorumu aşağıdaki denklemde verilmiştir (Şamkar Alpu ve Altan, 2011; Uyanık, Karatuğ ve Arslanoğlu, 2020).

$$y = X\beta + \epsilon \tag{2.3}$$

Burada y değişken vektörü, X, değişken matrisi, β parametre vektörüdür ve ϵ, hata vektörüdür.

Bağımsız değişkenler arasındaki çoklu bağlantı nedeniyle regresyon modelinde parametre tahminlerinde çakışma olacaktır. Bu dezavantajın üstesinden gelmek ve daha kesin regresyon parametresi tahminleri elde etmek için Ridge regresyonu küçük sapmaya neden olacaktır. Bayesian Ridge modellemesi, hiyerarşik veri yapısıyla başa çıkmadaki daha yüksek kabiliyeti nedeniyle tercih edilmektedir (Shi, Abdel-Aty ve Lee, 2016). Efendi (2017) farklı eşdoğrusallık düzeyleri bakımından Bayesian Ridge regresyon modelini başarılı bir şekilde simüle etmiştir.

2.2.2.6 Ekstra Ağaç Regresyon

Aşırı düzeyde rasgele ağaçlar olarak da adlandırılan Ekstra Ağaç Algoritması yeni bir makine öğrenme tekniğidir ve RF yönteminin bir uzantısı olarak Geurts, Ernst & Wehenkel (2006) tarafından geliştirilmiştir. Ekstra Ağaç Yöntemi, tahmin edicileri RF olarak eğitmek için rastgele alt küme kullanmaktadır. Aksine, ekstra ağaç regresyonunda rasgeleleştirme ile optimal kesme noktası da elde edilmektedir. Ekstra ağaç, her bir ağacı eğitmek için tüm eğitim veri kümesini kullanmaktadır, RF ise modeli eğitmek için önyükleme kopyası kullanmaktadır (Ahmad, Reynolds ve Rezgui, 2018; Aláiz-Moretón vd., 2019). Başka bir deyişle, RF'de yukarıdan aşağıya bölünme prosedürü rastgele prosedürle değiştirilmekte ve bu da ağacın yanlılığını artırarak varyansı azaltmaktadır. Ekstra ağaçtaki tahmin, tüm olasılık sınıflarının ortalaması alınarak ve "çoğunluk oylaması" olarak adlandırılan en yüksek olasılığa sahip sınıf seçilerek elde edilmektedir. Bu işlemin avantajları, karmaşık problemler için daha iyi sonuçlar vermesi ve hesaplama yükünün azaltılmasıdır (Patil ve Phalle, 2018).

Tablo 2.3. Sayısal Nitelikler için Ekstra Ağaçların Bölme Prosedürü

Düğüm S Bölünmüş
Girdi: düğüme karşılık gelir, yerel öğrenmenin alt kümesi S Çıktı: böl [$a<a_c$] DOĞRU, Bölünmüş S'yi durdur Hiçbir şey döndürme; K öğesini seçin, aksi takdirde; K için, nitelikler {a_1,…..a_k} K çizin; K seçimi için {s_1, … s_k} böler Rastgele bölme seç= s_i; (S, a_i), ∀=1,….K. Bölünmüş dönüş, Puan (s_i,S) =maks $_{i=1,\ldots K}$
Rastgele Bölme (S, a) Seçimi
Girdi: öznitelik ,S alt küme; Çıktı: bölme Maksimum a^s_{max} , Minimum a^s_{min} , S değeri Rastgele kesme noktası çizin; a_c rastgele kesme noktası, [a^s_{max} , a^s_{min}] tekdüze [$a<a_c$] dönüş bölümü
Böl Durdur
Giriş: S alt küme Çıktı: boole DOĞRU, $\|S\| < n_{min}$, TRUE, S'de sabit bir öznitelik DOĞRU, S cinsinden çıktı sabiti Aksi takdirde, YANLIŞ dönüş;

Tablo 2.3'teki algoritma, minimum örneklem boyutu (n_{min}) ve K gibi her bir düğümde rastgele seçilen sayıda özniteliği ele almaktadır. Topluluk modeli, birçok kez tüm orijinal öğrenme örneğini oluşturmaktadır. Yukarıdaki parametreler, hesaplama avantajlarını ve özerkliği en üst düzeye çıkarmak için varsayılan ayarlarla manuel olarak farklı problemlere uyarlanabilir (Nistane ve Harsha, 2018).

2.2.2.7 Doğrusal Destek Vektör Makineleri Regresyon

Doğrusal Destek Vektör Makineleri Regresyonu (DVM) yaygın kullanılan bir regresyon yöntemidir ve destek vektör sınıflandırmasının (DVS) genişletilmiş halidir (Vapnik, 1995). Bu yöntem, modellere ne kadar hatanın tolere edilebileceğini tanımlama ve

verilerle uyumlu doğruyu belirleme esnekliği sağlamaktadır. DVM, yapısal riski en aza indirmekte ve aşırı uyumluluk sorununu etkin bir şekilde çözmektedir (Lin, Lin, Zhou ve Yao, 2007). DVM'de verileri daha yüksek boyutlu bir uzaya eşlemek için çekirdek fonksiyonu kullanılmaktadır. Doğrusal DVM $w^T x_i$ 'nin en küçük y_i değerine yakın olduğu bir model (w) belirlemektedir. $\{(x_i, y_i)\}$, $x_i \in R^n, y_i \in R,\ i = 1, \dots, l$ gibi bir eğitim dizisi dikkate alındığında, DVM aşağıdaki gibi bir düzenli optimizasyon problemini çözmektedir:

$$\min_{w} f(w),\ where\ f(w) \equiv \frac{1}{2} w^T w + C \sum_{i=1}^{l} \xi_{\in}(w; x_i, y_i) \tag{2.4}$$

Yukarıdaki denklemde, C>0 düzenleme parametresidir ve $\xi_{\in}$, $(xi,\ yi)$ ile ilişkili ε-duyarsız kayıp fonksiyonudur (Ho ve Lin, 2012).

$$\xi_{\in}(w; x_i, y_i) = \begin{cases} max(|w^T x_i - y_i| - \epsilon, 0)\ \ or \\ max(|w^T x_i - y_i| - \epsilon, 0)^2 \end{cases} \tag{2.5}$$

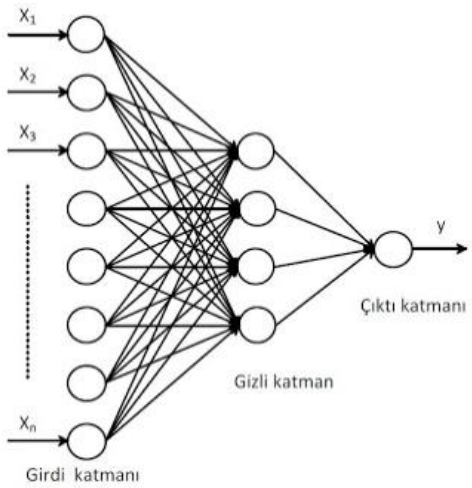
DVM, L2 kaybının türevlenebilir olduğu durumlarda sırasıyla L1-kayıp ve L2-kayıp DVM olarak her iki denklemi de kullanmaktadır. İlk denklem minimize edildiğinde, tahmin fonksiyonu $w^T\ x$ olmaktadır (Ho ve Lin, 2012). Ayrıca DVM'de bir bias (b) vardır ve tahmin fonksiyonu $w^T\ x + b$ (bias problem) olmaktadır. Burada b bir destek vektörünün yer değiştirmesiyle elde edilebilir (Lin, Lin, Zhou ve Yao, 2007). Büyük ölçekli doğrusal sınıflandırma üzerine yapılan araştırmalar, çoğu veri için performans üzerindeki ihmal edilebilir etkilerden dolayı genellikle b'yi ihmal etmektedir (Ho ve Lin, 2012).

2.2.3. Yapay Sinir Ağları

Makine öğrenmesinde kullanılan temel araçlardan biri yapay sinir ağlarıdır. Yapay sinir ağları, insanların öğrenme şeklini simüle etmeyi amaçlayan ve insan beyninden ilham alan sistemlerdir. Sinir ağları, girdi ve çıktı katmanlarından ve çoğu durumda girdileri çıktı katmanının kullanabileceği bir veriye dönüştüren gizli birim katmanından oluşmaktadır. Yapay sinir ağları, karmaşık örüntüleri tanımlamak için en uygun araçlardan biridir (Chouhan ve ark., 2020).

Haykin'e (1998) göre bir sinir ağı, deneysel bilgiyi depolamak ve bu bilgiyi kullanıma hazır hale getirmek için doğal bir eğilime sahip olan basit işlem birimlerinden oluşan,

büyük ölçüde paralel dağıtılmış bir işlemci olarak tanımlanmaktadır. Bu yapı insan beynine iki açıdan benzemektedir. Bilgi, ağ tarafından bir öğrenme süreciyle çevreden edinilmekte ve elde edilen bilgiyi depolamak için sinaptik ağırlıklar olarak bilinen internöron bağlantı güçleri kullanılmaktadır. Şekil 2.13'te gösterildiği gibi bir yapay sinir ağı, katmanlar halinde birleştirilmiş nöronların birbirine bağlanmasından oluşmaktadır. Her katman, sayısal ağırlıklı bağlantılar kullanarak birbirleriyle etkileşime giren düğümler veya nöronlar adı verilen bir dizi basit, nöron işleme elemanına sahiptir. Ağ, sırasıyla ikisi giriş ve çıkış katmanı olan n sayıda nöron katmanından oluşmaktadır. Birincisi dış sinyalleri alan ve ileten ilk ve tek katmandır. İkincisi ise hesaplamaların sonuçlarını gönderen son katmandır. İç kısımlardaki ağlar gizli katmanlardır ve alınan sinyallerden ilgili özellikleri veya örüntüleri çıkarmaktadırlar. Önemli olduğu düşünülen bu özellikler daha sonra çıktı katmanına yönlendirilmektedir. Gelişmiş sinir ağları, ilgili özellikleri veya örüntüleri ayırt etmede ağı mümkün olduğunca etkili kılmak için tasarlanmış birkaç gizli katmana, geri bildirim döngüsüne ve zaman geciktirme öğelerine sahip olabilir (Lazakis, Raptodimos ve Varelas, 2018).

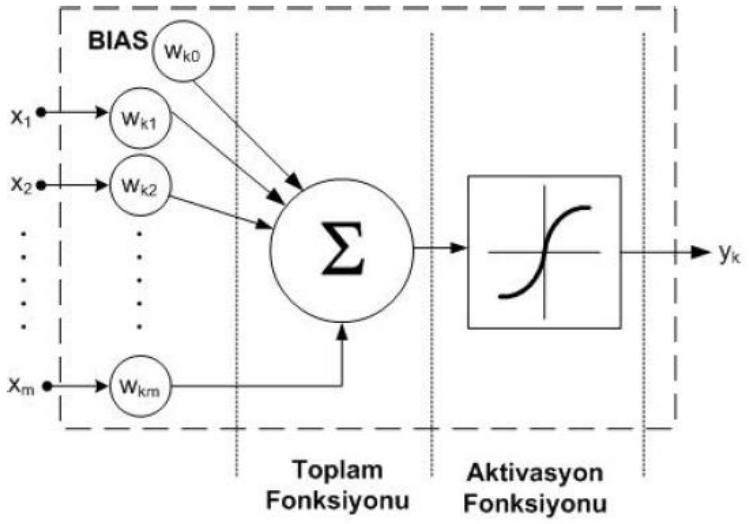


Şekil 2.13. Yapay Sinir Ağı Yapısı

Bu araştırmada kullanılacak Yapay Sinir Ağları (YSA), beyin gibi biyolojik sinir sistemlerinin bilgiyi işleme şeklinden esinlenen bir bilgi işleme paradigmasıdır. Bu paradigmanın en önemli unsuru, bilgi işleme sisteminin yeni yapısıdır (Zhongling, Ning ve Yan, 2005). Bu sistem, belirli problemleri çözmek için birlikte çalışan birbiriyle bağlantılı çok sayıda işlem elemanlarından (nöronlardan) oluşmaktadır. Sinir ağları, karmaşık veya kesin olmayan verilerden anlam çıkarma konusundaki güçlü yönüyle

desenler/örüntüler çıkarmak ve insanlar ya da diğer bilgisayar teknikleri tarafından fark edilemeyecek kadar karmaşık olan eğilimleri tespit etmek için kullanılmaktadır (Zhou ve Xu, 2011).

Bir yapay sinir ağında, girdi sinyalleri belirli katsayılar aracılığıyla ağırlıklandırılmış toplamı bir eşik fonksiyonundan geçirilmektedir. Her giriş değerleri w ile belirtilen ağırlık katsayıları ile çarpıldıktan sonra toplam fonksiyonuna gönderilir. "Bias" olarak adlandırılan ve giriş değerlerinden bağımsız olan sabit bir sayı da bu toplamda bulunmaktadır. Bias için genellikle "1" değeri kullanılmaktadır. Bu ağırlıklandırılmış toplam "aktivasyon fonksiyonu" olarak tanımlanan ve bir eşik fonksiyonu görevi üstlenen, genelde doğrusal olmayan bir fonksiyona tabi tutulmaktadır (Şekil 2.14) (Köroğlu, 2007).



Şekil 2.14. YSA'da Bir Nöron Modeli

Bir nöronun karakteristik yapısını belirten denklemde x_i girdi, y çıktı, $i=1, 2,\ldots,m$ arasındaki değerler, f aktivasyon (doğrusal olmayan dönüşüm) fonksiyonu, w_i ise ağırlıktır. YSA'da en sık kullanılan dönüşüm fonksiyonu ise sigmoid fonksiyonudur ve fonksiyonda "c" bir sabit sayıdır (Liu, 2007).

$$y = f[\sum_{i=1}^{m} w_i x_i] \quad \text{ve} \quad f(t) = \frac{1}{1+e^{ct}} \tag{2.6}$$

ReLu fonksiyonu, pozitif eksende 0 (sıfır) ile + ∞ (sonsuz) arasında bir değer, negatif eksende ise 0 (sıfır) değerini alır. Bu durum diğer aktivasyon fonksiyonlarına göre daha hızlı çalışmasını sağlar. Diğer bir aktivasyon fonksiyonu Swish, Google tarafından

geliştirilmiştir. ReLu'dan farklı olarak negatif eksende de değer alabilir. Bu aktivasyon fonksiyonu ReLu'dan daha iyi performans göstermesi sebebiyle çok katmanlı YSA'larda kullanılır.

Öğrenme ve adaptasyon, belirsiz ortamlarda sistemlerin performansını iyileştirmek amacıyla geri bildirim kontrol sistemleri için önemli becerilerdir. Uyarlanabilir sinir ağı kontrolü ile ilişkili bir diğer önemli konu ise, geribildirim kontrol sistemlerinin performansını iyileştirmek için "yaparak öğrenme" ve "öğrenilen bilgilerle yapma" anlamına gelen öğrenme becerisidir (Dai, Wang ve Wang, 2016).

2.3. Derin Öğrenme Modelleri

YSA'dan geliştirilen derin öğrenme, günümüzde makine öğrenmesi alanında yaygın kullanılan bir tekniktir. Derin öğrenme modelleri genellikle katmanları birbirine bağlayan hiyerarşik yapıları benimsemektedir. Bir alt katmanın çıktısı, basit doğrusal veya doğrusal olmayan hesaplamalarla yüksek bir katmanın girdisi olarak kabul edilmektedir. Bu modeller, verilerin düşük seviyeli özelliklerini yüksek seviyeli soyut özelliklere dönüştürebilir. Böylece derin öğrenme modelleri özelliklerin sunumu açısından diğer yöntemlere göre daha güçlüdür. Derin öğrenmenin kullanıcı deneyimi yerine veriye dayalı olması sayesinde kullanıcıya duyulan ihtiyacı azalttığı söylenebilir (Du vd., 2016). Derin öğrenme kapsamında kullanılan teknikler otomatik kodlayıcılar (autoencoders), evrişimsel sinir ağları (ESA) ve tekrarlayan sinir ağları (TSA) olarak sınıflandırılmaktadır.

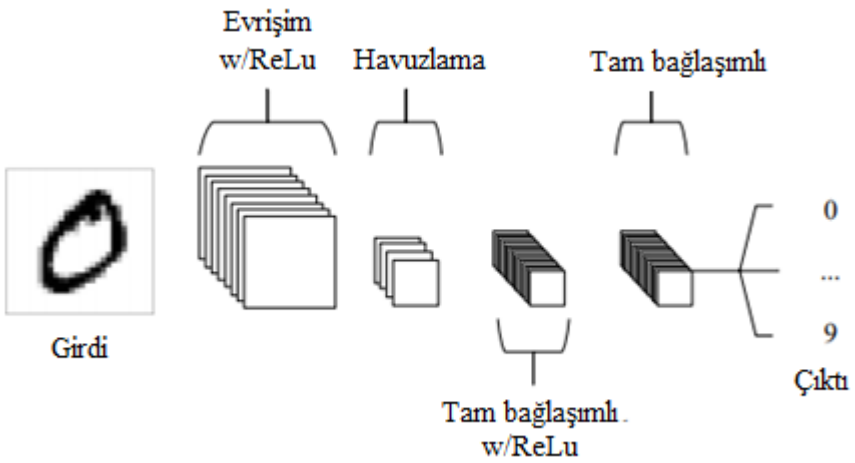
2.3.1 ESA Modeli

Son 10 yılda ESA'nın araştırma ve uygulama alanlarında artış gözlemlenmiştir. ESA'nın günümüzde üretilen büyük miktarlardaki verilerle çalışabilmesi bunu sağlamıştır. Uydular tarafından üretilen resimler, videolar ve diğer verilerle sınıflandırma işlemi, yüksek doğrulukla ESA kullanılarak gerçekleştirilir (Kurt, 2018).

ESA, canlıların doğal görsel algılama mekanizmasından esinlenerek ortaya atılan bir derin öğrenme mimarisidir. 1959 yılında Hubel ve Wiesel, hayvanların görsel korteksindeki hücrelerin, alıcı alanlardaki ışığı algılamaktan sorumlu olduğunu bulmuşlardır. Bu keşiften ilham alan Kunihiko Fukushima, 1980 yılında ESA'nın öncül

olarak kabul edilebilecek “neocognitron”u önermiştir. 1990 yılında LeCun ve arkadaşları, ESA’nın modern çerçevesini oluşturan eseri yayınlamışlar ve daha sonra mimariyi geliştirmişlerdir. Burada LeNet-5 adlı el yazısı rakamları sınıflandırabilen çok katmanlı bir yapay sinir ağı geliştirmişlerdir. Diğer sinir ağları gibi LeNet-5 de çoklu katmanlara sahiptir ve geri yayılım algoritması ile eğitilebilir. LeNet-5, orijinal görüntünün çeşitli gösterimlerini oluşturmakta, çok az ön işleme ile ham piksellerden görsel kalıpları doğrudan tanımayı mümkün kılmaktadır (Gu vd., 2018).

ESA’lar üç tip katmandan oluşmaktadır. Bunlar evrişim katmanları, havuzlama katmanları ve tam bağlaşımlı katmanlardır. Bu katmanların bir araya gelmesiyle ESA mimarisi oluşturulmuştur. Temel ESA mimarisi Şekil 2.15’te gösterilmektedir.



Şekil 2.15. ESA Mimarisi

Yukarıdaki ESA mimarisinin temel işlevselliği dört temel alanda sınıflandırılmaktadır (O’Shea ve Nash, 2015):

1. Diğer YSA formlarında olduğu gibi, giriş katmanında görüntünün piksel değerleri tutulmaktadır.

2. Evrişim katmanı, ağırlıklar ile giriş hacmine bağlı bölge arasındaki skaler çarpımı hesaplayarak, girişin yerel bölgelerine bağlı nöronların çıkışını belirlemektedir. Doğrultulmuş lineer birim (genellikle ReLu olarak kısaltılmaktadır), önceki katman tarafından üretilen aktivasyonun çıktısına sigmoid gibi bir aktivasyon fonksiyonunu uygulamayı amaçlamaktadır.

3. Havuzlama katmanı daha sonra verilen girdinin uzamsal boyutluluğu boyunca basitçe alt örnekleme gerçekleştirecek ve bu aktivasyon içindeki parametre sayısını daha da azaltacaktır.

4. Tam bağlaşımlı katmanlar daha sonra standart YSA'larda bulunan aynı görevleri yerine getirecek ve aktivasyonlardan sınıflandırma için kullanılmak üzere sınıf puanları üretmeye çalışacaktır. Performansı artırmak için bu katmanlar arasında ReLu kullanılması da önerilmektedir.

İyi bir derin öğrenme yöntemi olan evrişimli sinir ağı (ESA), daha az hesaplama kaynağı ve eğitim verisi kullanarak ESA parametrelerinin eğitilmesine yardımcı olmakta ve ortak filtreler kullanılarak daha az bağlantıyla oluşturulmaktadır. Arıza teşhis problemlerini çözmek için etkili bir araç olan ESA'lar, girdilerin temsili özellikleri otomatik olarak çıkarmak için eğitilebilir bir "özellik öğrenici" olarak yaygın bir şekilde kullanılmıştır ve ardından arızayı gerçekleştirmek için lojistik regresyon gibi bir makine öğrenmesi algoritması uygulanmıştır. Xia vd. (2018) arıza teşhisi için kaynaşmış çoklu sensör sinyallerinden arıza özelliklerini belirlemek için kullanılan ESA tabanlı bir yöntem önermişlerdir. Jiang vd. (2019) çok ölçekli sinyallerden temsili özellikleri çıkarmak ve hata sınıflandırmasını gerçekleştirmek için kullanılan çok ölçekli bir ESA modeli tasarlamışlardır. Yang vd. (2019) döner makinelerin arıza teşhisi için hiyerarşik sembolik analiz ve ESA'dan oluşan bir arıza teşhis yöntemi geliştirmişlerdir. Bununla birlikte bu ESA modellerinin girdi verilerinin oluşturulması gerekmekte ve sığ makine öğrenmesi algoritmalarının yalnızca sınırlı bir tanılama kapasitesi bulunmaktadır. Bu nedenle sinyal ön işleme olmadan doğrudan ham sinyallerden hata özelliklerini keşfetmek için uygun bir hata özelliği çıkarma yönteminin geliştirilmesi gereklidir (Li vd., 2021).

BÖLÜM 3 – BULGULAR VE TARTIŞMA

Çalışmada tasarlanan makine öğrenmesi modelleri AMD Ryzen 5 4600H işlemci 64 GB bellekli Windows 10 işletim sistemi bulunan bilgisayar üzerinden eğitilmiş ve test edilmiştir. Eğitilen modellerin performansını değerlendirmek için sistemin Ortalama Kare Hata (Mean Squared Error) (MSE), Korelasyon Katsayısı (R^2), Ortalama Mutlak Hata (Mean Absolute Error) (MAE), Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error) (MAPE), ve doğruluk değerlerine bakılmıştır. Modellerde kullanılan gerçek değerler ile tahmin edilen değerler arasındaki fark azaldıkça tahmin edilen değerlerin gerçek değerlere göre doğru tahmin edildiği bilinmektedir. Diğer başlıklarda, kullanılan modellerin analiz sonuçları ve bu sonuçların doğruluğunu değerlendirmek amacıyla hata metrikleri açıklanacaktır.

3.1. Modelin Değerlendirilmesinde Kullanılan Metrikler

Gaz türbini hata oranı için tahmin edilen değerler 4 farklı metrik kullanılarak değerlendirilmiştir. Bu metrikler Ortalama Kare Hata (MSE), Korelasyon Katsayısı (R^2), Ortalama Mutlak Hata (MAE) ve Ortalama Mutlak Yüzde Hata (MAPE)'dır.

3.1.1. Hata Metrikleri

MAE ve MSE, 0 ile ∞ arası değerler alabilir. Bu oranların hesaplamalarından elde edilen değerler ne kadar düşükse tahmin algoritması o kadar başarılıdır. Modellerin doğruluk oranı R^2 ile değerlendirilir. R^2 değeri bir (1)'e ne kadar yakınsa tahmin modelinin başarısı artar. Bu metriklerin formüllerinde bulunan semboller aşağıda açıklanmıştır.

y = $Gerçek\ Değerler$

$\overline{y}$ = $Gerçek\ Değerlerin\ Ortalaması$

$\hat{y}$ = $Tahmin\ Edilen\ Değerler$

n = $Veri\ Sayısı$

Akustik işaretlerin arıza durumu sınıflandırılması için tahmin edilen değerler doğruluk metriği kullanılarak değerlendirilmiştir. Doğruluk metriği gerçek pozitif, gerçek negatif, yanlış pozitif ve yanlış negatif değerlerinin oranlanması ile hesaplanır. 0 ile 1 arasında

değerler alır. Bu değerin 1'e yakın olması tahmin başarısının yüksek olduğu anlamına gelir.

3.1.1.1. Korelasyon Katsayısı (R^2)

Korelasyon katsayısı, değişkenler arasındaki ilişkiyi belirtmek için kullanılan katsayıdır. Formül 4.1 de gösterildiği şekilde hesaplanır ve -1 ile 1 arasında bir değer alır. Değerlerin negatif olması değişkenler arasında negatif bir ilişkinin olduğunu, pozitif olması değişken arasında pozitif bir ilişkinin olduğunu gösterir. Değerler 1 veya -1 olduğunda değişkenler arasında mükemmel bir ilişki vardır. Değerler 0'a yaklaştıklarında değişkenlerin aralarındaki ilişki de azalır.

$$R^2 = - \sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{(\overline{y} - y_i)^2} \tag{4.1}$$

3.1.1.2. Ortalama Mutlak Hata (MAE)

Ortalama Mutlak Hata (MAE), gerçek veriler ile tahmin edilen veriler arasındaki mutlak ortalama mesafeyi ölçer. MAE, gerçek değer ile tahmin edilen değer arasındaki ortalama dikey ve yatay mesafedir. MAE, regresyon ve zaman serisi problemlerinde değeri kolay yorumlanabilir olduğu için sıkça kullanılmaktadır. MAE, tüm tekil hataların ortalamada eşit olarak ağırlıklandırıldığı, yönlerini dikkate almadan bir dizi tahmindeki hataların ortalama büyüklüğünü ölçen doğrusal bir skordur. MAE değeri 0'dan ∞'a kadar değişebilir. Daha düşük değerlere sahip yani negatif yönelimli tahminleyiciler daha iyi performans gösterir.

$$MAE = \frac{1}{n} \sum_{i=0}^{n-1} |y_i - \hat{y}_i| \tag{4.2}$$

3.1.1.3. Ortalama Kare Hata (MSE)

Ortalama Kare Hata (MSE), gerçek veriler ile tahmin edilen veriler arasındaki kare ortalama mesafeyi ölçer. Burada büyük hatalar MSE'e göre daha iyi not edilir. MSE, bir regresyon eğrisinin bir dizi noktaya yakınlığını belirler. MSE, her zaman pozitif değerlidir, tahminleyicinin ve makine öğrenmesi algoritmasının performansını ölçer. MSE, tahminleyicilerin performansı açısından değerlendirildiğinde MSE değeri sıfıra yakın olanların daha başarılı olduğu söylenebilir.

$$MSE = \frac{1}{n} \sum_{i=0}^{n-1} (y_i - \hat{y}_i)^2 \tag{4.3}$$

3.1.1.4. Ortalama Mutlak Yüzde Hata (MAPE)

Ortalama mutlak yüzde hata (MAPE), regresyon ve zaman serileri modellerinde tahminlerin doğruluğunu ölçmek için sıkça kullanılmaktadır. Gerçek değerler arasında sıfır içerenler varsa, sıfır ile bölünme olacağı için MAPE hesaplanamaz. Çok düşük tahmin değerleri için yüzde hatası %100'ü geçemez, ancak çok yüksek tahmin değerleri olduğunda yüzde hatasının üst sınırı yoktur. MAPE, tahminleyicilerin doğruluğunu karşılaştırmak için kullanıldığında, tahminleri çok düşük olan bir yöntemi sistematik olarak seçmesi nedeniyle önyargılıdır. Bu ufak ama ciddi problemin, tahmin değerlerin gerçek değerlerine oranını bulan bir doğruluk ölçütü ile üstesinden gelinebilir. Bu yaklaşım geometrik ortalama açısından yorumlanabilen tahminlere yol açar.

$$MAPE = \frac{100}{n} \sum_{i=0}^{n} \frac{|y_i - \hat{y}_i|}{|\hat{y}_i|} \tag{4.4}$$

3.1.1.5. Doğruluk

Doğruluk, genel olarak modelin tüm sınıflarda nasıl performans gösterdiğini tanımlayan bir ölçümdür. Tüm sınıflar eşit öneme sahip olduğunda yararlıdır. Doğru tahmin sayısının toplam tahmin sayısına oranı olarak hesaplanır.

$$Doğruluk = \frac{Gerçek_{pozitif} + Gerçek_{negatif}}{Gerçek_{pozitif} + Gerçek_{negatif} + Yanlış_{pozitif} + Yanlış_{negatif}} \tag{4.5}$$

Formüldeki ifadeler aşağıda açıklanmıştır.

Gerçek Pozitif: Model pozitif örneği kaç kez pozitif olarak sınıflandırdığını gösterir.

Yanlış Negatif: Model, pozitif bir örneği kaç kez negatif olarak sınıflandırdığını gösterir.

Yanlış Pozitif: Model, negatif bir örneği kaç kez pozitif olarak sınıflandırdığını gösterir.

Gerçek Negatif: Model, negatif bir örneği kaç kez negatif olarak sınıflandırdığını gösterir.

3.2 Makine Öğrenmesi Yöntemleri ile Sonuçların Değerlendirilmesi

Çalışmanın bu bölümünde Doğrusal Regresyon, Rasgele Orman Regresyon, K-En Yakın Komşu Regresyon, Karar Ağaçları Regresyon, Doğrusal Destek Vektör Makineleri Regresyon, Ekstra Ağaç Regresyon, Bayesian Ridge Regresyon ve Yapay Sinir Ağları algoritmalarının veri kümesine uygulanarak analiz edilmesi ve elde edilen sonuçlara yer verilmiştir. Çalışmalar Google Colab yazılımı ile yürütülmüştür. Modelin eğitilmesi için Python programlama dilinin Scikit-learn kütüphanesi, keras kütüphanesi ve tensorflow kütüphanesi kullanılmıştır. Çalışmada kullanılan yöntemlerin parametrelerinin ayarlanması bilgisayar simülasyonları ile sağlanmıştır. Her makine öğrenimi tekniği için bir makine öğrenimi prosedürü uygulanmıştır. Modellerin doğru eğitimi için, veri kümesi eğitim olarak %75 ve test olarak %25 olarak bölünmüştür. Bu nedenle, veri kümesinin özelliklerini öğrenmek için toplam 11934 örnekten yaklaşık 9000 örnek rastgele belirlenmiştir ve modelleri eğitmek için kullanılmıştır. Modelleri test etmek için yaklaşık 3000 örnek kullanılmıştır.

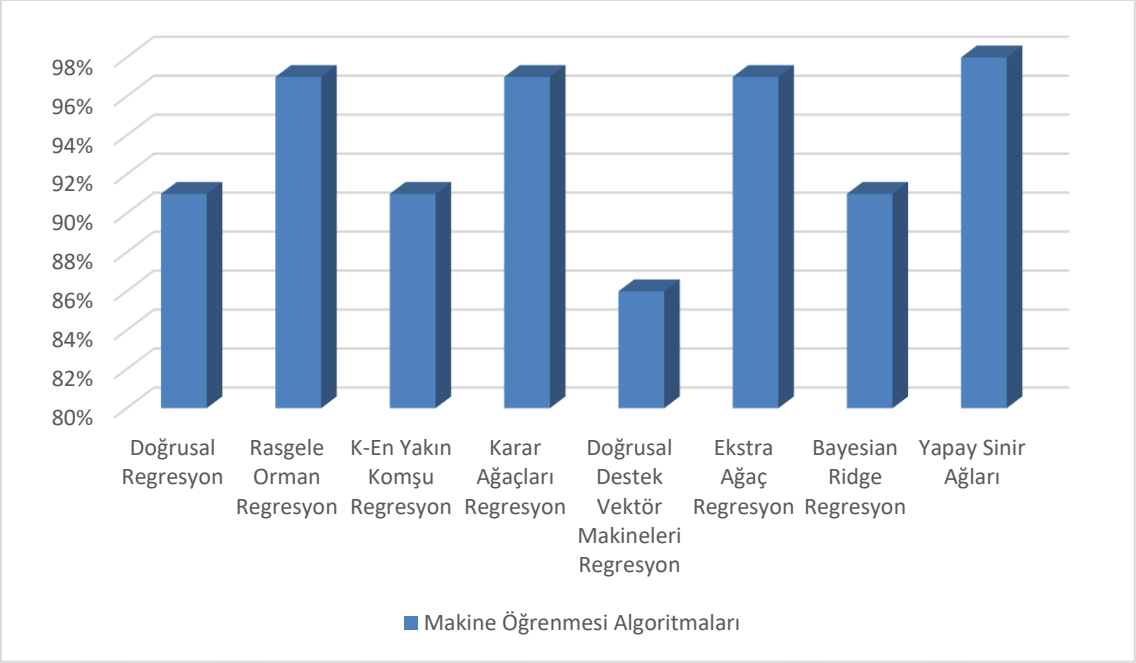
Modellerin eğitim sırasında kullanılan parametre değerleri Tablo 3.1'de verilmiştir.

Tablo 3.1. Regresyon ve YSA Modellerinin Eğitim Parametreleri

Model	Parametre 1	Parametre 2	Parametre 3
Doğrusal Regresyon	**-**	-	-
Karar Ağacı Regresyonu	**Random state=0**	Random state=20	Random state=30
K-NN Regresyon	**N neighbors=3**	N neighbors=10	N neighbors=20
Rastgele Orman Regresyon	**N estimators =1**	N estimators=2	N estimators=3
Bayesian Ridge Regresyon	**Compute score=True**	Compute score=False	-
Ekstra Ağaç Regresyon	**Random state=0**	Random state=20	Random state=40
Doğrusal DVM Regresyon	**Random state=10**	Random state=20	Random state=30'
YSA Modeli	**Input dim=16, activation='relu'**	Input dim=16, activation='sigmoid'	Input dim=16, activation='swish'

Tablo 3.1'de koyu olarak verilen parametre 1 değerleri çalışmada referans alınan değerlerdir. Parametre 1, parametre 2 ve parametre 3 değerleri ile modeller eğitilmiş optimum değer olarak parametre 1 seçilmiştir. Makine öğrenmesi algoritmalarından 8

tanesi karşılaştırılmıştır ve en iyi başarı oranı %98'lik bir doğruluk ile YSA modelinde elde edilmiştir. Tüm algoritmaların korelasyon matrisine göre karşılaştırılması Şekil 3.1'de verilmiştir. Grafikte yatay x ekseni, çalışmada kullanılan makine öğrenmesi algoritmalarını göstermektedir. Dikey y ekseni ise, bu algoritmaların doğruluk oranını göstermektedir.



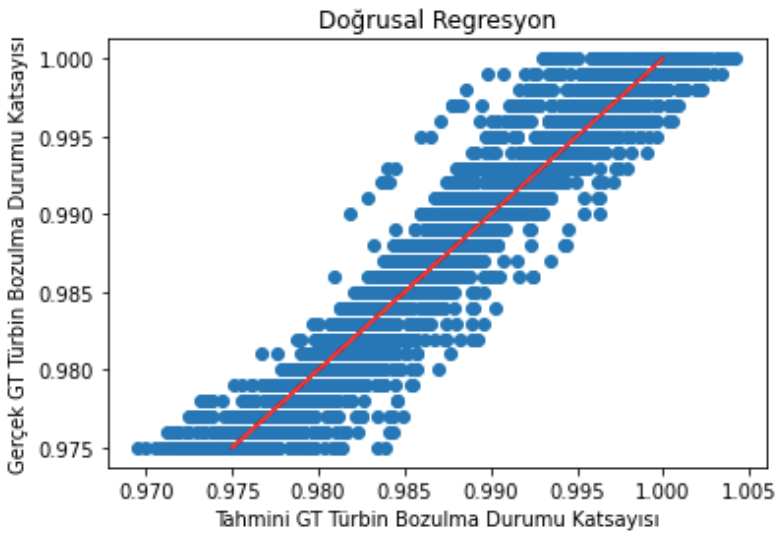
Şekil 3.1. Makine Öğrenmesi Algoritmaları

Doğrusal regresyon algoritmasını yönteminin formül ile belirtilmiş hali aşağıda gösterilmiştir.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \ldots + \beta_p X_{ip} + \varepsilon \tag{3.1}$$

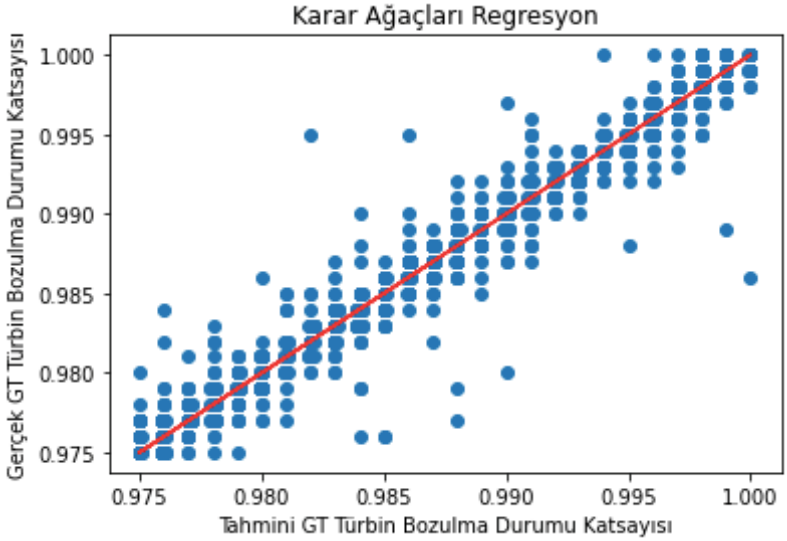
Bu formülde β_0 kesim noktası, β_1 doğrunun eğimi, ε hata terimidir.

Model eğitilirken doğrusal regresyon algoritma parametresi olarak farklı bir parametre uygulanmamıştır. Algoritmanın başarı oranı %91 olarak bulunmuştur. Bu modelin doğruluk grafiği Şekil 3.2'de verilmiştir. Modelin gerçek ve tahmin edilen arıza analizleri grafikte görselleştirilmiştir. Grafikte y ekseninde gösterilen noktalar, gerçek arıza analizi değerlerini gösterir ve x ekseninde gösterilen noktalar, tahmin edilen arıza analizi değerlerini gösterir. Kırmızı ile gösterilen çizgi düz bir çizgidir. Bu satırdaki gerçek değerler ve tahmin edilen değerler birbirine yakınsa, bu durum doğruluğun yüksek olduğunu gösterir.



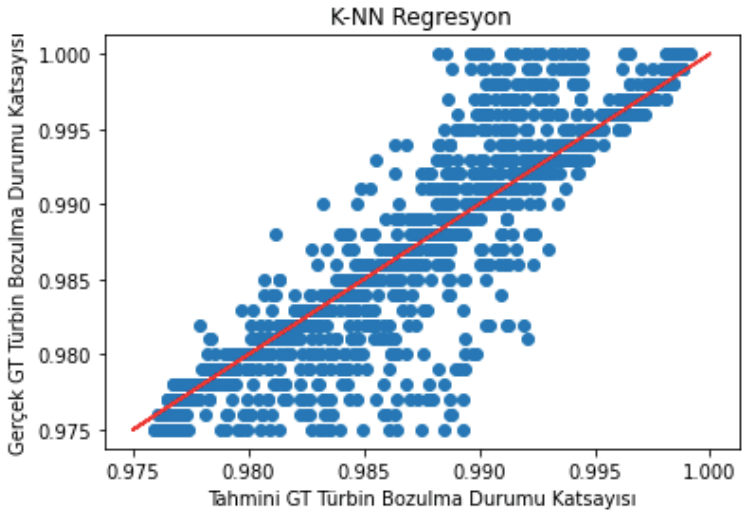
Şekil 3.2. Doğrusal Regresyon Analizine Göre Gerçek ve Tahmin Edilen Değerler

Karar Ağacı Regresyon algoritması kullanılan veri kümesinde en yüksek doğruluk değeri bulduğumuz algoritmalardan biridir. Karar ağacı algoritmasında random state değeri 0, 20 ve 30 olarak seçilmiştir. Sırayla, bu değerler için deneme yapılmıştır. Bu değerler arasından en iyi başarı oranını 0 değerinin verdiği gözlemlenmiştir. Çalışmada Karar ağaçları algoritması için Max_depth değeri kullanılmamıştır. Max_depth değeri karar ağacının azami büyüme uzunluğudur. Parametre olarak tanımlanmadığı durumda karar ağacı tüm düğümler sonuca ulaşana kadar büyümeye devam eder. Random_state değeri ağacın son bulma noktasıdır. Random state değeri modelin hızını arttırır ve aynı eğitim verisi ve parametreleri olduğunda her zaman aynı veriyi vermesine olanak sağlar. Karar ağacının random state 0 parametresi ile bulduğu doğruluk oranı %97'dir. Karar ağacı algoritmasının doğruluk grafiği Şekil 3.3'te verilmiştir. Grafikte gerçek değerler ile tahmin edilen değerler arasındaki mesafenin ne kadar az olduğu yani doğruluk oranının yüksek olduğu görülmektedir.



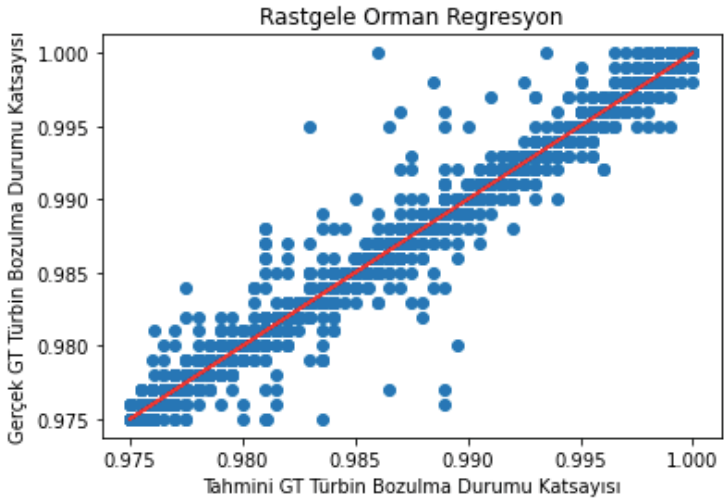
Şekil 3.3. Karar Ağacı Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

K-NN algoritmasında n_neigbours değeri komşu sayısını belirler ve n_neigbours değeri kadar komşu ile işlem yapıldığını ifade eder. K-NN algoritmasında n_neigbours değeri 3, 10 ve 20 olarak seçilmiştir. Sırayla, bu değerler için deneme yapılmıştır. Bu değerler arasından veri kümesine en uygun n_neighbours değerinin 3 olduğu tespit edilmiştir. Çalışmada K-NN algoritmasının n_jobs parametresi kullanılmamıştır. Bu değer algoritmada paralel olarak çalıştırılacak olan görev sayısını belirler. K-NN algoritmasının n_neigbours 3 parametresi ile bulduğu doğruluk oranı %91'dir. K-NN algoritmasının doğruluk grafiği Şekil 3.4'te verilmiştir. Grafikte bu algoritmada gerçek değerler ile tahmin edilen değerler arasındaki mesafenin büyüdüğünü etrafa saçılan değerler olduğu görülmektedir. Bu saçılmalar ve mesafe algoritmanın doğruluk oranını düşürmüştür.



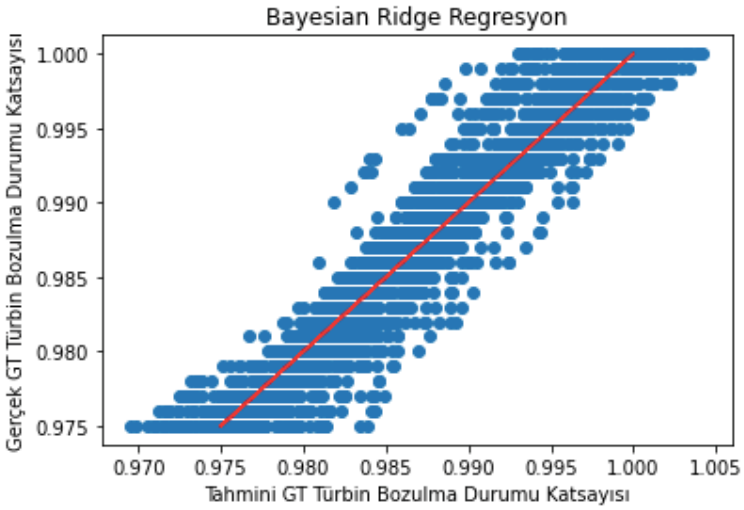
Şekil 3.4. K-NN Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

Rastgele ormanlar algoritması model üzerinde aşırı öğrenmeyi engellemek için birden fazla karar aracı oluşturarak bu ağaçların ortalamasını alır. Algoritmada n_estimators değeri temel veri kümesinden ne kadar alt veri kümesi oluşturulup, rastgele ormanlar algoritmasına verileceğini belirler. Rastgele ormanlar algoritmasında n_estimators değeri 1, 2 ve 3 olarak seçilmiş ve sırasıyla bu değerler için deneme yapılmıştır. Bu değerler arasından veri kümesine en uygun n_estimators değerinin 1 olduğu tespit edilmiştir. Çalışmada algoritmanın max_features parametresi kullanılmamıştır. Bu değer algoritmanın tek bir ağaç ile çalışmasını sağlar. Rastgele ormanlar algoritmasının n_estimators 1 parametresi ile bulduğu doğruluk oranı %97'dir. Algoritmanın doğruluk grafiği Şekil 3.5'te verilmiştir. Grafikte bu algoritmada gerçek değerler ile tahmin edilen değerler arasındaki mesafenin ne kadar az ve uyum içinde olduğu görülmektedir.



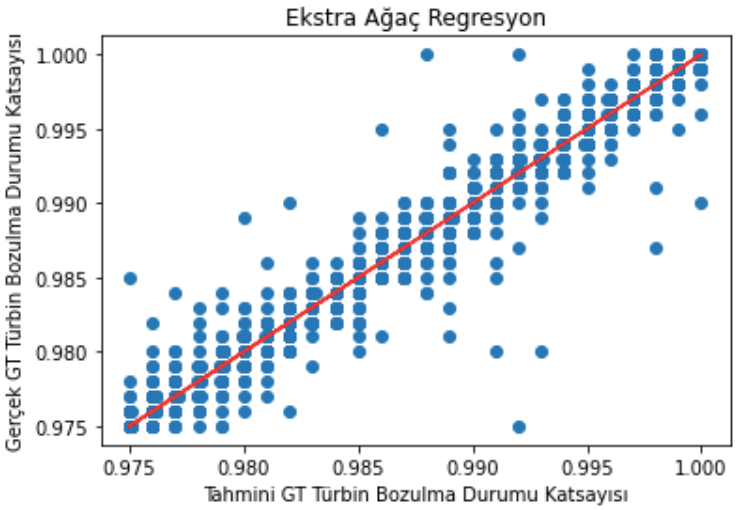
Şekil 3.5. Rastgele Orman Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

Bayesian Ridge algoritması, doğrusal regresyon algoritmasının farklılaşmış bir versiyonudur. Algoritmada compute_score değeri "True" ve "False" seçilmiştir. Bu değer model ağırlıklarının azalmasını sağlar. Sırayla, bu değerler için deneme yapılmıştır. Bu değerler arasından veri kümesine en uygun compute_score değeri "True" olduğu tespit edilmiştir. Bayesian Ridge algoritmasının bu parametre ile bulduğu doğruluk oranı %91'dir. Bayesian Ridge algoritması ile aynı doğruluk oranını vermiştir. Algoritmanın doğruluk grafiği Şekil 3.6'da verilmiştir. Grafikte gerçek değerler ile tahmin edilen değerler arasındaki mesafenin uzak olduğu görülmektedir.



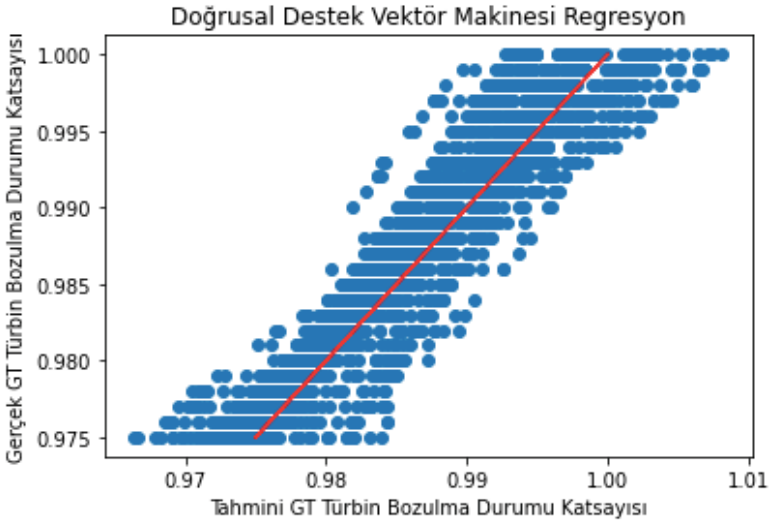
Şekil 3.6. Bayesian Ridge Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

Ekstra Ağaç algoritması, rastgele orman algoritmasının farklı bir çeşididir. Bu iki algoritma arasında modellerin eğitilmesi aynıdır fakat düğümlerin dallara ayrılması farklılık gösterir. Ekstra Ağaç algoritması karar kriteri olarak rastgele dallanma yapar. Algoritmada random_state değeri 0, 20 ve 40 seçilmiş ve sırasıyla bu değerler için deneme yapılmıştır. Bu değerler arasından veri kümesine en uygun random_state değeri 0 olduğu tespit edilmiştir. Ekstra Ağaç algoritmasının bu parametre ile bulduğu doğruluk oranı %97'dir. Algoritmanın doğruluk grafiği Şekil 3.7'de verilmiştir. Grafikte bu algoritmanın gerçek değerler ile tahmin edilen değerler arasındaki mesafesi görülmektedir.



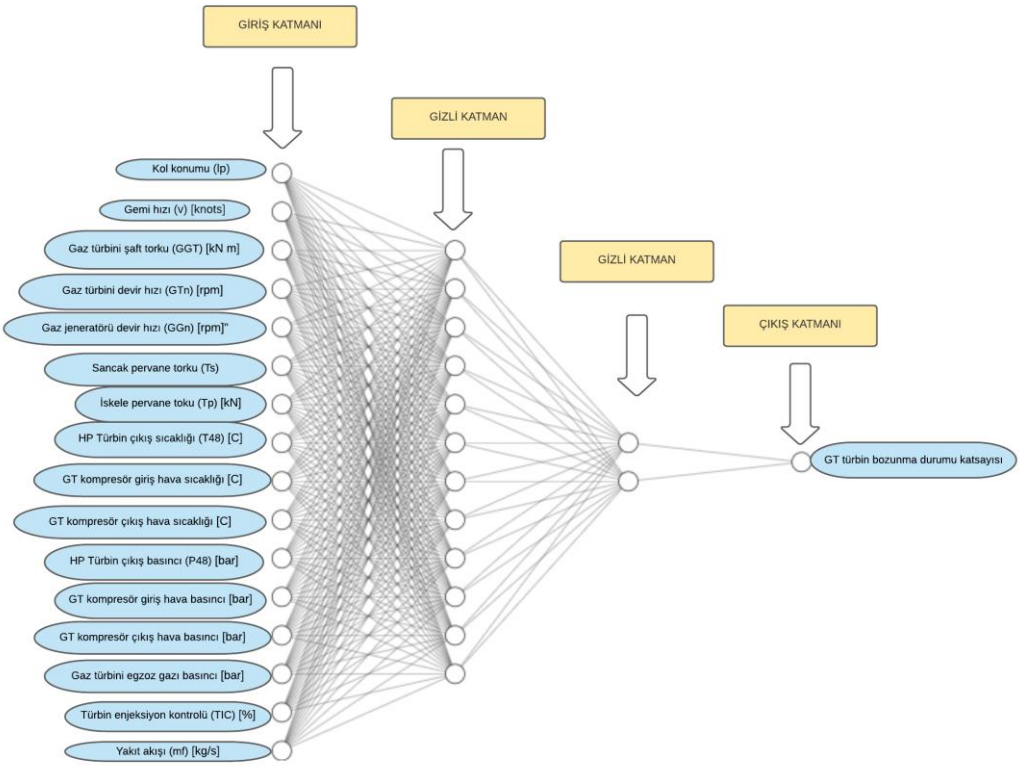
Şekil 3.7. Ekstra Ağaç Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

Doğrusal destek vektör makinesi algoritmasında model üzerinde random_state değeri 10, 20 ve 30 seçilmiştir. Sırayla, bu değerler için deneme yapılmıştır. Bu değerler arasından veri kümesine en uygun random_state değeri 10 olduğu tespit edilmiştir. Doğrusal destek vektör makinesi algoritmasının bu parametre ile bulduğu doğruluk oranı %86'dır. Algoritmanın doğruluk grafiği Şekil 3.8'de verilmiştir. Grafikte bu algoritmanın gerçek değerler ile tahmin edilen değerler arasındaki mesafesi görülmektedir.



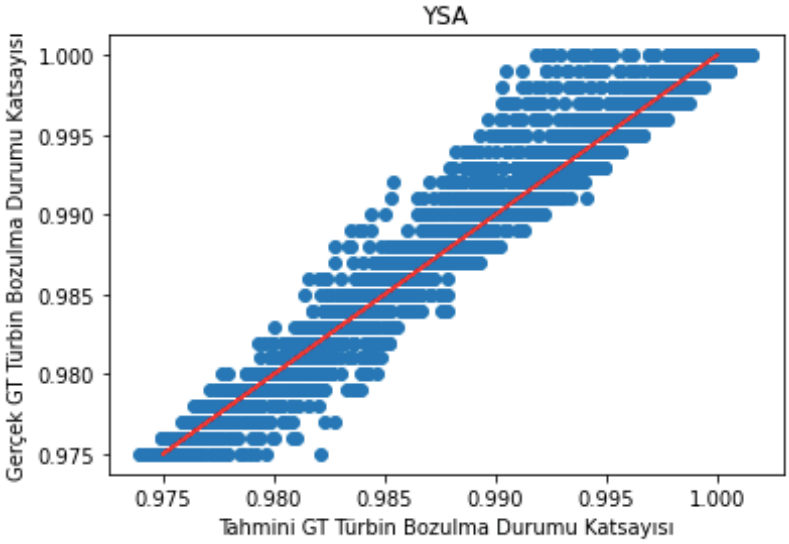
Şekil 3.8. Doğrusal DVM Regresyonuna Göre Gerçek ve Tahmin Edilen Değerler

YSA algoritmasında bir giriş katmanı, iki gizli katmanı ve bir çıkış katmanı kullanılmıştır. İki gizli katman kullanılmasının sebebi gizli katman sayısını değiştirdiğimizde elde edilen sonucun mevcuttaki sonuçtan daha kötü olmasıdır. Giriş katmanında ve gizli katmanlarında kullanılan nöron sayısı 32, aktivasyon fonksiyonu "ReLu"dur. Katmanlardaki nöron sayıları değiştirilerek performans ölçümleri yapılmıştır. Nöron sayıları arttırıldığında doğruluk sonuçları azalmıştır. Oluşturulan ağ Şekil 3.9'da gösterilmiştir.



Şekil 3.9. YSA Modeline İlişkin Sonuçlar

Bu denemelere ek olarak YSA algoritmasında model aktivasyon fonksiyonları “sigmoid” ve “swish” ile deneme yapılmıştır. Bu fonksiyonlar arasında veri kümesine en uygun aktivasyon fonksiyonu “ReLu” seçilmiştir. YSA algoritmasının bu parametre ile bulduğu doğruluk oranı %98’dir. Algoritmanın doğruluk grafiği Şekil 3.10’da verilmiştir. Grafikte bu algoritmanın gerçek değerler ile tahmin edilen değerler arasındaki mesafesi görülmektedir.



Şekil 3.10. YSA Modeline Göre Gerçek ve Tahmin Edilen Değerler

Yapılan çalışmaya ait tahmin edilen değerlerin karşılaştırılması Tablo 3.2'de verilmiştir. Bu karşılaştırma, R^2, MAE, MAPE ve MSE performanslarına bakılarak yapılmıştır. Tablo 3.2'de R^2 değeri 1'e yakın olan modellerin veri kümesini iyi tahmin edebildiği anlaşılmıştır. Çalışmada birden fazla yöntemle hata analizi tahmin edilerek modeller arasında en iyi performans gösteren yöntemin belirlenmesi amaçlanmıştır. Bu bağlamda klasik yöntem olarak Doğrusal Regresyon, Karar Ağacı Regresyon, K-NN Regresyon, Rastgele Orman Regresyon, Bayesian Ridge Regresyon, Ekstra Ağaç Regresyon, Doğrusal DVM Regresyon ve farklı bir model olarak YSA kullanılmıştır. R^2 değerleri incelendiğinde, tüm algoritmaların doğruluk değerleri %90'ın üzerinde tahmin edilmiştir. Algoritmaları birbirleriyle karşılaştırdığımızda YSA modelinin diğer algoritmalara göre daha iyi değer verdiği görülmüştür.

Tablo 3.2. Regresyon ve YSA Modellerine İlişkin MAE, MAPE, MSE ve R^2 Değerlerinin Karşılaştırması

Model	MAE	MSE	R^2	MAPE
Doğrusal Regresyon	0.0017	5.09	0.91	0.16
Karar Ağacı Regresyon	0.00052	1.075	0.97	0.053
K-NN Regresyon	0.0010	5.09	0.91	0.011
Rastgele Orman Regresyon	0.00038	5.27	0.97	0.038
Bayesian Ridge Regresyon	0.0016	5.09	0.91	0.16
Ekstra Ağaç Regresyon	0.00058	1.67	0.97	0.05
Doğrusal DVM Regresyon	0.0021	7.54	0.86	0.21
YSA Model	0.0008	9.59	0.98	0.87

Tablo 3.2'ye göre YSA %98 ile en yüksek tahmin oranına sahip olmuştur. Doğrusal DVM Regresyon %86 korelasyon katsayısı tahmini ile en düşük tahmin oranına sahip olmuştur. Çalışmada MAE değerleri incelendiğinde, analiz yapılan bütün algoritmalar için gerçek değer ile tahmin edilen değer arasındaki farkın çok az olduğu görülmektedir.

Bu tez çalışmasında YSA modeli ayrıntılı bir şekilde incelenmiş, uygulama alanları araştırılmış ve sık kullanılan bazı YSA algoritmaları veriler üzerinde test edilmiştir. Deneysel çalışmalarda ise en uygun olan YSA modeli kullanılarak tahmine yönelik bir model eğitilmiştir.

Sonuçlar karşılaştırıldığında YSA modelinin basitliği, modeli gerçekleştirme hızı ve yüksek doğruluk oranları ile diğer regresyon modellerine göre daha başarılı olduğu görülmüştür. Ayrıca YSA modeli, istatistik formüllerini kullanan regresyon modellerinden farklı olarak, veri kümesini kullanarak matematiksel hesaplamalar ile model geliştirmeye ihtiyaç duymaz. Veri kümesi içindeki ilişkileri ve bağlantıları eğitilerek öğrenir. Bir kez öğrendikten sonra model veri kümesindeki öğrenilmek istenilen bilgiyi hemen verir. Veri kümesi içindeki bilinmeyen ilişkisel bağlantıları açığa çıkarır. Bu sebeple YSA modeli açık, anlaşılabilir ve esnektir. Regresyon modelleri oluşacak hataya karşı çok müsaittir. Modelde oluşacak en küçük bir hata sonuçlarda büyük hatalara sebep olabilir. Fakat yapay sinir ağlarında nöronlardan birinin zarar

görmesi sonucu regresyon modellerindeki gibi çok etkilemez. Özellikle arıza teşhisinde makine arızasına uygun olarak eğitilmiş bir YSA modeli çok daha güzel sonuçlar verecektir.

3.3. ESA Yöntemi ile Sonucun Değerlendirilmesi

Bu çalışmada bir havuz pompa makinesinin çıkardığı normal ve anormal akustik işaretler üzerinden makinenin arızasız / arızalı çalışma durumu sınıflandırılmıştır. Bu tür bir sınıflandırma yapmak için literatürde son yıllarda yoğun olarak tercih edilen ESA algoritması kullanılmıştır. Bu şekilde makinenin kestirimci bakımı gerçekleştirilerek, motor durma noktasına gelmeden arızanın anlaşılması sağlanmaya çalışılacaktır. Yapılacak çalışmanın üreteceği çıktılar gemi makinelerinin arıza tahminlerinin yapılmasına olanak sağlayacağı düşünülmektedir.

Bu aşamada MIMII veri kümesi üzerinden alınan akustik sinyallere ait veriler Google Colab üzerinden Python yazılımı kullanılarak işlenmiştir. Burada sinyal öncelikle FFT algoritması kullanılarak frekans domenine dönüştürülmüştür. Fourier dönüşümü spektral analizde kullanılan en önemli dönüşümlerden birisidir. Bu dönüşüm sayesinde sinyallerin spektrumları analiz edilebilmekte, frekans domeninde sistemlerin ve özelliklerinin betimlenmesi sağlanabilmektedir (Jones vd., 1987).

Fourier Dönüşümü ile ses sinyalleri, sinüs ve kosinüs dalgalarının doğrusal birleşimine ayrışırlar. Fourier dönüşümü FD ile ifade edilir. Herhangi bir f (t) fonksiyonunun Fourier Dönüşümü (FD),

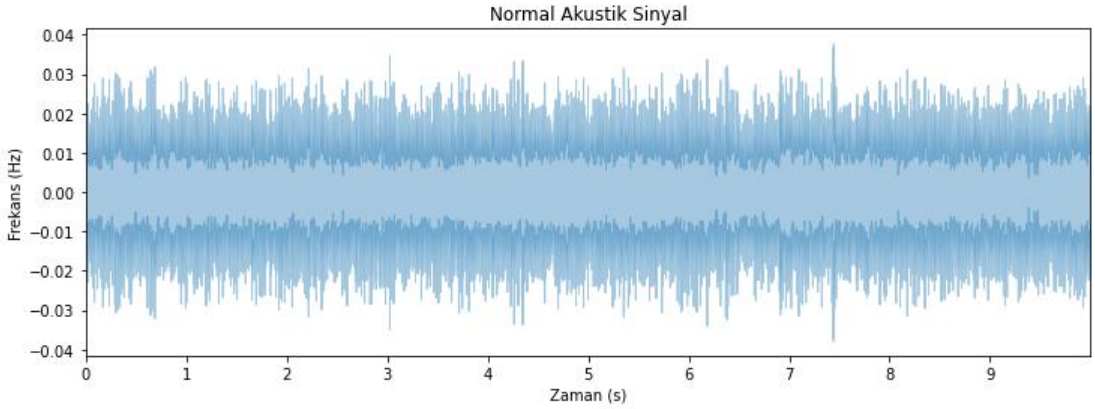
$$F\,(f) \;=\; \int_{-\infty}^{+\infty} x\,(t)\, e^{-i2\pi f t} dt$$

ile ifade edilir. (2.7)

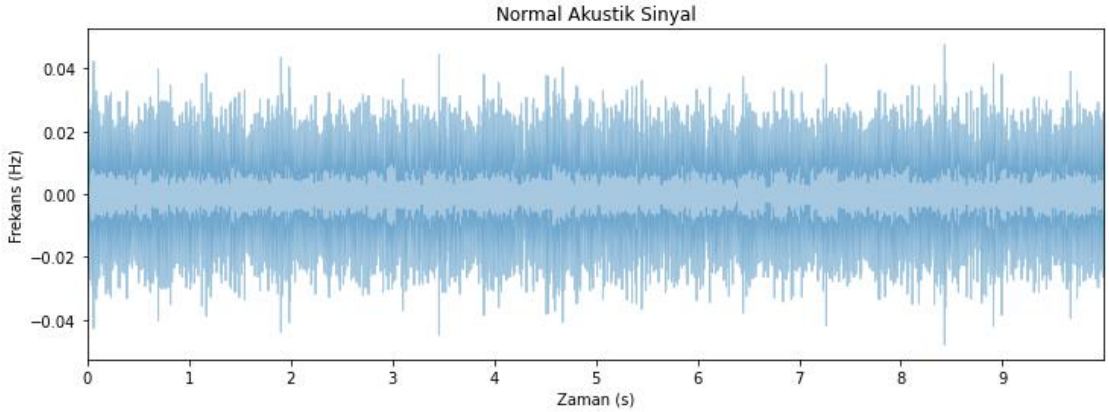
Burada;

- *F(f)*, frekans alanında Fourier Dönüşümünün çıktısıdır.
- *x(t)*, zaman alanı fonksiyonudur.
- *2πf*, saniyede radyan cinsinden frekanstır.

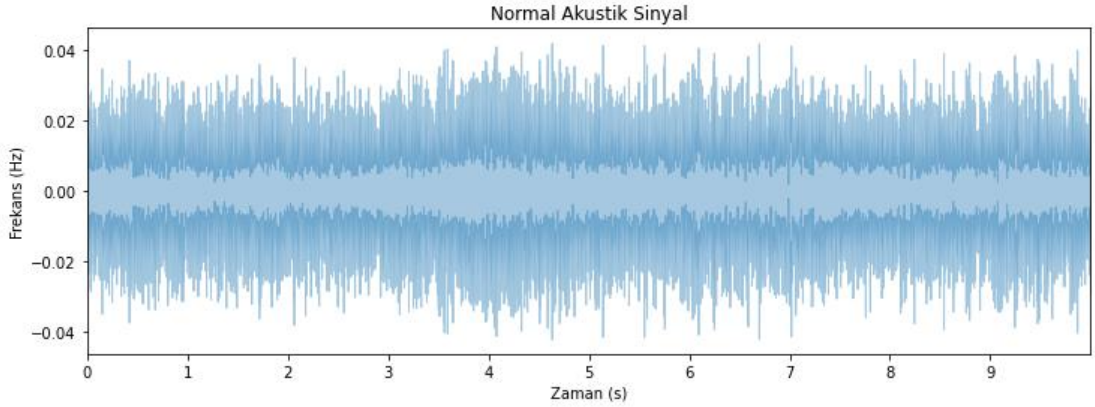
Fourier dönüşümü sonrası normal ve anormal akustik sinyallerin zaman-frekans domenine dönüşmüş örnek çıktıları Şekil 3.11'de ve Şekil 3.12'de verilmiştir.



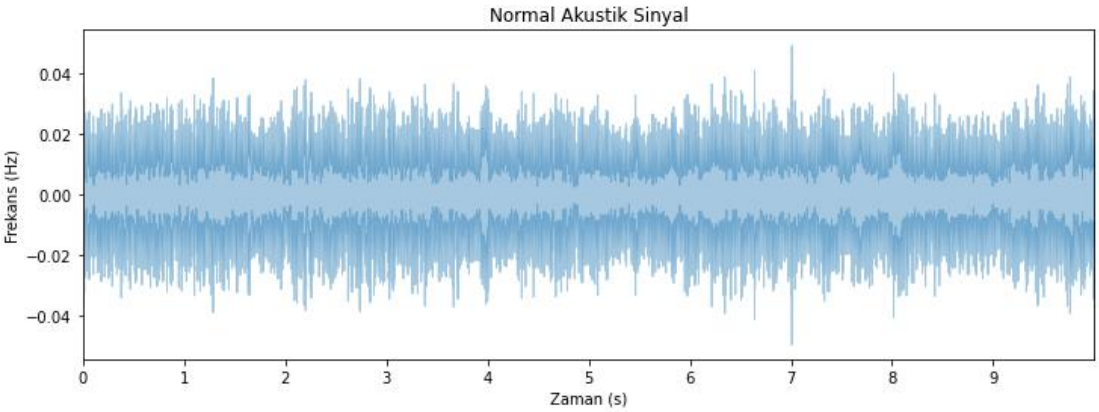
(a) Normal Akustik Sinyal



(b) Normal Akustik Sinyal

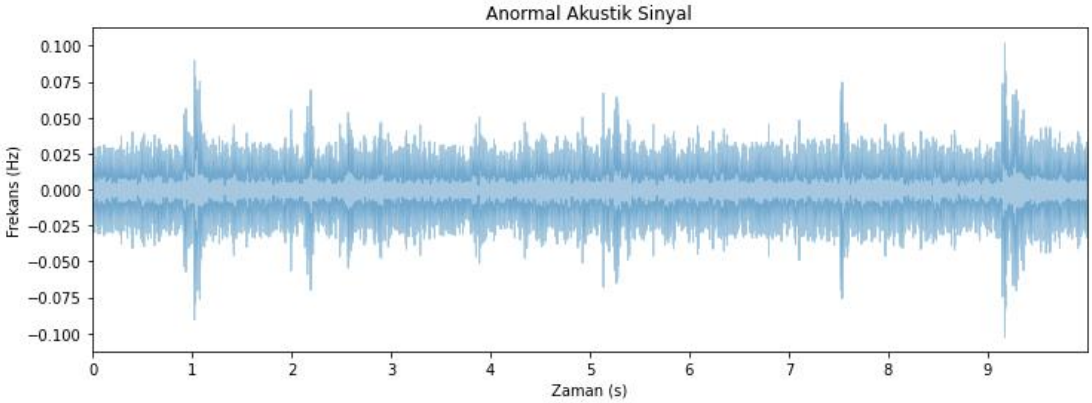


(c) Normal Akustik Sinyal

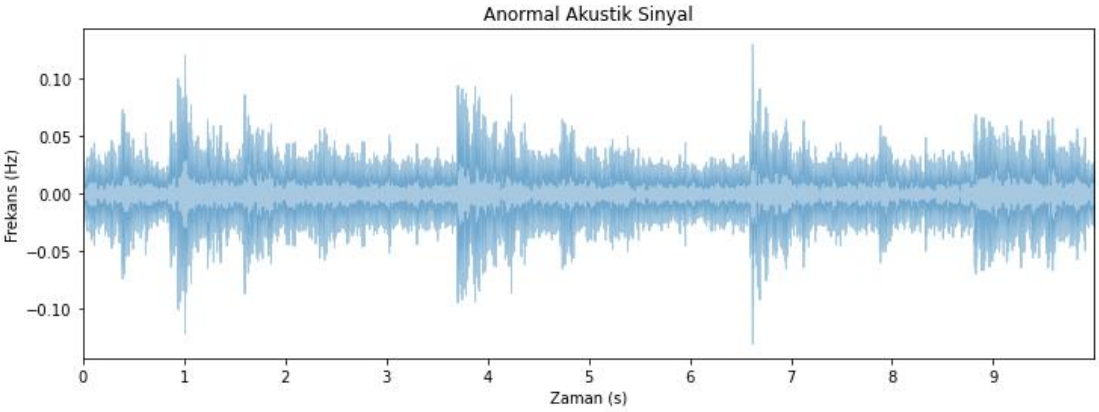


(d) Normal Akustik Sinyal

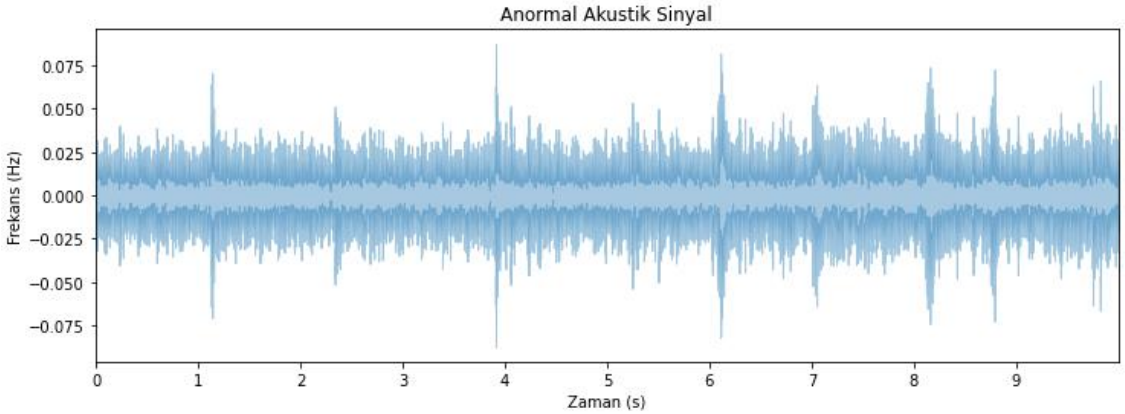
Şekil 3.11. Normal Akustik Sinyaller



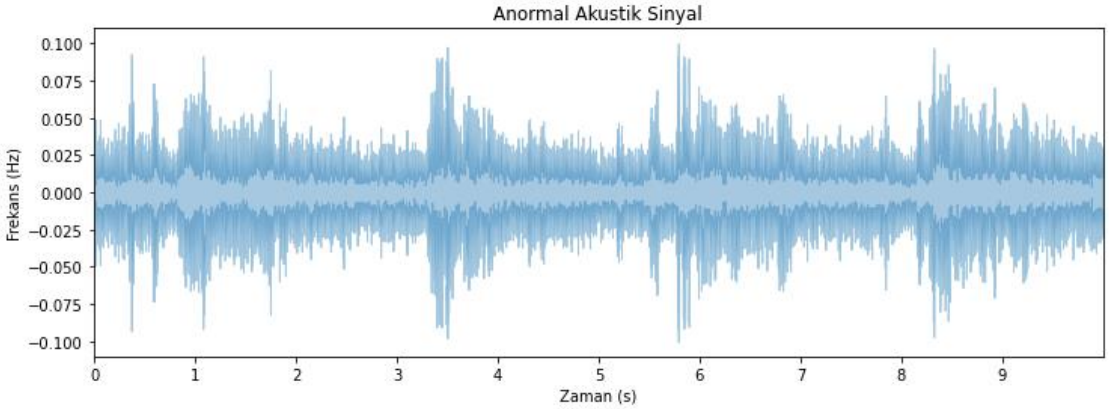
(a) Anormal Akustik Sinyal



(b) Anormal Akustik Sinyal



(c) Anormal Akustik Sinyal



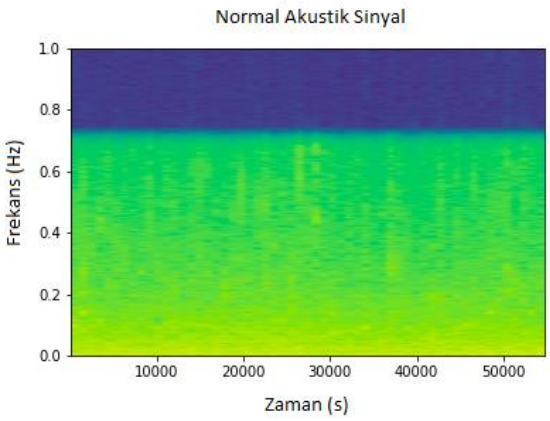
(d) Anormal Akustik Sinyal

Şekil 3.12. Anormal Akustik Sinyaller

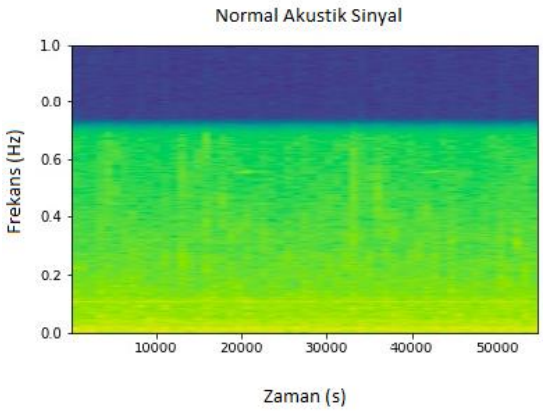
Elde edilen bu akustik sinyaller ayrıca zaman-frekans domenine dönüştürülmüş ve benzer spektrogram görüntüsü elde edilmiştir. Spektrogram, Fourier Dönüşümünün görselleştirilmiş halidir. Spektrogram grafiğinde giriş sinyalinin parçaları alınır ve her bir parçaya bir Fourier Dönüşümü uygulanır. Her parçanın belirli bir genişliği vardır ve bu parçaya bir Fourier Dönüşümü uygulanır. Her parça bir frekans dağılımına sahiptir. Zaman sinyalinde her parça için bir dizi frekans bileşeni elde edilir. Her bir parçadaki tüm bu frekans bileşenlerinin toplanması ve birlikte çizilmesi spektrogram grafiğini oluşturur.

Spektrogram grafiği, yatay eksenin sinyalin zamanını ve dikey eksenin frekans eksenini temsil ettiği bir 2D görsel ısı haritasıdır. Grafik üzerinde frekans bileşeninin büyüklüğü ne kadar yüksekse, renk o kadar açık olur. Spektrogram grafiğinin elde edilmesi için

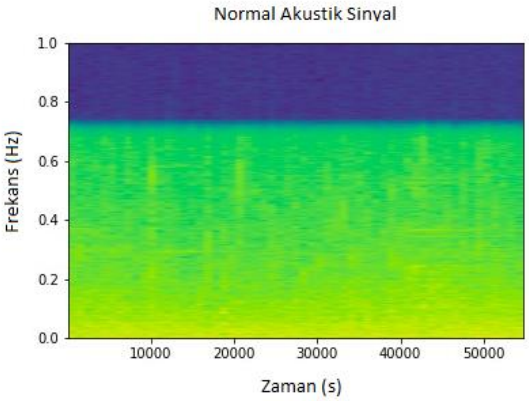
kullanılan parametreler NFFT, Fs, Noverlap, Fc, Scale (dB), Mod ve Side'dır. Burada NFFT, her bir parçanın FFT'sidir. N sinyalin uzunluğudur, bölünen blok uzunluğudur. Fs, sinyalin örnekleme frekansıdır. Noverlap, bloklar arasındaki örtüşme noktalarının sayısını tanımlar. Fc, y'nin merkez frekansıdır. Bir sinyal elde edildiğinde kullanılan frekans aralığını yansıtmak için grafiğin x uzantılarını kaydırır, ardından filtrelenir ve temel banda alt örneklenir. Scale (dB), değerlerin ölçeklendirilmesidir. 'dB', değerleri dB ölçeğine döndürür. Mode, ne tür bir spektrum kullanılacağını gösterir. Side, spektrumun hangi taraflarının döneceğini gösterir. Spektrogram grafiği örnek çıktıları aşağıda verilmiştir.



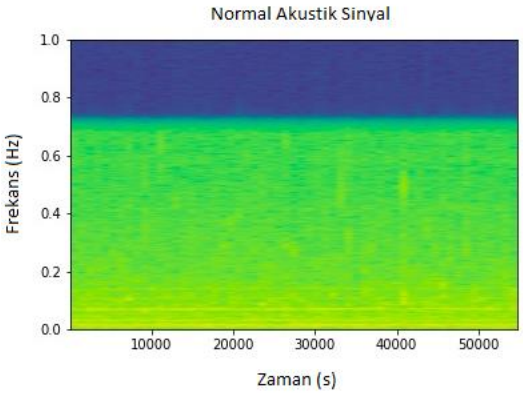
(a) Normal Ses Spektrogramı



(b) Normal Ses Spektrogramı

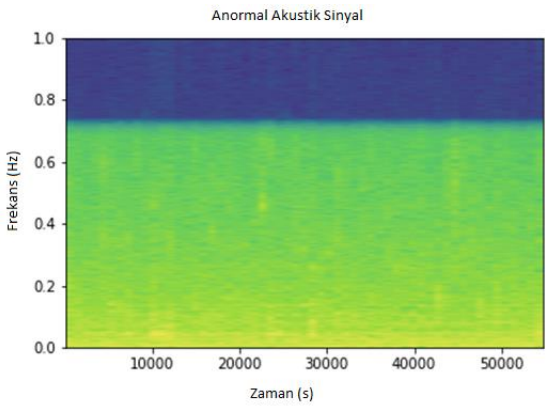


(c) Normal Ses Spektrogramı

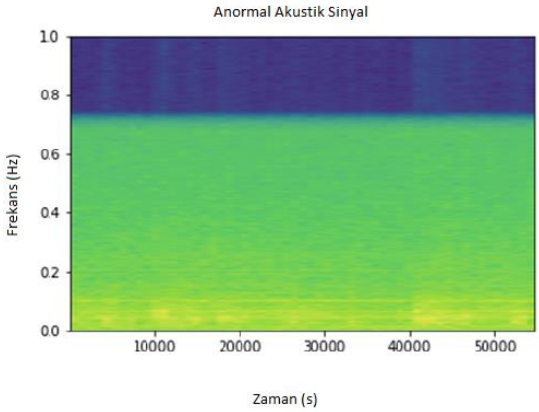


(d) Normal Ses Spektrogramı

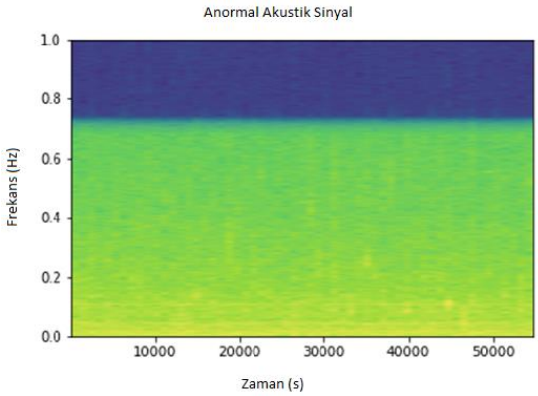
Şekil 3.13. Normal Ses Spektrogramı



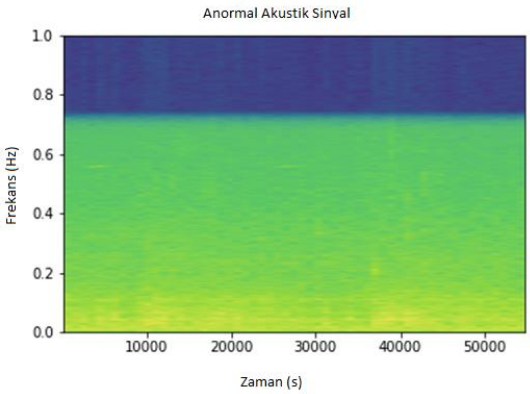
(a) Anormal Ses Spektrogramı



(b) Anormal Ses Spektrogramı



(c) Anormal Ses Spektrogramı



(d) Anormal Ses Spektrogramı

Şekil 3.14. Anormal Ses Spektrogramı

Akustik verilerden elde edilen spektrogram görüntülerin sınıflandırılması için Evrişimli Sinir Ağları (ESA) kullanılmıştır. ESA algoritması kullanılarak elde edilen spektrogram grafikleri üzerinden normal ve anormal akustik sinyallerin tespit edilmesi amaçlanmıştır. Spektrogram grafik görüntüleri eğitilmeden önce ön işleme yapılmıştır. Bu aşamada görüntü büyütme yapılmıştır. Model, çok resimle eğitilmek yerine daha az resimle ve modeli farklı açılarla çekilmiş resimleri değiştirerek eğitilmiştir.

0-255 arasındaki tüm piksel değerleri 0-1 arasına yeniden ölçeklendirilmiştir ve 0.2 yakınlaştırma uygulanmıştır. Büyütme işlemi için Keras'ın ImageDataGenerator sınıfı kullanılmıştır.

ImageDataGenerator sınıfı, büyük bir sayısal diziden ve görüntüleri içeren klasörlerden görüntüleri okumak için flow(), flow_from_directory() ve flow_from_dataframe() olmak üzere üç yönteme sahiptir. Bu çalışmada sadece flow_from_directory() kullanılmıştır. Modellerin doğru eğitimi için, veri kümesi eğitim olarak %80 ve test olarak %20 bölünmüştür. Gerçekleştirilen modelde sıralı katmanlardan oluşan sequential model kullanılmıştır. Modelde 1 girdi katmanı, 2 gizli katman, 1 flatten katman, 1 fully connected katman ve 1 çıkış katmanı bulunmaktadır. Ara katmanlar için "ReLu", çıkış katmanı için "softmax" aktivasyon fonksiyonu kullanılmıştır. Oluşturulan model Şekil 3.15'te verilmiştir.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_3 (Conv2D)            (None, 31, 31, 32)        896
_________________________________________________________________
average_pooling2d_3 (Average (None, 15, 15, 32)        0
_________________________________________________________________
activation_5 (Activation)    (None, 15, 15, 32)        0
_________________________________________________________________
conv2d_4 (Conv2D)            (None, 15, 15, 64)        18496
_________________________________________________________________
average_pooling2d_4 (Average (None, 7, 7, 64)          0
_________________________________________________________________
activation_6 (Activation)    (None, 7, 7, 64)          0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 7, 7, 64)          36928
_________________________________________________________________
average_pooling2d_5 (Average (None, 3, 3, 64)          0
_________________________________________________________________
activation_7 (Activation)    (None, 3, 3, 64)          0
_________________________________________________________________
flatten_1 (Flatten)          (None, 576)               0
_________________________________________________________________
dropout_2 (Dropout)          (None, 576)               0
_________________________________________________________________
dense_2 (Dense)              (None, 64)                36928
_________________________________________________________________
activation_8 (Activation)    (None, 64)                0
_________________________________________________________________
dropout_3 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_3 (Dense)              (None, 2)                 130
_________________________________________________________________
activation_9 (Activation)    (None, 2)                 0
=================================================================
```

Şekil 3.15. ESA Modeli

Oluşturulan modelde tahmin değeri doğruluk performansına bakılarak yapılmıştır. Şekil 3.16 incelendiğinde test verileri ile doğruluğun %92 olduğu görülmektedir. Test verileri için karşıtlık matrisi Şekil 3.17 ve Şekil 3.18'de gösterilmiştir.

```
model.evaluate_generator(test_set, steps=12)
```

```
[0.2681077718734741, 0.9244791865348816]
```

Şekil 3.16. ESA Modeli Doğruluk Değeri

Şekil 3.17 ve Şekil 3.18'de Filename sütunu gerçek değerleri, Predictions sütunu tahmin edilen değerleri gösterir. Burada örnek olarak sadece 20 satır gösterilmiştir. Şekil 3.17 incelendiğinde, bu modelde 10 anormal akustik sinyal görüntüden hepsi anormal olarak sınıflandırılmıştır. Benzer şekilde 3.18 incelendiğinde, 10 normal veriden 2 tane veri normal 8 tane veri anormal olarak sınıflandırılmıştır.

40 girişten 1-10 aralığındakiler Filtre

Filename	Predictions
abnormal/00000013.png	abnormal
abnormal/00000021.png	abnormal
abnormal/00000026.png	abnormal
abnormal/00000039.png	abnormal
abnormal/00000042.png	abnormal
abnormal/00000046.png	abnormal
abnormal/00000049.png	abnormal
abnormal/00000050.png	abnormal
abnormal/00000051.png	abnormal
abnormal/00000054.png	abnormal

Her sayfada 10 satır göster 1 2 3 4

Şekil 3.17. Anormal Verilere İlişkin Sinyal Görüntülerinin Sınıflandırılması

Şekil 3.18 incelendiğinde verilerin çoğu normal verilerdir ve anormal olarak tahmin edilmiştir. Bunlar modelde yanlış tahmin edilen değerlerdir.

Filename	Predictions
normal/00000627.png	abnormal
normal/00000632.png	abnormal
normal/00000633.png	normal
normal/00000638.png	abnormal
normal/00000640.png	abnormal
normal/00000643.png	abnormal
normal/00000647.png	abnormal
normal/00000649.png	abnormal
normal/00000652.png	normal
normal/00000655.png	abnormal

Her sayfada 10 satır göster 1 2 3 4

Şekil 3.18. Normal Verilere İlişkin Sinyal Görüntülerinin Sınıflandırılması

Bu çalışma ile ESA modelinin, arıza tahmini için uygun bir alternatif yöntem olduğu gösterilmiştir. Ortaya çıkan sonuç, akustik verilerin içsel özelliklerinin görüntü verilerinin özelliklerine benzer olduğunu göstermiştir. ESA modelinin, yüksek düzeyde ölçeklenebilir olduğu sonucu ortaya çıkmıştır. Genişletilmiş bir veri kümesi ile çok daha verimli sonuçlar alınabilir.

BÖLÜM 4 – SONUÇLAR

Bu çalışmada, gemi sistemlerinin arıza olasılıklarının tahmin edilmesine yönelik iki farklı analiz yapılmıştır. İlk analizde gemi makinelerinden elde edilen verilerle türbin bozulma durum katsayısını tahmin etmek için YSA'lar ve regresyon analizi yöntemleri kullanılmıştır. Gemi motor sisteminden toplanan veriler, regresyon ve YSA modellerinin girdilerini belirlemek için kullanılmış ve gemi makinelerinin arıza analizi için en uygun yöntem belirlenmiştir. İncelenen yöntemlerde kullanılan veri sayısı ve verilerin birbirleri ile olan ilişkisi, yöntemlerin sonuçlarının doğruluğunu önemli ölçüde etkilemiştir. YSA algoritması ile elde edilen sonuçlar, regresyon analizi ile elde edilen sonuçlara göre daha başarılı olmuştur. Bu sonuçlardan YSA modelinin birbiri ile ilişkili verilerle tahmin edilen modellerde regresyon modeline kıyasla daha iyi sonuçlar verdiği anlaşılmıştır. YSA'lar, diğer algoritmalara göre daha güvenilir sonuçlar verdikleri için tercih edilmektedir. Çalışma kapsamında YSA tasarımında gizli katman sayısı ve gizli katmandaki nöron sayısı gibi ağ parametreleri, sınırlı sayıda deney sonucunda belirlenmiştir. Ağ topolojisinde, gizli katmanların sayısının arttırılması, katmanlardaki nöronların sayısının değiştirilmesi ve ağda kullanılan transfer fonksiyonlarının değiştirilmesi gibi bazı değişikliklerle, YSA ile elde edilen tahminler daha da geliştirilmiştir. Sonuç olarak, YSA ile hazırlanan modelin kavramsal ve fiziksel tabanlı modellere kıyasla doğrusal olmayan gemi-makine arıza ilişkisini başarılı bir şekilde modelleyebildiği, yoğun veri ve kalibrasyon işlemlerine ihtiyaç duymadan hızlı bir modellemeye olanak sağladığı görülmüştür.

İkinci çalışmada ise bir motorunun çıkardığı akustik işaretlerin arıza durumu sınıflandırılmıştır. Çalışmada elde edilen görüntü verileri derin öğrenme algoritması ile eğitilmiştir. Eğitilen veriler test edildiğinde %92'lik doğruluk oranı bulunmuştur. Yapılan çalışmanın üreteceği çıktılar gemi sistemlerin arıza tahminlerinin yapılmasına olanak sağlayacağı düşünülmektedir. Ülkemizin coğrafi konumu göz önünde bulundurulduğunda denizcilik alanında yapılan çalışmalar sistemlerin arıza yapma olasılığını en aza indirerek hata-arıza tespitini hızlı ve önceden yapılması sağlanabilir.

KAYNAKLAR

Ahmad, M.W., Reynolds, J., & Rezgui, Y. (2018). Predictive modelling for solar thermal energy systems: A comparisonof support vector regression, random forest, extra trees and regression trees. *Journal of Cleaner Production*, 203, 810-821.

Aláiz-Moretón, H., Castejón-Limas, M., Casteleiro-Roca, J-L., Jove, E., Robles, L.F., & Calvo-Rolle, J.L. (2019). A Fault Detection System for a Geothermal Heat Exchanger Sensor Based on Intelligent Techniques. *Sensors*, 19, 2740; doi:10.3390/s19122740.

Ateş, B. (2020). Gemi Yapılarında Gerilme Yığılması Öngörülerinin Kaba Ağ Yapısı ve Makine Öğrenmesi İle Gerçekleştirilmesi. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Aydın, C. (2018). Makine Öğrenmesi Algoritmaları Kullanılarak İtfaiye İstasyonu İhtiyacının Sınıflandırılması. *Avrupa Bilim ve Teknoloji Dergisi*, 14, 169-175.

Bilgili, L. (2018). *Gemi Yaşam Döngüsünde Operasyonel Gaz Emisyonlarının Makine Öğrenmesi Yöntemiyle Tahmini.* Doktora Tezi, Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Cheliotis, M., Lazakis, I. ve Theotokatos, G. (2020). Machine learning and data-driven fault detection for ship systems operations. *Ocean Engineering*, 216, 107968.

Coraddu, A., Oneto, L., Ghio, A., Savio, S., Figari, M. ve Anguita, D. (2015). Machine learning for wear forecasting of naval assets for condition-based maintenance applications. *International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles*, 3-5 March, Italy.

Dai, S-L., Wang, M. ve Wang, C. (2016). Neural Learning Control of Marine Surface Vessels With Guaranteed Transient Tracking Performance. *IEEE Transactions On Industrial Electronics*, 63(3), 1717-1727.

Efendi, A. (2017). A Simulation Study on Bayesian Ridge Regression Models for Several Collinearity Levels. AIP Conference Proceedings 1913, 020031; https://doi.org/10.1063/1.5016665.

Elidolu, G. (2019). *Gemi Adamlarının Emniyetli Çalışma Davranışlarının Makine Öğrenmesi İle Tahminlenmesi.* Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3-42.

Gua, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X. Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354-377.

Hagenauer, J., & Helbich, M. (2013). Hierarchical self-organizing maps for clustering spatiotemporal data. *Int J Geogra Inform Sci.* 27: 2026–2042.

Haykin, S. (1998). Neural Networks: A Comprehensive Foundation, NJ, USA, Prentice Hall.

Ho, C-H., & Lin, C-J. (2012). Large-scale Linear Support Vector Regression. *Journal of Machine Learning Research*, 13, 3323-3348.

Horning, N. (2010). Random Forests: An algorithm for image classification and generation of continuous fields data sets. International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences, 9-11 December, Hanoi, Vietnam.

Huang, B. (2001). Detection of abrupt changes of total least square models and application in fault detection. IEEE Transactions on Control Systems Technology, 9(2), 357-367.

INCASS. (2015). Deliverable D5.4 'Data exchange'. Glasgow (UK): INCASSInspectionCapabilities for Enhanced Ship Safety.

Jiang, G., He, H., Yan, J., & Xie, P. (2019) "Multiscale convolutional neural networks for fault diagnosis of wind turbine gearbox," IEEE Trans. Ind. Electron., 66(4), 3196-3207.

Kamber, M., Pei, J. (2006). Data Mining. Morgan Kaufmann.

Kavzoğlu, T., Şahin, E.K., Çölkesen, İ., 2012. Heyelan duyarlılığınınincelenmesinde regresyon ağaçlarının kullanımı: Trabzon örneği.Harita Dergisi 147, 21–33.

Kimera, D. ve Nangolo, F.N. (2020). Predictive maintenance for ballast pumps on ship repair yards via machine learning. Transportation Engineering, 2, 100020.

Kobbacy, K.A. (2008). Complex System Maintenance Handbook (First Edition). New Jersey: Springer.

Köroğlu, S.A. (2007). Yapay Sinir Ağlarının Gemi Yapı Problemlerine Uygulanması. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Kurt, F. (2018). Evrişimli Sinir Ağlarında Hiper Parametrelerin Etkisinin İncelenmesi. Yüksek Lisans Tezi, Hacettepe Üniversitesi, Ankara.

Lazakis I, Ölçer A. (2015). Selection of the best maintenance approach in themaritime industry under fuzzymultiple attributive group decisionmaking environment. Proc Inst Mech Eng. 230(2), 297-309.

Lazakis, I., Raptodimos, Y. ve Varelas, T. (2018). Predicting ship machinery system condition through amalytical reliability tool and artificial neuron networks. Ocean Engineering, 152, 404-415.

Li, G., (2021). Convolutional Neural Network-Based Bayesian Gaussian Mixture for Intelligent Fault Diagnosis of Rotating Machinery. IEEE Transactions On Instrumentation and Measurement, 70, 3517410.

Li, W., Monti, A. ve Ponci, F. (2014). Fault Detection and Classification in Medium Voltage DC Shipboard Power Systems With Wavelets and Artificial Neural Networks. IEEE Transactions On Instrumentation and Measurement, 63(11), 2651-2665.

Liaw, A., & Wiener, M. (2002). Classification and regression by random Forest. *R News*, 2(3), 18-22.

Lin, K., Lin, Q., Zhou, C., & Yao, J. (2007). Time Series Prediction Based on Linear Regression and SVR. Third International Conference on Natural Computation, 24-27 August, Haikou, China.

Liu, X. (2007). Machinery Fault Diagnostics Based on Fuzzy Measure and Fuzzy Integral Data Fusion Techniques. PhD Thesis, Queensland University of Technology, School of Engineering Systems, Australia.

Lou, X. ve Loparo, K.A. (2004). Bearing fault diagnosis based on wavelet transform and fuzzy inference. Mechanical Systems and Signal Processing, 18, 1077-1095.

Makridis, G., Kyriazis, D. ve Plitsos, S. (2020). Predictive maintenance leveraging machine learning for time-series forecasting in the maritime industry. IEEE 23rd International Conference on Intelligent Transportation Systems, 20-23 September, Greece.

Mohanty, A.R. (2015). Machinery Condition Monitoring (First Edition). Taylor & Francis.

Moosavian, A., Ahmadi, H., Tabatabaeefar, A. ve Khazaee, M. (2013). Comparison of two classifiers; K-nearest neighbor and artificial neural network, for fault diagnosis on a main engine journal-bearing. Shock and Vibration, 20, 263-272.

Nguyen, T.T., Armitage, G., 2008. A survey of techniques for internettraffic classification using machine learning. IEEE communicationssurveys & tutorials 10, 56–76.

Nistane, V.M., & Harsha, S. (2018). Performance Evaluation of Bearing Degradation based on Stationary Wavelet Decomposition and Extra Trees Regression. *World Journal of Engineering,* doi:10.1108/wje-12-2017-0403.

O'Shea, K., Nash, R. (2015). An Introduction to Convolutional Neural Networks. arXiv preprint arXiv:1511.08458.

Öztanır, O. (2018). Makine Öğrenmesi Kullanılarak Kestirimci Bakım. Yüksek lisans tezi, Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.

Pal, P., Datta, R., Segev, A. ve Yasinsac, A. (2019). Condition Based Maintenance of Turbine and Compressor of a Codlag Naval Propulsion System Using Deep Neural Network. In: Dhinaharan Nagamalai ve ark. (Eds). Computer Science & Information Technology (ss.1-12).

Pal, M., Mather, P.M., 2003. An assessment of the effectiveness ofdecision tree methods for land cover classification. Remote sensingof environment 86, 554–565.

Pantelelis, N.G., Kanarachos, A.E. ve Gotzias, N. (2000). Neural networks and simple models for the fault diagnosis of naval turbochargers. Mathematics and Computers in Simulation, 51, 387-397.

Patil, S., & Phalle, V. (2018). Fault Detection of Anti-friction Bearing using Ensemble Machine Learning Methods. *IJE TRANSACTIONS B: Applications*, 31(11), 1972-1981.

Raza J, Liyanage JP. (2009). Application of intelligent technique to identify hidden abnormalities in a system: A case study from oil export pumps froman offshore oil production facility. J QualMaint Eng. 15:221–235.

Saatçioğlu, Ö.Y., Göksu, B., Yüksel, O. ve Gülmez, Y. (2017). Ship Engine Room Casualty Analysis by Using Decision Tree Method. Journal of ETA Maritime Science, 5(1), 59-68.

Safavian, S.R., Landgrebe, D., 1991. A survey of decision tree classi-fier methodology. IEEE transactions on systems, man, and cybernet-ics 21, 660–674.

Shi, Q., Abdel-Aty, M., & Lee, J. (2016). A Bayesian Ridge regression analysis of congestion's impact on urban expressway safety. *Accident Analysis & Prevention*, 88, 124-137.

Şamkar, H., Alpu, Ö., & Altan, E. (2011). Ridge Regresyonda M Tahmin Edicilerinin Kullanımı Üzerine Bir Uygulama Dokuz Eylül Üniversitesi İktisadi İdari Bilimler Fakültesi Dergisi, 26(1).

Tang, T. ve Yao, G. (2005). A Fault-tolerant Control MethodBased on Adaptive FNN for Ship Control System. International Conference on Control and Automation (ICCA2005) June 27-29, Budapest, Hungary.

Uyanık, T., Karatuğ, Ç., & Arslanoğlu, Y. (2020). Machine learning approach to ship fuel consumption: A case of container vessel. *Transportation Research Part D*, 84, 102389.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.

Vasylius, M., Didžiokas, R., Mažeika, P. ve Barzdaitis, V. (2008). The rotating system vibration and diagnostics. Scientific Journal Mechanika, 4(72), 54-58.

Whittaker, J.C., Thompson, R., & Denham, M.C. (2000). Marker-assisted selection using ridge regression. *Genet. Res., Camb.*, 75, 249-252.

Widodo, A., & Yang, B.S. (2007). Review support vector machine in machine condition monitoring and fault diagnosis. Mechanical Systems and Signal Processing, 21, 2560-2574.

Xia, M., Li, T., Xu, L., Liu, L., & de Silva, C. W. (2018). “Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks,” IEEE/ASME Trans. Mechatronics, 23(1), 101-110.

Xiao, C. (2015). Using Machine Learning for Exploratory Data Analysis and Predictive Models on Large Datasets (Master's thesis, University of Stavanger, Norway).

Xie, T., Wang, T., He, Q., Diallo, D. ve Claramunt, C. (2020). A review of current issues of marine current turbine blade fault detection. Ocean Engineering, 218, 108194.

Yang, B. S., Di, X., & Han, T. (2008). Random forests classifier for machine fault diagnosis. *J. Mech. Sci. Technol.*, 22(9), 1716-1725.

Yang, Y., Zheng, H., Li, Y., Xu, M., & Chen, Y. (2019). “A fault diagnosis scheme for rotating machinery using hierarchical symbolic analysis and convolutional neural network,” ISA Trans., 91, 235-252.

Zhongling, L., Ning, J. ve Yan, Z. (2005). The study on application of ANN in fault diagnosis of air conditioning system. Refrigeration and Air-conditioning, 5(1), 50-53.

Zhou, J. ve Xu, L. (2011). The Fault Diagnosis of Marine Engine Cooling System Based on Artificial Neural Network (ANN). Proceedings of the 2nd International Conference on Computer and Automation Engineering, Volume 2, February 26-28, Singapur, 186-189.

ÖZGEÇMİŞ