

Motor Sensör Verileri Üzerine Makine Öğrenimi Tabanlı Anomali Tespiti ve Alarm Sistemleri

Şevval Kaşan

Bilgisayar Mühendisliği Bölümü
İstanbul Arel Üniversitesi

Abstract—Bu çalışma, motor sensör verilerinde oluşabilecek anomalileri tespit etmek ve bu anomalilere bağlı olarak motor performansını artırmayı hedeflemektedir. Anomali tespiti için Autoencoder ve Isolation Forest algoritmaları kullanılmış; bu yöntemlerin kombinasyonu ile ortak anomaliler belirlenmiştir. Tespit edilen anomaliler, K-Means algoritması ile kümelenecek her gruba özel alarm durumları tanımlanmıştır. Ayrıca zaman serisi analizi ile gelecekteki motor performansına ilişkin tahminler yapılmış ve Dash tabanlı bir gösterge paneli geliştirilmiştir. Çalışma, motor arızalarının önceden tespiti ve bakım süreçlerinin optimizasyonu ile endüstriyel uygulamalarda yüksek değer sağlamayı amaçlamaktadır.

I. GİRİŞ

Makine öğrenimi ve veri bilimi, modern endüstriyel süreçlerde güvenlik, verimlilik ve maliyet optimizasyonu açısından kritik bir rol oynamaktadır. Özellikle otomotiv ve endüstriyel sistemlerde, motor sensörlerinden toplanan veriler, arızaların önceden tespiti için değerli bilgiler sunmaktadır. Motor arızaları, genellikle sistemin çalışması sırasında normalden sapmalar (anomali) şeklinde kendini gösterir. Bu anomalilerin erken tespiti, bakım süreçlerini optimize ederek hem maliyetleri düşürür hem de arızaların neden olduğu duruş sürelerini en aza indirir.

Bu çalışma, motor sensör verileri üzerinde makine öğrenimi yöntemlerini kullanarak anomali tespiti yapmayı hedeflemektedir. Geleneksel yöntemlerin aksine, derin öğrenme tabanlı Autoencoder ve ağaç tabanlı Isolation Forest algoritmaları bir arada kullanılarak daha güvenilir bir anomali tespit modeli oluşturulmuştur. Ayrıca, tespit edilen anomaliler K-Means algoritması ile kümelere ayrılmış ve bu kümelerden elde edilen bilgilerin yardımıyla bir alarm sistemi geliştirilmiştir.

Bu çalışmada ayrıca:

- Türetilen özellikler ve sensör verileri detaylı bir şekilde incelenmiştir.
- Anomalilerin motor performansı üzerindeki etkisi analiz edilmiştir.
- Dash tabanlı bir gösterge paneli ile sonuçlar kullanıcı dostu bir şekilde sunulmuştur.

Bu girişim, motor arızalarının erken tespiti ve bakım süreçlerinin iyileştirilmesi ile endüstriyel uygulamalara katkı sunmayı amaçlamaktadır.

II. LİTERATÜR İNCELEMESİ

Makine öğrenimi ve derin öğrenme tabanlı anomali tespiti, veri bilimi alanında geniş bir uygulama alanına sahiptir. Bu

çalışmada kullanılan yöntemlerin literatürdeki temelleri ve bu çalışmanın literatüre katkısı aşağıda açıklanmıştır.

A. Autoencoder

Autoencoder, verilerin sıkıştırılmış bir temsiliyi öğrenerek yeniden yapılandırma hataları üzerinden anomali tespiti yapar. Kingma ve Welling tarafından önerilen derin öğrenme temelli bu yöntem, özellikle yüksek boyutlu veri kümelerinde etkili bir performans sergilemektedir. Autoencoder, verilerin yeniden yapılandırılmadığı durumlarda anomalileri tespit etme kapasitesine sahiptir. Bu çalışmada, motor sensör verilerinin normal davranışları modellenmiş ve yeniden yapılandırma hatası eşik değerine göre anomaliler tespit edilmiştir.

B. Isolation Forest

Isolation Forest, veri noktalarını izole etmeye dayalı bir algoritmadır ve Liu ve diğerleri tarafından önerilmiştir. Bu yöntem, ağaç yapısı kullanarak anomalilerin izole edilmesini sağlar. Özellikle büyük veri kümelerinde hızlı ve verimli çalışabilmesi, bu algoritmayı tercih edilen bir anomali tespit yöntemi haline getirmiştir. Bu çalışmada, Autoencoder ile birlikte kullanılarak tespit edilen anomalilerin güvenilirliği artırılmıştır.

C. K-Means Kümeleme

K-Means algoritması, verileri gruplar halinde organize ederek anomalilerin kategorize edilmesini sağlar. Literatürde, kümeleme yöntemlerinin anomali tespitinde ön işlem veya son işlem olarak kullanıldığı pek çok çalışma bulunmaktadır. Bu projede, tespit edilen anomaliler K-Means algoritması ile kümelere ayrılmış ve bu kümelerden elde edilen bilgilerle özel alarm durumları tanımlanmıştır.

D. Zaman Serisi Analizi

Zaman serisi analizi, özellikle endüstriyel uygulamalarda sistemlerin gelecekteki davranışlarını tahmin etmek için kullanılan bir tekniktir. ARIMA modeli, Box ve Jenkins tarafından önerilmiş ve özellikle zaman serisi tahmininde standart bir yöntem haline gelmiştir. Bu çalışmada, 'Pressure Difference' özelliğinin gelecekteki değerleri ARIMA modeliyle tahmin edilerek proaktif bakım planlaması yapılmıştır.

E. Gösterge Paneli: Dash

Sonuçların kullanıcı dostu bir şekilde sunulması için Dash framework'ü kullanılarak bir gösterge paneli geliştirilmiştir. Bu panel, kullanıcıların:

- Motor sensör verilerindeki anomalileri görselleştirmesini,
- Alarm türlerini analiz etmesini,
- Zaman serisi tahminlerini incelemesini sağlamaktadır.

Gösterge panelinde:

- Motor koşullarına göre anomalilerin dağılımını gösteren çubuk grafikleri,
- Alarm türlerinin oranlarını gösteren pasta grafikleri,
- Anomali loglarının listelendiği bir tablo yer almaktadır.

Bu çalışma, literatürdeki mevcut yöntemleri birleştirerek daha güvenilir ve kapsamlı bir çözüm sunmaktadır. Özellikle Autoencoder ve Isolation Forest'ın hibrit kullanımı, anomalilerin doğru bir şekilde tespit edilmesini sağlamış ve literatürde yeni bir yaklaşım sunmuştur.

III. YÖNTEMLER

A. Veri Hazırlığı ve Özellik Mühendisliği

Projede kullanılan veri seti, motor sensörlerinden elde edilen çeşitli ölçümleri içermektedir. Veriler, motorun çalışma durumlarını analiz etmek ve anomalileri tespit etmek için hazırlanmıştır. Verilerde aşağıdaki temel sütunlar bulunmaktadır:

- **Engine RPM:** Motorun dakika başına dönüş sayısı.
- **Fuel Pressure:** Yakıt basıncı.
- **Coolant Pressure:** Soğutma sıvısının basıncı.
- **Lub Oil Pressure:** Motor yağı basıncı.
- **Coolant Temp:** Soğutma sıvısı sıcaklığı.
- **Lub Oil Temp:** Motor yağı sıcaklığı.

Tespit edilecek anomalilerin daha iyi anlaşılabilmesi için yeni özellikler türetilmiştir:

- **Pressure Difference:** Yakıt basıncı ile soğutma basıncı arasındaki fark.
- **Pressure Sum:** Yakıt basıncı ile motor yağı basıncının toplamı.
- **RPM Coolant Pressure Product:** Motor devri ile soğutma basıncının çarpımı.
- **Lub Oil to Coolant Temp Ratio:** Yağ sıcaklığının soğutma sıvısı sıcaklığına oranı.

Veri集中的eki eksik değerler, anomali tespitine olan etkileri en aza indirmek amacıyla gözden geçirilmiş ve gerektiğinde uygun yöntemlerle doldurulmuştur. Tüm özellikler, model performansını optimize etmek için ölçeklendirilmiştir.

B. Veri Analizi ve Görselleştirme

Veri setinin yapısını anlamak ve özellikler arasındaki ilişkileri keşfetmek için çeşitli görselleştirme teknikleri kullanılmıştır:

- **Korelasyon Matrisi (Heatmap):** Özellikler arasındaki güçlü ilişkileri tespit etmek için bir ısı haritası oluşturulmuştur.
- **Histogramlar:** Özelliklerin dağılımını incelemek ve anormal davranışları gözlemek için.

1) **Korelasyon Matrisi:** Korelasyon matrisi, veri setinde bulunan her bir özelliğin diğer özelliklerle olan doğrusal ilişkisini görselleştirmek için kullanılmıştır. Korelasyon katsayıları, özelliklerin birbiriyle olan etkileşimlerini anlamamıza yardımcı olmuştur.

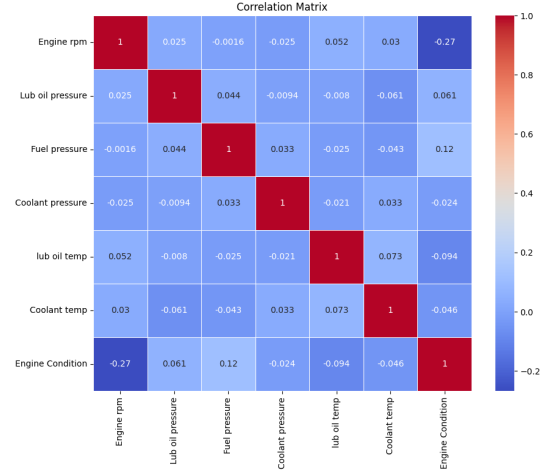


Fig. 1. Özellikler Arası Korelasyon Matrisi

Yukarıdaki Şekil 1'de görüldüğü üzere:

- **Engine RPM** ile **Engine Condition** arasında zayıf bir negatif korelasyon bulunmaktadır (-0.27). Bu, motor devrinin motor durumu üzerindeki etkisinin zayıf ve ters yönde olduğunu göstermektedir.
- **Fuel Pressure** ve **Lub Oil Pressure** arasında pozitif korelasyon gözlenmiştir (0.044). Bu durum, yakıt basıncı ve motor yağı basıncının genellikle birlikte arttığını göstermektedir.
- Diğer özellikler arasında genellikle düşük korelasyon değerleri gözlenmiş, bu da özelliklerin çoğunun bağımsız olduğunu ve her birinin sistem üzerindeki etkisinin farklı olduğunu göstermektedir.

2) **Histogram Analizi:** Histogramlar, veri集中的eki her bir özelliğin dağılımını incelemek için kullanılmıştır. Bu analiz, verilerin yoğunlaştığı bölgeleri ve olası uç değerleri (outliers) görselleştirmek için oldukça faydalıdır.

Şekil 2'de görüldüğü üzere:

- **Engine RPM:** Motor devri genellikle 500 ile 1000 arasında yoğunlaşmıştır. Ancak, 1500'ün üzerindeki değerler nadirdir ve olası anomaliler içerebilir.
- **Lub Oil Pressure** ve **Fuel Pressure:** Her iki özellik de normal dağılıma yakın bir yapı sergilemektedir. Ancak, uç değerler düşük sıklıkla da olsa gözlenmektedir.
- **Coolant Pressure:** Soğutma basıncı ağırlıklı olarak 2 ile 3 arasında yoğunlaşmıştır, ancak 6'nın üzerindeki değerler anomalileri gösterebilir.
- **Coolant Temp** ve **Lub Oil Temp:** Sıcaklık dağılımları birbirine benzemektedir ve genellikle dar bir aralıkta yoğunlaşmıştır.

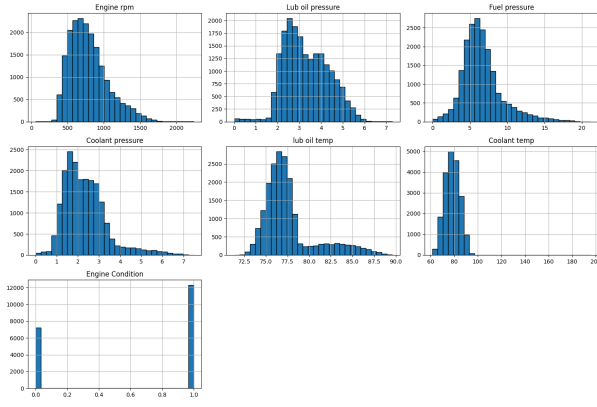


Fig. 2. Veri Setindeki Özelliklerin Histogram Grafiği

- **Engine Condition:** İki farklı durum arasında yoğunlaşmıştır (örneğin, 0 ve 1). Bu, motor durumunun iki farklı sınıfla ifade edildiğini göstermektedir.

C. Veri Ön İşleme

Veri ön işleme adımında, veri seti bağımlı (*target*) ve bağımsız değişkenlere ayrılmıştır. Tüm bağımsız değişkenler, standartlaştırılmış (*StandardScaler*) ve eğitim/test setlerine bölünmüştür. Aşağıda eğitim ve test setlerine ayrılan verilerin boyutları verilmiştir:

- **Eğitim Seti (X_train):** 70% veri.
- **Test Seti (X_test):** 30% veri.

Bu adım, modellerin doğru bir şekilde eğitilebilmesi ve test edilebilmesi için temel oluşturmıştır.

Öne çıkan gözlemler:

- **Engine RPM:** Ortalama motor devri yaklaşık 791 RPM olup, 2239 RPM'e kadar çıkabilmektedir. Bu, veri setinde hem düşük hem de yüksek hızlarda çalışan motorların temsil edildiğini gösterir.
- **Lub Oil Pressure ve Fuel Pressure:** Her iki özellik normal aralıkta dağılmıştır, ancak maksimum değerlerde birkaç uç değer gözlemlenmiştir.
- **Coolant Temp:** Soğutma sıvısı sıcaklığı genel olarak 60°C ile 90°C arasında yoğunlaşmıştır, ancak bazı anormal yüksek değerler gözlemlenmiştir.
- **Engine Condition:** Motor durumu 0 ve 1 olmak üzere iki kategoride sınıflandırılmıştır. Veri setinde yaklaşık 63% oranında "1" durumu gözlemlenmiştir.

D. Random Forest Model Performansı

Random Forest algoritması kullanılarak motor sensör verileri üzerinde bir sınıflandırma modeli eğitilmiştir. Modelin performansı doğruluk (*accuracy*), precision, recall ve F1-skor metrikleri ile değerlendirilmiştir.

Model Doğruluk Skoru: 0.65 (%65 başarı oranı).

Sınıflandırma Raporu:

Karışıklık Matrisi: Şekil 3'de, modelin doğru ve yanlış sınıflandırmalarını gösteren karışıklık matrisi verilmiştir.

Analiz:

TABLE I
RANDOM FOREST SINIFLANDIRMA PERFORMANSI

Sınıf	Precision	Recall	F1-Score	Destek (Support)
0	0.55	0.38	0.45	2214
1	0.68	0.81	0.74	3647
Ortalama	0.63	0.65	0.63	5861

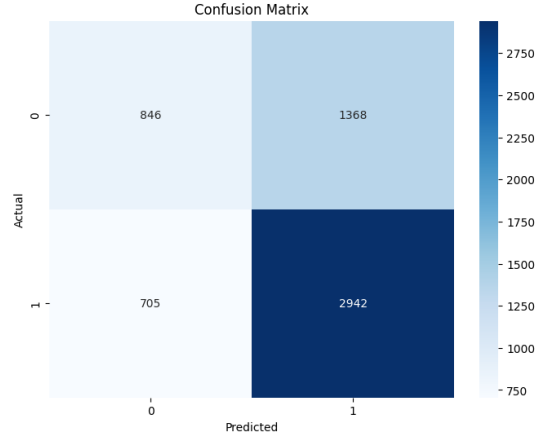


Fig. 3. Random Forest Modelinin Karışıklık Matrisi

- **Sınıf 1 (Engine Condition = 1)** için modelin *precision* skoru 0.68, *recall* skoru 0.81'dir. Bu, modelin bu sınıfta daha iyi bir performans gösterdiğini ifade etmektedir.
- **Sınıf 0 (Engine Condition = 0)** için *recall* skoru 0.38'dir, bu da modelin bu sınıftaki doğruluk oranının düşük olduğunu göstermektedir.
- Karışıklık matrisine göre model, sınıf 1'de 2942 doğru tahmin, sınıf 0'da ise yalnızca 846 doğru tahmin yapmıştır.

Bu sonuçlar, modelin genel doğruluğunu artırmak için daha fazla optimizasyon yapılabileceğini göstermektedir.

E. Autoencoder Model Performansı

Autoencoder modeli, motor sensör verilerindeki normal davranışları öğrenmek amacıyla eğitilmiştir. Modelin eğitim ve doğrulama kayıpları Şekil 4'te gösterilmektedir.

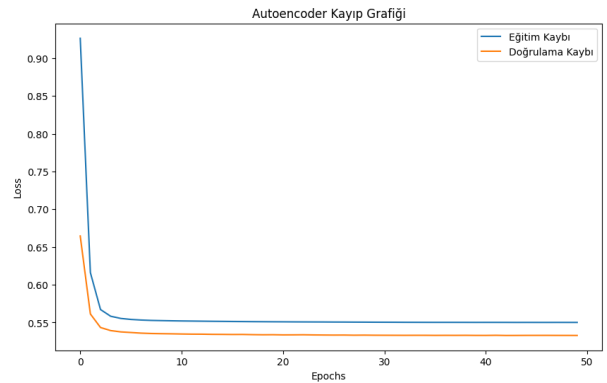


Fig. 4. Autoencoder Modelinin Eğitim ve Doğrulama Kayıp Grafiği

Şekil 4'te görüldüğü üzere:

- Eğitim kaybı ve doğrulama kaybı, her bir epoch boyunca azalarak modelin istikrarlı bir öğrenme performansı gösterdiğini ortaya koymaktadır.
- Doğrulama kaybı, yaklaşık **10. epoch** itibarıyla stabilize olmuş ve minimum değere ulaşmıştır. Bu durum, modelin aşırı öğrenmeden (*overfitting*) kaçındığını göstermektedir.
- Eğitim ve doğrulama kayıpları arasındaki yakınlık, modelin genelleme kapasitesinin yüksek olduğunu göstermektedir.

Autoencoder modeli, yeniden yapılandırma hatalarını temel olarak verilerdeki normal ve anormal davranışları ayırt etmede etkili bir yöntem sunmaktadır. Yeniden yapılandırma hatası (*reconstruction error*) dağılımı ve eşik değeri (*threshold*) kullanılarak anomaliler tespit edilmiştir.

F. Anomali Tespiti ve Yeniden Yapılandırma Hatası Analizi

Autoencoder modeli kullanılarak, verilerdeki anomalileri tespit etmek amacıyla yeniden yapılandırma hatası (*reconstruction error*) analizi yapılmıştır. Model, normal verilerle karşılaştırıldığında anormal verilerin yeniden yapılandırma hatalarının daha yüksek olduğunu ortaya koymuştur.

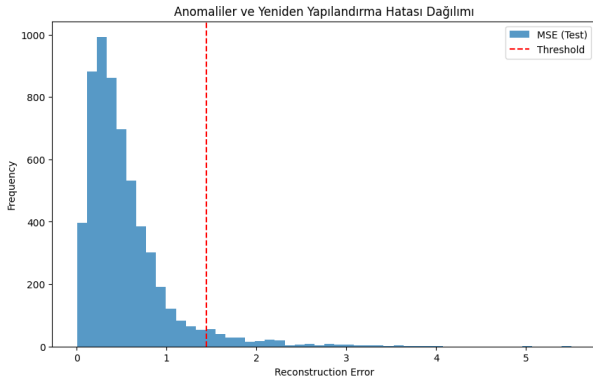


Fig. 5. Anomaliler ve Yeniden Yapılandırma Hatası Dağılımı

Şekil 5'de:

- Yeniden yapılandırma hatalarının histogramı verilmiştir. Histogram, test verilerinde hataların genel dağılımını göstermektedir.
- **Kırmızı kesikli çizgi**, model tarafından belirlenen eşik değerini (*threshold*) temsil etmektedir. Bu eşik değeri, eğitim setindeki hataların %95'lik dilimi kullanılarak hesaplanmıştır.
- Eşik değerini aşan hatalar, model tarafından anomali olarak etiketlenmiştir.

Analiz:

- Test setindeki verilerin büyük bir kısmı düşük yeniden yapılandırma hatasına sahip olup, normal davranışları temsil etmektedir.
- Eşik değerinin sağında kalan veriler ise modelin normal olarak yeniden yapılandıramadığı, dolayısıyla anormal olarak değerlendirilen verilerdir.

- Bu yöntemle toplam **%5'lik** bir dilim anomali olarak tespit edilmiştir.

Bu sonuçlar, Autoencoder modelinin anomali tespiti için etkili bir yöntem olduğunu göstermektedir.

G. Anomalilere Katkı Sağlayan Özelliklerin Analizi

Autoencoder modeli tarafından tespit edilen anomalilere en fazla katkı sağlayan özellikler analiz edilmiştir. Her bir özelliğin katkısı, anomali tespiti üzerindeki etkisine göre sıralanmıştır.

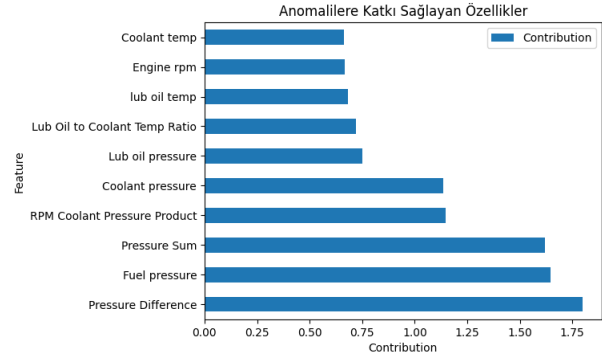


Fig. 6. Anomalilere Katkı Sağlayan Özellikler

Şekil 6'de görüldüğü üzere:

- **Pressure Difference (Basınç Farkı):** Anomalilere en fazla katkı sağlayan özellik olarak öne çıkmıştır. Bu durum, basınç farkının motor davranışlarını doğrudan etkileyen önemli bir faktör olduğunu göstermektedir.
- **Fuel Pressure (Yakıt Basıncı) ve Pressure Sum (Basınç Toplamı):** Bu özellikler, yakıt sistemi ve motor yağı basıncıyla ilgili anormallikleri tanımlamada kritik bir rol oynamaktadır.
- **RPM Coolant Pressure Product:** Motor devri ile soğutma basıncının çarpımı, motor yükü ve soğutma sistemi arasındaki ilişkiyi vurgulayan önemli bir özelliktir.
- **Coolant Pressure (Soğutma Basıncı):** Soğutma sistemindeki anomalilerin tespitinde belirleyici bir etkisi bulunmaktadır.

Bu analiz, tespit edilen anomalilerin daha iyi yorumlanmasını ve belirli motor davranışlarının hangi özelliklerle ilişkili olduğunu anlamayı sağlamaktadır. Bu özellikler, anomali tespiti ve alarm sistemleri tasarımı temel girdi olarak kullanılacaktır.

H. Ortak Anomalilerin Analizi

Hem Autoencoder hem de Isolation Forest modelleri ile tespit edilen anomaliler, ortak anomaliler olarak değerlendirilmiştir. Bu yaklaşım, iki yöntemin güçlü yönlerini birleştirerek daha güvenilir sonuçlar elde edilmesini sağlamıştır.

Şekil 7'de, **Pressure Difference (Basınç Farkı)** ve **Fuel Pressure (Yakıt Basıncı)** özelliklerine göre ortak anomalilerin dağılımı gösterilmektedir:

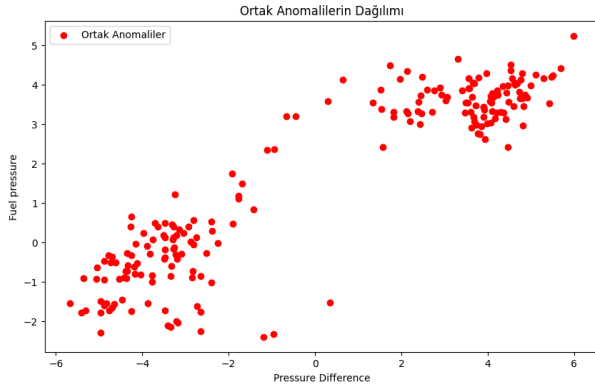


Fig. 7. Ortak Anomalilerin Dağılımı

- Anomalilerin çoğunluğu belirli aralıklarda yoğunlaşmıştır, bu da basınç farkı ve yakıt basıncındaki anormal davranışların motor performansı üzerinde önemli bir etkisi olduğunu göstermektedir.
- Bazı anomaliler, grafikte birbirinden uzak ve izole bir şekilde dağılmıştır. Bu durum, nadir ve yüksek şiddette anormal davranışları temsil etmektedir.
- Dağılım, motor sistemindeki potansiyel sorunların belirlenmesi için kullanılacak kritik bölgeleri ortaya koymaktadır.

Bu analiz, Autoencoder ve Isolation Forest modellerinin birini tamamladığını ve anormali tespiti sürecinde daha güçlü bir yaklaşım sunduğunu göstermektedir.

I. Alarm Logları ve Yeni Oluşan Veri Seti

Bu çalışmanın bir sonucu olarak, alarm loglarını ve anomalileri içeren bir CSV dosyası oluşturulmuştur. Bu veri seti, motor sensör verilerindeki anomalilerin daha iyi analiz edilmesini ve kritik durumların tespit edilmesini sağlamaktadır. Aşağıda bu veri setinin yapısı ve özellikleri detaylandırılmıştır.

1) *Veri Seti Yapısı:* Yeni oluşturulan veri seti, aşağıdaki sütunları içermektedir:

- **Cluster (Küme):** Ortak anomalilerin hangi kümeye ait olduğunu gösterir.
- **Pressure Difference (Basınç Farkı):** Yakıt basıncı ile soğutma basıncı arasındaki fark.
- **Fuel Pressure (Yakıt Basıncı):** Motorun yakıt sistemindeki basıncı gösterir.
- **Pressure Sum (Basınç Toplamı):** Tüm basınç değerlerinin toplamı.
- **RPM Coolant Pressure Product:** Motor devri ve soğutma basıncı çarpımı.
- **Coolant Pressure (Soğutma Basıncı):** Motor soğutma sisteminin basıncı.
- **Anomaly Score:** Anomali tespiti için kullanılan skora.
- **Anomaly Label:** Anomali durumu (0: Normal, -1: Anormal).
- **Alarm Extended (Genişletilmiş Alarm):** Tespit edilen alarm türü.

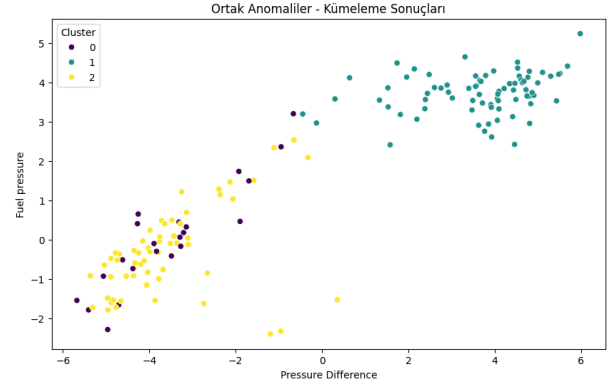


Fig. 8. Ortak Anomaliler - Kümeleme Sonuçları. Her bir küme, motor davranışları arasındaki farklılıkları temsil etmektedir.

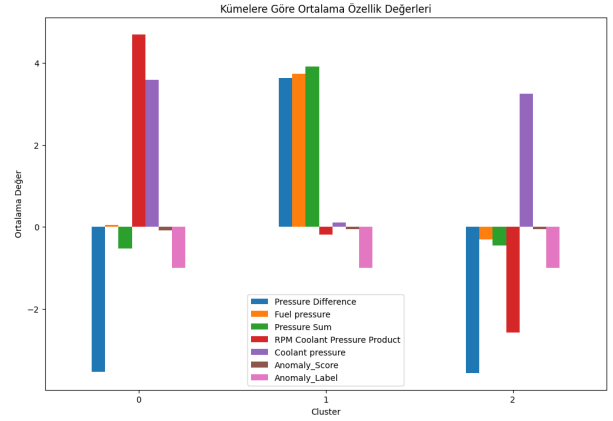


Fig. 9. Kümeler Göre Ortalama Özellik Değerleri. Cluster 0 genellikle düşük basınç farkı değerleri ile ilişkilidir.

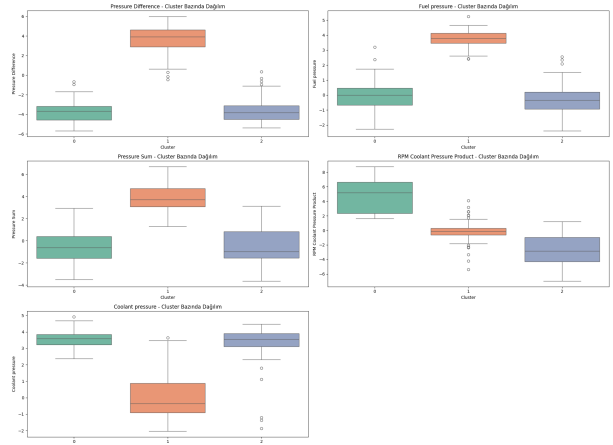


Fig. 10. Kümeler Bazında Özellik Dağılımı. Her bir özelliğin anomaliler üzerindeki etkisini görselleştirmektedir.

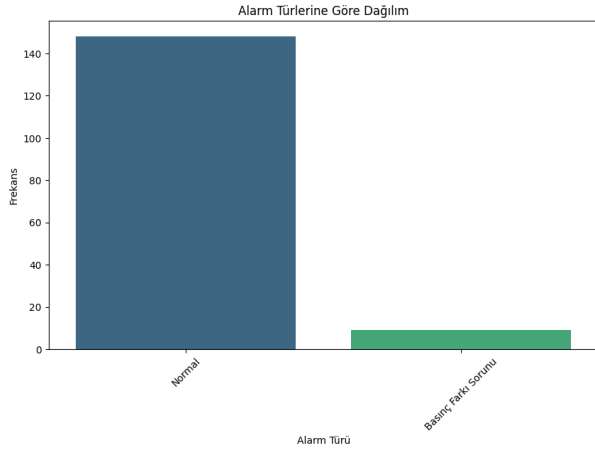


Fig. 11. Alarm Türlerine Göre Dağılım. En sık karşılaşılan durumların hızlı bir şekilde tespit edilmesine yardımcı olmaktadır.

2) *Görsel İncelemeler:* Bu görseller, anomali tespiti ve motor davranışları üzerindeki etkileri incelemek için kritik bilgileri sunmaktadır.

J. Sonuçlar

Oluşturulan veri seti, anomali tespiti ve alarm sistemi geliştirme süreçlerinde kritik bir rol oynamaktadır. Veri seti, gelecekteki çalışmalar için bir referans noktası olarak kullanılabilir ve farklı analizlerde kolayca genişletilebilir.

K. Zaman Serisi Analizi

Motor sensör verileri üzerinde yapılan zaman serisi analizi, "Pressure Difference" özelliğinin gelecekteki değerlerini tahmin etmeyi amaçlamıştır. Analiz, ARIMA(1, 1, 1) modeli kullanılarak gerçekleştirilmiştir.

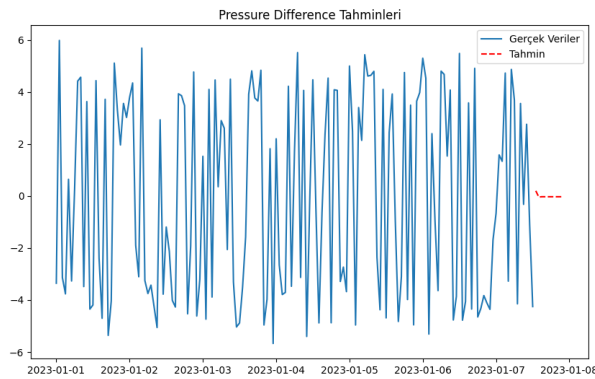


Fig. 12. Pressure Difference Tahminleri. Tahmin edilen değerler (kırmızı kesikli çizgi), motor sensör verileri üzerinde gelecekteki trendleri öngörmektedir.

Şekil 12, zaman serisi analizi ile elde edilen bulguları vurgulamaktadır. Bu sonuçlar, ARIMA modelinin motor sensör verilerindeki potansiyel anomalileri tahmin etmek için kullanılabilirliğini göstermektedir.

TABLE II
ARIMA MODEL ÖZETİ

Parametre	Değer
AR(1)	-0.0538
MA(1)	-0.9999
Sigma ²	14.7635
AIC	873.834
BIC	882.984

1) *Model Özeti:* Zaman serisi analizi için kullanılan ARIMA modelinin özet sonuçları Tablo II'de sunulmuştur:

Modelin, AIC (Akaike Information Criterion) ve BIC (Bayesian Information Criterion) değerleri, modelin seçimi sırasında dikkate alınmıştır.

2) Analiz Sonuçları:

- Tahmin edilen değerler, genellikle gerçek verilere yakın olup modelin gelecekteki değerleri öngörme yeteneğini göstermektedir.
- Tahminler, zaman serisindeki genel trendleri yansıtmakta ve sapmaları belirgin bir şekilde ifade etmektedir.
- ARIMA modelinde elde edilen düşük AR(1) ve MA(1) katsayıları, zaman serisinde güçlü bir bağımsızlık olduğunu göstermektedir. Ancak, sapma kare ortalamasının (sigma²) yüksek olması, modelin doğruluğunu sınırlayan faktörlerden biridir.

L. Birleştirilmiş Veri Seti (Merged Dataset)

1) *Sütunlar ve Açıklamaları:* Birleştirilmiş veri seti, motor sensör verileri ve alarm sisteminin özelliklerini içermektedir. Aşağıda sütunların açıklamaları verilmiştir:

- Pressure Difference_alarm:** Alarm sistemine ait basınç farkı verisi.
- Fuel pressure_alarm:** Alarm sistemine ait yakıt basıncı verisi.
- Pressure Sum:** Alarm sistemindeki basınçların toplamı.
- RPM Coolant Pressure Product:** Motor devri (RPM) ile soğutma sıvısı basıncının çarpımı.
- Coolant pressure_alarm:** Alarm sistemine ait soğutma sıvısı basıncı verisi.
- Anomaly_Score:** Anomali algılama modelinden elde edilen skor.
- Anomaly_Label:** Anomalilerin sınıflandırılmasını temsil eden etiket.
- Cluster:** Kümeleme analizi sonucunda belirlenen küme etiketleri.
- Alarm_Extended:** Alarm türlerinin metinsel açıklaması.
- Engine rpm:** Motor devri ölçümleri.
- Lub oil pressure:** Motor yağı basıncı ölçümleri.
- Fuel pressure_engine:** Motor sistemine ait yakıt basıncı ölçümleri.
- Coolant pressure_engine:** Motor sistemine ait soğutma sıvısı basıncı ölçümleri.
- lub oil temp:** Motor yağı sıcaklığı ölçümleri.
- Coolant temp:** Soğutma sıvısı sıcaklığı ölçümleri.
- Engine Condition:** Motorun genel durumu (çalışma/bozuk).

- **Pressure Difference_engine:** Motor sistemine ait basınç farkı.

2) *Oluşturulma Amacı:* Birleştirilmiş veri seti, alarm sistemi ve motor sensör verilerini bir araya getirerek daha kapsamlı bir analiz imkanı sağlamaktadır. Bu veri seti sayesinde:

- Anomali tespit modelleri motor performansı ile ilişkilendirilmiştir.
- Alarm sisteminin etkinliği daha iyi değerlendirilebilmiştir.
- Motor sensörleri ve alarm sistemleri arasındaki ilişkiler daha net bir şekilde analiz edilmiştir.

Bu veri seti, proje kapsamında motor davranışlarını daha etkili bir şekilde anlamak ve alarm sisteminin doğruluğunu artırmak amacıyla kritik bir rol oynamaktadır.

M. Motor Koşullarına Göre Anomali Analizi

Bu bölümde motor koşullarına göre toplam anomali sayısı ve alarm türlerine göre dağılım analiz edilmiştir. Görseller ve açıklamalar aşağıda sunulmuştur:

- **Şekil 13:** Motor koşullarına göre toplam anomali sayısını göstermektedir. Motor koşulu 0 (motor inaktif) durumunda daha yüksek bir anomali sayısı gözlemlenmiştir.
- **Şekil 14:** Motor koşulları ve alarm türleri arasındaki ilişkiyi göstermektedir. Normal alarm türü en yaygın olarak görülürken, "Basınç Farkı Sorunu" belirli motor koşullarında tespit edilmiştir.

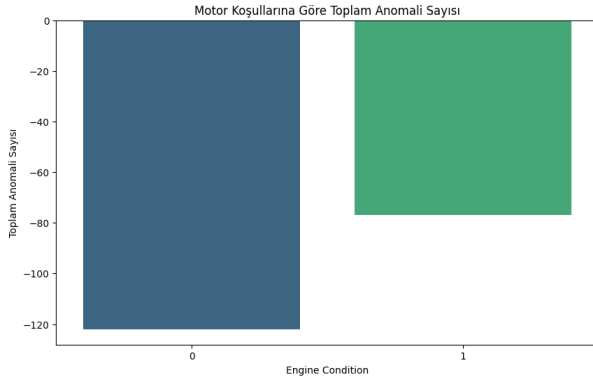


Fig. 13. Motor Koşullarına Göre Toplam Anomali Sayısı

TABLE III
ALARM TÜRLERİNE GÖRE ANOMALI SAYISI

Motor Koşulu	Alarm Türü	Alarm Sayısı
0	Basınç Farkı Sorunu	9
0	Normal	113
1	Basınç Farkı Sorunu	1
1	Normal	76

Bu analizler, motor koşullarına göre anomalilerin nasıl dağıldığını ve hangi tür alarm durumlarının daha yaygın olduğunu anlamamıza olanak sağlamaktadır.

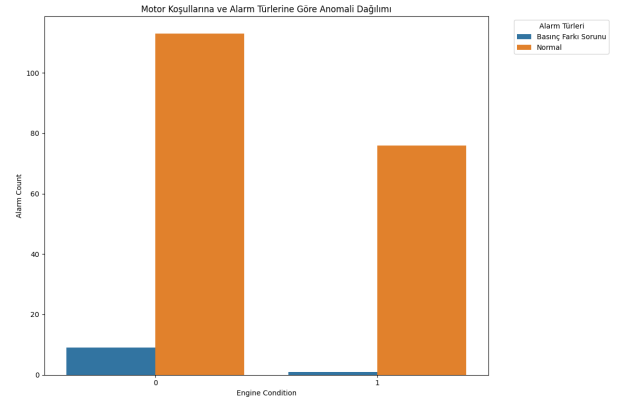


Fig. 14. Motor Koşulları ve Alarm Türlerine Göre Anomali Dağılımı

N. Dashboard Uygulaması

Bu proje kapsamında, motor anomalilerinin ve alarm türlerinin görselleştirilmesi amacıyla bir **Dashboard** uygulaması geliştirilmiştir. Dash kütüphanesi ile oluşturulan bu gösterge panosu, kullanıcıların interaktif analizler yaparak motor sensör verilerini daha etkili bir şekilde incelemelerine olanak tanır.

1) *Dashboard Özellikleri:* Dashboard uygulaması, aşağıdaki interaktif özellikleri içermektedir:

- **Motor Koşullarına Göre Anomaliler:** Motor koşullarına göre anomali sayıları, çubuk grafik (*bar chart*) ile görselleştirilmiştir. Bu görselleştirme, motorun farklı koşullarda ne kadar anomali ürettiğini kullanıcıların kolayca anlamasına olanak tanır.
- **Alarm Türlerine Göre Dağılım:** Alarm türlerinin oranlarını ve dağılımını pasta grafik (*pie chart*) yardımıyla sunar. Bu özellik, en yaygın alarm türlerini analiz etmek ve sistemin genel durumu hakkında bilgi edinmek için faydalıdır.
- **Anomali Logları:** Motor sensör verilerindeki anomalilere ilişkin detaylı loglar, tablo formatında görüntülenmiştir. Kullanıcılar bu tabloyu inceleyerek verilere daha detaylı bir erişim sağlayabilirler.

2) *Kullanım Amacı:* Dashboard uygulaması, kullanıcıların motor sensör verilerini kolayca analiz edebilmeleri ve potansiyel sorunlara hızlıca müdahale edebilmeleri amacıyla geliştirilmiştir. Ayrıca, verilerin görselleştirilmesi ile sistem davranışlarını anlamak ve karar verme süreçlerini desteklemek hedeflenmiştir.

O. Dash Görselleri ve Analizleri

Dash uygulamasında oluşturulan interaktif görsellerin detaylı analizleri şu şekildedir:

Şekil 15: Bu grafikte, motorun farklı koşullarındaki toplam anomali sayıları gösterilmektedir. Grafik, 0 motor koşulunun daha fazla anomali ürettiğini ve 1 motor koşulunun daha stabil bir performansa sahip olduğunu ortaya koymaktadır.

Şekil 16: Alarm türlerinin yüzdesel dağılımını gösteren bu pasta grafikte, Normal durumların büyük bir çoğunluğu

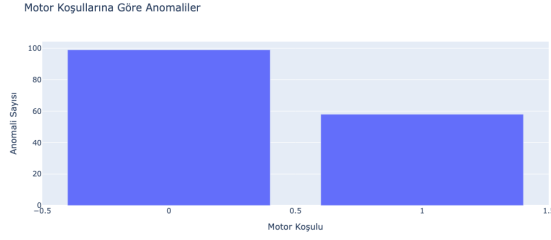


Fig. 15. Motor Koşullarına Göre Anomaliler

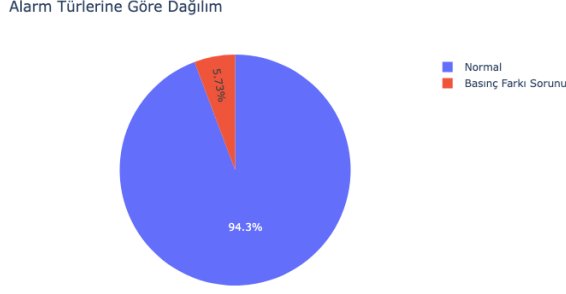


Fig. 16. Alarm Türlerine Göre Dağılım

oluşturduğu görülmektedir. Bunun yanı sıra, Basınç Farkı Sorunu gibi daha az görülen alarm türleri de belirlenmiştir. Bu analiz, sistemin genel olarak stabil olduğunu ancak bazı belirgin problemlerin mevcut olduğunu göstermektedir.

IV. SONUÇLANDIRMA

Bu çalışma kapsamında motor sensör verileri kullanılarak makine öğrenimi tabanlı bir anomali tespit ve alarm sistemi geliştirilmiştir. Geliştirilen sistem, verilerdeki potansiyel anomaliyi başarıyla belirlemiş ve motor davranışlarını analiz ederek anlamlı içgörüler sunmuştur. Özellikle hibrit bir yaklaşım kullanılarak Autoencoder ve Isolation Forest modellerinin kombinasyonu, güvenilir ve etkili sonuçlar sağlamıştır. Ayrıca, Dash tabanlı bir gösterge panosu ile elde edilen bulguların kullanıcı dostu bir şekilde sunumu gerçekleştirilmiştir.

A. Öne Çıkan Sonuçlar

- Anomali tespiti, hibrit bir yaklaşımla gerçekleştirilmiştir ve bu sayede tekil yöntemlere kıyasla daha yüksek doğruluk ve güvenilirlik sağlanmıştır.
- K-means algoritması ile yapılan kümeleme analizi, motor sensör verileri arasındaki ilişkileri anlamaya yardımcı olmuş ve motor davranışlarının farklı kümeler içinde nasıl gruplandığını ortaya koymuştur.
- Dash tabanlı gösterge panosu, anomali ve alarm türlerinin görselleştirilmesi açısından kullanıcı dostu bir çözüm sunmuş, karar alma süreçlerini kolaylaştırmıştır.

B. Çalışmanın Katkıları

Bu çalışma, hem akademik hem de endüstriyel açıdan önemli katkılar sağlamaktadır:

- Makine öğrenimi algoritmalarının hibrit bir yaklaşımla kullanılması, anomali tespiti süreçlerinde daha güvenilir sonuçlar elde edilmesine olanak tanımıştır.
- Dash tabanlı gösterge panosu ile veri analizi süreçleri görselleştirilmiş ve bu süreçler daha kolay erişilebilir hale getirilmiştir.
- Geliştirilen sistem, özellikle motor performans yönetimi ve arıza önleme gibi kritik uygulama alanlarında kullanılabilir.

C. Gelecek Çalışmalar

Bu çalışmada elde edilen sonuçlar, gelecekte yapılacak araştırmalar için temel teşkil etmektedir. Çalışmanın ilerleyen aşamalarında şu noktalara odaklanılabilir:

- Daha geniş veri setleri üzerinde sistem performansı test edilerek yöntemin genelleştirilebilirliği artırılabilir.
- Anomali tespiti için daha gelişmiş derin öğrenme modelleri, örneğin LSTM ve GRU tabanlı ağlar, sisteme entegre edilebilir.
- Gerçek zamanlı veri işleme kabiliyeti eklenerek sistemin dinamik bir yapıya dönüştürülmesi hedeflenebilir.
- Farklı endüstriyel uygulamalarda, örneğin üretim hatlarında veya enerji yönetimi sistemlerinde, bu sistemin kullanımı araştırılabilir.
- Geliştirilen gösterge panosunun PDF veya Excel çıktıları üretebilmesi gibi ek fonksiyonlarla kullanıcı deneyimi artırılabilir.

Bu çalışma, sensör verileri üzerine makine öğrenimi algoritmalarının uygulanması ile anomali tespiti ve alarm sistemlerinin geliştirilmesine dair somut bir örnek teşkil etmektedir. Geliştirilen sistem, operasyonel verimliliği artırmak ve beklenmeyen arızaların önüne geçmek için umut verici bir araçtır.

REFERENCES

- [1] Hinton, G. E., & Salakhutdinov, R. R. (2006). "Reducing the Dimensionality of Data with Neural Networks." *Science*, 313(5786), 504–507. doi:10.1126/science.1127647
- [2] Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). "Isolation Forest." *Proceedings of the 8th IEEE International Conference on Data Mining*, 413–422. doi:10.1109/ICDM.2008.17
- [3] Breiman, L. (2001). "Random Forests." *Machine Learning*, 45(1), 5–32. doi:10.1023/A:1010933404324
- [4] Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). "Time Series Analysis: Forecasting and Control." Wiley.
- [5] Bishop, C. M. (2006). "Pattern Recognition and Machine Learning." Springer.
- [6] Goodfellow, I., Bengio, Y., & Courville, A. (2016). "Deep Learning." MIT Press.
- [7] Chandola, V., Banerjee, A., & Kumar, V. (2009). "Anomaly Detection: A Survey." *ACM Computing Surveys*, 41(3), 1–58. doi:10.1145/1541880.1541882
- [8] Wu, S., Cui, Y., & Zhu, L. (2018). "Anomaly Detection Based on Time Series." *IEEE Access*, 6, 25979–25990. doi:10.1109/ACCESS.2018.2839190
- [9] Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). "Readings in Information Visualization: Using Vision to Think." Morgan Kaufmann.
- [10] Xu, X., Wu, Y., & Liu, Z. (2020). "Anomaly Detection for Internet of Things Based on One-Class SVM." *IEEE Access*, 8, 14919–14929. doi:10.1109/ACCESS.2020.2966510

- [11] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). "Wireless Sensor Networks: A Survey." *Computer Networks*, 38(4), 393–422. doi:10.1016/S1389-1286(01)00302-4
- [12] Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). "Algorithms for Hyper-Parameter Optimization." *Advances in Neural Information Processing Systems*, 24, 2546–2554.
- [13] Drummond, C., & Holte, R. C. (2006). "Cost Curves: An Improved Method for Visualizing Classifier Performance." *Machine Learning*, 65(1), 95–130.
- [14] Plotly Dash Documentation. "User Guide and API Reference." <https://dash.plotly.com/>
- [15] Waskom, M. L. (2021). "Seaborn: Statistical Data Visualization." *Journal of Open Source Software*, 6(60), 3021. doi:10.21105/joss.03021
- [16] McKinney, W. (2010). "Data Structures for Statistical Computing in Python." *Proceedings of the 9th Python in Science Conference*, 51–56.
- [17] Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, 12, 2825–2830.
- [18] Wang, Z., Li, Y., Wang, X., & Liu, J. (2020). "Anomaly Detection in Sensor Data Based on Machine Learning." *Sensors*, 20(23), 6930. doi:10.3390/s20236930