Rao Ikram Bari

Unsupervised Anomaly Detection in Industrial Engine Under Varying Load Conditions

School of Technology and Innovations

Master’s Programme in Computing Sciences

Sustainable and Autonomous Systems

Vaasa 2026

Acknowledgments

This thesis was completed at the University of Vaasa as part of the Master's Programme in Sustainable and Autonomous Systems, in collaboration with Nome Oy, where the experimental work and industrial evaluation were conducted.

I gratefully acknowledge the financial support provided by the Technology Industries of Finland Centennial Foundation through the AI Thesis Grant Programme, without which this research would not have been possible.

I would like to express my sincere gratitude to Professor Mourad Oussalah from the University of Oulu for his invaluable guidance, constructive feedback, and continuous encouragement throughout this research. I am equally thankful to Mr. Juha Hautala at Nome Oy for his practical insights and generous support during the experimental phase of this work. I would also like to extend my appreciation to Professor Mohammed Elmusrati from the University of Vaasa for his supervision and academic mentorship throughout my master's studies.

Above all, I dedicate this work to the memory of my late mother, whose unconditional love and silent prayers remain my greatest strength. I would also like to express my deepest gratitude to my father, Abdul Bari, my brothers Ehsan Bari, Inam Bari, and Ahtisham Bari, and my sisters for their unwavering love, prayers, and support throughout this journey.

VAASAN YLIOPISTO
School of Technology and Innovations

Author:	Rao Ikram Bari
Thesis title:	Unsupervised Anomaly Detection in Industrial Engine Under Varying Load Condition
Degree:	Sustainable and Autonomous Systems
Supervisor:	Mohammed Elmusrati
Co-Supervisor:	Mourad Oussalah (University of Oulu)
Instructor:	Juha Hautala (Nome Oy)
Year of graduation:	2026 **Number of pages:** 73

ABSTRACT:
Reliable identification of abnormal behavior in industrial machines is important for ensuring operational reliability of the systems. Traditional condition monitoring methods rely on manually derived features and fixed thresholds, which are not suitable in varying operating conditions. To address these limitations, an anomaly detection architecture has been proposed for acoustic signals in an industrial engine using unsupervised learning. Acoustic data has been collected, where raw audio recordings have been aligned with the engine data using timestamps. A structured feature extraction pipeline has been developed, where time-domain and spectral characteristics have been extracted from each audio segment. Four unsupervised anomaly detection methods have been implemented and evaluated: K-Means clustering, Local Outlier Factor, Isolation Forest, and an Autoencoder. A load-based modeling method has been used, where each anomaly detection model has been trained separately for each engine load condition. Samples with anomaly scores above the 95th percentile of the score distribution are treated as anomalous. In addition, the original decision boundaries of Isolation Forest and Local Outlier Factor are also evaluated to examine how different thresholding strategies influence the detection results. Findings of the thesis demonstrate that anomaly detection in acoustic signals can be carried out effectively through unsupervised learning methods, without relying on labeled fault data. The developed framework opens several directions for future investigation, including validation on real fault data, applicability to different kinds of industrial machinery, and integration into real-time condition monitoring systems operating under continuously changing conditions.

Keywords: Anomaly Detection, Condition Monitoring, Unsupervised Learning, Acoustic Signals, K-Means Clustering, Local Outlier Factor, Isolation Forest, Autoencoder

Contents

List of Figures

List of Tables

Abbreviations

AI	Artificial Intelligence
AE	Acoustic Emission
CBM	Condition-Based Maintenance
CM	Condition Monitoring
FT	Fourier Transform
HHT	Hilbert–Huang Transform
IoT	Internet of Things
IF	Isolation Forest
LOF	Local Outlier Factor
LSTM	Long Short-Term Memory
MFCC	Mel-Frequency Cepstral Coefficients
ML	Machine Learning
RMS	Root Mean Square
RPM	Revolutions Per Minute
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
TAN	Total Acid Number
TBN	Total Base Number
TF	Time-Frequency
TFR	Time-Frequency Representation
WCSS	Within-Cluster Sum of Squares
WT	Wavelet Transform
ZCR	Zero Crossing Rate

1 Introduction

Industrial machinery and equipment are dependent on continuous and reliable operation, where unexpected failures can lead to downtime and economic consequences (Micheal, 2026). In areas like energy generation and manufacturing, concerns like material degradation and overload can lead to unforeseen failures and downtime. (Ismail, Abdelmoti, Basu, Berini, & Naouss, 2025). As a result, machine breakdown can pose significant challenges in daily operation and can have a considerable impact on productivity (Jalayer, Kaboli, Orsenigo, & Vercellis, 2022).

Condition monitoring is an advanced diagnostic approach used to identify faults in the early stages, before they develop into serious failures (Surucu, Gadsden, & Yawney, 2023). These systems utilize sensor data to capture the dynamic behavior of machines over time (H. Liu, Xia, Williams, Sun, & Hongsheng, 2022). In particular, vibration and acoustic signals have proven to be highly informative, as they reflect the mechanical behavior and underlying physical interactions of machine components, making them effective for fault detection (Lv, Zhao, Zhao, Li, & Ng, 2022).

1.1 Background and Motivation

Traditional fault detection approaches are highly dependent on signal processing techniques and expert-based feature extraction (Goodarzi, Schütze, & Schneider, 2023) (Chaudhary, Thakur, & Joshi, 2025). However, industrial machinery usually works in noisy environments, where external disturbances and environmental noise are often mixed with monitoring signals. This reduces the quality of the collected data and makes fault detection more difficult (Y. Lei et al., 2018). As a result, the accuracy and reliability of traditional approaches often decrease in real industrial applications.

Another key limitation of conventional methods is their dependence on expert knowledge for feature design (Nie, Geng, & Liu, 2026). As machines operate under varying conditions, these approaches become complex and less scalable. In a practical situation, there are other operating conditions like noise variation in operating machines or multiple faults occurring at the same time, making the problem more complex. This often results in the failure to detect the distributional changes and, therefore, the overall diagnostic performance decreases (Nie et al., 2026).

The recent developments of artificial intelligence (AI) address the problems of early fault detection and enable the implementation of preventive measures prior to substantial loss of performance. (Vlachou, Karakatsanis, & Efstathiou, 2025). In particular, anomaly detection techniques have gained attention in condition monitoring and fault detection, enabling automated and efficient analysis using sensor data (M. Huang, 2023). These methods model the distribution of normal operating conditions and identify deviations from this distribution as potential anomalies (Bountzis, Kavallieros, Tsikrika, Vrochidis, & Kompatsiaris, 2025). Techniques such as clustering and density-based methods enable the identification of abnormal behavior because they can identify groups of data points that differ significantly from the majority.(Shu Fuhnwi, Agbaje, Oshinubi, & Peter, 2023).

However, challenges remain in ensuring robustness, interpretation, and consistent performance under varying operating conditions. Motivated by such challenges, this thesis focuses on developing an AI-based anomaly detection architecture for an industrial engine that uses vibration and acoustic signals. The aim of this thesis is to develop and evaluate an AI-based anomaly detection framework for detecting abnormal patterns in engine vibration signals under different load conditions. The study examines how different unsupervised anomaly detection methods behave as engine load changes affect the signal characteristics, and evaluates how reliably these methods can identify anomalies in the presence of noise.

1.2 Problem Statement

Industrial machinery often operates under changing conditions, where factors such as load, speed, and environment affect how the system behaves. Because of this, the signals collected from sensors like vibration and acoustics can vary even during normal operation. This makes it challenging to clearly differentiate between normal and actual faults. Traditional condition monitoring methods depend on manually designed features and fixed thresholds, which may not adapt well to different operating conditions and can lead to unreliable results.

Another challenge is that many anomaly detection approaches assume that the data follow a consistent pattern. In real industrial settings, however, machine behavior changes over time and across different conditions, so this assumption does not always hold. As a result, models trained on mixed data may incorrectly flag normal behavior as abnormal or miss early signs of faults. This challenge becomes complex because labeled fault data is often unavailable, making supervised approaches difficult to apply in real industrial environments.

For these reasons, there is a need for a more flexible and reliable approach that can learn normal machine behavior under different conditions without relying on labeled data. In particular, it is important to explore how unsupervised anomaly detection methods can be applied to vibration and acoustic signals, and how their performance can be evaluated consistently across varying operating conditions. Solving this problem can improve the reliability and practical effectiveness of condition monitoring systems in industrial environments.

1.3 Research Questions

Research Question 1 (RQ1): What anomaly detection approaches reported in the literature are suitable for acoustic based industrial condition monitoring?

Research Question 2 (RQ2): How can anomaly detection methods be developed and evaluated for industrial engine acoustic signals under varying load conditions?

RQ1 focuses on the identification of anomaly detection approaches reported in the literature on industrial condition monitoring to analyze acoustic signals. The purpose of this question is to establish the methodological foundation of the study by selecting unsupervised methods that are suitable for detecting abnormal behavior in industrial engine data under different operating conditions.

RQ2 examines how these anomaly detection methods can be implemented and evaluated for industrial engine acoustic signals collected under different engine load conditions. The objective of this question is to assess how consistently the selected methods can identify abnormal behavior despite variations in signal characteristics caused by changing operating conditions and measurement noise commonly present in industrial environments.

1.4 Contribution of Thesis

This thesis presents a data-driven anomaly detection framework for analysing acoustic and vibration signals collected from an AGCO Power 49 LFTN-D5 turbocharged common rail diesel engine under varying load conditions.The study develops a complete pipeline, starting from raw audio recordings to feature extraction and the use of unsupervised machine learning methods for detecting abnormal behavior. One of the key contributions of this work is the use of a load-based modeling strategy, where different operating conditions are analyzed separately to better represent changes in engine behavior. In addition, multiple anomaly detection techniques are implemented and compared using a consistent percentile-based evaluation approach to ensure fair performance assessment. The study also provides insights into how signal characteristics change under varying operating conditions and how these changes affect detection performance.

1.5 Thesis Structure

This thesis consists of six chapters. Chapter 1 introduces the research problem and the research questions addressed in this study. Chapter 2 presents the literature review related to condition monitoring, signal analysis, machine learning, and anomaly detection. Chapter 3 explains the implementation process, including data collection, preprocessing, feature extraction, and model development. Chapter 4 presents the obtained results, followed by a discussion of the findings in Chapter 5. Finally, Chapter 6 concludes the thesis and discusses the limitations and possible future work.

2 Literature Review

This chapter reviews the theoretical background and previous studies related to condition monitoring and anomaly detection in industrial systems. This study follows a constructive research approach McGregor (2018), focusing on developing a practical solution to a real-world problem and evaluating its applicability in a specific domain. In this thesis, the constructive research approach is reflected in the development of an anomaly detection framework based on acoustic signals for industrial condition monitoring. The literature discussed in this chapter helps identify relevant techniques, examine their limitations, and establish the research context for the proposed framework, particularly under varying operating conditions.

2.1 Condition monitoring system

The Condition monitoring system plays a key role in asset management (AM). They are used to predict potential failures or to enable proactive actions that help prevent them (Attou & Ahmed, 2009). Condition monitoring (CM) is a predictive maintenance approach that continuously observes the health of assets and systems, while also identifying faults and anomalies based on real-time data collected from respective sources (IBM, 2026). Similarly, condition-based maintenance (CBM) is a maintenance approach where decisions are made based on information collected through condition monitoring (CM). In general, it consists of three main stages, as shown in 1.

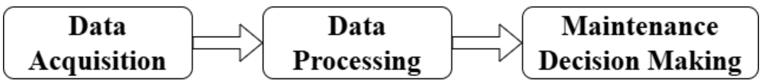

Figure 1. CBM Pipeline, redrawn from (Jardine et al., 2006).

In industrial systems, faults that remain unnoticed and occur unexpectedly can lead to increased maintenance costs, reduced production efficiency, and poor equipment performance L. Lei, Li, Zhang, Wu, and Yu (2025). Traditional maintenance approaches, such as reactive maintenance and preventive maintenance, are not always effective Benhanifia, Cheikh, Oliveira, Valente, and Lima (2025). Reactive maintenance is applied after a failure has occurred, which can result in service disruption and production losses Gholipour, Zare, Vaziri Sarashk, and Gholipour (2025). On the other hand, preventive maintenance is performed at fixed intervals, which may result in unnecessary servicing even when the machine is fully functional Meddaoui, Hain, and Hachmoud (2023). These limitations increase operational costs and reduce the efficiency of the system Benhanifia et al. (2025). In comparison, CM systems continuously analyze the machine's health using real-time sensor data during normal operation, identify potential issues or faults at early stages, and prevent unexpected failures Hassan, Panduru, and Walsh (2024). By observing changes in the signal characteristics, including fluctuations in amplitude, frequency, and statistical behavior, these systems can identify the signs of abnormal operation before they turn into a critical fault, ahead of time (Hassan et al., 2024).

2.1.1 Methods

Condition monitoring techniques are used to monitor the condition and performance of rotating equipment such as reciprocating machines, turbines, electric motors, presses, and internal combustion engines. These techniques are also used for monitoring standby systems to ensure reliable operation. Generally, two main approaches are used in condition monitoring: trend monitoring and condition checking (Nithin et al., 2022).

Trend monitoring It involves continuously measuring and analyzing data to identify patterns that indicate deterioration of the equipment. In this approach, a specific parameter is selected as an indicator of the equipment's health, and its changes are monitored over time to evaluate performance. Observing these trends, it becomes possible to predict when the asset may reach a critical condition or fail. Trend monitoring is commonly used in applications such as tracking engine performance (IBM, 2026).

Condition checking It is carried out by taking periodic measurements using precise indicators on the equipment while it is operating. The data collected during these checks is then analyzed to evaluate the current condition of the machine and identify any potential issues (Nithin et al., 2022).

2.1.2 Applications

Power Systems In recent years, the use of power equipment in the power industry has increased significantly. Because of this, maintaining the safe and reliable operation of such equipment has become an important challenge. To overcome this challenge, online condition monitoring has been widely used to detect faults and support maintenance in power generation, transmission, distribution, and consumption systems. Condition monitoring also offers several advantages, including reducing repair and maintenance costs by identifying potential faults early. It also helps to improve the safety and reliability of the power supply by reducing the chances of serious failures. It also helps limit the extent of damage, reduces repair efforts, and enables better fault diagnosis by identifying the root causes of failures (Guo, Dong, Yang, & Wang, 2019).

Wind turbines Unexpected failures in wind turbines can cause costly repairs and long downtime periods, increasing maintenance costs and the overall cost of energy production. Because of this, early condition monitoring and fault diagnosis are important for preventing serious damage to turbine components. Various CM techniques, including vibration analysis, oil analysis, acoustic emission, ultrasonic testing, etc., are used to detect and diagnose faults effectively in wind turbines (Dao, Staszewski, Barszcz, & Uhl, 2018).

Automobiles Some automobile companies use CBM systems to monitor engine oil quality based on the condition of engine components. These systems help estimate the remaining useful life or mileage of the oil by analyzing parameters such as viscosity, total acid number (TAN), and total base number (TBN). This approach allows oil to be used for its full useful life instead of being replaced at fixed intervals, improving efficiency and reducing unnecessary maintenance (Prajapati, Bechtel, & Ganesan, 2012)

Machinery Industries use various machines, most of which have rotating parts that must operate efficiently to maintain production. To avoid downtime, machines must be maintained in good condition through effective maintenance. Condition monitoring helps to track machine health using sensors and data analysis. This information is used to plan maintenance and prevent failures. Proper maintenance is essential, as it depends on the type of machine and the impact of faults on overall operations (Mohanty, 2014).

Industry 4.0 Industry 4.0 represents the modernization of industrial automation through the use of smart sensors, interconnected systems, and Internet of Things (IoT) technologies. The growing demand for improved efficiency and reliability has also accelerated the development of cyber-physical systems (CPS) and IoT-based technologies. These systems improve efficiency, reliability, and real-time decision-making in modern industries (Shahzad & O'Nils, 2018). To improve machine reliability, CBM is widely used for real-time monitoring of machine health. It helps predict failures in advance, reducing downtime and maintenance costs. This approach also improves machine performance and supports better maintenance decision-making (Cocconcelli, Capelli, Cavalaglio Camargo Molano, & Borghi, 2018).

2.2 Signal Characteristics

Signals extracted from sensors, such as vibration and acoustic measurements, contain useful information about the machine's internal behavior (Alharbi et al., 2023). These signals provide physical data, from which predictive maintenance planning is carried out effectively (Karabacak & Gürsel Özmen, 2022). Under normal circumstances, signals usually follow consistent patterns, where the presence of faults inserts irregularities, such as variation in the signals (Goel, Ghosh, Kumar, & Akula, 2015). By carefully analyzing the signals, we can detect the category or type of fault (Goel et al., 2015).

2.2.1 Vibration Signals

Vibration signals are widely used because they are highly sensitive to physical contact within machine components, making them efficient for fault detection, such as imbalance, misalignment, and bearing defects (Goel et al., 2015). For safe operation of the system, sensor data is continuously extracted and monitored to ensure that captured data can be transmitted to a base station without compromising signal quality (Bhuiyan et al., 2017). These signals can generally be categorized into stationary and non-stationary types.

Stationary Signals Stationary signals are those whose statistical properties remain constant over time, such as periodic vibrations produced by a damaged bearing. These signals are easier to analyze as they do not change their behavior significantly over time. They can be effectively studied using spectral analysis techniques, particularly those based on the Fourier Transform, to identify frequency components and detect faults (Goel et al., 2015).

Non-Stationary Signals Non-stationary signals are temporary in nature and usually last for a short duration compared to the overall observation time. These signals are often generated by sudden events, such as damaged bearing raceways or the development of cracks in a component. Since their characteristics change over time, they require advanced analysis methods (Goel et al., 2015). Machines operating in industries usually function under non-stationary conditions, where the system might experience changes over time due to fluctuations in speed, load, and various external factors, which present challenges for traditional monitoring systems that typically operate in stationary environments (A. Kumar, Wy lomańska, Zimroz, Xiang, & Antoni, 2025). Time-frequency techniques such as the Short-Time Fourier Transform (STFT) and the Hilbert–Huang Transform (HHT) are commonly used to analyze these signals and detect faults effectively (Goel et al., 2015).

2.2.2 Acoustic Signals

Acoustic-based monitoring techniques provide an alternative approach for machine monitoring, since acoustic sensors can be used in different situations without direct contact with the machine. These sensors capture acoustic signals that contain useful information about the machine's operating condition. However, acoustic signals are often affected by environmental noise and changes in machine operating conditions(Jombo & Zhang, 2023). These signals originate from various sources, including dislocation activity, crack information, and fluid leakage. Piezoelectric sensors are usually connected to the preamplifier to capture acoustic emission waves, which are then transformed to electrical waveforms and then assessed using digital signal processing (Olejnik & Desta, 2025).

2.3 Traditional Approaches

Traditional approaches for condition monitoring are based on stable system operation and continuous or periodic measurement of process parameters. When the system operates under normal conditions, its behavior follows consistent patterns that are reflected in the monitored parameters. Any deviation from this normal behavior can be identified through trend analysis or other evaluation methods, which helps to detect potential failures at an early stage. These approaches involve techniques such as vibration analysis, wear debris analysis, thermal imaging, noise monitoring, visual inspection, and environmental pollution monitoring. Each of these methods focuses on the observation of different aspects of the system, enabling effective monitoring and diagnosis within standard operating conditions (Davies, 2012).

2.3.1 Vibration Analysis

Machines consist of multiple moving components that naturally produce sound and vibrations during operation. Each component generates a distinct vibration pattern depending on its condition and design. As the condition of a machine component changes, its vibration signal also changes. These variations in vibration patterns can indicate the early stages of a fault, allowing it to be detected and addressed before a complete failure occurs (S. Kumar, Lokesha, Kumar, & Srinivas, 2018).

Advantages and Disadvantages Vibration signal analysis offers several advantages and some limitations. It allows real-time monitoring of machines, and many well-established signal processing techniques can be used for effective analysis. However, the method also has certain drawbacks, including sensitivity to noise interference and the proper placement of vibration sensors to ensure accurate results (Mohd Ghazali & Rahiman, 2021).

2.3.2 Wear Debris Analysis

Wear debris detection is an important part of oil particle analysis, as it helps detect early signs of abnormal wear and predict possible machine failure, which is especially important for marine and aircraft engines (L. Liu et al., 2019). Significant research has been conducted in this area, which has resulted in many positive outcomes. Offline laboratory testing and online testing are generally used to detect metal particles in oil (Yang, Cao, & Yu, 2023).

Offline Testing Offline testing involves collecting oil samples from the machine and analyzing them in a laboratory environment. Because the samples are analyzed separately after they are collected, this approach is not suitable for continuous real-time monitoring of machine condition. Common techniques used in offline oil analysis include ferrography and spectroscopy(Yang et al., 2023).

Online Testing online detection methods allow continuous and real-time monitoring, and the results directly show the working condition of the equipment (Yang et al., 2023). The methods developed for online testing are inductive sensors, color extraction, optical, and acoustic detection (L. Liu et al., 2019).

2.3.3 Thermal Imaging

Thermal imaging is a non-destructive technique that is commonly used for motor inspection and fault analysis. It is applied in several areas, including energy management, industrial maintenance, and medical sciences. In industrial motor monitoring, thermal imaging helps to detect temperature differences in the motor surface. Such temperature changes indicate faults such as broken rotor bars, shorted coils, insulation defects, and overvoltage conditions. By examining thermal images, both the location and nature of a fault can be identified effectively (Glowacz, 2021).

2.4 Signal Processing Techniques

Traditional signal analysis techniques are usually developed under the assumption that signals remain stationary over time. These methods typically analyze signals either in the time domain or in the frequency domain separately, which limits their ability to capture detailed characteristics from both domains together. As a result, they are less effective for analyzing non-stationary signals that commonly occur in industrial applications (Feng, Liang, & Chu, 2013). The analysis of non-stationary signals has become important, particularly for fault diagnosis in rotating machinery systems. To address this challenge, time–frequency (TF) analysis methods are widely used because they make it possible to analyze signal behavior in both the time and frequency domains at the same time. A well-localized time–frequency representation (TFR) is especially useful for identifying fault-related patterns and signal characteristics (Yu, 2021).

Fourier Transform (FT) and Short-Time Fourier Transform (STFT) are commonly used signal analysis techniques that rely on predefined basis functions to analyze signals. FT is effective for frequency analysis, but it cannot show how signal features change over time. STFT improves this by providing time–frequency analysis using a fixed window, but it has a limited ability to detect impulsive faults. In contrast, Wavelet Transform (WT) offers more flexibility with different basis functions, making it more suitable for extracting fault features (Chen et al., 2016).

However, due to the limitation of the Heisenberg uncertainty principle (Heisenberg, 1927), TF analysis methods, such as the short-time Fourier transform (STFT) and the wavelet transform (WT) often struggle to provide highly concentrated results when analyzing signals that change rapidly over time. (Yu, 2021).

2.5 Machine Learning

Machine learning (ML) refers to the field of study that focuses on developing algorithms and statistical models that allow computers to perform tasks by learning patterns from data rather than relying on specific instructions (Hamishebahar, Li, & Guan, 2021). In industrial applications, machine learning is widely used for anomaly detection, fault prediction, and automatic pattern recognition using both historical and real-time data sources (Kumari, 2026). The focus is to support intelligent condition monitoring and quality control capabilities that are comparable to human decision making (Rossetti, 2018). Modern industries increasingly rely on advanced monitoring and sensing systems that continuously generate large volumes of operational data. These large datasets are then utilized to develop machine learning-based intelligent monitoring and expert systems (Rossetti, 2018).

2.5.1 Machine Learning Approaches

Machine learning (ML) has become increasingly important with the emergence of Industry 4.0, which focuses on smart factories where machines and humans communicate through interconnected systems, sensors, and the internet. To support these advances, industries require intelligent and cost-effective models. ML provides effective techniques to solve complex problems and is generally categorized into supervised and unsupervised learning approaches (Das, Bagci Das, & Birant, 2023).

Supervised learning In supervised learning, the algorithm develops a mathematical model from the data set that includes both inputs and expected outputs(Hamishebahar et al., 2021). These expected outputs are usually provided by experts (Araghizad, 2024).

Supervised machine learning algorithms are commonly divided into two main categories: regression and classification. Regression methods use labeled data to predict continuous trends, such as how a potential fault may develop over time. However, classification methods use labeled data to assign inputs to predefined categories, for example, distinguishing between normal and faulty operating conditions based on system status indicators (Ouadah, Zemmouchi-Ghomari, & Salhi, 2022).

Kuncan (2020) evaluated and compared several machine learning techniques, including k-Nearest Neighbors (KNN), Artificial Neural Networks (ANN), Support Vector Machines (SVM), Linear Discriminant Analysis (LDA), Logistic Regression (LR), and Gray relational analysis (GRA) for fault detection.

Unsupervised Learning The algorithm in unsupervised learning develops a model from the data set, which only has inputs and no expected outputs or previously known outputs (Hamishebahar et al., 2021)(Araghizad, 2024). Among unsupervised approaches, an autoencoder is also one method that is fundamentally a customized neural network architecture, based on an encoder and decoder (L. Huang, Pan, Liu, & Gong, 2023).

When an autoencoder is used for system health monitoring, the monitored data gradually moves away from the normal operating range as the condition of the system degrades.

Greater deviation usually indicates a higher level of degradation. Since the autoencoder is trained using normal operating data, it can reconstruct normal samples more accurately. Therefore, abnormal data or data collected after system deterioration generally produce higher reconstruction errors. Due to this, it is more rational to employ an autoencoder to build a condition indicating system depending on system monitoring data (L. Huang et al., 2023).

2.6 Anomaly Detection Methods

2.6.1 Overview of anomaly detection

Anomaly detection has become an important part of modern condition monitoring systems, especially in industrial applications where early detection of abnormal machine behavior is important for maintaining system reliability and reducing maintenance costs. In practical applications, obtaining sufficient labeled fault data is difficult because failure events occur infrequently and accurate annotation of faults is often challenging. For this reason, unsupervised anomaly detection approaches have gained significant attention, as they learn the normal operating behavior of a system and identify deviations from that behavior as potential anomalies (Chandola, Banerjee, & Kumar, 2009; L. Lei et al., 2025; Pimentel, Clifton, Clifton, & Tarassenko, 2014).

Generally, anomaly detection is a process of identifying unexpected patterns in data (Chandola et al., 2009). In industrial condition monitoring, such anomalies are commonly associated with faults, degradation processes, or abnormal operating conditions. For sensor-based systems that include vibration and acoustic monitoring, anomalies are associated with changes in signal characteristics (i.e, variations in amplitude, frequency distribution, or statistical properties) (Alharbi et al., 2023; Jombo & Zhang, 2023). As such deviations occur before critical failures, anomaly detection can be significant in facilitating predictive maintenance strategies in this context (Jardine et al., 2006; L. Lei et al., 2025).

Building upon the general concept, anomaly detection can be further understood from a modeling perspective as the task of distinguishing normal observations from those that exhibit statistically significant deviations.

The objective of anomaly detection is to assign an anomaly score to each observation. Data points that differ significantly from the normal data distribution are expected to produce higher scores. A threshold can then be applied to these scores to determine whether a sample should be considered normal or anomalous (Pang, Shen, Cao, & Hengel, 2021; Ruff et al., 2021). This general formulation is widely applicable across different anomaly detection approaches, including distance-based, density-based, isolation-based, and reconstruction-based methods.

Anomalies are categorized into three main groups: *point anomalies*, *contextual anomalies*, *collective anomalies*. Point anomalies refer to single observations that differ significantly from the normal behavior of the majority of the dataset. Contextual anomalies, on the other hand, are observations that may be considered anomalous depending on temporal or operational conditions. Collective anomalies represent groups of observations that are anomalous when considered together, despite individual observations in the group not appearing abnormal in isolation (Chandola et al., 2009). In industrial time-series data, such as acoustic or vibration signals, both contextual and collective anomalies are particularly relevant, as abnormal behavior may emerge over a sequence of observations rather than at a single point in time.

In sensor-based monitoring systems, anomaly detection is closely related to changes in signal characteristics. These variations are subtle and are often affected by operating conditions such as load and speed, which complicates the distinction between normal variability and true anomalies (Jombo & Zhang, 2023; A. Kumar et al., 2025). Consequently, effective anomaly detection methods should be able to capture both global and local deviations in the data while remaining robust to such variations.

2.6.2 Categories of methods

Based on the availability of labeled data and the learning paradigm used, anomaly detection methods can be broadly categorized into three main types: *supervised*, *semi-supervised*, and *unsupervised* anomaly detection approaches (Chandola et al., 2009; Pang et al., 2021; Pimentel et al., 2014).

Supervised Anomaly Detection It relies on labeled datasets that include both normal and faulty samples. In this approach, anomaly detection is treated as a classification problem, where the model is trained to differentiate between normal and faulty behavior (Chandola et al., 2009). Commonly used supervised techniques include support vector machines (SVM). When sufficient data are available, these methods can provide high detection performance. However, their applicability in industrial condition monitoring is often limited due to the scarcity of labeled fault data and the difficulty of capturing all possible failure modes in advance (L. Lei et al., 2025; Pimentel et al., 2014). Moreover, as these models are trained on labeled data, they may struggle in generalizing to unseen or novel anomalies, which are common in complex machinery systems.

Semi-supervised Anomaly Detection These methods address this limitation by requiring only labeled data from the normal operating condition. In semi-supervised learning, a model is trained explicitly on normal data to learn its underlying structure, and any important deviation from this learned representation is considered anomalous (Pimentel et al., 2014). Such methods are useful in industrial contexts where normal operating data are abundant and easier to obtain compared to fault data. In particular, one-class SVMs and certain models based on neural networks fall into this category. Although semi-supervised methods are generally more practical than fully supervised approaches, they still assume that the training data are free from anomalies, which is often difficult to guarantee in real-world situations.

Unsupervised Anomaly Detection Unlike supervised and semi-supervised approaches, unsupervised anomaly detection methods function without any labeled data and aim to identify anomalies directly from the inherent structure of the dataset. These methods assume that normal data points occur more frequently and form consistent patterns, while anomalies are rare and deviate from these patterns (Chandola et al., 2009; Pang et al., 2021). Unsupervised approaches are especially suitable for industrial condition monitoring because labeled data are often limited, and machine operating conditions may vary over time. These methods include clustering-based techniques, which group similar observations and identify outliers based on distance measures; density-based approaches, which detect anomalies in low-density regions; isolation-based models, which separate anomalies through recursive partitioning; and reconstruction-based methods, such as autoencoders, which detect anomalies based on reconstruction error (Pang et al., 2021; Ruff et al., 2021).

Each of these approaches captures different aspects of abnormal behavior and has been successfully applied in industrial fault detection tasks. For example, clustering and distance-based methods are effective in identifying global deviations in feature space, whereas density-based methods are more sensitive to local irregularities. Isolation-based methods, such as Isolation Forest, are particularly efficient for high-dimensional data, and reconstruction-based models are well-suited to identify complex nonlinear relationships in sensor signals (Ahmad, Styp-Rekowski, Nedelkoski, & Kao, 2020; F. T. Liu, Ting, & Zhou, 2008a; Ruff et al., 2021). In the context of sensor-based monitoring systems that involve acoustic and vibration signals, unsupervised methods have become popular for their flexibility and ability to operate under limited prior knowledge of fault conditions (Alharbi et al., 2023; Jombo & Zhang, 2023). As a result, unsupervised anomaly detection provides a practical and scalable framework to identify abnormal behavior in such environments.

2.6.3 Anomaly detection approaches

Most unsupervised anomaly detection methods in industrial condition monitoring, particularly for sensor-based data such as acoustic and vibration signals, include distance-based, density-based, isolation-based, and reconstruction-based methods (Chandola et al., 2009; Pang et al., 2021). These approaches differ in how they define normal behavior and measure deviations, making them suitable for detecting anomalies in complex datasets.

Distance-based methods assume that normal data samples are grouped close together in the feature space, while abnormal samples tend to lie farther away from these groups. Among these methods, K-means clustering is one of the most widely used approaches because of its simplicity and computational efficiency. In this method, anomaly scores are usually calculated using the distance between each data point and its assigned cluster centroid.(Chandola et al., 2009; MacQueen, 1967). Such methods have been successfully applied in industrial monitoring tasks, including fault detection in rotating machinery and manufacturing systems, where deviations in feature distributions indicate abnormal behavior Kateris et al. (2014); L. Lei et al. (2025).

Density-based methods, such as the Local Outlier Factor (LOF), detect anomalies by comparing the local density of a data sample with that of its neighboring samples. A sample is considered anomalous when it is located in a region with much lower density than the surrounding points. (Breunig, Kriegel, Ng, & Sander, 2000). Unlike global distance-based methods, density-based approaches are more effective at detecting local anomalies that might otherwise remain unnoticed. LOF has been widely used in sensor-based condition monitoring applications, particularly in vibration and acoustic signal analysis, where small local changes in the feature space can indicate the presence of early-stage faults (Alharbi et al., 2023; Karabacak & Gürsel Özmen, 2022).

Isolation-based anomaly detection methods, particularly the Isolation Forest algorithm, detect anomalies by repeatedly partitioning the data and isolating individual observations. In comparison to other approaches that primarily model normal behavior, Isolation Forest focuses directly on separating anomalous samples(F. T. Liu et al., 2008a). Isolation Forest has demonstrated better performance in high-dimensional and large-scale datasets. In industrial condition monitoring applications, it has been widely used with sensor data to identify abnormal behavior without relying on strong assumptions about the underlying data distribution (A. Kumar et al., 2025; L. Lei et al., 2025). Its computational efficiency and scalability make it particularly suitable for real-time monitoring systems.

Reconstruction-based methods, such as autoencoders, are a powerful class of techniques that use neural networks to learn compact representations of normal operating data. These models are trained to reconstruct the original input data, and anomalies are detected using the reconstruction error, since abnormal patterns are generally more difficult to reconstruct accurately (Ahmad et al., 2020; Ruff et al., 2021). Autoencoder-based approaches have been widely applied in industrial condition monitoring tasks, including vibration and acoustic signal analysis, because they are capable of learning complex non-linear relationships within the data (Ahmad et al., 2020; Chen et al., 2016). Their ability to handle high-dimensional feature spaces makes them well-suited for detecting subtle anomalies in signals that contain complex temporal and spectral patterns.

In practical industrial environments, especially those involving acoustic signals, anomalies can appear in different forms depending on the operating conditions and the type of fault present. Acoustic emissions are influenced by mechanical interactions, load variations, and environmental noise, resulting in complex and non-stationary signal behavior (Jombo & Zhang, 2023; A. Kumar et al., 2025). As a result, no single anomaly detection approach is sufficient to capture all types of abnormal behavior. Distance-based methods are effective in identifying global shifts in feature distributions, density-based methods capture local irregularities, isolation-based approaches provide efficient detection in high-dimensional spaces, and reconstruction-based models enable the modeling of complex signal structures.

2.6.4 Challenges in anomaly detection

Although there have been significant advances in data-driven anomaly detection techniques, their application in practical industrial condition monitoring systems remains challenging. Factors that contribute to these challenges include inherent characteristics of sensor data, the complexity of machine behavior, and the practical limitations of deploying learning-based models in dynamic environments.

Major challenges are the lack of labeled data, particularly for anomalous operating conditions. As failure events are relatively rare and often unpredictable in industrial systems, it is difficult to obtain comprehensive labeled datasets that cover all possible fault scenarios (L. Lei et al., 2025; Pimentel et al., 2014). This limitation makes unsupervised and semi-supervised learning techniques more practical for these applications. However, even in these settings, the assumption that training data is entirely free from anomalies may not always hold, potentially affecting model reliability.

Another major challenge is the non-stationary nature of sensor signals. In practical environments, signals obtained from machines are influenced by changing operating conditions such as load, speed, and environmental factors. Such variations cause shifts in the statistical properties of the data over time, making it difficult for anomaly detection models to distinguish between normal variability and actual faults (Feng et al., 2013; A. Kumar et al., 2025). This issue is particularly relevant in acoustic and vibration signals data, where changes in operating regimes can lead to substantial variations in signal characteristics even under normal conditions. Furthermore, a model developed using one set of operating conditions may not work reliably when applied to new environments or other machines. This is particularly relevant in applications involving acoustic signals, where environmental noise and measurement variability can further complicate the detection process (Alharbi et al., 2023; Jombo & Zhang, 2023). Therefore, ensuring consistent performance across varying conditions remains an open challenge.

The problem of context-dependent anomalies is another relevant problem. It is common for an observation to be considered anomalous in a certain operating condition while being treated as normal in a different condition. For example, changes in engine load can largely alter the amplitude and frequency distribution of acoustic signals. Consequently, anomaly detection methods that do not explicitly consider operating context may generate false positives or fail to detect true anomalies (Chandola et al., 2009; Jombo & Zhang, 2023). This issue highlights the significance of incorporating contextual information or designing models that are robust to such variations.

As modern condition monitoring systems often depend on a large number of extracted features, including statistical, spectral, and cepstral characteristics, the high dimensionality and complexity of feature representations from sensor data are another challenge. Although these features provide rich information about system behavior, they introduce challenges related to the curse of dimensionality, where distance and density measures become less reliable (Pang et al., 2021; Ruff et al., 2021). As a result, the performance of conventional anomaly detection approaches based on similarity measures or local data distribution may be adversely affected.

Most methods generate a continuous anomaly score that later needs to be transformed into a normal or anomalous classification through the use of a threshold value. In practice, selecting an appropriate threshold is non-trivial, especially in the absence of labeled data. Therefore, various thresholding strategies can lead to different detection outcomes, affecting both false positive and false negative rates (Pimentel et al., 2014). This issue becomes more prominent when comparing multiple anomaly detection methods, as each method may produce scores with different distributions and scales.

In summary, anomaly detection in industrial condition monitoring is affected by multiple interconnected challenges, including limited labeled data, non-stationary signals, context-dependent behavior, high-dimensional feature spaces, and threshold selection issues. Addressing these challenges requires careful selection and evaluation of detection methods, particularly in environments characterized by varying operating conditions. These limitations also highlight the need for systematic analysis of anomaly detection techniques under realistic scenarios, which forms the basis for the research gap addressed in this thesis.

2.7 Research Gap

The reviewed studies show that different unsupervised methods for anomaly detection, including approaches based on similarity, local density, isolation mechanisms, and reconstruction techniques, have been applied in industrial monitoring applications using vibration and acoustic signals. Despite these contributions, several practical and methodological limitations in the existing literature remain unaddressed, which form the basis for the research presented in this thesis.

One important limitation identified in the reviewed studies is that many anomaly detection models are designed without properly considering changes in operating conditions, even though acoustic and vibration signals are highly affected by these variations. When measurements collected under multiple operating conditions are merged into one learning framework, the similarity between normal operating behaviors can make it challenging to distinguish actual faults from variations caused by changing conditions. In addition, several unsupervised techniques for anomaly identification have been presented in previous studies; a systematic and fair comparison across these methods under varying operating conditions is unavailable. In the absence of labeled fault data, differences in default decision thresholds across methods make direct comparison unreliable. Despite the growing recognition of acoustic signals as an important source of condition information, a complete unsupervised methodology for industrial monitoring under different varying conditions, structured feature extraction, and condition-specific modeling remains unexplored.

This thesis addresses these limitations by introducing a condition-specific modeling approach in which an independent model is developed for each engine operating condition. In addition, a consistent percentile-based thresholding strategy is applied to provide fair and comparable evaluation across different anomaly detection methods and load conditions.

3 Implementation

3.1 Proposed Methodology

Fault detection is an important part of predictive maintenance, as it is widely used in industrial applications. Detecting faults early can help prevent severe machine failures and reduce operational risks. In general, any behavior that differs from the normal operating condition of a system may indicate a possible fault or anomaly. (Amruthnath & Gupta, 2018). In this thesis project, a comprehensive anomaly detection framework is developed to identify abnormal behavior in industrial engine acoustic and vibration signals under different operating conditions. Since labeled fault data are not available, unsupervised learning methods are used throughout the study. Multiple anomaly detection techniques are evaluated within the same processing pipeline to examine how they respond to changes in engine load conditions.

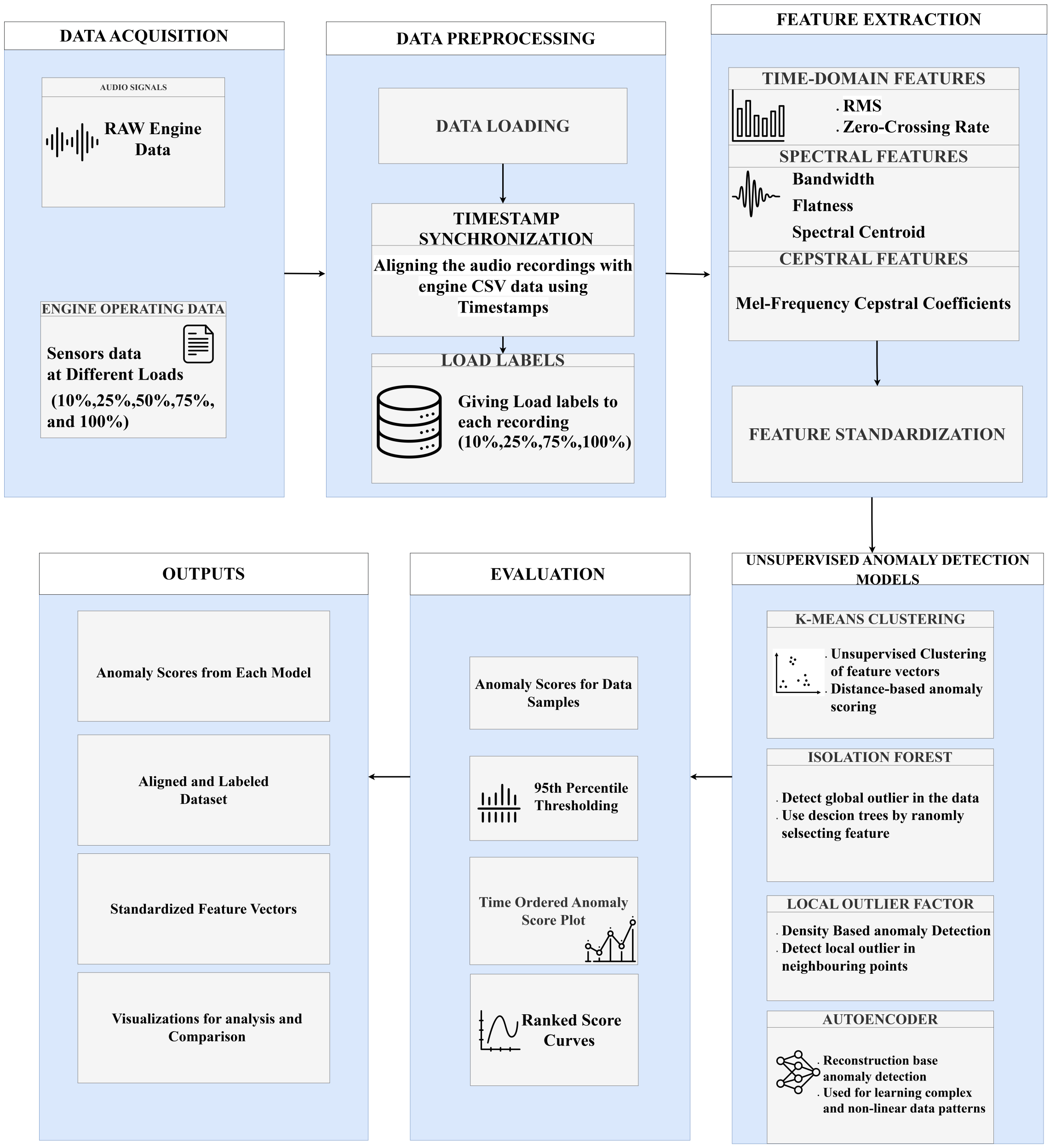

Figure 2. Overview of the proposed anomaly detection framework for an industrial engine.

Figure 2 presents the overall anomaly detection framework used in this study. The framework includes data acquisition, pre-processing, and timestamp synchronization, feature extraction, feature standardization, unsupervised anomaly detection, and evaluation stages. Time-domain, spectral, and cepstral features extracted from the engine acoustic signals are provided as input to K-Means, Isolation Forest, Local Outlier Factor (LOF), and Autoencoder models. The comparison of different unsupervised anomaly detection methods addresses RQ1, while the evaluation of model behavior under different engine load conditions addresses RQ2.

The proposed pipeline is applied separately for the load level (10%, 25%, 50%, 75% and 100%) of each engine. This is important because the acoustic and vibration characteristics of the engine change significantly with different operating loads. Combining data from all load levels into a single model can introduce unnecessary variability and reduce detection performance. By analyzing each load condition separately, the model can learn the normal behavior specific to that load, which improves its ability to detect deviations. As a result, the proposed anomaly detection framework becomes more reliable and suitable for real industrial applications.

3.2 Data Description

Data acquisition was performed using the AVL Indiset 642 indicating system, a widely used instrument for in-cylinder diagnostics. The system measures crank angle with a resolution of 0.1 crankshaft degrees, enabling accurate tracking of the combustion process (Turku University of Applied Sciences, 2024). The acquired data were further analyzed using AVL IndiCom v2.3 software, which provides comprehensive visualization and extraction of key combustion parameters. During data acquisition, pressure data were recorded from four cylinders to analyze overall engine behavior.

Initial data analysis was performed to understand engine behavior under different operating conditions. Figure 3 illustrates the relationship between engine speed, torque, and power. It can be observed that torque remains relatively stable at lower RPM values and gradually decreases as engine speed increases, while power increases with speed, reaching a peak before declining at higher RPM.

Figure 4 shows the variation of engine speed (RPM) and torque over time. The RPM follows a step-like pattern, indicating operation under different load conditions, while

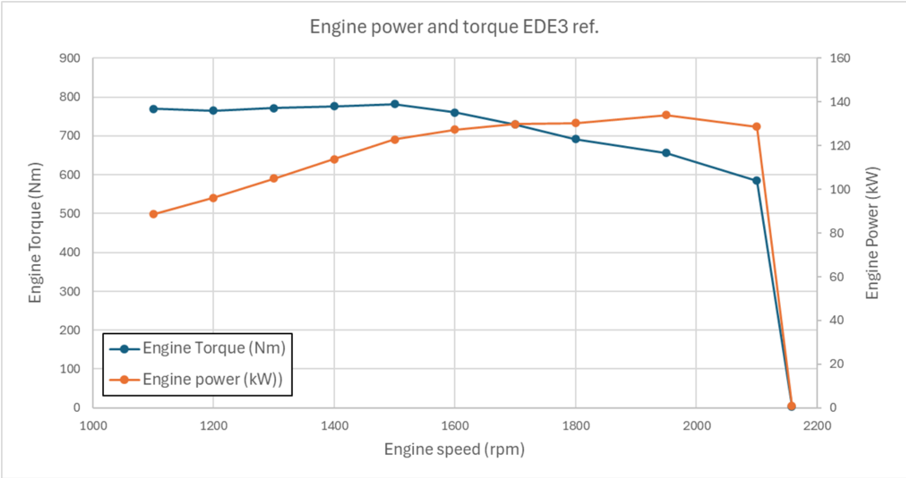

Figure 3. Engine power and torque characteristics as a function of engine speed (RPM) based on the EDE3 reference data.

torque exhibits fluctuations, particularly during transitions between operating states. When RPM stabilizes, torque also becomes stable, whereas sudden changes in speed lead to irregular torque behavior. The variations in the graphs make the data suitable for further analysis using anomaly detection methods.

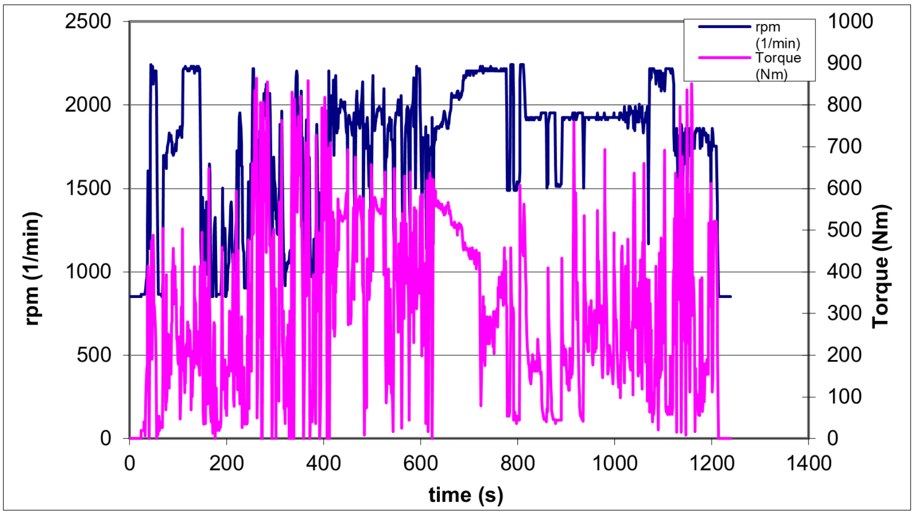

Figure 4. Time-series representation of engine speed (RPM) and torque over time during the experiment.

3.3 Data Pre-processing and Construction

The dataset used in this work includes audio recordings and engine measurement data collected under different engine load conditions. It is composed of three main components: audio files, metadata files, and CSV files, which contain time-series engine parameters. The audio data shows engine sound recordings, where each corresponds to approximately 3 seconds of engine operation. Each audio recording is accompanied by a JSON metadata file that provides additional information, including device specifications and recording parameters. The CSV files contain time-stamped measurements of engine parameters, such as load, RPM, and temperature, for different operating conditions.

Since the audio recordings and engine log data were stored separately without explicit labels to link them, a preprocessing step was required to establish a reliable association between each audio segment and its corresponding engine operating condition. A timestamp-based alignment process was used to achieve this. Each audio file name contains multiple timestamp fields. Among these, the *prefix-ms* timestamp was identified as the most reliable representation of the actual recording time. Timestamps, including start, midpoint, and end times, were evaluated during the preprocessing stage. The *prefix-ms* timestamp was consistent in providing the best alignment with the engine log data, achieving sub-second differences in most cases.

The engine measurement data from the CSV files were first standardized and converted into a consistent timestamp format in milliseconds. Each audio file was then matched with the nearest CSV based on the *prefix-ms* timestamp. The time difference between the audio timestamp and the corresponding CSV timestamp was computed for each match. To ensure reliable alignment, only matches with a time difference less than or equal to 1500ms were retained. This filtering step removed poorly aligned or uncertain matches, which makes the dataset more reliable and consistent. In practice, majority of the valid matches exhibited sub-second alignment, typically within 500ms.

After the matching procedure, each audio segment was assigned the corresponding engine load label obtained from the CSV data. The dataset includes multiple operating conditions, specifically load levels of 10%, 25%, 75%, and 100%, representing different engine regimes. The labeling step is important, as it facilitates the analysis of acoustic behavior under varying operating conditions. The final dataset consists of 2920 audio samples, with reliable time alignment and corresponding load labels. The structured dataset forms the bases for subsequent analysis, where each audio segment is associated with a specific operating condition. Accurate alignment ensures that the extracted features and anomaly detection results represent how the acoustic behavior changes with engine load.

3.4 Feature Extraction

Feature extraction is an important stage of the proposed methodology, where raw engine audio recordings are converted into structured feature representations that can be used for machine learning models. Each recording is processed as a fixed-duration segment of 3 seconds with a sampling frequency of 92 kHz. To maintain consistency across all samples, audio signals longer than the target duration are truncated, while shorter signals are zero-padded to maintain a uniform length. The audio signals are first converted to a mono format by averaging multiple channels when necessary. In addition, all recordings are resampled to 92kHZ so that the entire dataset follows a consistent sampling frequency.

The extracted features capture both the time and frequency related behavior of the engine signals. In industrial systems, anomalies commonly appear as variations in energy patterns, frequency components, or signal behavior. Therefore, the analysis uses a combination of time-domain spectral and cepstral features to capture these changes.

Root Mean Square (RMS) Energy

RMS is a time-domain feature that indicates the overall energy present in a signal. (Zhu, Nostrand, Spiegel, & Morton, 2014)(Caesarendra & Tjahjowidodo, 2017):

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x[n]^2} \tag{1}$$

RMS reflects the intensity of engine vibrations and is highly sensitive to abnormal operating conditions. Changes in RMS values often indicate variations in mechanical behavior or combustion irregularities.

Zero Crossing Rate (ZCR)

ZCR indicates how often the signal crosses the zero-amplitude axis (Wan, Lin, & Gao, 2024):

$$\mathrm{ZCR} = \frac{1}{2N}\sum_{n=1}^{N-1} |\mathrm{sgn}(I(n+1)) - \mathrm{sgn}(I(n))| \tag{2}$$

The sign function is given by:

$$\mathrm{sgn}(I(n)) = \begin{cases} 1, & \text{if } I(n) > 0 \\ 0, & \text{if } I(n) = 0 \\ -1, & \text{if } I(n) < 0 \end{cases} \tag{3}$$

where N indicates the total number of samples, while $I(n)$ represents the signal amplitude at sample n.

Mel-Frequency Cepstral Coefficients (MFCCs)

MFCC describes the overall spectral characteristics of a signal and represents its frequency content in a compact form (Prawin & Anbarasan, 2021):

$$c_{\text{frame}}(m) = \frac{2}{N_f} \sum_{i=0}^{N_c - 1} \log\left[\bar{X}(i)\right] \cos\left(\frac{\pi}{N_c}\left(m + \frac{1}{2}\right) i\right), \quad 0 \le m \le N_c - 1 \tag{4}$$

where $\bar{X}(i)$ represents the Mel-spectrum, $c_{\text{frame}}(m)$ denotes the MFCC of a selected frame, and m and N_c represent the coefficient index and the total number of MFCC coefficients, respectively.

Spectral Features

Additional spectral features are extracted to describe the signal energy distributed in the frequency domain. These features include bandwidth, spectral centroid, roll-off, and flatness, which characterize the center of mass, spread, high-frequency energy, and noisiness of the signal, respectively. For each feature, statistical values including mean and standard deviation are calculated across entire audio segment. The computed statistics are then concatenated into a fixed-size feature vector representing each audio sample. The final feature representation captures both the average behavior and variations within the signal, making it suitable for anomaly detection tasks.

Feature Standardization After feature extraction, the features exhibit different numerical scales and distributions. This can negatively affect model performance, particularly for distance-based methods such as K-Means clustering.

To address this, feature standardization is applied using z-score normalization (Al-Faiz, Ibrahim, & Hadi, 2019):

$$z(ij) = \frac{a(ij) - \mu}{\alpha} \tag{5}$$

This transformation standardizes the features so that each feature follows the same scale, with a mean of zero and a variance of one, allowing each feature to contribute equally during model training. In this implementation, standardization is performed using the *StandardScaler* from the Scikit-learn library.

The standardization process is applied separately for each load condition to maintain consistency within each operating regime. This step is essential for ensuring stable and reliable anomaly detection across varying industrial conditions.

3.5 Models

All models operate on the same set of extracted features. To capture variations in engine behavior under different operating conditions, the models are applied at each load level. The details of these methods are described as follows:

3.5.1 K-Means Clustering

K-Means is a widely used clustering method introduced by J. MacQueen (MacQueen, 1967). It is commonly applied in unsupervised learning, pattern recognition, and data clustering tasks. The method separates the dataset into K clusters based on similarities among features. It works by reducing the distance between data samples and the centroid of their assigned cluster, which helps group similar observations together effectively (Lima et al., 2010).

The K-means algorithm groups the data into k clusters by reducing square-error criterion, commonly referred as the within-cluster sum of squares (WCSS), which is expressed as:

$$E = \sum_{i=1}^{k} \sum_{p \in C_i} \|p - m_i\|^2 \tag{6}$$

where p represents an individual data sample, C_i denotes to the i^{th} cluster, and m_i corresponds to the centroid of that cluster. Equation 6 calculates the overall squared distance between the data samples and their assigned centroids. Reducing this value helps produce clusters that are more compact and clearly separated from each other.(Chauhan & Shukla, 2015).

In this thesis, K-means clustering is used for features extracted from engine audio signals to detect abnormal behavior under normal operating conditions. Since engine sound changes with different load levels, clustering is performed separately for each condition so that the model can learn specific patterns. The model uses K=3 clusters, where each data sample is assigned to the closest centroid based on Euclidean distance (Danielsson, 1980). This distance is used as the anomaly score, where higher values represent stronger deviation from normal operating behavior. To identify anomalous samples, a threshold is selected using the 95th percentile of the obtained anomaly score.

3.5.2 Local Outlier Factor (LOF)

The Local Outlier Factor (LOF) is a density-based anomaly detection method proposed by Breunig et al. (Breunig et al., 2000). It detects outliers by evaluating how isolated a data sample is relative to its neighboring samples. This method determines how strongly a sample deviates from the density of its local neighborhood. The calculation of LOF values follows two main stages. First, the nearest neighbors of each sample are identified using a selected MinPts range. Next, the LOF score compares the local density of a sample with the densities of the surrounding neighboring samples. Higher LOF score indicates that a sample is more isolated from nearby data points and is therefore more likely to be considered an outlier(Breunig et al., 2000). The LOF score for a data sample p is calculated using Equation 7:

$$LOF_{MinPts}(p) = \frac{\sum_{o \in N_{MinPts}(p)} \frac{lrd_{MinPts}(o)}{lrd_{MinPts}(p)}}{|N_{MinPts}(p)|} \tag{7}$$

Equation 7 calculates the average relationship between the local density of data sample p and the densities of its neighboring samples determined by MinPts parameter . Based on this score, each data point is evaluated to determine whether it is an outlier or a normal instance (Alghushairy, Alsini, Soule, & Ma, 2021).

In this study, feature representations obtained from engine audio signals are evaluated using the Local Outlier Factor (LOF) approach to identify abnormal system behavior. Since the statistical characteristics of the signals change under different engine load conditions, the LOF-based analysis is carried out separately for each load level. The neighborhood size is selected adaptively according to the amount of data available for each load condition, allowing for more stable local density estimation. The LOF algorithm generates anomaly scores by measuring local density deviations, where larger scores indicate more abnormal observations. For anomaly identification, both the built-in LOF classification and a percentile-based thresholding method are employed to ensure reliable comparison with other detection techniques.

3.5.3 Isolation Forest

Isolation Forest is an unsupervised anomaly detection method introduced by Liu et al. that detects anomalous samples by isolating data points using a group of randomly generated decision trees, called isolation trees. Data samples that can be separated with fewer splits are considered more likely to be anomalies because they differ significantly from the majority of the data. The Isolation Forest approach is computationally efficient and mainly depends on two key parameters: the number of isolation trees and sub-sampling size used during model construction (F. T. Liu, Ting, & Zhou, 2008b). The average path length for a dataset with n samples is calculated as:

$$c(n) = 2H(n-1) - \frac{2(n-1)}{n} \tag{8}$$

where H represents the harmonic number, which can be approximated as:

$$H \approx \ln + 0.5772156649 \tag{9}$$

Using the normalized path length, the anomaly score corresponding to data sample d can be calculated as:

$$S(d, n) = 2^{-\frac{E(h(d))}{c(n)}} \tag{10}$$

where $E(h(d))$ represents the average path length of the instance in all isolation trees. Higher anomaly scores indicate a higher probability of anomalous behavior (Xu, Wang, Meng, & Zhang, 2017).

In this thesis, Isolation Forest method is used on standardized feature vectors obtained from engine audio signals for anomaly detection across different engine load conditions. Separate models are trained for each operating state to capture variations in the data distribution, while z-score normalization (Al-Faiz et al., 2019) is performed prior to model fitting. The algorithm identifies anomalies based on isolation path lengths, where observations with higher anomaly scores are considered abnormal. To ensure reliable detection and fair comparison with other methods, both the default Isolation Forest prediction and a percentile-based thresholding approach are utilized.

3.5.4 Autoencoder

The Autoencoder is an unsupervised neural network–based learning method introduced by Geoffrey Hinton and Ruslan Salakhutdinov in 2006. It mainly consists of two components: an encoder and a decoder. The encoder converts the input data into a compact lower-dimensional representation, whereas the decoder attempts to reconstruct the original input from this compressed representation. During the training process, model learns by minimizing the reconstruction error, and samples with large reconstruction errors are considered potential anomalies (Hinton & Salakhutdinov, 2006). In this architecture, both the encoder and the decoder are implemented using Long Short-Term Memory

(LSTM) networks to learn temporal relationships present in the data. The encoder can be expressed as:

$$h_t = \phi_e(x_t, h_{t-1}) \tag{11}$$

x_t corresponds to the input sequence at time step t, h_{t-1} refers to the hidden state from the previous step ϕ_e represents the encoder mapping function. The decoder subsequently rebuilds the original sequence using the compressed encoded representation as follows:

$$h'_t = \phi_d(h_t, h'_{t-1}), \qquad x'_t = \sigma(h'_t) \tag{12}$$

where ϕ_d represents the decoder function and x'_t is the reconstructed output. The reconstruction error is calculated using the mean squared error (MSE):

$$RE = \|x_t - x'_t\|^2 \tag{13}$$

Higher reconstruction error values indicate anomalies (Ahmad et al., 2020).

In this work, the Autoencoder approach is used on standardized feature vectors extracted from engine audio signals for anomaly detection under different load conditions. Separate models are trained for each operating state to capture load-specific data characteristics. The network is composed of fully connected encoding and decoding layers that learn a compact latent representation of input data, while the training process is carried out using the Adam optimization algorithm. After the model has been trained, the reconstruction error for each sample is determined and used as an anomaly score, where higher reconstruction errors indicate abnormal behavior. A percentile-based thresholding approach is then used to identify anomalous observations consistently across all load conditions.

3.6 Load-based modeling

The models in this study are applied separately for each load condition. The design choice is motivated by the observation that the statistical properties of acoustic signals vary with engine load. Changes in operating conditions affect signal energy, frequency distribution, and temporal behavior, resulting in different feature distributions for each load level. Training a single model on combined data could introduce unnecessary variability and reduce detection sensitivity. By modeling each load condition independently, the methods can learn load-specific patterns of normal behavior, leading to more reliable and consistent anomaly detection.

3.7 Evaluation

Since the dataset does not contain ground truth labels for anomalous conditions, traditional supervised evaluation metrics like accuracy or recall cannot be applied. Therefore, evaluation is based on the analysis of anomaly score distributions, detection behavior, and consistency between different models and load conditions.

All models generate a continuous anomaly score for each audio segment. To enable consistent comparison, a unified percentile-based thresholding strategy is adopted as the primary evaluation method. Specifically, the 95th percentile of the anomaly score distribution is calculated separately for every load condition, and samples exceeding this threshold are classified as anomalies. This ensures a consistent anomaly ratio between models and allows for fair comparison independent of their internal scoring mechanisms.

Moreover, model-specific decision rules are also evaluated. Isolation Forest and Local Outlier Factor provide native anomaly predictions, which are analyzed alongside the percentile-based results. This enables the assessment of how different thresholding strategies influence anomaly detection behavior.

Ranked anomaly score plots are used to analyze the distribution of anomaly scores. These plots present scores in ascending order, allowing a clear visualization of how anomalies separate from normal samples. A sharp increase at the higher ends of the distribution curve indicates a strong deviation, while smoother transitions suggest weaker separation. These plots provide insight into the sensitivity of each model to extreme values. Furthermore, the temporal behavior of anomaly scores is examined using time-ordered plots. These plots show how anomaly scores evolve across sequential audio segments, allowing identification of temporal patterns and clusters of anomalous behavior. Moving averages are included to highlight underlying trends, while detected anomalies are marked to show their distribution over time. This analysis is particularly useful for understanding whether anomalies occur as isolated events or as sustained deviations.

For quantitative evaluation, anomaly detection ratios are analyzed under different load conditions. This comparison highlights how each method responds to changes in operating regimes. The results show that different models produce varying anomaly ratios under native thresholding, while the percentile-based approach maintains a consistent detection rate. Further insight is obtained through statistical analysis of anomaly scores. Summary statistics, including mean, standard deviation, and range, are calculated for each model and load condition. These statistics provide information about the spread and variability of anomaly scores.

Research Question 1 focuses on identifying suitable anomaly detection approaches for acoustic signals in industrial condition monitoring. The application of multiple methods, including clustering-based, density-based, isolation-based, and reconstruction-based approaches, demonstrates that a diverse set of techniques can be effectively applied to this problem. Research Question 2 focuses on how these methods can be developed and evaluated under varying operating conditions. By applying all models separately across load levels and using a unified evaluation framework based on percentile thresholding and comparative analysis, this study provides a systematic approach to modeling machine behavior under different operating regimes.

4 Results

This section presents the findings of the anomaly detection methods performed using engine acoustic data collected under different load conditions. The purpose is to study how changing operating conditions affect acoustic signals and to evaluate how effectively different anomaly detection approaches can identify abnormal behavior.

Results are presented in four sections. First, raw acoustic signals are analyzed to observe how different engine load conditions effect the signal characteristics. Second, the extracted features are analyzed to see how these changes appear in the feature space. Third, anomaly detection outcomes from different models are discussed using anomaly scores and detection performance. Finally, the important findings are summarized in relation to the objectives of the study.

4.1 Signal and Data Behavior

To develop a basic understanding of engine behavior, acoustic signals are initially analyzed in the time domain under different load conditions, as shown in figure 5.

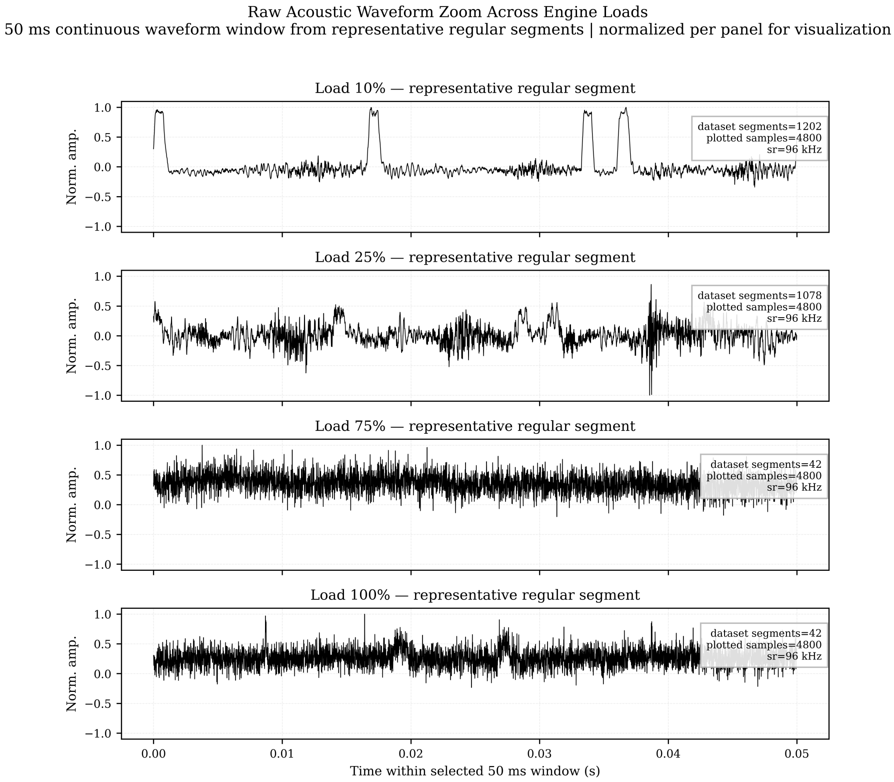

Figure 5. Raw waveform comparison across loads.

The acoustic signals recorded under lower load conditions, particularly at 10% and 25%, show uneven temporal behavior with sudden spikes in amplitude and less consistent activity, suggesting higher variability in engine operation. As the load increases to 75% and 100%, the signals become smoother and more regular, with continuous oscillations and stable amplitude patterns that indicate steady mechanical performance. In order to further examine these differences, the variation in autoencoder reconstruction error is analyzed across all load conditions, as shown in Table 1.

Table 1. Autoencoder reconstruction error statistics across different engine load conditions.

Load (%)	Samples	Mean Score	Std. Dev.	Min–Max
10	1202	0.086	0.023	0.039–0.174
25	1078	0.094	0.024	0.035–0.201
75	42	0.084	0.016	0.049–0.111
100	42	0.084	0.019	0.050–0.130

The findings show that the acoustic signals exhibit a higher variation in lower engine loads, with standard deviation values close to 0.023 and 0.024 at load conditions of 10% and 25%. In comparison, higher load conditions produce more stable signals, as reflected by the reduced standard deviation of nearly 0.016 at a load of 75%. This reduction in variability indicates that the engine operation becomes more consistent as the load increases. Overall, the results confirm that changes in engine load strongly affect the behavior of the acoustic signal, which also affects the performance of the anomaly detection models.

4.2 Analysis of Feature Space

After analyzing the acoustic signals, the extracted feature space is studied to understand how changes in signal behavior are represented in the statistical features used by the anomaly detection models.

At lower load conditions, the extracted features show greater variability, leading to a more scattered feature space. This reflects the unstable and irregular behavior observed in the time-domain analysis, where normal operating patterns are less clearly defined, making anomaly identification more difficult. However, higher load conditions result in a denser and more organized feature space, indicating a more stable and consistent engine behavior. To further validate these findings, the standard deviation of anomaly scores is analyzed across different models and load conditions, as shown in Table 2.

Table 2. Standard deviation of anomaly scores across different models and load conditions.

Load (%)	Autoencoder	Isolation Forest	K-Means	LOF
10	0.023	0.025	1.506	0.049
25	0.024	0.035	2.590	0.069
75	0.016	0.013	0.949	0.014
100	0.019	0.014	0.779	0.021

The results demonstrate that the variability in the feature space decreases as the engine load increases. Methods such as Isolation Forest and Autoencoder produce higher standard deviation values at lower loads and smaller values at higher loads, indicating improved stability in the extracted features. K-Means and LOF also follow a similar pattern, although their score ranges differ because each model uses a different anomaly scoring method. These results suggest that the feature space becomes more compact and organized with higher loads, while lower loads produce a more dispersed structure. Consequently, anomaly detection becomes more effective in higher load conditions, where the distinction between normal and abnormal samples is clearer.

4.3 Anomaly Score Distribution and Detection Behavior

This subsection describes the anomaly detection results obtained from each model using ranked anomaly score distributions and time-sequence visualizations. The results are presented through quadrant graphs, with each quadrant representing a load condition (10%, 25%, 75%, and 100%), enabling clear comparison of model behavior across different operating conditions. In addition, time-ordered anomaly score plots are shown in Figures 6, 7, 8, and 9 to illustrate the variation of anomaly scores in consecutive audio segments for each model.

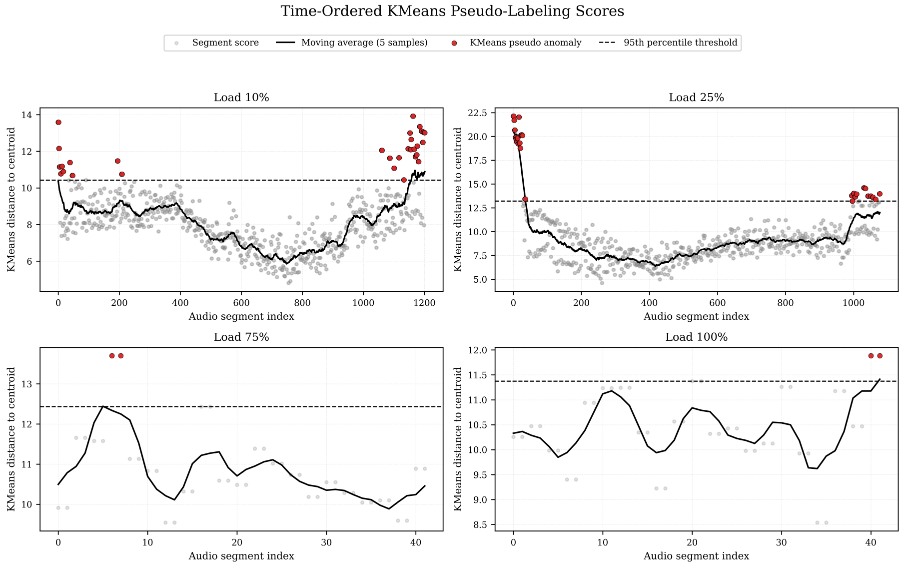

Figure 6. Time ordered anomaly score plots for K-Means across loads.

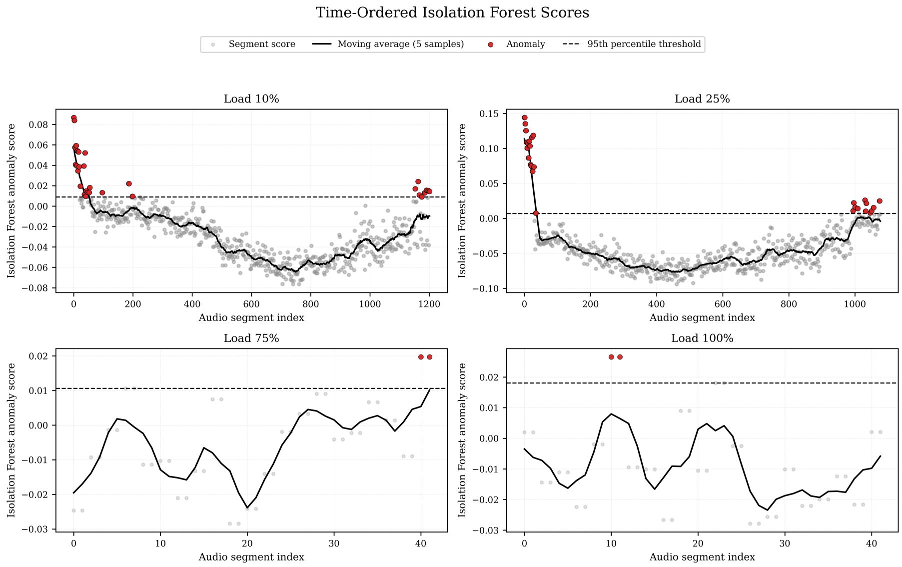

Figure 7. Time ordered anomaly score plots for Isolation Forest across loads.

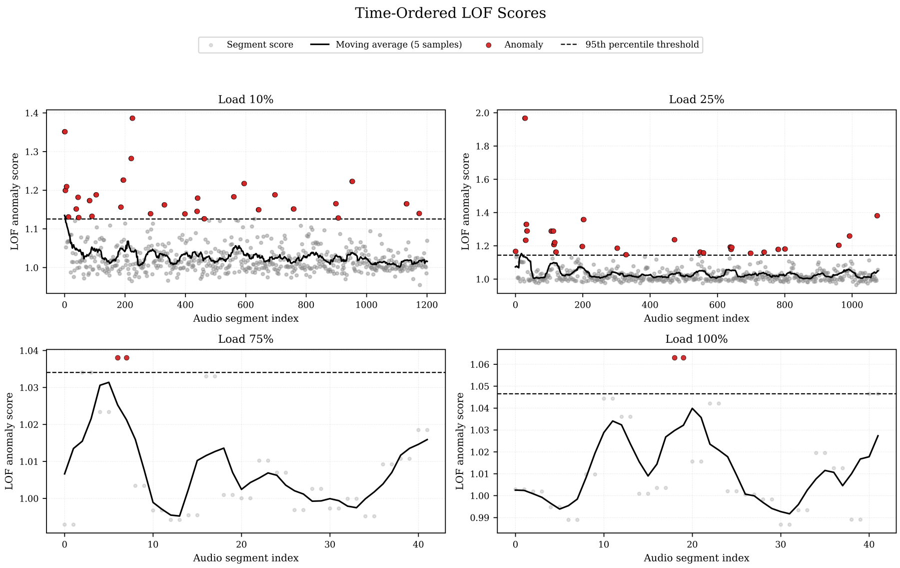

Figure 8. Time ordered anomaly score plots for LOF across loads.

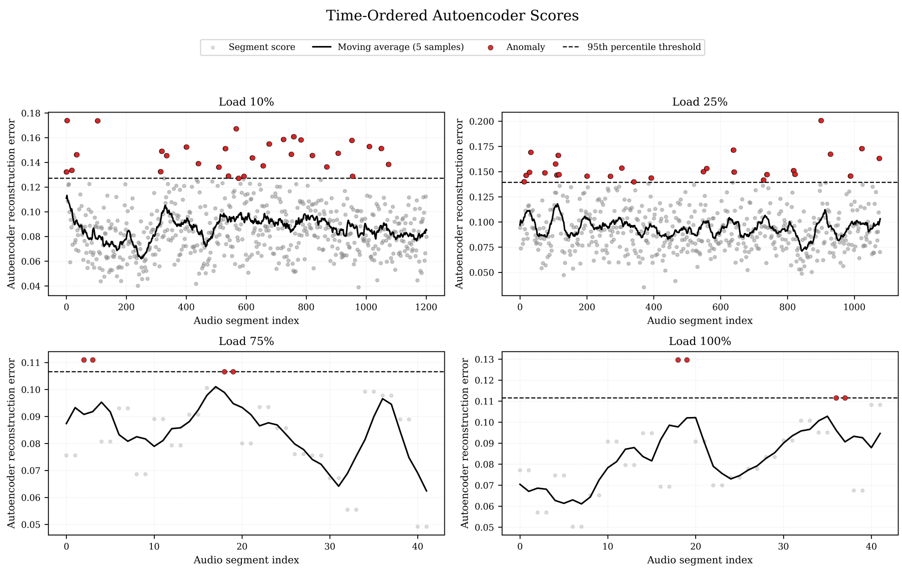

Figure 9. Time ordered anomaly score plots for Autoencoder across loads.

At lower load conditions (10% and 25%), all models show higher fluctuations in anomaly scores, with anomalies appearing as scattered peaks due to irregular acoustic behavior. At higher loads (75% and 100%), the anomaly scores become smoother and more stable, with clearer localized deviations. K-Means produces gradual score changes due to its distance-based approach, while Isolation Forest generates sharper peaks due to its sensitivity to data separation. LOF shows smaller variations in many cases, whereas the autoencoder captures anomalies through smoother reconstruction error transitions.

Ranked anomaly score plots provide further insight into score distribution and separation as shown in Figures 10, 11, 12, and 13.

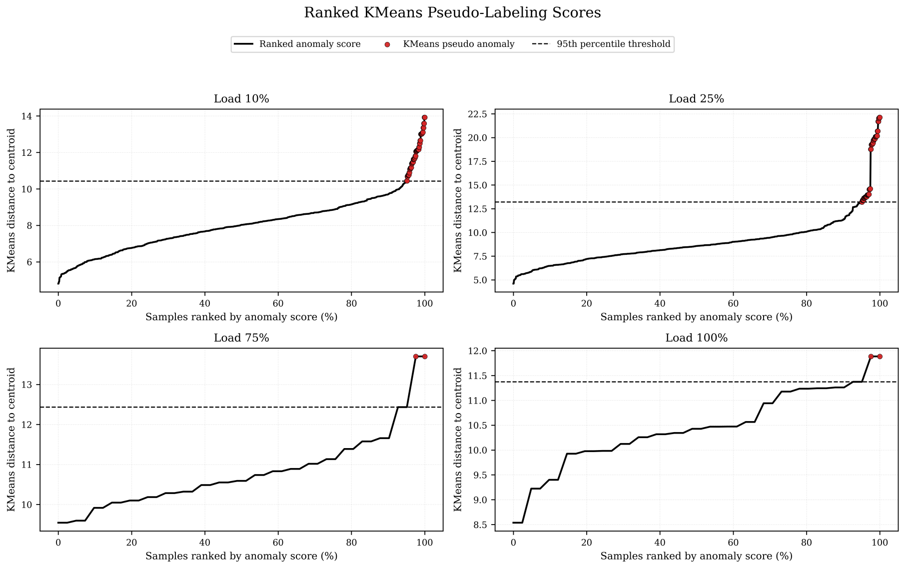

Figure 10. Ranked anomaly score distributions for K-Means across loads.

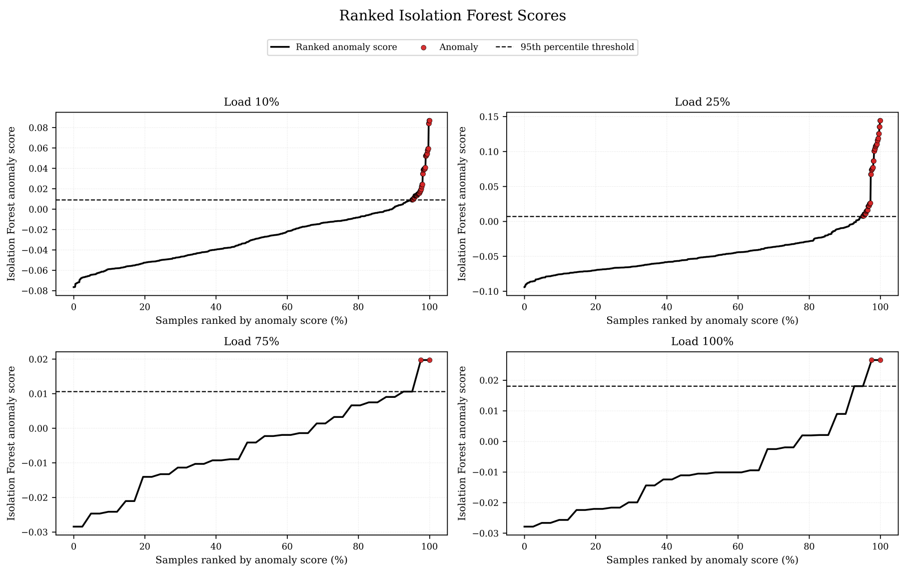

Figure 11. Ranked anomaly score distributions for Isolation Forest across loads.

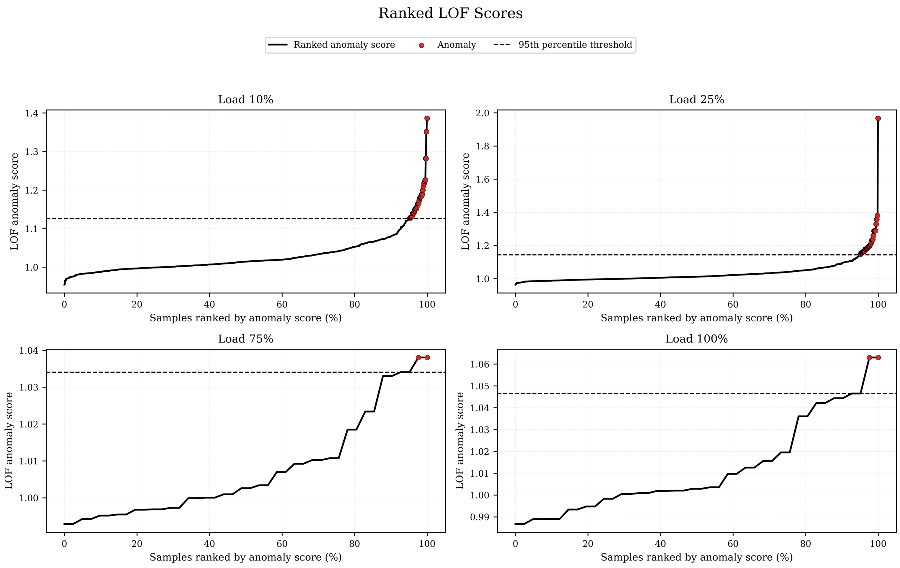

Figure 12. Ranked anomaly score distributions for LOF across loads.

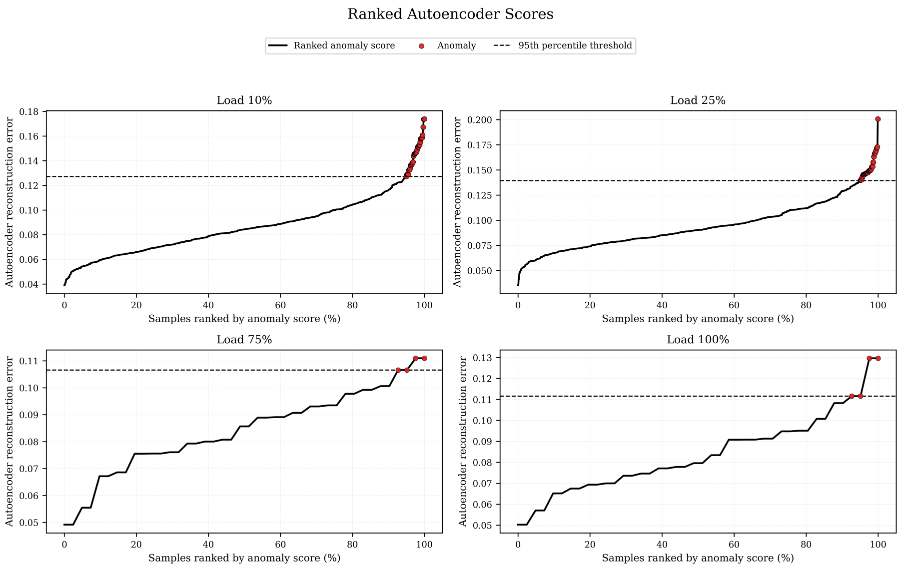

Figure 13. Ranked anomaly score distributions for Autoencoder across loads.

The ranked anomaly score plots indicate that anomalous samples mostly appear at the higher end of the score distribution. K-Means displays a gradual increase in scores, while Isolation Forest shows a sharper rise, indicating a stronger separation between normal and anomalous samples. LOF exhibits limited score variation with weaker separation, whereas the autoencoder produces a smooth but clear increase in reconstruction error toward higher anomaly scores.

In addition to the percentile-based evaluation, the native decision of Isolation Forest and Local Outlier Factor analyzes the influence of different thresholding strategies on detection behavior. The results are presented through quadrant graphs for each model, with each quadrant representing a load condition (10%, 25%, 75%, and 100%). Time-ordered anomaly score plots at the native threshold are shown in Figures 15 and 16 to illustrate the variation of anomaly scores across consecutive audio segments under the native decision boundary

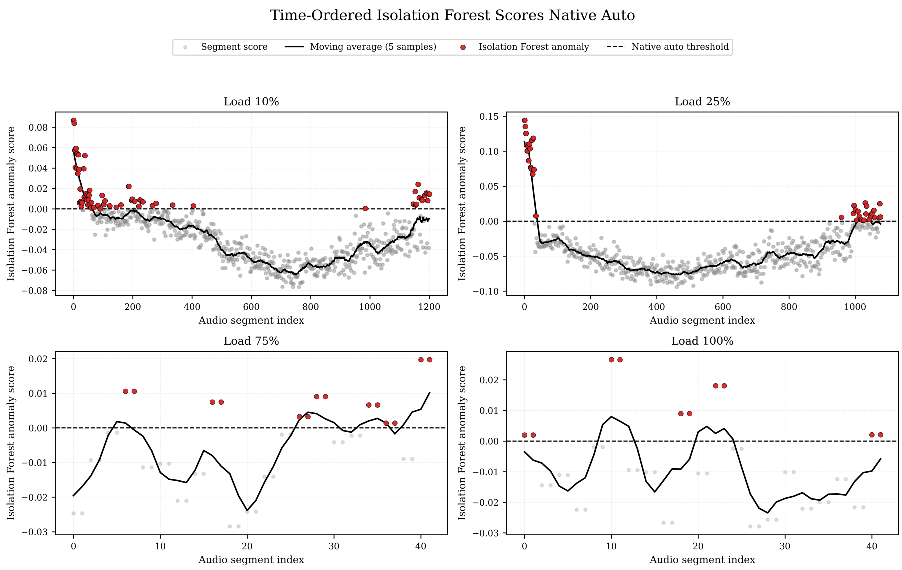

Figure 14. Time ordered anomaly score plots at auto native threshold for Isolation Forest across loads.

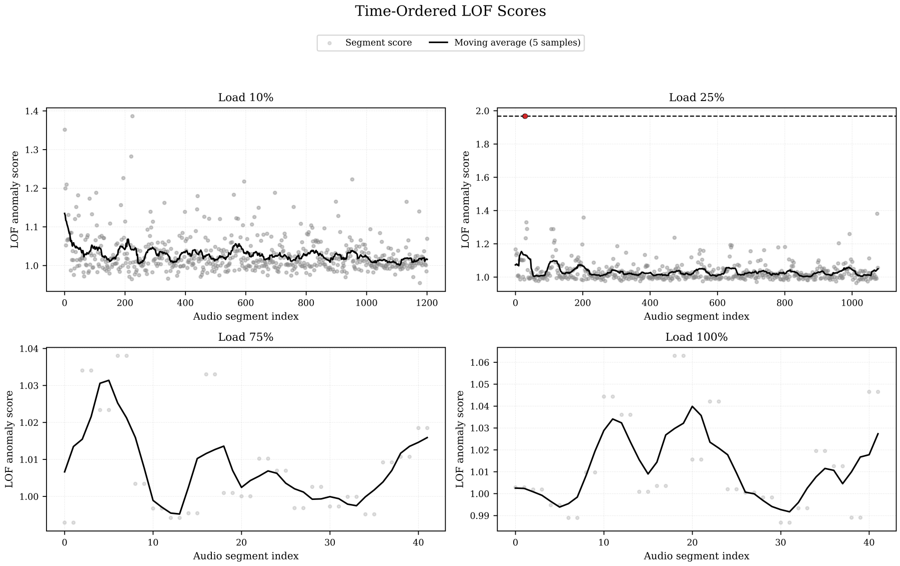

Figure 15. Time ordered anomaly score plots at auto native threshold for LOF across loads.

Under the native threshold, Isolation Forest identifies a higher number of anomalies at lower load conditions (10% and 25%), where a large proportion of samples exceed the native decision boundary due to the irregular nature of signals at these operating regimes. At higher load conditions (75% and 100%), the number of detected anomalies decreases significantly, with only a small number of isolated deviations identified above the threshold. Whereas Local Outlier Factor under the native threshold produces a different detection pattern. At lower load conditions, the LOF scores remain relatively stable with limited separation between normal and anomalous samples. At higher load conditions, the LOF scores show smooth variations with small deviations above the native threshold.

Ranked anomaly score plots at the native threshold provide further insight into score distribution under the native decision mechanism, as shown in Figures 16 and 17

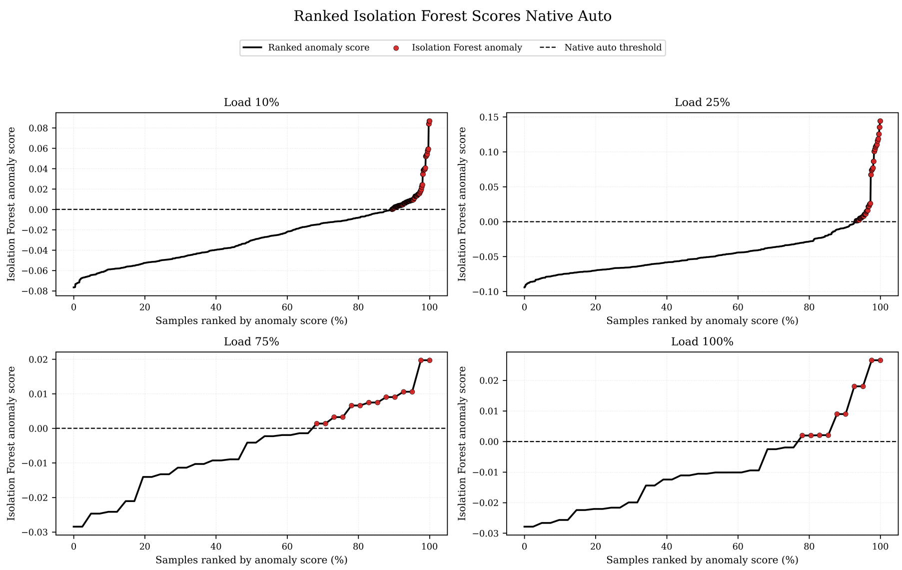

Figure 16. Ranked anomaly score distributions at auto native threshold for Isolation Forest across loads.

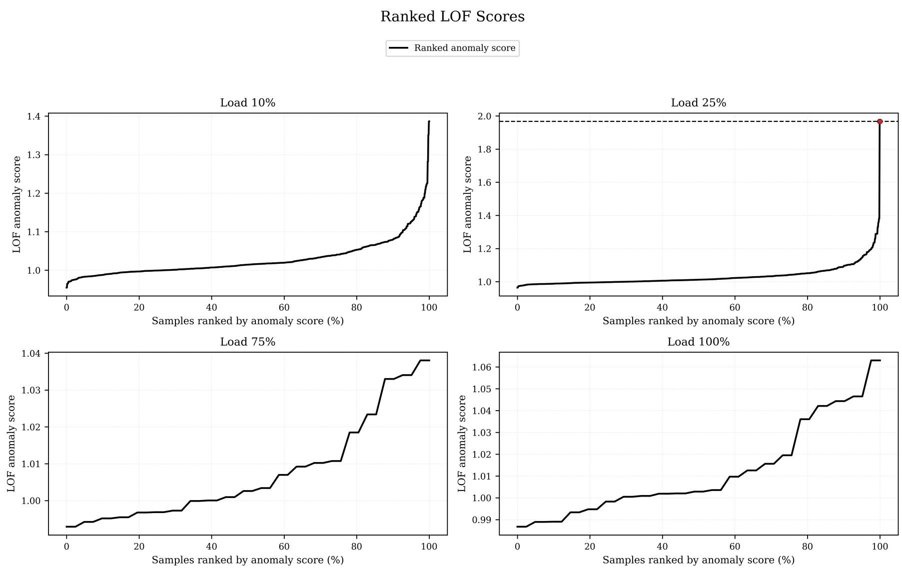

Figure 17. Ranked anomaly score distributions at auto native threshold for LOF across loads.

The ranked score distributions under the native threshold highlight detection characteristics between the two models. For Isolation Forest, an upward curve is observed toward the upper end of the score distribution across all load conditions, where detected anomalies are concentrated beyond the native decision boundary. The degree of separation between normal and anomalous samples is particularly evident at lower load conditions, where the native threshold captures a larger proportion of flagged samples. At higher load conditions, the overall score range is limited, and fewer samples are identified as anomalous. While LOF score distribution follows a uniform behavior across all load conditions. The native LOF boundary identifies only a small number of abnormal samples under low load conditions, as the overall score spread remains narrow.

The effect of thresholding is analyzed through anomaly detection ratios as shown in Figure 18.

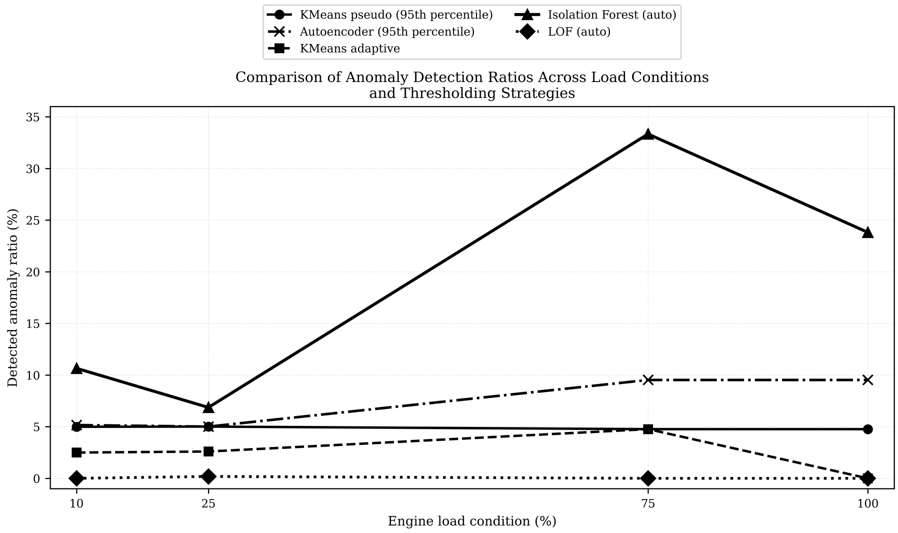

Figure 18. Comparison of anomaly detection ratios across models and loads.

The results indicate that percentile-based thresholding provides a more balanced and consistent anomaly ratio across all load conditions, making model comparison more reliable. In comparison, native thresholding causes noticeable differences in anomaly detection behavior among the models. The Isolation Forest detects a higher number of anomalies at intermediate loads, while the LOF identifies very few anomalies under most conditions. Overall, the analysis indicates that the choice of model and operating load significantly influences the anomaly detection performance, with higher loads producing more stable and clearly separable patterns.

4.4 Summary of Results

The results of this chapter show that engine load plays an important role in influencing acoustic signal characteristics and anomaly detection outcomes. Lower load conditions produce more variable signals, resulting in a scattered feature space and less stable anomaly score distributions, whereas higher loads lead to more structured signals, compact feature representations, and smoother anomaly patterns.

The evaluated models respond differently to abnormal behavior, as clustering methods capture global deviations, density-based approaches identify local irregularities, isolation-based models focus on data separability, and autoencoders detect deviations through reconstruction error. These observations suggest that each method captures different characteristics of anomalies, making multiple approaches valuable for analysis. Furthermore, the comparison of thresholding methods indicates that percentile-based thresholding provides a more consistent framework for model comparison, while native thresholding produces varying detection behavior due to differences in score distributions across models and load conditions.

These results directly contribute to Research Question 2 by illustrating the effectiveness of anomaly detection methods under varying engine load conditions. By analyzing each load condition separately and applying a consistent evaluation framework, the study establishes a systematic method to understand machine behavior across multiple operating regimes. Furthermore, the results also support Research Question 1 by demonstrating that different approaches for identifying anomalies, including methods including clustering-based, density-based, isolation-based, and reconstruction-based approaches, can be successfully applied to acoustic data for industrial condition monitoring.

5 Discussion

The findings of this study highlight an important challenge in industrial acoustic anomaly detection that is often simplified in the existing literature, which is the assumption that machine behavior can be represented using a single stable distribution across all operating conditions. According to the results of this study, this assumption is not well suited for industrial acoustic signals, where the operating load substantially changes the statistical structure of the data. Rather than acting as a stationary process, the acoustic response of the engine evolved across operating regimes, causing changes not only in signal behavior but also in feature distribution and anomaly score characteristics. This has important implications for industrial condition monitoring, because anomaly detection methods trained under mixed operating conditions may incorrectly interpret normal operational variability as abnormal behavior. Therefore, the findings reinforce the importance of operating condition aware modeling, which has been identified as a growing challenge in modern condition monitoring systems that handle non stationary industrial data (Feng et al., 2013; A. Kumar et al., 2025). In this context, the proposed load based framework provides a practical strategy for reducing distributional overlap and improving the consistency of unsupervised anomaly detection in acoustic monitoring applications.

The results provide insight into RQ1 by showing that different anomaly detection approaches do not identify the same type of abnormal behavior, even when applied to the same acoustic feature space. This is an important observation because anomaly detection methods are often evaluated primarily in terms of overall detection capability rather than the nature of the deviations they capture. The behavior observed in this study suggests that clustering based, density based, isolation based, and reconstruction based methods emphasize different structural properties of the data. Distance based approaches such as K Means were more influenced by global distributional changes, whereas Isolation Forest responded more strongly to separable deviations within the feature space. Reconstruction based modeling using the autoencoder appeared more sensitive to broader structural inconsistencies in learned feature representations, while

LOF depended heavily on local density consistency. These differences indicate that abnormal behavior in industrial acoustic signals cannot be treated as a single uniform phenomenon. Instead, anomalies may emerge as global shifts, local irregularities, or deviations from learned signal structure, depending on the underlying mechanical process and operating condition. This interpretation aligns with recent industrial monitoring studies that emphasize the complementary nature of unsupervised anomaly detection methods in complex sensor environments (Ahmad et al., 2020; Jombo & Zhang, 2023; L. Lei et al., 2025). The findings therefore support the idea that combining multiple anomaly detection perspectives provides a more reliable understanding of machine behavior than relying on a single detection mechanism.

RQ2 focused on how anomaly detection methods can be developed and evaluated under varying operating conditions. One of the most important contributions of this study is the demonstration that the evaluation strategy itself significantly affects the interpretation of anomaly detection performance. The comparison between native thresholding and percentile based thresholding showed that different models produce substantially different anomaly ratios when relying solely on their internal decision mechanisms. This makes direct comparison difficult because differences in anomaly counts may reflect threshold behavior rather than genuine differences in detection capability. By introducing a unified percentile based evaluation framework, the study isolates the behavior of the models from the variability introduced by threshold selection. This provides a more controlled basis for comparative analysis across operating conditions and detection methods. Therefore, the findings address an important limitation in existing anomaly detection studies, where evaluation procedures are often inconsistent between models and datasets (Chandola et al., 2009; Pimentel et al., 2014). Furthermore, the results demonstrate that reliable anomaly detection in industrial applications depends not only on selecting suitable algorithms but also on designing evaluation methodologies that account for operating condition variability, score distribution differences, and the absence of labeled fault data.

6 Conclusion and Future Work

6.1 Conclusion

This thesis investigates the use of unsupervised learning approaches for identifying anomalies in an industrial engine under different operating conditions. The work was inspired by the need to assess the condition of the engine in situations where labeled fault data are unavailable, which is a limitation in practical industrial environments. Instead of classifying the fault types, the study focused on modeling the normal behavior in different load conditions, and any deviation from the learned behavior was considered a potential anomaly.

A processing pipeline was created, consisting of time-based alignment of audio recordings with the engine data, extraction of time domain, spectral features, and use of four different unsupervised anomaly detection methods, including K-Means, Isolation Forest, Local Outlier Factor, and Autoencoder. A condition-specific approach was used, for which the model was trained separately on each load condition to learn load-specific patterns. To ensure consistent evaluation across all models and load conditions, a percentile-based thresholding was used. The results show that the signal characteristics vary with the engine load, low load conditions resulted in more variable feature representation, while higher load conditions resulted in a more structured and stable representation of the signals.

Overall, this thesis showed that unsupervised learning approaches are suitable for industrial condition monitoring for detecting abnormal engine behavior across varying operating conditions.

6.2 Limitations

Although the study shows positive outcomes but has some shortcomings that need to be recognized. Lack of labeled fault data makes it challenging to apply standard evaluation metrics. Also, there was an imbalance in the availability of the sample size between the load conditions, with fewer samples available at higher loads than at the low load condition, which can impact the stability of the model training and threshold estimation at these conditions. The study depends on the vibration signals, not the other characteristics, like temperature measurements, which can provide additional diagnostic information and improve the detection abilities.

6.3 Future Work

The proposed methodology provides a foundation for several research directions. Validation against actual fault data would allow a better assessment of the detection performance. The developed approach could be applied to other types of industrial machinery and engines, which would help evaluate its broader practical applicability. Addition of different sensor modalities, like temperature measurements, along with acoustic and vibrational features, could enhance the feature representation and provide improved detection sensitivity.

AI-Generated Content Acknowledgment

(ChatGPT-OpenAI) was used for language refinement, text editing, and grammatical improvement. All technical content, implementation, results, and interpretations presented in this thesis are the original work of the author.

Bibliography

Ahmad, S., Styp-Rekowski, K., Nedelkoski, S., & Kao, O. (2020). Autoencoder-based condition monitoring and anomaly detection method for rotating machines. In *2020 ieee international conference on big data (big data)* (pp. 4093–4102).

Al-Faiz, M., Ibrahim, A., & Hadi, S. (2019, 02). The effect of z-score standardization (normalization) on binary input due the speed of learning in back-propagation neural network. *Iraqi Journal of Information Communications Technology*, *1*, 42-48.

Alghushairy, O., Alsini, R., Soule, T., & Ma, X. (2021). A review of local outlier factor algorithms for outlier detection in big data streams. *Big Data and Cognitive Computing*, *5*(1). https://doi.org/10.3390/bdcc5010001

Alharbi, F., Luo, S., Zhang, H., Shaukat, K., Yang, G., Wheeler, C. A., & Chen, Z. (2023). A brief review of acoustic and vibration signal-based fault detection for belt conveyor idlers using machine learning models. *Sensors*, *23*(4). https://doi.org/10.3390/s23041902

Amruthnath, N., & Gupta, T. (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In *2018 5th international conference on industrial engineering and applications (iciea)* (p. 355-361).

Araghizad, A. E. (2024). *Machine learning-based modeling and monitoring of machining processes and tool wear* (PhD Dissertation). Sabancı University.

Attou, A. K., & Ahmed, Q. (2009). Asset management practices at qatargas. In H. E. Alfadala, G. Rex Reklaitis, & M. M. El-Halwagi (Eds.), *Proceedings of the 1st annual gas processing symposium* (Vol. 1, p. 286-296). Amsterdam: Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-53292-3.50036-5

Benhanifia, A., Cheikh, Z. B., Oliveira, P. M., Valente, A., & Lima, J. (2025). Systematic review of predictive maintenance practices in the manufacturing sector. *Intelligent Systems with Applications*, *26*, 200501. https://doi.org/https://doi.org/10.1016/j.iswa.2025.200501

Bhuiyan, M. Z. A., Wu, J., Wang, G., Chen, Z., Chen, J., & Wang, T. (2017). Quality-guaranteed event-sensitive data collection and monitoring in vibration sensor networks. *IEEE Transactions on Industrial Informatics*, *13*(2), 572-583.

Bountzis, P., Kavallieros, D., Tsikrika, T., Vrochidis, S., & Kompatsiaris, I. (2025). A deep one-class classifier for network anomaly detection using autoencoders and one-class support vector machines. *Frontiers in Computer Science*, *Volume 7 - 2025*. https://doi.org/10.3389/fcomp.2025.1646679

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). Lof: identifying density-based local outliers. In *Proceedings of the 2000 acm sigmod international conference on management of data* (pp. 93–104).

Caesarendra, W., & Tjahjowidodo, T. (2017, 09). A review of feature extraction methods in vibration-based condition monitoring and its application for degradation trend estimation of low-speed slew bearing. *Machines*, *5*, 21.

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

Chaudhary, A., Thakur, A., & Joshi, I. C. (2025, March–April). A review bearmath bearing fault diagnostics using machine learning. *International Journal for Multidisciplinary Research (IJFMR)*, *7*(2), IJFMR250239641.

Chauhan, P., & Shukla, M. (2015). A review on outlier detection techniques on data stream by using different approaches of k-means algorithm. In *2015 international conference on advances in computer engineering and applications* (p. 580-585).

Chen, J., Li, Z., Pan, J., Chen, G., Zi, Y., Yuan, J., … He, Z. (2016). Wavelet transform based on inner product in fault diagnosis of rotating machinery: A review. *Mechanical Systems and Signal Processing*, *70-71*, 1-35. https://doi.org/https://doi.org/10.1016/j.ymssp.2015.08.023

Cocconcelli, M., Capelli, L., Cavalaglio Camargo Molano, J., & Borghi, D. (2018). Development of a methodology for condition-based maintenance in a large-scale application field. *Machines*, *6*(2). https://doi.org/10.3390/machines6020017

Danielsson, P.-E. (1980). Euclidean distance mapping. *Computer Graphics and image processing*, *14*(3), 227–248.

Dao, P. B., Staszewski, W. J., Barszcz, T., & Uhl, T. (2018). Condition monitoring and fault detection in wind turbines based on cointegration analysis of scada data. *Renewable Energy*, *116*, 107-122. (Real-time monitoring, prognosis and resilient control for wind energy systems) https://doi.org/https://doi.org/10.1016/j.renene.2017.06.089

Das, O., Bagci Das, D., & Birant, D. (2023). Machine learning for fault analysis in rotating machinery: A comprehensive review. *Heliyon*, *9*(6). https://doi.org/10.1016/j.heliyon.2023.e17584

Davies, A. (2012). *Handbook of condition monitoring: techniques and methodology*. Springer Science & Business Media.

Feng, Z., Liang, M., & Chu, F. (2013). Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples. *Mechanical Systems and Signal Processing*, *38*(1), 165-205. (Condition monitoring of machines in non-stationary operations.) https://doi.org/https://doi.org/10.1016/j.ymssp.2013.01.017

Gholipour, Y., Zare, M., Vaziri Sarashk, M., & Gholipour, Y. (2025, 07). A comprehensive review of maintenance strategies: From reactive to proactive approaches. *CENTRAL ASIA AND THE CAUCASUS*, *26*, 70-83.

Glowacz, A. (2021). Ventilation diagnosis of angle grinder using thermal imaging. *Sensors*, *21*(8). https://doi.org/10.3390/s21082853

Goel, S., Ghosh, R., Kumar, S., & Akula, A. (2015). A methodical review of condition monitoring techniques for electrical equipment. *International Journal of Electrical Engineering and Technology*, *6*(4), 1–12.

Goodarzi, P., Schütze, A., & Schneider, T. (2023, 08). *Comparing automl and deep learning methods for condition monitoring using realistic validation scenarios*.

Guo, C., Dong, M., Yang, X., & Wang, W. (2019). A review of on-line condition monitoring in power system. In *2019 ieee 8th international conference on advanced power system automation and protection (apap)* (p. 634-637).

Hamishebahar, Y., Li, H., & Guan, H. (2021, 01). Application of machine learning algorithms in structural health monitoring research. In (p. 219-228).

Hassan, I., Panduru, K., & Walsh, J. (2024, 01). An in-depth study of vibration sensors for condition monitoring. *Sensors*, *24*, 740.

Heisenberg, W. (1927). Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik*, *43*(3), 172–198.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507.

Huang, L., Pan, X., Liu, Y., & Gong, L. (2023). An unsupervised machine learning approach for monitoring data fusion and health indicator construction. *Sensors*, *23*(16). https://doi.org/10.3390/s23167239

Huang, M. (2023). *Anomaly detection for condition monitoring in robot systems* (Degree Project). Uppsala University.

IBM. (2026). *Condition monitoring*. https://www.ibm.com/think/topics/condition-monitoring. (Accessed: 2026-05-02)

Ismail, L., Abdelmoti, A., Basu, A., Berini, A., & Naouss, M. (2025, 09). *A systematic review of digital twin-driven predictive maintenance in industrial engineering: Taxonomy, architectural elements, and future research directions.*

Jalayer, M., Kaboli, A., Orsenigo, C., & Vercellis, C. (2022, 03). Fault detection and diagnosis with imbalanced and noisy data: A hybrid framework for rotating machinery. *Machines*, *10*, 237.

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, *20*(7), 1483-1510. https://doi.org/https://doi.org/10.1016/j.ymssp.2005.09.012

Jombo, G., & Zhang, Y. (2023, 01). Acoustic-based machine condition monitoring—methods and challenges. *Eng*, *4*, 47-79.

Karabacak, Y. E., & Gürsel Özmen, N. (2022). Common spatial pattern-based feature extraction and worm gear fault detection through vibration and acoustic measurements. *Measurement*, *187*, 110366. https://doi.org/https://doi.org/10.1016/j.measurement.2021.110366

Kateris, D., Moshou, D., Pantazi, X.-E., Gravalos, I., Sawalhi, N., & Loutridis, S. (2014). A machine learning approach for the condition monitoring of rotating machinery. *Journal of Mechanical Science and Technology*, *28*(1), 61–71.

Kumar, A., Wy lomańska, A., Zimroz, R., Xiang, J., & Antoni, J. (2025). Critical challenges and advances in vibration signal processing for non-stationary condition monitoring. *Advanced Engineering Informatics*, *65*, 103290.

https://doi.org/https://doi.org/10.1016/j.aei.2025.103290

Kumar, S., Lokesha, M., Kumar, K., & Srinivas, K. R. (2018, jun). Vibration based fault diagnosis techniques for rotating mechanical components: Review paper. *IOP Conference Series: Materials Science and Engineering*, *376*(1), 012109. https://doi.org/10.1088/1757-899X/376/1/012109

Kumari, S. M. (2026). Development of a digital twin-based intelligent monitoring system for industrial machinery using edge computing and machine learning in industry 4.0 environments. *Global Journal of Engineering and Innovative Research (GJEIIR)*, *6*(2), E324.

Kuncan, M. (2020). An intelligent approach for bearing fault diagnosis: Combination of 1d-lbp and gra. *IEEE Access*, *8*, 137517-137529.

Lei, L., Li, W., Zhang, S., Wu, C., & Yu, H. (2025). Research progress on data-driven industrial fault diagnosis methods. *Sensors*, *25*(9). https://doi.org/10.3390/s25092952

Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to rul prediction. *Mechanical Systems and Signal Processing*, *104*, 799-834. https://doi.org/https://doi.org/10.1016/j.ymssp.2017.11.016

Lima, M. F., Zarpelão, B. B., Sampaio, L. D. H., Rodrigues, J. J. P. C., Abrão, T., & Proença, M. L. (2010). Anomaly detection using baseline and k-means clustering. In *Softcom 2010, 18th international conference on software, telecommunications and computer networks* (p. 305-309).

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

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

Liu, H., Xia, M., Williams, D., Sun, J., & Hongsheng, Y. (2022, 07). Digital twin-driven machine condition monitoring: A literature review. *Journal of Sensors*, *2022*, 1-13.

Liu, L., Chen, L., Wang, S., Yin, Y., Liu, D., Wu, S., … Pan, X. (2019). Improving sensitivity of a micro inductive sensor for wear debris detection with magnetic powder surrounded. *Micromachines*, *10*(7). https://doi.org/10.3390/mi10070440

Lv, Y., Zhao, W., Zhao, Z., Li, W., & Ng, K. K. (2022). Vibration signal-based early fault prognosis: Status quo and applications. *Advanced Engineering Informatics*, *52*, 101609. https://doi.org/https://doi.org/10.1016/j.aei.2022.101609

MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297). Berkeley, CA: University of California Press.

McGregor, C. (2018). Using constructive research to structure the path to transdisciplinary innovation and its application for precision public health with big data analytics. *Technology Innovation Management Review*, *8*(8).

Meddaoui, A., Hain, M., & Hachmoud, A. (2023). The benefits of predictive maintenance in manufacturing excellence: a case study to establish reliable methods for predicting failures. *The International Journal of Advanced Manufacturing Technology*, *128*, 3685–3690.

Micheal, D. (2026, 03). Predictive maintenance systems: Leveraging machine learning for industrial reliability and cost reduction.

Mohanty, A. R. (2014). *Machinery condition monitoring: Principles and practices*. CRC Press.

Mohd Ghazali, M. H., & Rahiman, W. (2021). Vibration analysis for machine monitoring and diagnosis: A systematic review. *Shock and Vibration*, *2021*(1), 9469318. https://doi.org/https://doi.org/10.1155/2021/9469318

Nie, Q., Geng, J., & Liu, C. (2026). A review of fault diagnosis methods: From traditional machine learning to large language model fusion paradigm. *Sensors*, *26*(2). Retrieved from https://www.mdpi.com/1424-8220/26/2/702

Nithin, S., Hemanth, K., Shamanth, V., Shrinivas Mahale, R., Sharath, P., & Patil, A. (2022). Importance of condition monitoring in mechanical domain. *Materials Today: Proceedings*, *54*, 234-239. (5th International Conference on Advanced Research in Mechanical, Materials and Manufacturing Engineering-2021) https://doi.org/https://doi.org/10.1016/j.matpr.2021.08.299

Olejnik, P., & Desta, Y. D. (2025). Friction-induced interactions: acoustic emissions, vibrations, and wear – a multiscale review. *Nonlinear Dynamics*, *113*, 24101–24140.

Ouadah, A., Zemmouchi-Ghomari, L., & Salhi, N. (2022). Selecting an appropriate supervised machine learning algorithm for predictive maintenance. *The International*

Journal of Advanced Manufacturing Technology, *119*, 4277–4301.

Pang, G., Shen, C., Cao, L., & Hengel, A. V. D. (2021, March). Deep learning for anomaly detection: A review. , *54*(2). https://doi.org/10.1145/3439950

Pimentel, M. A., Clifton, D. A., Clifton, L., & Tarassenko, L. (2014). A review of novelty detection. *Signal processing*, *99*, 215–249.

Prajapati, A., Bechtel, J., & Ganesan, S. (2012). Condition based maintenance: a survey. *Journal of Quality in Maintenance Engineering*, *18*(4), 384–400.

Prawin, J., & Anbarasan, R. (2021). A novel mel-frequency cepstral analysis based damage diagnostic technique using ambient vibration data. *Engineering Structures*, *228*, 111552. https://doi.org/https://doi.org/10.1016/j.engstruct.2020.111552

Rossetti, D. (2018). *Advanced machine learning techniques for condition monitoring in industrial engineering applications* (PhD Dissertation). Università Politecnica delle Marche.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., … Müller, K.-R. (2021). A unifying review of deep and shallow anomaly detection. *Proceedings of the IEEE*, *109*(5), 756–795.

Shahzad, K., & O'Nils, M. (2018). Condition monitoring in industry 4.0-design challenges and possibilities: A case study. In *2018 workshop on metrology for industry 4.0 and iot* (p. 101-106).

Shu Fuhnwi, G., Agbaje, J., Oshinubi, K., & Peter, O. (2023, 04). An empirical study on anomaly detection using density-based and representative-based clustering algorithms. *Journal of the Nigerian Society of Physical Sciences*, 1364.

Surucu, O., Gadsden, S. A., & Yawney, J. (2023). Condition monitoring using machine learning: A review of theory, applications, and recent advances. *Expert Systems with Applications*, *221*, 119738. https://doi.org/https://doi.org/10.1016/j.eswa.2023.119738

Turku University of Applied Sciences. (2024). *Early detection of extreme engine events (ede3).* (Accessed: 2026-04-07) Retrieved from https://www.turkuamk.fi/en/project/early-detection-of-extreme-engine-events-ede3/

Vlachou, V. I., Karakatsanis, T. S., & Efstathiou, D. E. (2025). Recent advances of artificial intelligence methods in pmsm condition monitoring and fault diagnosis in elevator systems. *Applied System Innovation*, *8*(5). Retrieved from https://www.mdpi.com/2571-5577/8/5/154

Wan, Y., Lin, S., & Gao, Y. (2024). Pipeline and rotating pump condition monitoring based on sound vibration feature-level fusion. *Machines*, *12*(12). https://doi.org/10.3390/machines12120921

Xu, D., Wang, Y., Meng, Y., & Zhang, Z. (2017). An improved data anomaly detection method based on isolation forest. In *2017 10th international symposium on computational intelligence and design (iscid)* (Vol. 2, p. 287-291).

Yang, S., Cao, N., & Yu, B. (2023). Wear debris measurement in lubricating oil based on inductive method: A review. *Measurement and Control*, *56*(7-8), 1422-1435. https://doi.org/10.1177/00202940231159117

Yu, G. (2021). A multisynchrosqueezing-based high-resolution time-frequency analysis tool for the analysis of non-stationary signals. *Journal of Sound and Vibration*, *492*, 115813. https://doi.org/https://doi.org/10.1016/j.jsv.2020.115813

Zhu, J., Nostrand, T., Spiegel, C., & Morton, B. (2014, September). Survey of condition indicators for condition monitoring systems. In *Proceedings of the annual conference of the phm society* (Vol. 6). PHM Society.