الجمهـــوريـــــة الجزائـريـــــة الديـمـقـراطيـــــة الشـعـبيـــــة

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

وزارة التعليــم العالـــي والـبحـــث العلمـــي

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

المدرسة الوطنية العليا للتكنولوجيا والهندسة- عنابة-

ECOLE NATIONALE SUPERIEURE DE TECHNOLOGIE ET D'INGENIERIE - ANNABA

Département Génie Industriel

MEMOIRE

En vue d'obtention du diplôme d'INGÉNIEUR D'ETAT

Domaine : Science et Technologie

Filière : Génie Industriel

Spécialité : Maintenance et Fiabilité des Systèmes Industriels

Présenté par

HOUSSEM EDDINE BELAID
SOHEIB ABDESLAM OSMANI

AI-Driven Predictive Maintenance: Detecting Defective Bearings in Rotating Machines

Encadré par
Dr. Nour Elislem KARABADJI
ENSTI Annaba

Co-encadré par
Dr. Ilyes Khelf
ENGIE France

Membres du jury :

Pr. Messaoud DJEGHABA	Président	ENSTI Annaba
Dr. Salim KEBIR	Examinateur	ENSTI Annaba
Dr. Tarek MESSIKH	Examinateur	ENSTI Annaba

Année 2024

ACKNOWLEDGEMENTS

First of all, we thank Allah the Almighty for giving us the courage and the patience to complete this work.

We would love to offer our gratitude to our dear parents and every person close to our hearts who cared, guided and supported us this whole journey.

We would like to thank in particular our mentor Mr. KARABADJI Nour ElIslem, for the competent help he has given us, for his patience and encouragement. His critical eye was invaluable in structuring the work and for improve the quality of the different sections. We also appreciate the help of Mr. KHELF Ilyes for his help and guideness in providing the necessary information, alongside to our internship engineers Mr. LEKSIR Fares and Mr. MOUATES Hamza who showed a lot of support during the internship.

The jury members find here the expression of our sincere thanks for the honor they do us by taking the time to read and evaluate this work.

We would also like to thank the teaching and administrative team of ENSTI for their efforts to provide us with excellent training. We are also grateful to our friends and colleagues for the camaraderie, and their ever-enthusiastic supports and cheerfulness.

Finally, we would like to thank all those who have contributed far to the realization of this work.

Abstract

This dissertation addresses the critical challenge of detecting defective bearings in industrial rotating machines to prevent downtime and costly repairs. The diagnostic process begins by collecting vibration signals from both normal and faulty bearings. These signals are being analyzed to extract essential features such as Root Mean Square (RMS), Kurtosis, Skewness, and Frequency Domain characteristics, which collectively form a comprehensive dataset.

To analyze this dataset, a variety of machine learning algorithms are employed: decision trees for their interpretability, random forests to enhance accuracy and mitigate overfitting, Support Vector Machines (SVM) for robustness in high-dimensional spaces, and neural networks for their ability to capture complex patterns in the data. These algorithms are integrated into a unified Python application that streamlines both training and testing phases, featuring an intuitive interface for data input, algorithm percentage, and classification result visualization.

The experimental results demonstrate significant advancements facilitated by AI techniques in analyzing large volumes of data. Specifically, the algorithms effectively identify and classify faulty bearings with high accuracy and reliable predictions. This capability not only enhances the speed and precision of fault detection but also provides a scalable solution applicable across various industrial applications.

In conclusion, this dissertation underscores the transformative potential of machine learning and artificial intelligence in industrial maintenance. The developed Python application, leveraging multiple machine learning algorithms, serves as a robust tool for early detection of defective bearings, thereby significantly improving overall reliability and efficiency in industrial operations.

Keywords: Bearings, Rotating machines, Vibration analysis, RMS, Kurtosis, Machine learning algorithms, Decision tree, Random forest, SVM, Neural network, Python application, AI methods, Accuracy.

ملخص

تتناول هذه الأطروحة التحدي الخطير المتمثل في اكتشاف المحامل المعيبة في الآلات الدوارة الصناعية لمنع التعطل والإصلاحات المكلفة. تبدأ عملية التشخيص بجمع إشارات الاهتزاز من المحامل الطبيعية والمعيبة. يتم تحليل هذه الإشارات لاستخراج الميزات الأساسية مثل Root Mean Square (RMS) و Kurtosis و Skewness وخصائص مجال التردد، والتي تشكل مجتمعة مجموعة بيانات شاملة.

لتحليل مجموعة البيانات هذه، يتم استخدام مجموعة متنوعة من خوارزميات التعلم الآلي: أشجار القرار لإمكانية تفسيرها، والغابات العشوائية لتعزيز الدقة والتخفيف من التركيب الزائد، وآلات ناقلات الدعم (SVM) للقوة في المساحات عالية الأبعاد، والشبكات العصبية لقدرتها على التقاط أنماط معقدة في البيانات. تم دمج هذه الخوارزميات في تطبيق بايثون موحد يبسط مراحل التدريب والاختبار، ويتميز بواجهة بديهية لإدخال البيانات، ونسبة الخوارزمية، وتصور نتائج التصنيف.

تُظهر النتائج التجريبية تطورات كبيرة سهلتها تقنيات الذكاء الاصطناعي في تحليل كميات كبيرة من البيانات. على وجه التحديد، تحدد الخوارزميات وتصنف بشكل فعال المحامل الخاطئة بدقة عالية وتنبؤات موثوقة. لا تعزز هذه القدرة سرعة ودقة اكتشاف الأخطاء فحسب، بل توفر أيضًا حلاً قابلاً للتطوير قابل للتطبيق عبر التطبيقات الصناعية المختلفة.

في الختام، تؤكد هذه الأطروحة على الإمكانات التحويلية للتعلم الآلي والذكاء الاصطناعي في الصيانة الصناعية. يعمل تطبيق Python المطور، الذي يستفيد من خوارزميات التعلم الآلي المتعددة، كأداة قوية للكشف المبكر عن المحامل المعيبة، وبالتالي تحسين الموثوقية والكفاءة بشكل كبير في العمليات الصناعية.

كلمات مفتاحية: المحامل، الآلات الدوارة، تحليل الاهتزازات، RMS، Kurtosis، خوارزميات التعلم الآلي، شجرة القرار، الغابة العشوائية، SVM، الشبكة العصبية، تطبيق بايثون، طرق الذكاء الاصطناعي، الدقة.

Résumé

Ce mémoire aborde le défi critique de la détection des roulements défectueux dans les machines tournantes industrielles pour éviter les temps d'arrêt et les réparations coûteuses. Le processus de diagnostic commence par la collecte des signaux de vibration des roulements normaux et défectueux. Ces signaux sont analysés pour extraire des caractéristiques essentielles telles que le carré moyen de racine (RMS), le kurtosis, l'asymétrie et les caractéristiques du domaine de fréquence, qui forment collectivement un ensemble de données complet.

Pour analyser cet ensemble de données, une variété d'algorithmes d'apprentissage automatique sont utilisés : des arbres de décision pour leur interprétabilité, des forêts aléatoires pour améliorer la précision et atténuer le sur-ajustement, des machines vectorielles de soutien (SVM) pour la robustesse dans des espaces dimensionnels et des réseaux de neurones pour leur capacité à capturer des motifs complexes dans les données. Ces algorithmes sont intégrés dans une application Python unifiée qui rationalise les phases de formation et de test, avec une interface intuitive pour la saisie de données et la visualisation des résultats de classification.

Les résultats expérimentaux démontrent des avancées significatives facilitées par les techniques d'IA dans l'analyse de grands volumes de données. Plus précisément, les algorithmes identifient et classent efficacement les roulements défectueux avec une grande précision et des prévisions fiables. Cette capacité améliore non seulement la vitesse et la précision de la détection des défauts, mais fournit également une solution évolutive applicable à diverses applications industrielles.

En conclusion, ce mémoire souligne le potentiel transformateur de l'apprentissage automatique et de l'intelligence artificielle dans la maintenance industrielle. L'application Python développée, utilisant plusieurs algorithmes d'apprentissage automatique, sert d'outil robuste pour la détection précoce des roulements défectueux, améliorant ainsi considérablement la fiabilité et l'efficacité globales dans les opérations industrielles.

Mots clés: Roulements, machines rotatives, analyse des vibrations, RMS, Kurtosis, algorithmes d'apprentissage automatique, arbre de décision, forêt aléatoire, SVM, réseau neuronal, application Python, méthodes d'IA, précision.

Abbreviations & Acronyms

Acronym	Definition
n	Number of rolling elements (balls or rollers)
d	Diameter of the rolling element
D	Pitch diameter (diameter of the circle that passes through the centers of the rolling elements)
β	Contact angle (angle between the line connecting the centers of the rolling element and the raceway, and the radial plane)
f	Rotational frequency of the shaft (in Hz)
RPM	Revolutions Per Minute (unit of rotational speed)
BPFO	Ball Pass Frequency Outer Race
BPFI	Ball Pass Frequency Inner Race
BSF	Ball Spin Frequency
FTF	Fundamental Train Frequency
E.P.S	Engineering and Production Services
AI	Artificial intelligence
RMS	Root Mean Square
ML	Machine Learning
ANN	Artificial Neural Networks
DL	Deep Learning
SVM	Support Vector Machines
PCA	Principal Component Analysis
PM	Preventive Maintenance
CBM	Condition Based Maintenance
IFDP	Intelligent Fault Diagnosis and Prognosis

SVD	Singular Value Decomposition
VMD	Variational mode decomposition
IMF	Intrinsic Mode Function
LSTM	Long Short-Term Memory
FFT	Fast Fourier Transform
CSV	Comma-Separated Values
RFC	Random Forest Classifier
DTC	Decision Tree Classifier
SVC	Support Vector Classifier
NN	Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network
MCC	Matthew's correlation coefficient

Tables list

Figures list

Summary

General Introduction:

Rotating machinery forms the backbone of various industrial sectors, they have to be reliable and efficient enough so that operations run smoothly and for cost-effective solutions in maintenance procedures. However, fault detection and diagnosis in rotating machinery remain a critical challenge that the maintenance engineers and the industry specialists encounter on a regular basis.

Vibration analysis is the most common predictive maintenance technology employed in monitoring rotating machinery conditions. The technique measures the machine vibrations and analyzes these vibrations to find the signs of potential failure. Yet, this process is often very time-consuming and complex, and it's hard for human experts to manually recognize all possible signs of failure every time. It can be used to increase the accuracy of predictive maintenance by employing artificial intelligence in predictive maintenance. Large amounts of vibration data are analyzed by the AI algorithm, and patterns that reflect potential failures are identified, this allows detecting the problem even earlier and at lower maintenance costs. The present dissertation discusses the applications of machine learning techniques applicable for the detection of faulty signals in rotating machinery, drawing on insights gathered from data collected during our internship in an industrial field.

1. Description of the theme and its importance

Many industrial processes mainly use rotating machines. Machine reliability and efficiency are very well understood for their continuous operation and cost reduction in maintenance. However, the automatic detection and diagnosis of faults occurring in such machines are always challenging tasks for researchers. Vibration analysis happens to be one of the most commonly used predictive maintenance technologies applied in monitoring the health of rotating machinery. Essentially, it involves signal capture and consequent analysis to spot the very first signs of malfunction. While all of this offers a great deal, manual analysis of signals tends to be highly consultant-dependent, laborious, and highly error-prone. There is the need to have efficient solutions. One major step in this area has been the use of artificial intelligence. As they can process huge volumes of vibration signaling data, AI algorithms can identify complex patterns related to failures. Apart from ensuring fault detection more efficiently and accurately. A system like that would bring reduction of both manual interventions and maintenance costs. This dissertation discusses machine learning techniques for the automatic detection of rotating machine faults, using full data collected during an industrial internship.

2. The problematic

Although there are considerable advances in sensor technology and data analytics, automatic exploitation of data for fault localization in rotating machinery is very challenging. Machines produce highly complex signals that are exposed to a variety of noise sources, environmental factors, and operational conditions. For automatic differentiation between normal operational behavior and the first signs of degradation or impending failure, advanced analysis methods far beyond traditional diagnostic are necessary. Also, the great volume and complexity of data from rotating machinery constitute very serious challenges for its automatized processing and interpretation. Traditional analysis methods usually cannot capture the high dimensionality, non-linearity, and heterogeneity of the data, so their performance is inefficient and leads to false alerts. Therefore, there is a pressing need for innovative solutions that can automatically extract meaningful insights from the data, enabling more effective and timely decision-making in maintenance practices.

3. The manuscript

This dissertation is divided into four chapters, each of them contributes to the progress of fault detection in rotating machinery by combining machine learning approaches and domain-specific knowledge. Throughout the research, we use a variety of approaches, tools, and datasets to look into the complexities of signal processing, bearing analysis, and machine learning.

In Chapter 1, we write a short description about our internship's companies, which we used to gather some data. Then setting up the reasons behind working this theme and its importance in the field and what it will add in the future. In Chapter 2, we carry out an in-depth analysis of signals and bearings. We explore the key concepts behind signals emitted by rotating machinery using both theoretical and practical research. Furthermore, we dive into the complexities of bearings, recognizing their importance as essential parts in spinning machinery. We look at many bearing indicators such as kurtosis, RMS, and skewness, which are useful measurements for evaluating the health and performance of bearings. We use MATLAB to conduct analysis of bearing frequencies and build codes for extracting relevant information from raw sensor data. Focusing later on the development and application of machine learning techniques for fault detection in our matrix. We explore through WEKA and Python the utility of machine learning algorithms in identifying patterns and anomalies indicative of impending failures, to train predictive models capable of discerning between normal operating conditions and fault states. We refine our models to achieve high levels of accuracy and reliability in fault detection tasks.

In Chapter 3, we move from theory to practice, we examine the limitations and complexities imposed by machine learning applications for fault detection by performing experiments in an industrial setting. From data collection and organizing to model deployment and performance evaluation. Furthermore, we explain the actual issues and decisions associated with choosing and executing appropriate algorithms and approaches.

Finally, in Chapter 4, we undertake an extensive evaluation of the findings and provide critical commentary on the efficacy and implications of our approach. We examine the ability of our prediction algorithms in detecting defects in rotating machinery using both quantitative measures and qualitative assessments. In addition, we discuss the impact of those results for predictive maintenance and make recommendations for future research and advancement.

CHAPTER 1

CHAPTER 1 Company Description, Motivation, Problematic

1.1 Description of the company's internship

In the interest of practical experience and industry insights, both my friend and I set out on separate internship journeys inside different companies, each providing interesting perspectives and important learning opportunities.

1.1.1 Skikda Port Enterprise

The Skikda Port Enterprise, abbreviated as E.P.S, is a Public Economic Company. It operates as a joint-stock company governed by laws and regulations related to corporate autonomy. Its social capital amounts to 9,000,000,000 Algerian dinars (DA). Skikda port stands as a crucial maritime center, boasting an array of essential infrastructure to support different cargo and passenger needs. With offices including a traveler station pleasing up to 1,200 travelers and a dedicated terminal capable of handling 500 vehicles, the port ensures smooth travel and logistics operations. Also, its quayside facilities are equally amazing, featuring a high-capacity compartment able to take care of 23 million tons of oil shipments, alongside another flexible quay designed for products, with a capacity of 3.7 million tons and space for 132,000 containers, displaying its flexibility and importance as a crucial logistics hub.

The port is organized into three departments: the methods office, the equipment department, and the maintenance department.

The Maintenance Department is responsible for maintaining and servicing the equipment managed by the E.P.S. This includes rolling stock, forklifts, cranes, and related accessories. Additionally, the department handles procurement and stock management for spare parts, tires, oils, and lubricants, among other items.

Figure 1: Container crane

1.1.2 Iris Tires Setif

Iris Tires Setif is an Algerian company specializing in the production of tires for heavy and light vehicles. Founded in 2008, the company quickly became a significant player in the national market through its commitment to quality and innovation. Iris Tires employs more than 500 people and has a modern production site equipped with cutting-edge technical equipment. The company is ISO 9001 and ISO 14001 certified and stands out for its commitment to quality and respect for the environment. Focused on innovation and quality, the company produces a wide range of tires adapted to different conditions and vehicles.

Many steps participate in the creation of the tires. It is a complex production which combines the constraints of heavy industry with the requirements of high technology:

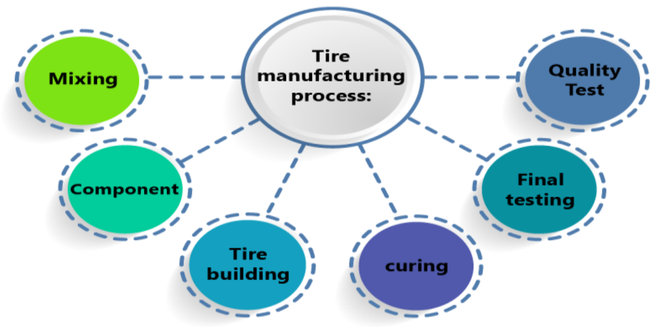


Figure 2: Manufacturing

The maintenance service:

The department is composed of a team of engineers and technicians in different fields, as shown in the following figure:

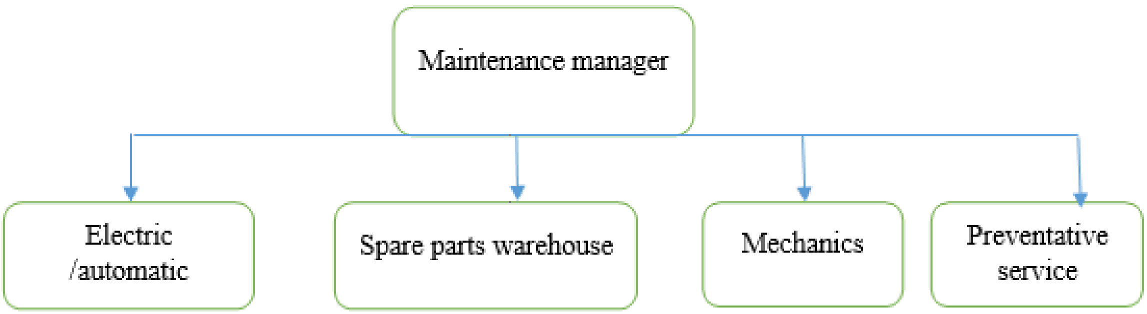


Figure 3: Maintenance service

1.2 Motivation

Our interest in this internship can be explained by several factors. First of all, we are passionate about artificial intelligence and its potential application in the industrial field. WE believe AI can play an important role in improving predictive maintenance of rotating machinery, which can help reduce maintenance costs and improve machine availability. This technology has the potential to transform how defects are detected and addressed, thereby reducing downtime and improving operational efficiency

We are convinced that our knowledge in AI and vibration analysis can contribute to the improvement of production processes. Finally, this internship will allow us to acquire valuable skills and work within a dynamic and professional team.

We aim to contribute valuable experiences to the field of maintenance and engineering. Through serious examination and innovative approaches, our research seeks to address basic gaps in information and clear the way for impactful advancements in detecting the faulty vibration signal. With faithful assurance, we set out on this journey of discovery, motivated by the belief that our work has the potential to create a meaningful contrast in industrial engineering.

1.3 Problematic

Rotating machines, such as motors, pumps and fans, cranes, conveyor belts, are essential components of many industrial processes. However, these machines are prone to failures which can lead to costly production downtime. Predictive maintenance is a key strategy for preventing these failures by identifying potential problems before they arise.

Vibration analysis is defined as a process for measuring the vibration levels and frequencies of machinery and then using that information to analyze how healthy the machines and their components are.it can be difficult for human experts to identify all the potential signs of failure.

Artificial intelligence (AI) can be used to automate vibration analysis and improve the accuracy of predictive maintenance. Artificial intelligence algorithms can learn large amounts of vibration data to identify patterns indicating potential failures. This helps detect problems earlier and reduce maintenance costs.

CHAPTER 2

CHAPTER 2 Literature review & Bibliographic researches

2.1 Introduction

Rotating machinery is integral to various industrial applications. The reliable operation of these machines is critical because their failure can result in significant economic losses, production downtime, and potential safety hazards. Therefore, early detection and diagnosis of faults are crucial to prevent unexpected breakdowns and extend operational life.

In this chapter, we review the literature on methods used to detect defects in rotating machinery through bearing vibration analysis, as bearings are a common component in almost all rotating machinery. Vibration analysis is widely recognized as a powerful tool in the industry, with numerous methods developed to identify different types of defects, such as Fourier transform and wavelet transform, have been widely used to identify and characterize fault signatures. However, these methods often require significant domain expertise and may not effectively handle the complex and nonlinear nature of real-world vibration data. On the other hand, machine learning (ML) has emerged as a powerful tool for fault diagnosis. Its algorithms can automatically learn patterns and features from large volumes of data, making them well-suited for analyzing the complex vibration signals associated with machinery faults. By leveraging ML techniques, researchers and practitioners can develop more accurate and efficient fault detection systems that adapt to various operational conditions and fault types.

This literature review aims to provide a comprehensive overview of the current state of research on using machine learning for fault detection in rotating machinery. It highlights key studies and methodologies which employ Artificial Intelligence approaches for fault detection and classification with a primary focus on bearing faults.

2.2 Related works

Karabadji et al. in **[1]** proposed an improved decision tree construction method which was applied to fault diagnosis in rotating machines. The method focuses on attribute selection and data sampling strategies to address the limitations of decision trees construction process. The authors introduced a new attribute selection criterion that prioritizes attributes based on their relevance to fault detection, improving the accuracy and robustness of the decision tree, and avoiding the pruning phase. The study demonstrated that the improved decision tree outperformed standard building one in terms of classification accuracy and fault detection reliability. This was particularly evident in scenarios with noisy and imbalanced data, where traditional methods often struggle.

Karabadji et al. in **[2]** further refined their decision tree construction method by introducing advanced data sampling and attribute selection strategies. This study emphasized a systematic approach to sampling, ensuring that the data used for training the decision tree captures the full range of operational conditions, including rare but critical fault states. They further proposed a new kind of advanced attribute selection strategy, which harmonizes statistical and domain-specific knowledge in selecting the most informative features. This hybrid approach enhances not only the performance of the decision tree but also its ability to be generalized with respect to different kinds of rotating machinery and fault conditions.

Thuan, N.D & Hong, H.S. in **[3]** The quick development of machine learning methods results in enlarging the demand for data. The data acquisition for bearing fault diagnosis is time-consuming and with complicated processes. The existing datasets are targeted only for one type of bearing, which limits the practical application of them. Therefore, the goal of the current work is to propose a diverse dataset for ball bearing fault diagnosis based on vibration. Almost half of the motor failures are due to bearing faults, which may cause massive production standstills and expensive repairs. Early detection of those bearing faults is critical for equipment health. Vibration signals are common practice in detecting faults, which often contain very useful information. Machine learning methods, especially deep learning models, have high accuracy. However, the data acquisition necessary to train them remains challenging. Existing datasets are focused on one target bearing, which motivated the creation of a more diverse dataset for ball bearing fault diagnosis.

Singh, V., et al. in **[4]** highlight the critical role of fault diagnosis in industrial machinery as part of the Industry 4.0 revolution. Despite its complexity, artificial intelligence (AI) techniques have become essential for condition monitoring of mechanical and electrical machines due to their rapid computation, high accuracy, and robust performance. AI reduces the need for reliance on experienced personnel with expert knowledge. In this paper, they review successful applications of AI-based fault diagnosis techniques in various industrial machinery over the past two decades.

While vibration analysis is widely used for predictive maintenance in high-speed rotating machines, it can be inadequate for detecting bearing faults due to their subtle nature, especially in early failure stages with low signal-to-noise ratios.

Zoungrana, W. B., et al. in **[5]** developed a system that analyzes vibration signals of rotary machines through time/frequency data captured by sensors. The goal was to use machine learning algorithms to detect abnormal behavior and identify different operation modes for Industry 4.0 predictive maintenance, even though current methods are not yet sufficiently efficient.

Lee, Y. E., et al. in **[6]** address shaft misalignment, a common cause of mechanical vibration and shaft failure. This study proposes an AI-based machine learning approach to detect defects caused by shaft axis misalignment in rotating machinery. Unlike traditional methods that focus on the time domain, this approach transforms vibration data from the time domain to power spectrum component values in the frequency domain. These power spectrum values are then classified using principal component analysis (PCA). Based on the results, defects are identified using a support vector machine (SVM) as part of the machine learning technique.

Their findings demonstrate that defects can be rapidly detected using only vibration data from normal and abnormal states. Furthermore, the proposed pre-processing method allows the machine learning framework to effectively diagnose not only shaft defects but also faults in other rotating machines, such as motors and engines.

Mang, Huan, et al. in **[7]** conduct a comprehensive investigation into advanced signal processing techniques and feature extraction methods customized for bearing fault diagnosis using vibration data. By delving into the effectiveness of innovative signal processing algorithms, their objective is to enhance the precision and dependability of fault detection systems. Their research illuminates the critical role of feature engineering in machine learning-based fault diagnosis.

2.3 The relation of my work

Based on these cited works, wherein their contributions are interesting, and their approaches are helpful in using machine learning techniques for further development of fault analysis methods for rotating machinery. Thus, we can have a clearer view on the aims and objectives of choosing Ai techniques for detecting the defect in these signals, and coming up with an effective predictive maintenance which is more accurate for whatever fault occurrence.

2.4 Conclusion

AI-based technologies are used in maintaining the faults and diagnosing various kinds of industrial machines. In such technologies, computation is fast, accuracy is high, and performance is robust, with reduced dependency on skilled and experienced personnel having specialized knowledge. Various AI fault diagnosis techniques have successfully been applied over the last two decades in different types of industrial machines, which mainly include induction motors, bearings, gearboxes, and centrifugal pumps. The proposed method of AI-based audio inspection can be applied to sound, including vibration, for analysis, and the identified different kinds of wear and tear or damage of the rotating equipment will indirectly diagnose faults in an efficient and automated way. Applications of AI in condition monitoring, can detect faults in time so as to avoid catastrophic failure and minimize economic loss in the industry.

CHAPTER 3

CHAPTER 3 Solutions and Experimentations carried out

3.1 Introduction (Background and Motivation)

Fault detection, diagnosis, and failure prognosis have received a considerable amount of attention during the past years, especially in industrial processes due to the negative impact of unexpected shutdowns, as well as the need for the development of more effective maintenance strategies. The latter has been marked by a shift in maintenance paradigm from Preventive Maintenance (PM) to Condition Based Maintenance (CBM). CBM relies heavily on the existence of monitoring parameters that can reflect the condition of the system. Depending on the system, different measurements can carry potential information for its condition. For the case of systems involving bearings (a typical construction is presented in Figure 4), the most widely used systems for fault detection and diagnosis involve vibration monitoring, even though other measures can enhance the performance, requiring, however, more expensive equipment [8] [9].

Therefore, numerous feature extraction methods have been proposed. Some of them rely on time analysis of the monitored signal, while others involve frequency analysis, since the occurrence of the fault is followed by the occurrence of characteristic components in the frequency spectrum.

Figure 4: Elements of a rolling bearing

3.1.1 Setting up the solution

Intelligent Fault Diagnosis and Prognosis (IFDP) procedures for rotating machinery, covering data acquisition, signal processing, feature extraction, fault detection, and prognosis. They highlight challenges like operational variability and machinery complexity that affect diagnostic model accuracy. Future directions include hybrid models, robust algorithms, and the use of big data for real-time monitoring and predictive maintenance [10].

Fault-diagnosis method using Singular Value Decomposition (SVD) and Variational Mode Decomposition (VMD) entropy-based feature extraction with machine learning. Vibration signals are decomposed into intrinsic mode functions (IMFs) via VMD, then significant singular values are extracted using SVD. The entropy of these values serves as features for fault classification using machine learning models like Support Vector Machines (SVM) [11].

Some learning techniques, like NNs, have proven to be good in fault diagnosis in rotating machinery. While CNNs can extract the spatial features from the vibration signal hierarchically and automatically, RNNs, especially LSTM, can capture the time series dependence efficiently. Research evidence indicates that CNNs can handle spatial characteristics to achieve high diagnostic accuracy, and RNNs can model temporal data to improve fault prognosis and maintenance efficiency [12].

3.1.2 Dataset description

At the company during our internship, there was a lack of adequate vibration data we needed for this specific theme, showing the value of our solution for these companies where they will start giving this kind of data much importance for future uses. After doing some deep research, we found a trustworthy open-source named vibrobox [13] that provides the useful dataset required.

The dataset contains 40 signals of vibration of rolling bearing 6213. Shaft rotation speed was 973 RPM for all observations. Half of signals contain vibration of health bearing, and the other half records contain vibration of an isolated bearing with severe defect of outer ring and incipient defect of inner ring. Vibration acceleration was recorded using B&R sensor (Figure 5) fixed using a stud with sample frequency 96000 Hz (Figure 6). Resonant frequency is 10 kHz.

Figure 5: Acceleration sensor (B&R Automation)

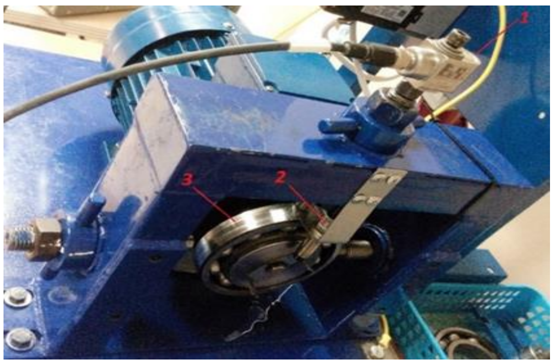


Figure 6: Testing stand: 1 – accelerometer; 2 – tachometer; 3 – diagnosed bearing

3.2 Definition of Bearings and Their Role in Machinery

Bearings are mechanical components that constrain relative motion and reduce friction between moving parts of a machine to achieve smooth operation. They support rotating shafts and reduce the friction between the surfaces in contact.

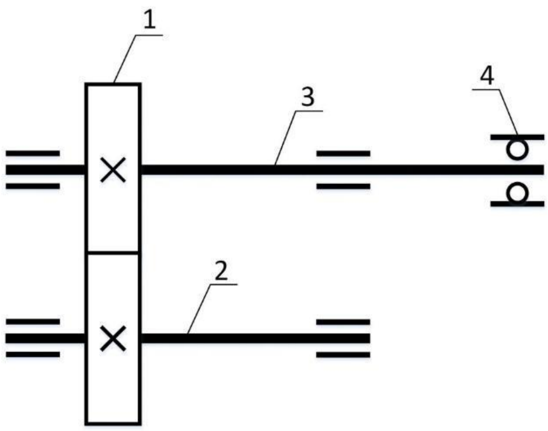
Bearings are forms of mechanical elements that restrict relative motion and reduce friction between moving parts of a machine in order to give a machine efficiency. They support rotating shafts and reduce the friction between the surfaces in contact.

Bearings are a critical element of machinery, as they support the frictionless movement of parts (types of bearings in Figure 7), which leads to a decrease in wear, improvement in performance, and a longer lifespan of equipment. The basic working principles of bearings involve load distribution, friction reduction, alignment maintenance, and support for rotational motion [14].

Figure 7: Bearings types

A schematic plan for a machine is essential as it offers a detailed representation of components and their interconnections, aiding in troubleshooting, maintenance, and repairs.

The schematic plan of the machine used in this dataset is resumed in Figure 8.



1	gearing001	Gear (z_1 = 56, z_2 = 34)
2	shaft002	Shaft
3	shaft001	Shaft
4	bearing001	Rolling bearing 6220

Figure 8: Schematic plan of machine

3.2.1 Bearing frequencies

They are specific vibration frequencies associated with the different components of a bearing during its operation. Monitoring these frequencies helps in diagnosing bearing conditions and predicting failures. (figure 9, Table 1)

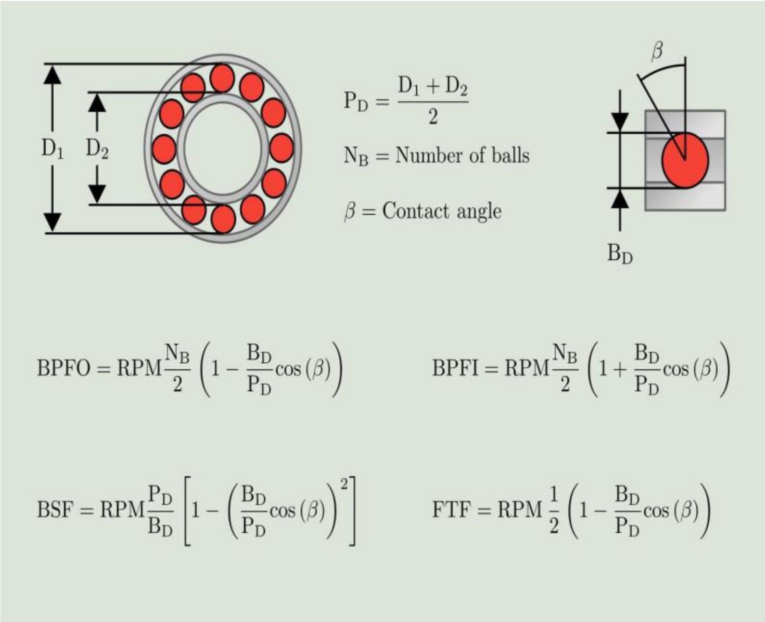


Figure 9: Rolling element bearing frequencies

BPFO	Ball Pass Frequency Outer Race
BPFI	Ball Pass Frequency Inner Race
BSF	Ball Spin Frequency
FTF	Fondamental Train Frequency

Table 1: *bearing frequencies*

To use bearing frequencies in signal analysis, monitor the vibration spectrum of the machinery and identify peaks at the calculated bearing frequencies (BPFO, BPFI, BSF, and FTF). Peaks at these specific frequencies indicate potential defects in the corresponding bearing components, allowing for early diagnosis and preventive maintenance. (Figure 10)

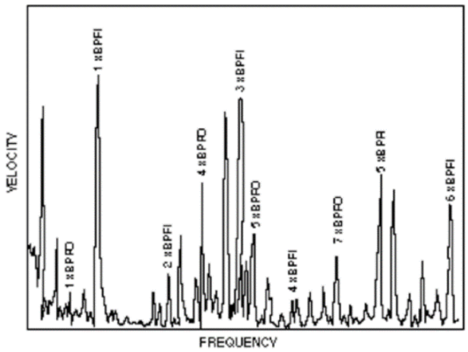


Figure 10: peaks that appear synchronous to x1 RPM but are non-synchronous

3.2.2 Importance of Vibration Analysis

Monitoring of bearing vibrations is key for predictive maintenance because it helps in early detection of the wearing processes and potential failures. Thus, timely interventions prevent catastrophic breakdowns, which reduce repair costs and downtime of machines [15].

3.3 Introduction to MATLAB

MATLAB is a strong environment for data analysis, signal processing, and its features provide extensive libraries with inbuilt functions for numerical computations, plotting visualization and developing algorithms. Its applications and use have pervaded through engineering or scientific programs because it is crucial in handling complex mathematical operations, modeling simulations, and the analysis of large datasets in a convenient manner, hence a necessary tool for both researchers and engineers [16].

3.3.1 Signal Factors and Their Importance

a. Root Mean Square (RMS):

- **Definition:** RMS finds the magnitude of a varying signal by taking the square root of the mean of the squares of the signal values.

$$RMS = \sqrt{\frac{1}{n}\sum_{i} {x_i}^2}$$

- **Importance:** RMS provides a measure of the overall energy in the vibration signal, in other words, how intense the mechanical activity is, and if there are areas or potential faults in the system [17].

b. Kurtosis :

- **Definition:** It is a measure of how much of the data lies in the tails and peaks of a probability distribution compared to the normal distribution.

$$Kurt = \frac{\mu_4}{\sigma^4}$$

- **Importance:** Higher kurtosis values are often indicators of open faults or other deviations as they have irregular shapes, mainly bearing defects [18].

c. Skewness :

- **Definition:** It is a measure of the asymmetry of a probability distribution around its mean.

$$\tilde{\mu}_3 = \frac{\sum_i^N (X_i - \bar{X})^3}{(N-1)*\sigma^3}$$

- **Importance:** Skewness is a measure that allows the operator to observe irregularities in the functioning reflected by the data that is overloaded to one side more than the middle, away from the normal operations, and thus detects faults or failures, or anomalies [19].

3.3.2 Why These Factors ?

These factors are key indicators for analyzing vibration signals. Normal signals exhibit low RMS, low kurtosis, and minimal skewness, while faulty signals show higher factors's deviations.

3.4 Fourier Transform (FFT)

The Fourier Transform is a mathematical method that changes a time-domain signal into its frequency components for analysis of the signal's frequency spectrum. Frequency domain analysis is very important for bearing fault detection, because it identifies specific frequency patterns and anomalies related to different types of bearing defects that may usually remain hidden in a time domain analysis [20].

3.4.1 Implementing FFT in MATLAB

After we upload our raw signal in an audio format (.wav), we apply the “audioread” function. Then we plot the vector to get the noisy signal (Figure 11), we use any filtering function (we used butter) to get a filtered signal (Figure 12) to remove unwanted noise or specific frequency components, enhancing the clarity and quality of the desired signal.

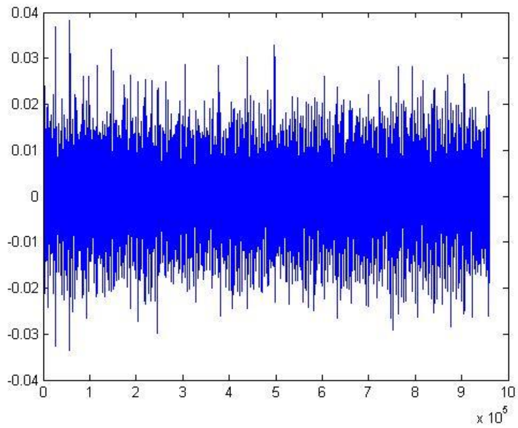


Figure 12: Noisy signal

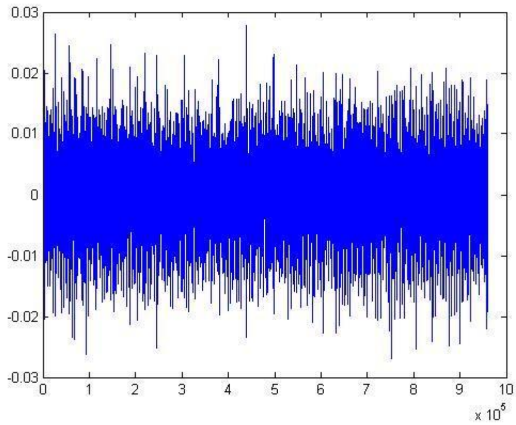


Figure 11: Filtered signal

After filtering the data and applying fft to locate the hidden information and patterns,by applying the function we can clearly 0see the difference between a normal (Figure 13) and faulty signal (Figure 14) amplitudes

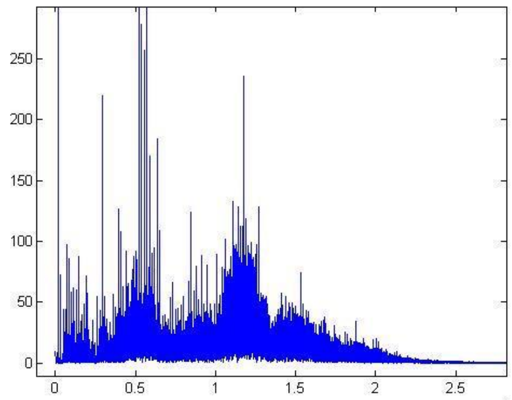


Figure 13: healthy bearing signal

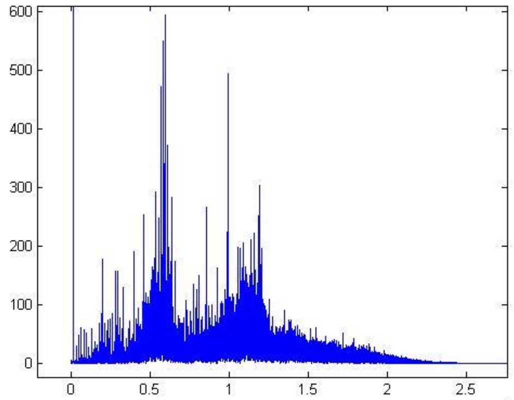


Figure 14: faulty bearing signal

3.5 Envelope Analysis

a. Definition:

Envelope analysis is a technique that extracts the modulation from a signal by demodulating it to isolate the slowly varying envelope (Figure 15), highlighting changes in amplitude over time [21].

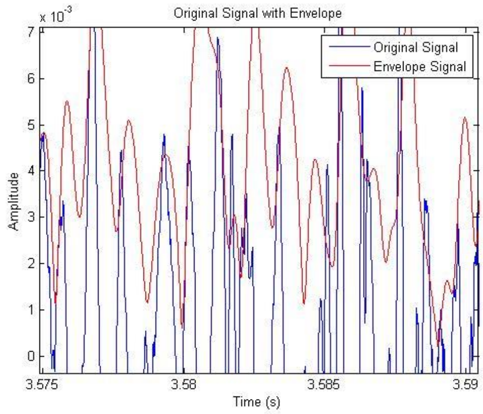


Figure 15: Close view of the envelope graph

b. Importance:

Envelope analysis is quite effective in finding fault frequencies in bearings as the amplitude modulation factor of the fault and its harmonics is evidently amplified in comparison to the sum signal, which in this way enhances the clarity, intelligibility, and visibility of the fault signal over noise.

c. Transition from Time Domain to Frequency Domain:

Moving to frequency-domain analysis from the time-domain requires the application of techniques such as Fourier Transform to change the signal based on time into frequency components to be able to identify certain frequency peaks corresponding to the fault frequencies.

d. How Envelope Analysis Enhances Fault Detection:

Envelope analysis improves fault detection by filtering out the modulation generated by the faults. It enhances the envelope of the fault signal in the spectrum and renders it to be more distinctive and separated from the background noise, to achieve an accurate fault diagnosis and prognosis. Following is the code example:

```
% Filter the signal (example: low-pass Butterworth filter)
fc = 2000; % Cut-off frequency (Hz)
[b, a] = butter(6, fc/(Fs/2), 'low');
filtered_signal = filtfilt(b, a, x);

% Compute the envelope signal using Hilbert transform
envelope_signal = abs(hilbert(filtered_signal));

% Compute the one-sided envelope spectrum
N = length(envelope_signal);
envelope_spectrum = fft(envelope_signal);
one_sided_envelope_spectrum = abs(envelope_spectrum(1:N/2+1)); % One-sided FFT
frequencies = Fs*(0:(N/2))/N;
```

Figure 16: Envelope Matlab code

We can clearly see the difference with applying envelope on a normal (Figure 17) and faulty (Figure 18) signal, especially in term of density, also the envelope of a normal bearing signal remains relatively stable, while a faulty bearing signal's envelope displays distinct amplitude fluctuations indicative of fault frequencies.

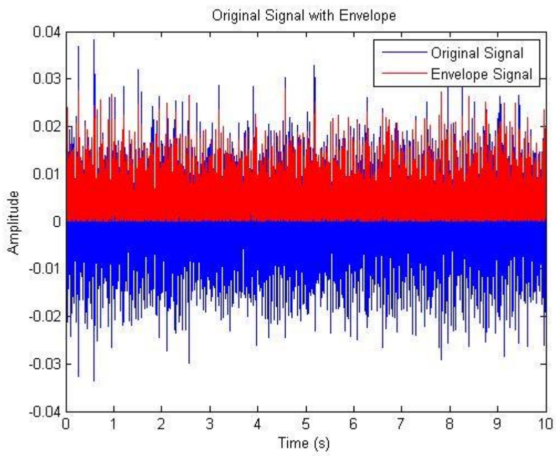


Figure 17: Envelope of healthy bearing signal

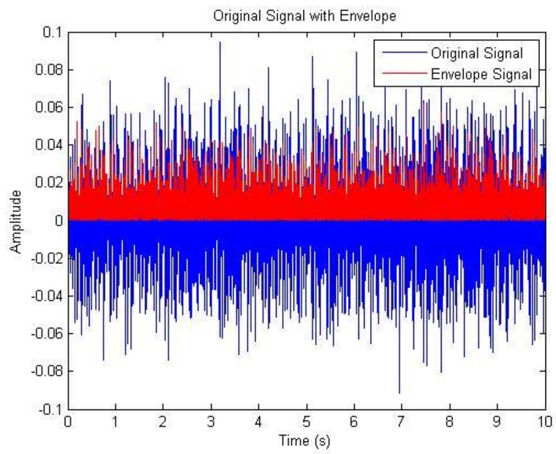


Figure 18: Envelope of faulty bearing signal

3.5.2 Code Development and Challenges

Modular approach:

Passing through many steps to find the best and most accurate function from the official website of Matlab, we managed to get these factors functions written in one single matrices (Figure 19)

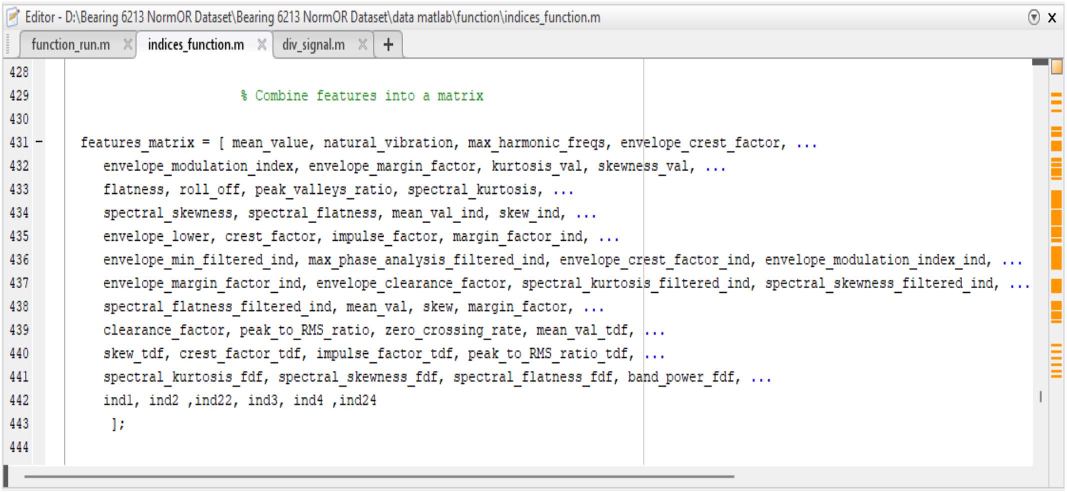


Figure 19: factors function

We faced some challenges during the development of these functions. Starting from keeping the right order of running the code, to handling large signals length compared to limited datasets samples. So the optimal solution was to divide each signal into 10 parts (Figure 20) to gives 10 times dataset for the final matrix, then we combined all these codes in one single code (Figure 21).

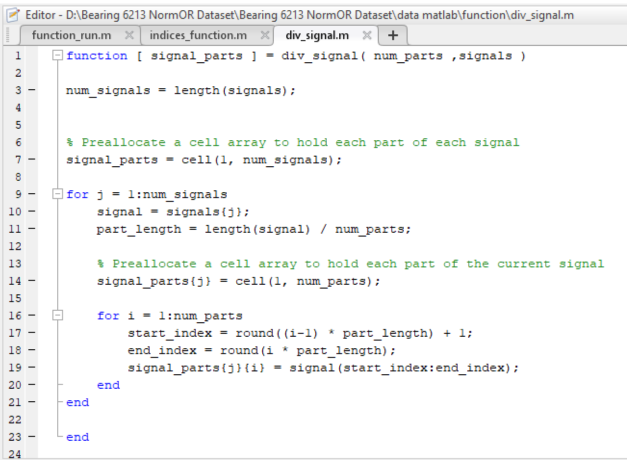


Figure 20: Code of signals division

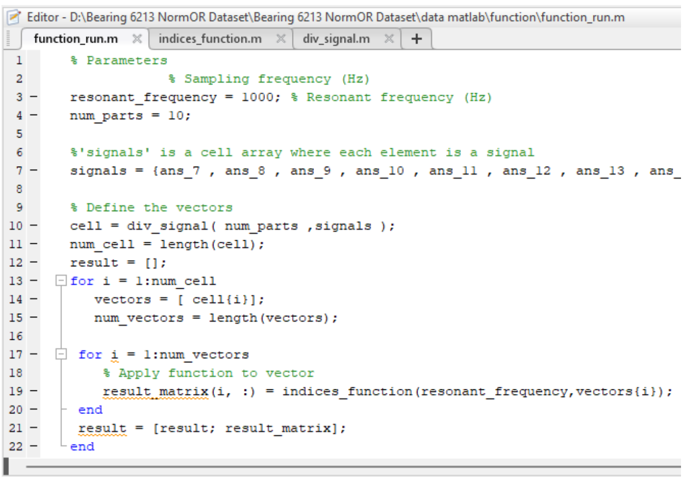


Figure 21: Code that run all functions

3.5.3 Dataset result in csv format

After applying the code and do all the necessary work, we can finally get our dataset ready:

kurtosis_val_f	skewness_val	envelepe	...	peak_to_RMS	spectral_kurt	spectral_skew	BPFO	BPFI	class
31132.0716	159.837537	4.966228	...	8.080014	14.887448	3.407387	0.000129	3.60E-05	faulty
34626.9072	172.317289	3.761079	...	5.447879	14.347949	3.351541	0.000101	3.30E-05	normal
33471.1167	168.430236	3.639393	...	5.190673	14.304512	3.347938	7.00E-05	5.00E-05	normal
34093.6322	170.525287	4.04928	...	7.019363	14.217246	3.302458	0.000109	2.30E-05	normal
35508.7058	175.557205	3.769427	...	5.98207	14.334132	3.40012	0.000101	3.40E-05	normal
31370.2455	160.517887	6.848026	...	9.644706	13.20367	3.00801	2.90E-05	4.90E-05	faulty
34677.5632	172.699397	3.742084	...	5.247459	14.348925	3.343052	8.70E-05	1.70E-05	normal
34629.9005	172.536492	3.332883	...	5.713925	13.645919	3.263371	6.10E-05	5.50E-05	normal
33347.1916	168.024809	3.700889	...	5.55993	14.315443	3.334097	7.60E-05	9.00E-05	normal
34066.3512	170.362644	3.1186	...	5.364225	13.848244	3.202492	0.000123	5.60E-05	normal
33700.3392	168.954914	3.683848	...	7.671962	14.81066	3.409813	3.00E-05	3.40E-05	faulty
30358.1118	157.216978	4.665132	...	7.062409	14.94017	3.455349	0.000186	4.60E-05	faulty
33298.8624	167.599343	3.853457	...	5.804515	14.904349	3.376896	0.000106	3.90E-05	faulty
34408.9124	171.943555	3.889764	...	5.04	14.275963	3.3	8.70E-05	2.10E-05	normal
32480.5085	164.815253	5.023917	...	6.346193	15.108049	3.432567	9.80E-05	4.50E-05	faulty
33991.3053	170.217747	3.562094	...	5.394485	14.101469	3.295234	5.40E-05	1.70E-05	normal
33533.2866	168.250928	3.954141	...	5.902219	15.007401	3.381355	2.40E-05	5.60E-05	faulty
34068.1807	170.417596	4.229826	...	7.034896	14.072011	3.27513	6.90E-05	6.20E-05	normal
34168.5866	170.600094	4.676243	...	7.480687	14.447287	3.24565	5.10E-05	4.70E-05	faulty
33132.313	166.948089	4.084054	...	8.722206	14.62035	3.341988	3.60E-05	2.50E-05	faulty
34176.2703	170.748035	4.077696	...	6.065798	14.296103	3.238927	0.000127	3.20E-05	faulty
33748.7254	169.156326	4.692582	...	6.105282	14.943334	3.392184	3.00E-05	3.30E-05	faulty
34122.4656	170.494899	5.188646	...	6.069234	14.301599	3.330963	0.00017	2.60E-05	normal
34067.4937	170.487592	3.863662	...	6.253562	14.209387	3.311542	0.00011	5.30E-05	normal
34000.6354	170.140466	3.907851	...	6.614213	13.953842	3.243648	2.90E-05	2.30E-05	normal
33481.3468	168.161705	4.078036	...	7.898118	14.135237	3.253416	2.30E-05	2.60E-05	faulty

Figure 22: dataset in csv format

3.6 Results and Analysis

Looking at the range of results obtained for each indicator, we can tell how highly accurate they are, and how valid and ready to use for the next step.

The code automates the process to produce the matrix efficiently. While the final matrix is saved in CSV format, adding to the last column an empty indicator called ‘class’ to be filled later with the correct class corresponding to the signal either normal or faulty to create a new dataset.

3.7 Conclusion

In conclusion, this chapter successfully demonstrates the execution of advanced signal processing techniques to analyze vibration bearing signals with MATLAB. We found crucial features such as RMS, kurtosis, and skewness employing open-source data and enhanced problem detection with FFT and envelope analysis. The modular approach to code development, along with integrating these methods into a comprehensive script, resulted in greatly enhanced diagnostic accuracy. Splitting the data into manageable segments made for more efficient processing. The findings show the efficacy of these techniques in preventive fault diagnosis for rotating machinery. Future research can focus on improving these methods and investigating new data sources.

CHAPTER 4

CHAPTER 4 The proposed AI solutions to recognize faults

4.1 Introduction

Automating the process of detecting faults in turning machinery is crucial to ensure the reliability and efficiency of rotating machinery. This chapter brings proposed solutions and experiment results for the detection of default using machine learning techniques. Our target was to develop an automated system that will detect the 'Bearing' either 'Normal' or 'Faulty' by looking at the vibration patterns of the dataset.

4.2 Tools used

For this stage of our work, we chose Python for coding, which is a high level programming language known for its simplicity and readability. It is the most extensively employed language in machine learning due to the huge number of libraries and frameworks provided by it, for handling data processing, model building, and carrying out evaluation, which shortens the machine learning workflow very quickly. These are Python's libraries used for our model: (Figure 23)

final code.py

```
import tkinter as tk
from tkinter import ttk
from tkinter import filedialog
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.svm import SVC
from sklearn.neural_network import MLPClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, matthews_corrcoef
```

Figure 23: libraries used

4.3 Methods used:

4.3.1 Random Forest Classifier

Definition: Random Forest Classifier (RFC) is an ensemble learning method for classification, regression, and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) of the individual trees. It is particularly useful for its robustness against overfitting and its ability to handle a high-dimensional feature space effectively [22].

Importance in Fault Detection: Random forests are highly effective in carrying out fault detection due to their ability to handle large datasets with higher dimensionality and give good accuracy. They also help in identifying the most important features contributing to fault detection, an important aspect of understanding the underlying mechanics of bearing faults [22].

4.3.2 Decision Tree Classifier

Definition: Decision Tree Classifier is a classification algorithm that is a part of the family of supervised learning algorithms. It is a flowchart-like structure in which each internal node represents a test on an attribute (a feature), where each branch represents an outcome of the test, and each leaf node represents a class label [23].

Importance in Fault Detection: Decision trees are an easily interpretable and visual way of understanding the decision-making process in fault detection. They can work with both numerical and categorical data surprisingly fast, giving clear insights for the fault detection process, for small to medium-sized datasets. Furthermore, their computational cost is relatively small compared to other types of classifiers [23].

4.3.3 Support Vector Classifier (SVC)

Definition: Support Vector Classifier, used for classification problems, is one of many versions of Support Vector Machines. Basically, what it does is to find a hyperplane in the N-dimensional space that distinctly classifies the data points. To better generalize the model to unseen data, a principle is set, which is known as margin maximization of the data points [24].

Importance in Fault Detection: Due to the ability of SVC to work in high-dimensional spaces or where the number of dimensions exceeds the number of samples, it becomes effective in fault detection among others. This algorithm is further hard to overfit, especially in cases of high feature dimensional spaces that characterize vibration signal data space [24].

4.3.4 Neural Network

Definition: Neural Network is a computational model developed with the motivation of imitating the working of a human brain formed of layers of interconnected nodes (neurons). A neural network can learn complex to simple patterns and representations from the data by using the process of back-propagation. Neural networks work wonders while implementing them on large datasets, as well as for intricate patterns [25].

Importance in Fault Detection: Neural networks, especially deep learning models, have proved to be particularly effective in fault detection, because of their ability to learn the features directly from the raw data on their own. This is of explicit importance in bearing signal analysis, where they extract useful patterns embedded in vibration signals, leading to effective and accurate equations for reliability and proper fault detection [25].

4.4 Evaluation of Proposed Solutions

We applied 10-fold cross-validation in our dataset analysis to assess the performance and generalization ability of machine learning models. This provides an estimate in which the data is split into 10 subsets (folds), whereby the general model is trained on 9 of them and tested on 1, thus reducing the possible risk of overfitting. Providing a robust estimate of model performance over different subsets of the data hence reduces the risk of overfitting, which eventually gains reliability in assessing how the model will perform on previously unseen data.

The evaluation of these algorithms is defined based on their accuracy, precision, recall, and F-measure, MCC. Accuracy is the proportion of correct predictions, precision is the proportion of true positives among predicted positives, recall is the proportion of true positives among actual positives, F1-score is the balance between precision and recall, and MCC measures the correlation between predicted and actual classifications. To better explain their job, we apply their formulas on an example of the confusion matrix in this figure below:

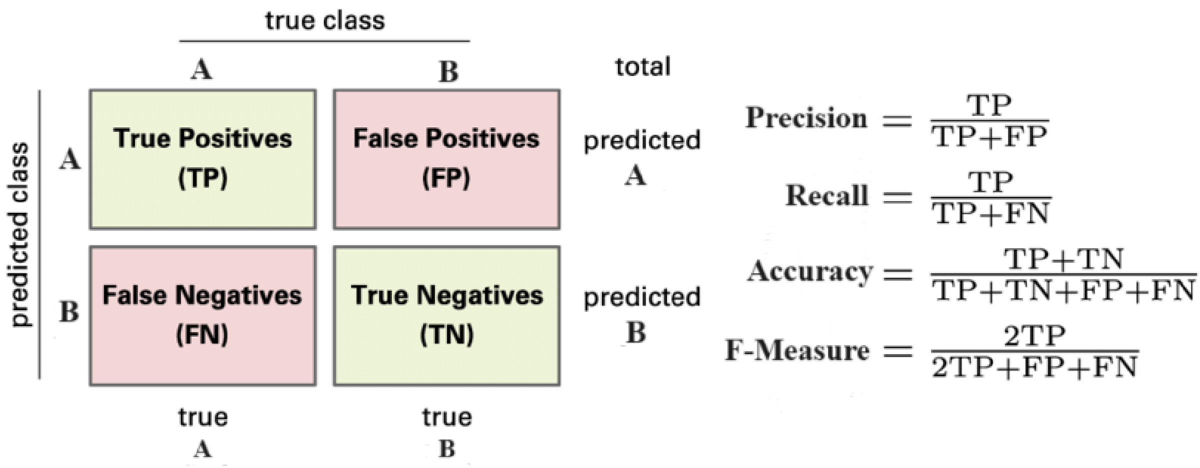


Figure 24: confusion matrix

Now after setting everything up, the work on the application starts, and by making the most accurate lines in the code, finding the best features for the algorithms, we managed to create an application that tests any dataset based on the training dataset it learned. That it gives readable results in a fast way and even for many lines in the test datasets. It describes all the necessary evaluation details for each algorithm, and also identifies each class with a high prediction. It only requires a dataset in a CSV format and the class names ‘class’ for it to work well (Tab.2).

Algorithm	Random Forest	Decision Tree	SVM Metrics	Neural Network
Accuracy	99.72%	98.83%	99.44%	99.72%
Precision	99.72%	98.36%	99.52%	99.61%
Recall	99.72%	98.89%	99.39%	99.72%
F1-Score	99.72%	98.33%	99.44%	99.82%
MCC	0.99	0.97	0.98	0.99

Table 2: evaluation pourcentages of the results

We created an organized interface in Python's application using Tkinter library (Figure 25) (Figure 26)

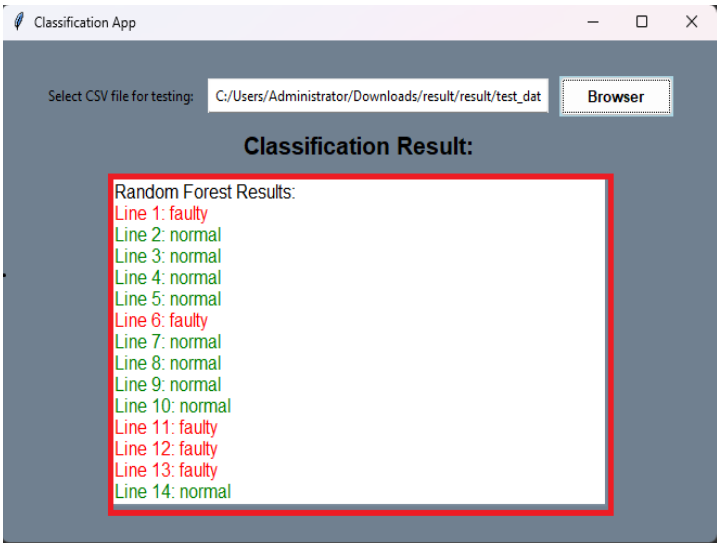


Figure 25: interface of prediction

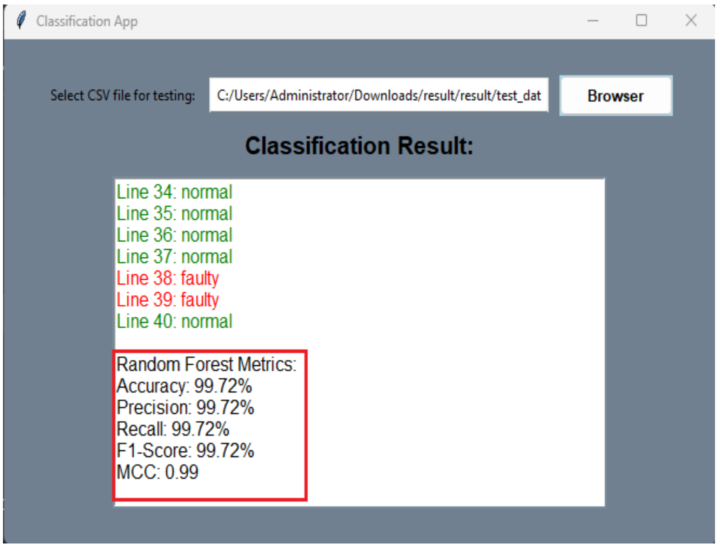


Figure 26: algorithm evaluation pourcentage

The performance of random forest was the most predictive one without the need for any extra parameters, however after adding the right settings to the others they succeeded in detecting the faulty signals among the testing dataset.

The results are easy to read by anyone, there is no need to do a whole tiring work just to detect whether this signal is faulty or not and takes more time and energy and maybe not even get it right at the end.

4.5 Analysis of Experimental Results

The experimental results showed that all four methods were effective: decision trees and random forests with ensemble learning; SVM-based on kernel methods; and RNNs. Each of the methods achieved high accuracy in rotating machinery fault detections based on the dataset and, therefore, we can use all the results together for the verification. The obtained results hint that an ensemble of traditional techniques and deep learning could provide the intended objective for reliable fault detection solutions with industrial implementation. These methods would be implemented in the industry fields to prevent downtime and high maintenance costs since the safety and productivity of the working process will also be enhanced.

4.6 Discussion of Results

At this point in the review of our results, we can see that for most tested cases, SVMs, along with decision trees, RNNs and random forests, have provided good performance. Each method presented very good performance in the identification of faults in rotating machinery, meaning that methods are ready for real-world applications. However, the strengths and weaknesses of each approach have to be recognized. These results are relatively robust for SVMs, decision trees, and random forests, only slightly falling behind RNNs when the patterns are difficult. The recognition of this, leads us to find space for further work on the optimization of parameters and hybrid approaches. Going forward, we are focused on the optimization of the methodology for the detection of faults in order to satisfy the demands of industrial applications.

4.7 Other datasets

In order to ensure the reliability of the application and the algorithms used, we applied them on other 4 datasets with different features or types of defects.

Applying our matlab code on Dataset 1 and Dataset 4 to get the same features as our main dataset, while the classes here are 4 (inner race defect, outer, ball and normal). Dataset 2 is the same in classes but has only the main features such as RMS, Kurtosis, and Skewness. Dataset 3 is about the lubrication of the bearing and how features like temperature will define the class either normal or got less or excess of grease.

We put the data in our application for train and test using same method just to make sure of the test for better results and we got these satisfying results in the table below:

Algorithm/ dataset		Dataset 1	Dataset 2	Dataset 3	Dataset 4
Random forest	Accuracy	0.9899	0.9474	0.9966	0.8958
	Precision	0.9903	0.9473	0.9831	0.8968
	Recall	0.9899	0.9474	0.9831	0.8958
	F1-Score	0.9974	0.9473	0.9831	0.8959
	MCC	0.9848	0.9291	0.9743	0.8188
Decision tree	Accuracy	0.9899	0.962	0.9967	0.9583
	Precision	0.9901	0.9622	0.9772	0.962
	Recall	0.9899	0.962	0.9767	0.9583
	F1-Score	0.9898	0.962	0.9781	0.9583
	MCC	0.9847	0.9488	0.9639	0.9302
SVM	Accuracy	0.9667	0.9576	0.9417	0.9667
	Precision	0.9858	0.9875	0.9721	0.9432
	Recall	0.9667	0.9576	0.9417	0.9667
	F1-Score	0.9693	0.9377	0.9693	0.9522
	MCC	0.9804	0.9794	0.9654	0.9475
Neuronal network	Accuracy	0.9848	0.9509	0.9417	0.9767
	Precision	0.9883	0.9743	0.9624	0.9758
	Recall	0.9848	0.9509	0.9617	0.9667
	F1-Score	0.9721	0.9485	0.9582	0.9647
	MCC	0.9526	0.9461	0.9754	0.9567

Table 3: diffrent datasets's results

[1] We can clearly notice how accurate methods are on multiple datasets, and the only thing that should be edited is the dataset for training, we simply change it from this part of the code (Figure 27).

```python
def load_dataset(self):
    # Load your dataset CSV file
    self.dataset = pd.read_csv("train_data.csv")  # Replace with your dataset file path
```

Figure 27: loading dataset for training

4.8 Confirmation and tips

The application runs perfectly and smoothly, with a simple and trustworthy appearance. Keeping some work with matlab and some with python was required since the code that extract features works better on matlab but there is almost no work to do except loading the data in. mat format and adding them the code, will provide the required matrix, after saving it in CSV format, the normal signal of the datasets should be named as ‘normal’.

4.9 Conclusion

Conclusively, this chapter showed the assessment achieved by the proposed solution of detecting defaults in rotating machinery. It further established that the machine learning tools, such as, decision trees, random forest, CNNs, and RNNs, performed better when it comes to accuracy and effectiveness compared to traditional techniques. In addition, challenges for future research and development shall be seen in the application of these methods under realistic boundary conditions and the development of more sophisticated and efficient methods for fault detection: either in simplifying them further, making the path shorter, increasing the predictive power to 100%, or through the parameter optimization process, or even exploring hybrid approaches to make them reliable and adaptive over a wide range of operating conditions. By meeting these challenges, we will realize more efficient fault detection strategies in reducing downtime, improving safety, and ultimately boosting productivity in industrial applications.

GENERAL CONCLUSION AND PERSPECTIVES

5.1 General conclusion and perspective

We have successfully addressed the major challenges faced by industrial maintenance processes in industrial companies, such as unexpected downtime, poor resource management, and high costs, in this dissertation. This phenomenon is well represented in an industry because of the absence of an appropriate management and maintenance system, which results in exorbitant costs incited from costly production interruptions. Our project focuses on the development of innovative approaches in the optimization of preventive maintenance processes through the implementation of a complete solution integrating machine learning techniques.

In Chapter 3, we focused on advanced signal processing techniques to analyze vibration bearing signals using MATLAB. We extracted crucial features like RMS, kurtosis, and skewness from open-source data and enhanced problem detection with FFT and envelope analysis. By developing a modular code structure and splitting the data into manageable segments, we achieved greater diagnostic accuracy and efficiency.

In Chapter 4, we presented the results from experiments performed to evaluate proposed solutions for the rotating machinery fault-detection problem. The experiments proved that the proposed SVMD entropy-based solution, decision trees, random forests, CNN, and RNN solutions are more efficient and outperform traditional ones in terms of accuracy. This really proves the superiority of advanced techniques in real-world scenarios and opens the door for further research along some pathways like simplification of the methods, parameter optimization, and the use of hybrid methodology.

All these approaches are suitably validated with extensive experiments and simulations with multiple bearing datasets discussed into chapter 4, which justifies the potential of the proposed methodology over existing approaches in terms of achieving satisfactory theoretical results compared to experiment.

5.2 Achievement of project objectives

We have successfully implemented a Computerized Maintenance Management System (CMMS) that digitizes the processes involved in maintenance, thereby simplifying and modernizing the workflow. We developed a predictive maintenance application meant to predict rotating machine failures by analyzing only the bearing's signals, through the integration of artificial intelligence techniques, specifically machine learning methods. Such proactive monitoring enables maintenance teams to eliminate problems before they lead to costly accidental downtime or complicated, unnecessary extra work.

The financial and operational impacts of our system demonstrate all-around improvements in maintenance. This predictive system has led to a considerable reduction of unplanned downtimes, improved operational efficiency of the maintenance team, and reduction of maintenance-related costs.

5.3 Importance of the project

The relevance and importance of this project to the field of maintenance management cannot be overemphasized. The proposed solution allows one to cope with priorities of the industry, it facilitates the work, reduces downtimes, cuts operational costs by optimizing maintenance work, and improves the efficiency of maintenance teams. The use of AI in predictive maintenance also prolongs the lifespan of equipment, improves company image, and enhances competitiveness in the market.

5.4 Future directions

Looking forward, there are various directions for further research and optimization to enhance the capabilities of our maintenance system. Future works should focus on :

- Real-world implementation: Deploying the developed system in various industrial settings to validate its effectiveness across different environments and equipment types.
- Refinement of techniques: Continuously training machine learning models to improve the detection of equipment failures in terms of both accuracy and capability, and especially under complex and variable operating conditions.
- Hybrid approaches: Combining multiple machine learning techniques to collect sets of strengths and improve the robustness of fault detection.
- Advanced predictive analytics: More advanced predictive analytics should be integrated to forecast not only close failures but also maintenance needs and trends that may emerge in more extended periods of time.
- Scalability: The design of the system shall scale to handle the massive data and extensive networks of equipment in the extensive spatial area of a large industrial plant.

Further innovations in our approach will continue to incrementally increase the reliability and efficiency of maintenance processes to eventually embody a sustainable increase in the productivity of industrial operations. Our project paves the way for a future in which maintenance management becomes one of the most proactive applications of advanced AI technologies, moving predictably.

References

[1] KARABADJI, Nour El Islem, SERIDI, Hassina, KHELF, Ilyes, et al. Improved decision tree construction based on attribute selection and data sampling for fault diagnosis in rotating machines. Engineering Applications of Artificial Intelligence, 2014, vol. 35, p. 71-83.

[2] KARABADJI, Nour El Islem, KHELF, Ilyes, SERIDI, Hassina, et al. A data sampling and attribute selection strategy for improving decision tree construction. Expert Systems with Applications, 2019, vol. 129, p. 84-96.

[3] THUAN, Nguyen Duc et HONG, Hoang Si. HUST bearing: a practical dataset for ball bearing fault diagnosis. BMC research notes, 2023, vol. 16, no 1, p. 138.

[4] SINGH, Vikas, GANGSAR, Purushottam, PORWAL, Rajkumar, et al. Artificial intelligence application in fault diagnostics of rotating industrial machines: A state-of-the-art review. Journal of Intelligent Manufacturing, 2023, vol. 34, no 3, p. 931-960.

[5] ZOUNGRANA, Wend-Benedo, CHEHRI, Abdellah, et ZIMMERMANN, Alfred. Automatic classification of rotating machinery defects usingz machine learning (ML) algorithms. In: Human Centered Intelligent Systems: Proceedings of KES-HCIS 2020 Conference. Springer Singapore, 2021. p. 193-203.

[6] LEE, Yong Eun, KIM, Bok-Kyung, BAE, Jun-Hee, et al. Misalignment detection of a rotating machine shaft using a support vector machine learning algorithm. International Journal of Precision Engineering and Manufacturing, 2021, vol. 22, no 3, p. 409-416.

[7] MENG, Huan, ZHANG, Jiakai, ZHAO, Jingbo, *et al.* A multi-scale feature extraction and fusion method for bearing fault diagnosis based on hybrid attention mechanisms. *Signal, Image and Video Processing*, 2024, p. 1-11.

[8] MCFADDEN, P. D. et SMITH, J. D. Vibration monitoring of rolling element bearings by the high-frequency resonance technique—a review. Tribology international, 1984, vol. 17, no 1, p. 3-10.

[9] AL-GHAMD, Abdullah M. et MBA, David. A comparative experimental study on the use of acoustic emission and vibration analysis for bearing defect identification and estimation of defect size. Mechanical systems and signal processing, 2006, vol. 20, no 7, p. 1537-1571.

[10] DAS, Oguzhan, DAS, Duygu Bagci, et BIRANT, Derya. Machine learning for fault analysis in rotating machinery: A comprehensive review. Heliyon, 2023.

[11] ZHANG, Lijun, ZHANG, Yuejian, et LI, Guangfeng. Fault-Diagnosis Method for Rotating Machinery Based on SVMD Entropy and Machine Learning. Algorithms, 2023, vol. 16, no 6, p. 304.

[12] ZHANG, Siyu, LEI, S. U., JIEFEI, G. U., et al. Rotating machinery fault detection and diagnosis based on deep domain adaptation: A survey. Chinese Journal of Aeronautics, 2023, vol. 36, no 1, p. 45-74.

[13] KECHIK, Daniil. Bearing 6213 norm/or dataset. Mendeley Data [online]. 19 October 2020. [Accessed 15 June 2024]. Available from: https://data.mendeley.com/datasets/fbf6y8m4mv/1

[14] OBERG, Erik, JONES, Franklin D., HORTON, Holbrook L., et al. Machinery's Handbook. New York: Industrial Press, 1914.

[15] SHIROISHI, J. Y. S. T., LI, Yawei, LIANG, Steven, et al. Bearing condition diagnostics via vibration and acoustic emission measurements. Mechanical systems and signal processing, 1997, vol. 11, no 5, p. 693-705.

[16] MathWorks. (2023). MATLAB: The language of technical computing. Retrieved from https://www.mathworks.com/products/matlab.html

[17] COHEN, Fred E. et STERNBERG, Michael JE. On the prediction of protein structure: the significance of the root-mean-square deviation. Journal of molecular biology, 1980, vol. 138, no 2, p. 321-333.

[18] RUPPERT, David. What is kurtosis? An influence function approach. The American Statistician, 1987, vol. 41, no 1, p. 1-5.

[19] ARNOLD, Barry C. et GROENEVELD, Richard A. Measuring skewness with respect to the mode. The American Statistician, 1995, vol. 49, no 1, p. 34-38.

[20] RANDALL, Robert Bond. Vibration-based condition monitoring: industrial, automotive and aerospace applications. John Wiley & Sons, 2021.

[21] MCFADDEN, P. D. et SMITH, J. D. Model for the vibration produced by a single point defect in a rolling element bearing. Journal of sound and vibration, 1984, vol. 96, no 1, p. 69-82. 69-82.

[22] BREIMAN, Leo. Random forests. Machine learning, 2001, vol. 45, p. 5-32.

[23] QUINLAN, J.. Ross . Induction of decision trees. Machine learning, 1986, vol. 1, p. 81-106.

[24] WIDODO, Achmad et YANG, Bo-Suk. Support vector machine in machine condition monitoring and fault diagnosis. Mechanical systems and signal processing, 2007, vol. 21, no 6, p. 2560-2574.

[25] GOODFELLOW, Ian, BENGIO, Yoshua, et COURVILLE, Aaron. Deep learning. MIT press, 2016.