DE-42 (MTS) Abdullah, Ibtisam

Condition Based Monitoring of Ball Bearings Using Machine Learning

COLLEGE OF
ELECTRICAL AND MECHANICAL ENGINEERING
NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY
RAWALPINDI
2024

COLLEGE OF ELECTRICAL AND MECHANICAL ENGINEERING

DE-42 MTS
PROJECT REPORT

Condition Based Monitoring of Ball Bearings Using Machine Learning

Submitted to the Department of Electrical Engineering
in partial fulfillment of the requirements
for the degree of

Bachelor of Engineering
in
Mechatronics
2024

Sponsoring DS:
Dr. Tayyab Zafar

Submitted By:
NS Muhammad Abdullah

NS Ibtisam Ahmed Kayani

Acknowledgements

First and foremost, we express our gratitude to Allah for granting us the strength and determination to undertake and accomplish this research work. We would also like to extend our heartfelt appreciation to our supervisors, Dr. Tayyab Zafar and Dr. Ahmad Rauf Subhani, for their unwavering support throughout the project. Their invaluable technical guidance and moral encouragement were instrumental in our success. They provided extensive expertise in model selection, feature engineering, and deployment, ensuring the project's progress and efficiency.

Additionally, we would like to acknowledge the contributions of Sir Usman Asad, whose expertise in mechanical design greatly influenced the project's direction. His feedback and guidance were crucial in optimizing the design and functionality of the test rig. We would also like to express our heartfelt appreciation to our families for their unwavering support and encouragement throughout this journey. Their moral support and prayers have been a constant source of strength and motivation.

Lastly, we extend our sincere gratitude to our friends for their unwavering moral support. Their presence and encouragement have played a significant role in our achievements. We would like to acknowledge all individuals who have contributed, directly or indirectly, to the successful completion of this Proof-of-Concept project. Their support and guidance have been invaluable, and we are truly grateful for their contributions.

NS Muhammad Abdullah

NS Ibtisam Ahmed Kayani

Abstract

Ball bearings are used in rotary machines such as turbines, generators, automobiles, and electric motors just to name a few. They operate in harsh conditions, heavy loads and shocks which deteriorate their health and if not monitored timely, can lead to costly damage. With the advancements in technology, various sensors are integrated to monitor machines' health. An important health indicator of such machines is vibration signal data which provides meaningful insight into a variety of mechanical faults. Traditionally, bearing faults are analyzed in time-frequency domain which is incapable to classify the types of faults accurately. In this project, we used vibration signal data acquired by Korea Advanced Institute of Technology and ourselves and coupled it with machine learning algorithms to better understand the principles of these algorithms and how they perform. Additionally, data acquired by our setup is tested on these algorithms for validation.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Rotatory machines constitute a large portion of machines under use by the industry. Power generation plants, automotive industry, mining plants, agricultural equipment, material processing, and product manufacturing industries depend on electric motors, generators, pumps, and turbines for operation. For smooth transfer of power and motion, bearings are deployed in these machines, and they undergo heavy loads operating in harsh environments. Consequently, 41 percent of the failures occur in bearings (Figure. 1) [1]. Hence, it is of utmost importance to monitor the health of these machines.

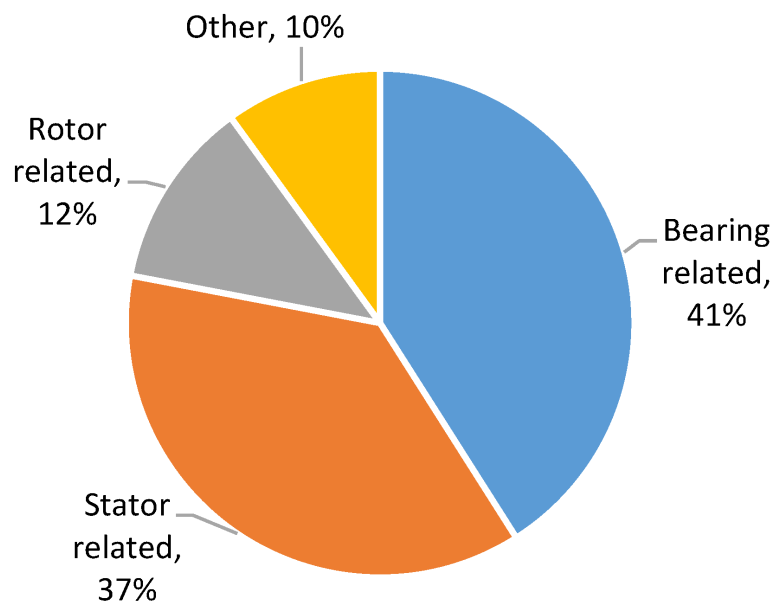

Figure 1. Percentage of failures by rotating machine components

Predictive maintenance and condition monitoring go hand in hand for machine maintenance and these two methods have proven to be the most useful ones for increasing reliability and life of the machines. Sensors provide data in real time which is analysed using model-based diagnosis systems and computer software. Microcomputer systems have decreased the footprint of the monitoring systems making health monitoring of the machines a less tedious procedure. Industry 4.0 has brought in the technical advancements of artificial intelligence coupled with traditional data acquisition techniques to modernize condition monitoring and fault diagnosis [2].

Industries across the globe rely on numerous integrated sensors which provide useful insights into the optimal operation and health of the machine. This raw data coming straight out of the machine contains rich information and features which can be processed and analysed using AI for accurate fault diagnosis. Organizations are now turning towards AI for less cost, better performance, and future proof solutions [3].

The primary goal of this project is to devise a system to study raw vibration signal data acquired from the bearings using a custom hardware and propose an efficient, modern artificial intelligence based system for fault classification of ball bearings. These enhancements are expected to bring maintenance costs down, reduce downtimes, increase both machine and worker safety, extend life and improve reliability of the machines. The objectives of the project are stated below.

- Develop a test rig for data acquisition from faulty ball bearings
- Preprocess and extract features from the data for training of machine learning algorithms
- Compare the performance of various machine learning algorithms such as Random Forest Classifier, Logistic Regressor, KNN, Support Vector Machine [4], and Multi Layer Perceptron and to validate the performance of the selected algorithm

It is important to establish the fact that condition-based monitoring has three essential stages: health monitoring, fault diagnosis and fault classification. The scope of this study is fault classification rather than the establishment of the whole condition monitoring system. Fault diagnosis and health monitoring are suggested as future work to formulate a complete condition-based monitoring system.

Initially, to understand the behaviour of the machine learning algorithms, vibration signal data from Korea Advanced Institute of Science and Technology will be used for training and validation of the model. Then, vibrational data will be acquired from an experimental setup developed by us and accuracy of the previously trained models will be tested. If the models fail to classify faults at an acceptable rate, the new data acquired from our setup will be used to retrain the model and the performance shall be reevaluated.

The thesis focuses specifically on the classification of the 6205 ball bearing faults that occur at the inner and outer races of the bearings using vibration signal data. The research includes the design and implementation of the experimental setup, feature extraction from vibration signals, and training, validation and evaluation of machine learning models.

The first chapter of the thesis introduces the reader to the problem, defines the approach towards the solution and states the objectives of the project. Chapter two provides a background of fault diagnosis methods for ball bearings. Chapter three discusses the dataset, features extracted from the data and an overview of machine learning algorithms. Chapter four details the hardware setup used for the data acquisition and the results based on the dataset are briefed upon in chapter five. Lastly, chapter six concludes the thesis with suggestions for future improvements.

CHAPTER 2: LITERATURE REVIEW

This chapter details the work done previously by researchers in the field of fault diagnosis. Various techniques that were used in the past such as Finite Element Analysis (FEA), mathematical models and signal processing have been tested with artificial intelligence to improve their accuracy and reduce complexity.

2.1 Bearing

A bearing is a mechanical device that constrains the relative motion of parts in a machine to only one desired direction. Bearings also help in reducing friction between moving parts. Their design help in free rotatory or linear motion along some fixed axis. They might also help in minimizing motion by controlling acting forces on parts of the machine. The name is derived from "to bear," as they allow one part to bear another.

By relative motion, bearings allow:

- Linear motion (drawers)
- Radial rotation (shafts)
- Spherical rotation (ball and socket joints)
- Hinge motion (doors)

There are several types of bearings plain bearings being the most common of them. Other types include rolling element, jewel, fluid, magnetic, and flexure bearings.

Rolling element bearings are classified into two categories by the element used for rotation (Fig. 2):

- Roller bearings
- Ball bearings

Ball and roller bearings have two races, outer and inner with the rolling elements in between them along a grooved path known as raceways.

Figure 2. Types of bearings

Bearings are named according to a set of specified rules showing their series, type, dimensions, design, and seal types. For example, for our project, we have decided to use 6205-2RS ball bearings as they are used more commonly than other types with their vibrational signal data available readily in substantial amounts. The digit six represents deep grooved ball bearings, digit two representing diameter series 2, 05 indicating a diameter of 25mm and 2RS showing that the bearing is sealed with rubber on both sides. Figure below helps in understanding the nomenclature better.

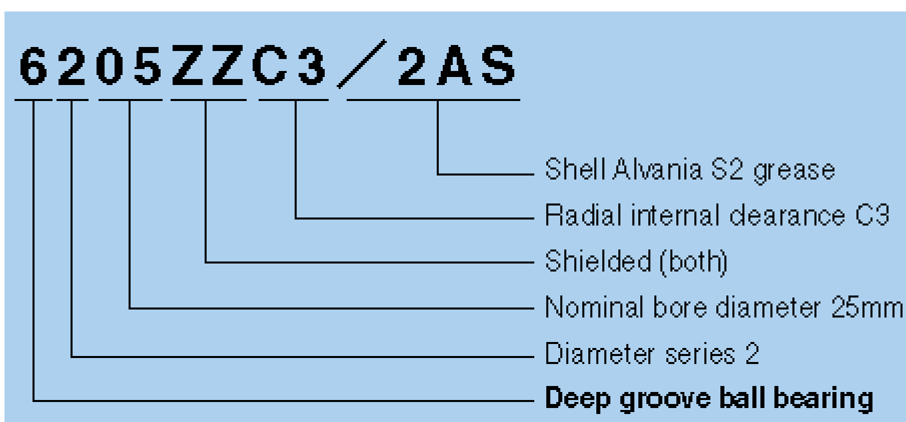

Figure 3. Nomenclature of bearings

To minimize friction, bearings are lubricated with oil or grease. Friction is a major factor governing the life cycle of a bearing and if timely lubrication is neglected, friction can cause severe thermal and mechanical degradation of the bearing. Lubrication can help in:

- Prolonging bearing fatigue life
- Reducing friction and wear

- Carrying heat away from within bearing due to friction
- Preventing corrosion and contamination by dirt

Grease, highly refined mineral oils, and synthetic oils are used to lubricate bearings. Mineral and synthetic oils generally have a wider operating temperature range as compared to greases. Oils are also combined with additives to enhance their rust preventive and oxidation inhibiting properties.

2.2. Types of maintenance strategies

A machine is a physical system that uses energy or forces, manipulates or converts them into other forms of energy to perform actions on the environment. From a wheel to combustion engines, motors to automobiles, airplanes, ships, rockets, robots, and computers, all these machines have proved to be the backbone of modern technology and advancements.

Machine maintenance is the act of keeping machines running smoothly with minimal downtimes. Machine maintenance can be done at regular intervals at scheduled times or in emergencies which are caused randomly and are out human control.

Maintenance can be categorized distinctively based on the approach used.

2.2.1. Reactive Maintenance

Reactive maintenance refers to emergency repairs as this sort of maintenance is done after the machine has failed and is one of the oldest approaches. This maintenance is rushed, unplanned, and unscheduled and is often referred to as “fighting fires.”

Run-to-fail maintenance is like reactive maintenance, with just the difference that the machine is let run to failure deliberately. There is usually contingency for such machines and skilled labour along with spare parts are ready to get the machines up and running again as soon as possible. Run-to-failure approach can be useful when the repairs are cheaper in comparison to scheduled maintenance plans. 61 percent of the manufacturing facilities used run-to-failure method for maintenance [5] as shown by Pelliccione in their study.

2.2.2. Routine Maintenance

Routine maintenance, as the name implies, is a regular, scheduled maintenance done at regular intervals and ongoing basis. Worn out parts might be tested, replaced, and lubricated based on their condition.

2.2.3. Corrective Maintenance

Corrective maintenance usually consists of smaller, non-invasive inspections only fixing parts, such as realignment of a shaft, before a complete failure occurs.

2.2.4. Preventive Maintenance

Preventive maintenance is a regular, scheduled maintenance done to avoid unexpected failures before they occur. Preventive maintenance can be split into time-based and usage-based maintenance. Time-based maintenance is done after regular intervals, for example at last day of every month. Usage-based maintenance is carried out after certain units consumed, like checking engine oil viscosity after every five hundred miles. 78 percent of the manufacturing facilities deploy a preventive maintenance plan.

2.2.5. Condition-Based Maintenance

Condition-based maintenance carried based on real-time data collected from the machine using sensors indicating actual health and condition of the machine. Condition-based monitoring is discussed in detail later in this report and lays the foundation for this project.

2.2.6. Predictive Maintenance

Predictive maintenance builds on the foundation laid by condition-based monitoring for predicting failures in real-time. Remaining useful life prediction is a major benefit of predictive maintenance.

2.2.7. Prescriptive Maintenance

Artificial Intelligence and Machine Learning techniques can be used to automate maintenance process. AI can prescribe optimal solutions against failures in real-time using data acquired from sensors.

A combination of above-mentioned plans can also be used for better and more efficient maintenance according to our needs and applications.

2.3. Condition-Based Monitoring

As reviewed by Surucu et al. [6], progress made in the field of microcontrollers and other computer systems have made it even more practical to acquire and analyse data for predictions with the help of not only traditional mathematical models, software, artificial intelligence, and machine learning algorithms.

Condition-based monitoring is deployed to prevent downtime, asset failures, and unnecessary practices by monitoring machine performance metrics in real-time by integration of various sensors with the machine itself. Condition-based monitoring ensures maintenance is carried out only when required above a certain threshold avoiding unnecessary costs and downtimes in the wake of routine checkups.

A big advantage of this monitoring technique is its non-invasive nature as sensors are mounted on the outer surface of the machine saving time by avoiding disassembly and giving health indicators merely by sensory data. This data can also be used to study randomness, repeatability and abnormality in the operating nature of the machine.

2.4. Fault Diagnosis of Ball Bearings

Various methods can be used for fault diagnosis of the bearings making use of sensors, signal processing [7], as shown by Gangasar and Tiwari in their paper, and artificial intelligence. Fault diagnosis of ball bearings is critical in determining the health and performance of rotatory machines. Some commonly used methods are stated below.

2.4.1. Ultrasonic Testing

Ultrasonics have been used in marine technology for deep sea exploration [8], in medical field for non-invasive diagnosis [9], and detection of cracks which is helpful in not only fault diagnosis of the bearings but also in estimating the size of the crack, as proved by Komura et al, [10].

2.4.2. Acoustic Emission Analysis

Mba in their paper showed that they used the acoustic emissions for crack size estimation [11], the analysis of the high-frequency sound waves produced by the initiation and growth of cracks or defects in bearings. These acoustic signals are captured using highly sensitive, high-frequency, MEMS based microphones.

2.4.3. Thermal Analysis

Choudhary et al, proved that the convolutional networks can be used for bearing fault diagnosis using thermal images [12]. Worn out or defective bearings generate excessive heat due to friction that is higher than a normal and healthy bearing. Temperature sensors mounted on the bearing bracket can be used for temperature monitoring or infrared thermography can be used for detecting abnormal hotspots on the surface of the bearings which might indicate potential faults.

2.4.4. Model-Based Estimation

Nelson et al, studied and analysed the dynamics of a rotor using FEA approach [13]. Model-based estimation uses mathematical modelling like state estimation techniques or Finite Element Analysis (FEA) of the bearing system to estimate the life and health. FEA enables detailed simulation of the bearing under varying loads and conditions predicting stress distribution, deformation and potential failures. Multibody Dynamics analyses interaction of components of the machinery with the bearing, thus providing a comprehensive overview of the operational environment.

2.4.5. Vibration Analysis

Raw data obtained from the sensors is in time domain by default and it can be transformed into frequency domain by using Fourier transform. Hilbert Huang Transform (HHT) transforms the time domain data to intrinsic mode functions which are capable of analysing non-linear and non-stationary signals, as demonstrated by Atoui et al, [14].

Vibration analysis is the most effective and commonly used technique for fault diagnosis in bearings. Accelerometers are mounted on the bearing bracket and vibrations are recorded and analysed using mathematical models, computer software, or machine learning techniques. The data obtained gives useful insight into features and patterns produced by the faults which are incredibly useful in successful diagnosis and classification. Vibrational analysis can be done in time, frequency, time-frequency domains and using wavelet techniques.

2.4.6. Machine Learning

A combination of vibrational analysis and machine learning algorithms is a very viable solution for fault diagnosis, especially fault classification as some machine learning algorithms is far better at classifying data than traditional model-based and signal processing techniques, shown by M. He and D. He in their paper, "Deep Learning Based Approach for Bearing Fault Diagnosis" [15].

Machine learning has become a supportive tool and tends to replace traditional fault diagnosis methods as it is data-driven, requires less extensive knowledge of the domain, and less experience. It is also less time consuming as the algorithms monitor the data in real time with minimal interference with the operation of the machine.

2.5. Summary

This chapter reviews the work done by researchers in the field of fault diagnosis using various methods such as signal processing and artificial intelligence, exploring the possibility and effectiveness of the proposed methods for better and more efficient fault diagnosis. In the next chapter, methodology for utilizing artificial intelligence for fault classification will be stated and a basis for experimentation will be laid down.

CHAPTER 3: METHODOLOGY

The methodology section describes the systematic approach to research the faulty bearings, the methods to classify these faults and developing machine learning techniques to classify the faults.

3.1. 6205 Ball Bearing Vibration Dataset

The dataset used for the training and validation of the machine learning is acquired from Korea Advanced Institute of Science and Technology by Jung et al [16]. Their hardware comprised of a SIEMENS three phase, four pole, three horsepower induction motor driven at 380 volts and 60 Hz at 1770 rated rpm, torque meter, gearbox, rotors, hysteresis brake, bearing housing A and bearing housing B. A total of four PCB35234 accelerometers were installed on two housings, in x and y directions.

The bearings were damaged artificially by means of fault seeding method with fault sizes of 0.1 mm, 1 mm and 3 mm, each on inner and outer races of the bearings. This gives us a total of seven classes, six of which are faulty, and one is a healthy bearing. The vibration data was logged at a rate of 25.6 kHz, three load conditions of 0 Nm, 2 Nm, and 4 Nm. The vibration data is saved in MATLAB file format in measure of g-forces, which is the force equivalent to gravitational pull of the Earth (1g = 9.8 m/s^2).

This dataset uses standardized NSK bearing (NSK 6205 DDU) with a ball diameter (d) of 7.90 mm, a pitch diameter (D) of 38.5 mm, contact degree angle (θ) of zero degrees, and the number of balls (N) is 9. Therefore, the shaft frequency (fs) is 50.17 Hz, fundamental train frequency (FTF) is 19.94 Hz, ball pass frequency inner (BPFI) is 272.07 Hz, ball pass frequency outer (BPFO) is 179.43 Hz, and ball spin frequency (BSF) is 234.19 Hz.

To extract features from KAIST dataset, a sliding window of 500 elements and a stride of 42 was used giving us approximately 35,570 feature samples.

3.2. Feature Extraction

Feature extraction techniques [17] are essential in providing extensive and structured information about the data to the machine learning algorithms for effective usage. The feature extraction techniques we used in our project are detailed next and they are all in time domain. We decided to extract following six features from two data channels, x and y axes which gives us a total of twelve features.

3.2.1. Mean

The most, normally known as the average, is a statistic that gives an image of the middle or common value in a collection. The absolute deviation is the total of values in the dataset divided by the total number of the values in the dataset. The mean is often used to compute how the data that was examined is centrally located while comparing different groups or samples.

$$\text{Mean} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \text{(Eq. 1)}$$

Where:

- μ is the mean.
- n is the number of observations in the data set.
- x_i represents each individual value in the data set.

3.2.2. Kurtosis

Kurtosis sets out the shape of a distribution for the dataset's value. It is an indicator if a curve is pointed or even or flat compared to the normal distribution curve. The high kurtosis means a lot of data in the distribution, and this implies that the statistics are skewed and there are more outliers among the frequency. The kurtosis coefficient lower than 1 implies that the distribution has fewer points around the mean, and therefore, there are fewer outliers. Curtailment can help get the fact that variability and data are homogeneously distributed. Kurtosis is regarded as one of the most important feature in classifying faulty bearings [18].

$$\text{Kurtosis} = \frac{n(n+1)}{(n-1)(n-2)(n-3)} \sum_{i=1}^{n} (\frac{x_i-\bar{x}}{s})^4 - \frac{3(n-1)^2}{(n-2)(n-3)} \quad \text{(Eq. 2)}$$

3.2.3. Skewness

Skewness, which is a factor that indicates the degree of the dataset's value distribution, is a concept that measures the asymmetry of data's value in a set. A greater number of positive scores for skewness implies that the data have been skewed to the right, and this is characterized by a larger tail at the end of the distribution on the right of the data line. However, if the coefficient is less than zero, it implies that the distribution is left skewed, since a longer tail is located on the left side of the distribution. Skewness helps our determination about whether our data is skewed or no skewed when around the mean.

$$\text{Skewness} = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} (\frac{x_i-\bar{x}}{s})^3 \quad \text{(Eq. 3)}$$

3.2.4. Root Mean Square (RMS)

Determining the RMS of a dataset, which is simply the square root of the squares' average values, is a fundamental operation in statistics. Within the subject areas of engineering and physics, it is often associated with this "effective" value of a variable quantity, which would account for both positive and negative values. On the one hand, RMS finds its application in data processing and analytical tasks. It provides the measure of either total magnitude of the data, or simply the total intensity.

$$\text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2} \quad \text{(Eq. 4)}$$

3.2.5. Minimum and Maximum Values

Data set shows the highest value by being represented as max value, and lowest value could be represented as min value. These values are the extremes, and they show where the highs and lows are saved in the data. These limits define the boundaries or limits of the data. By this fact, we can identify the outliers or the data points that are out of the range of the normal statistics that can give out some misleading results or analysis.

3.3. Machine Learning Algorithms

Machine learning is the technique through which computer systems learn to predict real world phenomenon based on the data collected. These machine learning algorithms are better performant when they are provided with rich data features. Features can be defined as the description of a data point which explain characteristics of the said data point with more detail. Machine learning algorithms can predict continuous as well as classify data into distinctive types. We have decided to use five machine learning algorithms which are best for classification purposes as our application requires classification of faulty bearings. Classification is simply categorization of data based on previously provided information. The input to a machine learning algorithm consists of the data features and its output is a prediction. The algorithms are explained in detail below for better understanding.

3.3.1 Random Forest

Random Forest - a machine learning method, works by linking a variety of decisions trees during the training phase and forming a powerful prediction model [19]. At any tree in the forest the data and type of characteristics will be selected in random way for the tree's training that will help to reduce overfitting and at the same time improve the performance of the generalization. A variety of methods such as taking statements or averages of the trees' findings are used in making predictions - the algorithm can either work for classification problems or it can determine the probabilities of certain events occurring. Random Forest becomes a good choice for many applications involving multidimensional data and noise and/or with missing values. Besides calculating feature importance that can unveil feature relevance and thus key elements that influence the prediction a model finds. They usually take part in such activities like assigning items to classes, regression, and simple data analysis.

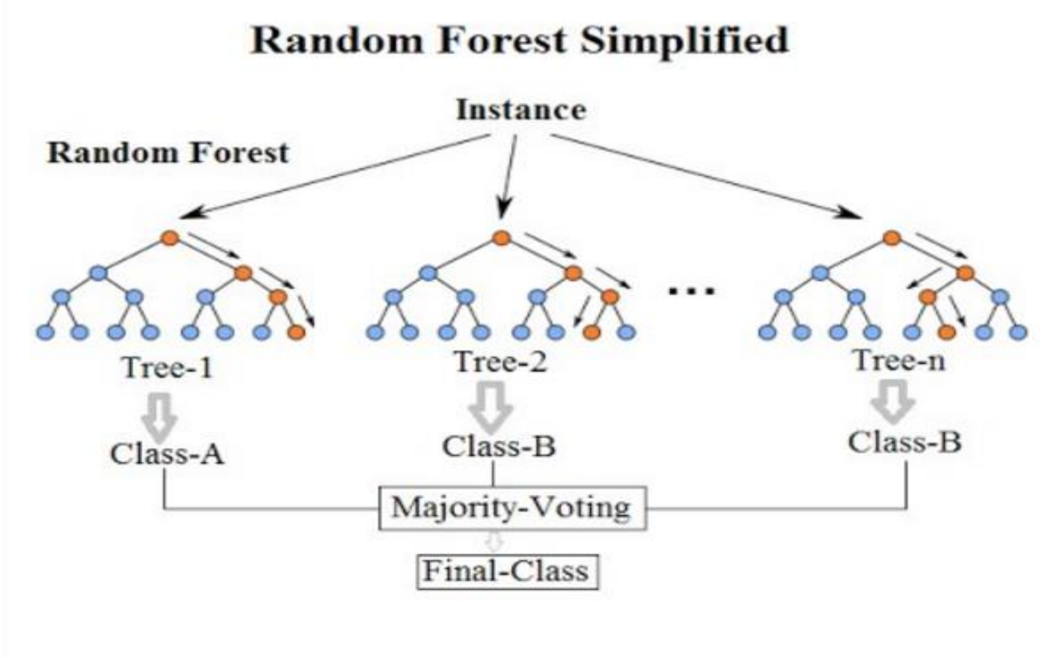

Figure 4. A simple representation of Random Forest

3.3.2. Support Vector Machines (SVM)

Support Vector Machines (SVM) are one type of transversal learning models, which belong to the supervised learning branch and are used for classification and regression tasks [20]. SVM is divided into the process whereby the optimal plane is found which divide points with different classes and maximum margin using it. To put it in another way, this margin is the space between the closest distances of these two different classes, which is necessary to make the decisions more effective. In doing so, it can deal with nonlinear as well as linear decision boundaries using kernel procedures, which thus projects the input data into higher-dimensional spaces. SVM is very suitable in the data that has high-dimensional spaces, particularly in those where the number of features is greater than the number of samples. It is perhaps never fitting too closely to the data and is very popular used in tasks such as images, texts, and spotting extraordinary items. One of the most cherished advantages an SVM has is that it fights the process of overfitting and makes it more suitable for categories of complex data. SVMs are working across many domains as image classification, text sorting and differentiating of anomalies.

The decision boundary in SVM is given by the equation:

$$w^t.x + b = 0$$

where ‘w’ is the weight vector, ‘x’ is the input vector, and ‘b’ is the bias term.

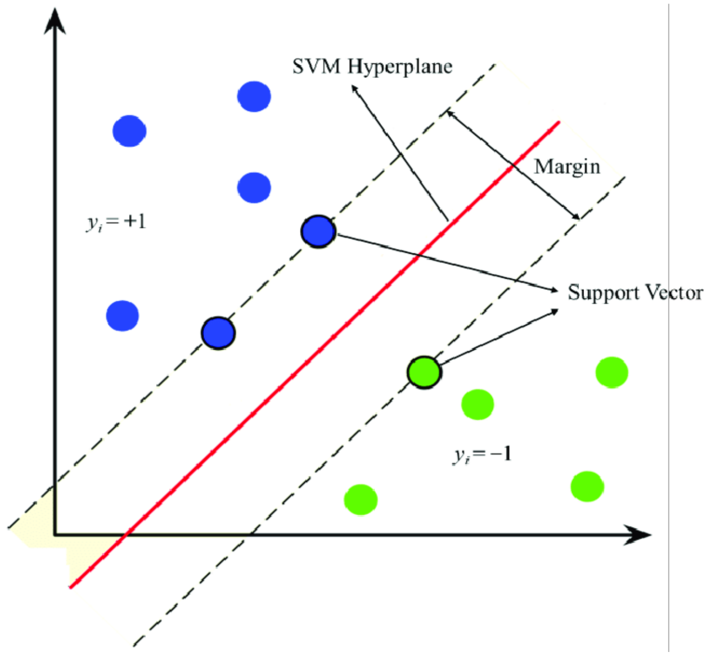

Figure 5. Support Vector Machine and their hyperplanes

3.3.3. Logistic Regression

Logistic Regression is a computational approach which is used for the sorting things into two groups, among the computing methods, it is a statistical method for the two category classification tasks, where the output variable exists in categories with only two possible results (yes or no). g., true/false, yes/no, 0/1). The logistic regression is a linear model, which uses the logistic (sigmoid) function to increase the saturation growth of inputs to the corresponding probabilities of a particular class [21]. It produces the likelihood by applying models on input data and select the highest threshold to make prediction. Logistical regression is easy to understand, runs on weak computing systems and works correctly when the decision line between groups is completely straight. It is good performance for the output features and feature input variables that are linear. Logistic Regression has been popular beyond the health sector as in it is applied in healthcare. g.), but also notaries (in document certificate issues), finance (so in project finance), microbiology (in infectious diseases), genetics (in clan research), insurance (in risk assessments) and tourism (in mapping activities). g. Here, the specific competencies clearly state case evidence assessment, credit risk assessment), marketing (as well as others). g. In the coming months, the AI funnel will not only nurture leads but will along with and social sciences. g., survey analysis). On the other hand, binary logistic regression presents the constraint of performing binary classification tasks and a linear decision boundary supposition.

In logistic regression, probability of the positive class is given by the logistic function:

$$p(y = 1|x) = \frac{1}{1 + e^{-z}}$$

where ‘w’ is the weight vector, ‘x’ is the input vector, and ’b’ is the bias term.

3.3.4. K-Nearest Neighbours (KNN)

One of the simplest and powerful methods of classification and regression is KNN (k-Nearest Neighbour) algorithm [22]. In KNN, the prediction is based on the majority class (for classification) or the average value (for regression) of the K nearest neighbours in the training dataset. Therefore, there is a lack of data to train the model when you have a large amount of data from several categories simultaneously. Distance as used in this context refers to the magnitude of how far two points are from each other using the Euclidean metric to determine how equally points go as similar points. And K adds the nuance of the model weighting between proportionality to the data versus not falling into overfitting. KNN doesn't concern any suppositions about how the information is spread out it only considers the nearby point for a future prediction. It is within the scale of medium to large data sets and effective with derailed data. KNN is a quite easy to grasp and work with algorithm which is probably one of the main canons produced for being a basic machine learning task and baseline model. Though, in contrast, the KNN model slowdown in terms of prediction time—as compared to the model-based approaches, KNN algorithm is sensitive to the choice of distance metric and can suffer from the curse-of K.

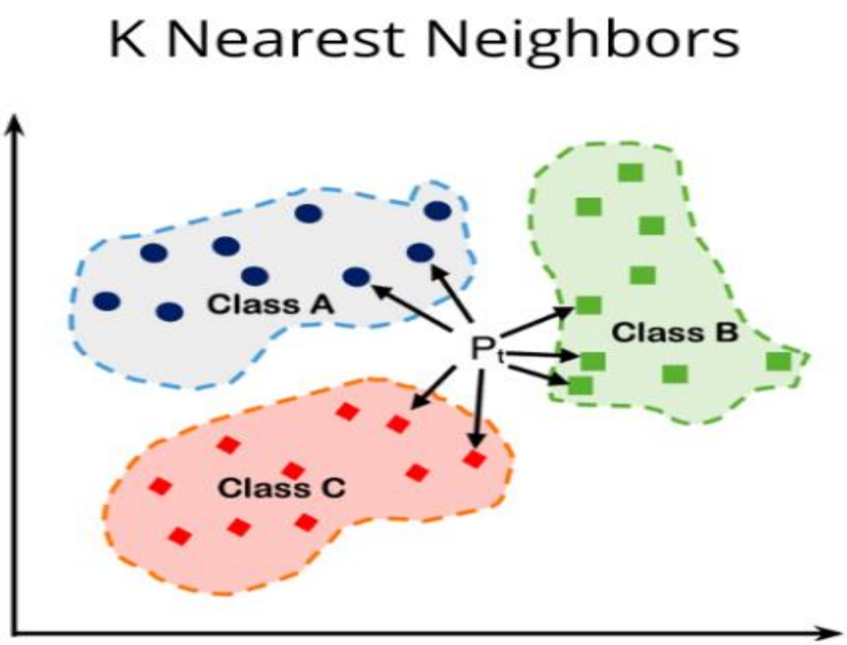

Figure 6. A visual representation of K- Nearest Neighbours

3.3.5. Multi-Layer Perceptron (MLP)

The Multilayer Perceptron (MLP) is one of the most widespread types of neuro networks which have several hidden layers in addition to the input layer. Without rules, there would be no picture. Each "thinking unit" or which is exactly called “neuron” then performs a weighted sum of inputs. But, after this, a special rule (like a light switch turning on or off) is used to decide how active it should be. In this way, the data are passed through the layers head to tail, thereby, MLP finds relations in multi-layered, complex data. MLP functions well if it is to be applied in a sort to organize into categories, make predictions, or spot the patterns. MLPs (Multi-layer Perceptron) are used for supervised learning problems such as classification & regression [23]. Every MLP neuron accumulates DA with various weights and adds them with a special function that indicates how much it needs to be activated. g., we want to build neural networks with layers of nodes (sigmoid, ReLU, etc.), and feed the obtained result as input for the subsequent layers. Using MLPs, one can learn different complex non-linear relationships in data. Because of this, the function can be approximated by a single hidden unit if there are enough of them. These algorithms work by backpropagating the error through all the layers to the input layer, and by using the gradient descent optimization algorithm the increase in accuracy and decrease in errors is achieved. They can approximate any continuous function by means of a network even the data and training if they have enough. MLPs researchers have gained success in domain such as image recognition, natural language processing and financial forecasting. On the contrary, the training and using of MLPs require substantially large amounts of computational resources, and they could easily get fitted if not properly developed.

The output of a neuron in an MLP is calculated using the formula:

$$y = f\,(w^t.x + b)$$

where ‘f’ is the activation function (e.g., sigmoid or ReLU), ‘w’ is the weight vector,’ x’ is the input vector, and ‘b’ is the bias term.

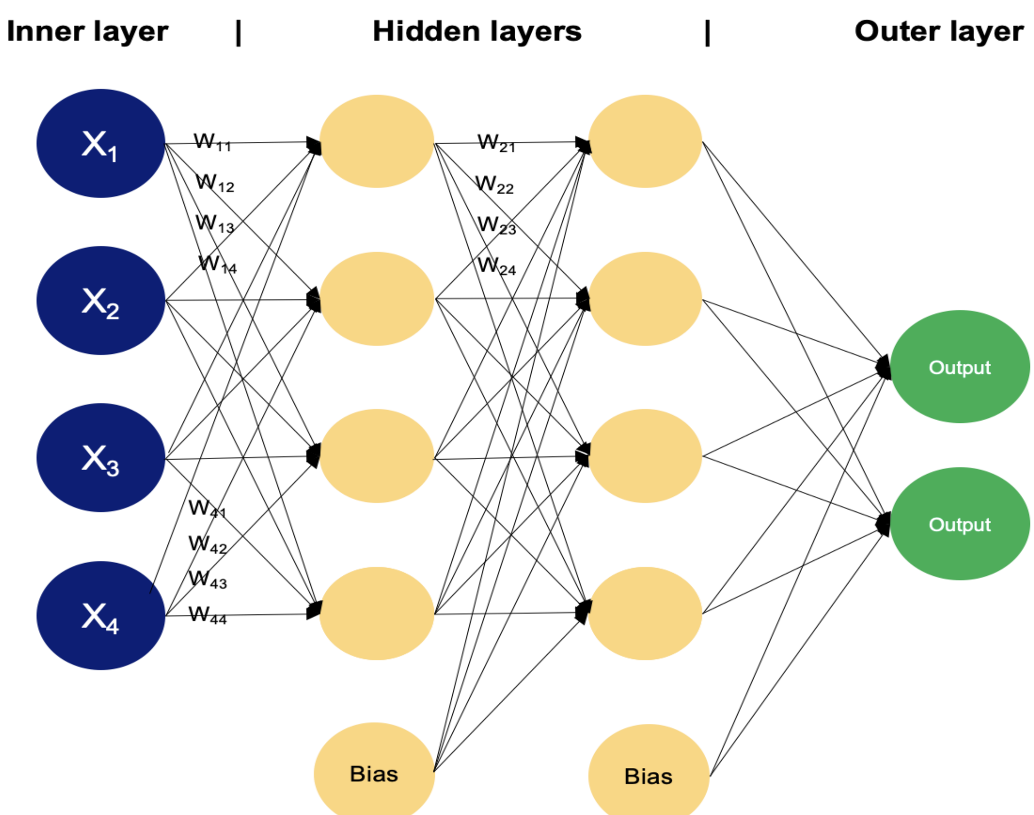

Figure 7. Multi-Layer Perceptron and their hidden layers

3.4. Summary

This chapter expands on the tools and techniques used for manipulating the dataset for feature extraction and machine learning algorithms that will be implemented for fault classification. Common statistical features in the time domain are extracted from the data and five machine learning algorithms are chosen for training on the mentioned features.

Next chapter will detail the specifications of the test rigs developed for data acquisition and will explain the datalogging procedure.

Chapter 4: Experimental Setup

This chapter will expand on the design and development of the test rig used for data acquisition. Hardware description will delve deep into the details of each component, the reason for choosing such components, and the design criteria behind the test rig. Following this, data collection process will be explained, detailing the types of faults and how these faults were induced. Feature extraction techniques applied on the data will also be detailed, highlighting the rationale behind selection of specific features. Finally, the chapter will conclude by giving an overview of the machine learning algorithms used for classification, along with training and evaluation procedures opted for performance assessment.

4.1 Test Rig Hardware

The hardware for this project was designed to simulate the rotatory motion of the bearings in keeping the conditions as close to real world as possible. The test rig comprises of the following components:

A DC motor is an electric device that converts electrical energy to mechanical energy in the form of rotatory motion. DC motors operate on direct current and usually have high speed but low torque. They must be coupled with a gearbox to increase their capability of handling higher loads, but it decreases their speed limit. Most DC motor operate north of 3,000 RPM at just 12 volts and can go up to 20,000 RPM at 24 volts. To keep track of their speeds, an encoder is coupled with the motor mounted on the back of the rotor shaft. The encoder generates pulses which can be counted by means of electronic devices and using simple mathematical relations, speed can be calculated.

In addition to voltage supply, to provide the flexibility of controlling the speed, 3.3 volts or 5 volts Pulse Width Modulated signal is applied to the motor. The duty cycle of the signal determines the speed.

For our project, we are using 24 volts, brushless DC DSE38BE27-001 motor. It comes with a 100 Pulses per Revolution GP1A30R optical encoder. At 24 volts, its rated speed is 4,400 RPM. It has a good balance of torque and speed.

A motor driver is an electronic circuit that is used to control and manage the operations of a motors. It plays the role of an interface between low-power control signals of a microcontroller and high-power requirements of the motor. It supplies power and pulse-width modulated signal to precisely control the speed and direction of the motor. To ensure safe and efficient operation, they include features such as thermal shutdown, overcurrent protection, and fuses. The circuit is built using high power transistors or MOSFETs, logic controllers, diodes and other electronic components.

BTS7960 is a powerful and versatile dual H-Bridge configuration motor driver. Due to its high current capabilities of up to 43 amperes, it is widely used in robotics and even electric vehicles as well. It operates in the range of 6 volts to 27 volts and supports up to 25kHz pulse-width modulated signals. It has built in protections against overtemperature, undervoltage, overcurrent, short circuit, and has an overvoltage lockout.

A shaft is a mechanical component used for transmission of torque and rotational motion between other components of the machine. Typically cylindrical in shape, shafts are designed to support and transfer torque while maintaining structural integrity under loads and stresses in wide ranges. Shafts are commonly found in engines, gearboxes, differential systems, axle-wheel pairs, etc. They are made from high-strength materials like aluminium and steel to withstand high loads. Shafts may include features like splines, keyways, and tapers to secure gears, pulleys, and other components to facilitate effective and precise mechanical connections.

The design of the entire test rig is built around the shaft, its length and critical speed of the shaft resulting due to its length. Critical speed of the shaft is defined as the speed beyond which it tends to vibrate transversely. First critical speed of the shaft can be calculated using the equation 7.22 from Shigley's Mechanical Engineering Design book.

$$\boldsymbol{\omega} = \left(\frac{\pi}{l}\right)^2 \sqrt{\frac{EI}{A\rho}} \qquad \text{(Eq. 5)}$$

Since we are concerned with the vibrations produced by the bearings only, any vibrations produced by other components, especially the shaft, would add unwanted noise to the data acquired. So, to avoid whirling of the shaft following critical speeds for a steel shaft of 8 millimetres diameter with various lengths were calculated which are listed in Table 1 below.

Table 1. Lengths and critical speeds for the shaft

Length (mm)	Critical Speed (RPM)
150	2670
175	1957
200	1499
225	1184
250	955
275	792
300	668

A length of 250 millimetres is suitable as it would provide us a reasonable cushion of 955 RPM and a moderate length of the shaft.

Motor couplings are mechanical devices used to connect two shafts together and transmit power between them. They compensate for misalignments, axial movement, and dampen vibrations.

The DC motor used in this project is coupled with the 8 millimetres diameter shaft using a 5x8 millimetres flexible coupling.

6205-2RS belongs to the family of deep grooved ball bearings as evident by the digit six. It has a bore of 25mm, outer diameter of 52mm and width of 15mm to 19mm. The suffix 2RS shows that it is sealed with rubber on both sides. These bearings find their application in transmissions, motors, bikes, pumps, etc.

This bearing has been used by several universities for research purposes such as CASE Western Reserve University (CWRU), Paderborn University, Korean Advanced Institute of Science and Technology (KAIST), Huazhong University of Science and Technology (HUST) to name a few. As a result, this bearing has large datasets available openly for researchers all over the world.

Bearing brackets, also called bearing housings, are used to hold and support the bearings within an assembly or machinery. These brackets ensure proper alignment and stability by providing a secure mounting surface for the bearing. They are tailored to accommodate different types and sizes of bearings. They are made from highly durable and strong materials such as cast iron or steel to ensure stability and integrity of the bearing assembly.

The bracket for 6205 bearing used in this project is modular in nature, that is, bearings can be replaced easily by merely taking off the top half of the bracket.

An accelerometer is a sensor used to measure acceleration forces in various directions. It senses changes in velocity or orientation of the object, giving data on movement, vibration and tilt. Accelerometers are being used in smartphones, fitness trackers, automotive systems, industrial equipment and navigation. Accelerometers are usually based on Micro Electro-Mechanical Systems Technology (MEMS), but modern accelerometers use piezoelectric materials as sensing element as well for better frequency response.

ADXL345 is a MEMS-based accelerometer provided by Analog Devices. It is used in modern smartphones and fitness trackers for measuring vibration patterns (single-tap, double-tap gestures, etc.) and tracking workout activity, respectively. For its well-crafted specifications, which are listed in Table 2 below, it used in this project as the sensor to measure the vibrations from faulty bearings.

Table 2. Specifications of the ADXL345 accelerometer

Specification	Value
Acceleration Range	±2g, ±4g, ±8g, ±16g
Sensitivity	Adjustable (4 mg/LSB to 256 mg/LSB)
Output Data Rate (ODR)	Selectable up to 3200 Hz
Resolution	Up to 13-bit
Operating Voltage	2.0V to 3.6V (Typical 3.3V)
Supply Current (Measurement)	25 µA at 100 Hz ODR
Supply Current (Standby)	0.1 µA
Digital Interface	SPI, I2C
Operating Temperature Range	-40°C to +85°C
Package	Small, low-profile LGA or LCC
Features	Tap/double-tap detection, free-fall detection, low power consumption

Although the range is shown to be up to 16g, the accelerometer is capable of handling shocks of up to 10,000g. The output data rate of 3200 Hz means that 3200 data points are being output every second and the sensor can pick up mechanical frequencies of up to 1600 Hz.

ESP32 is a versatile and feature rich System-on-Chip (SOC) developed by Espressif Systems providing a dual-core processor, a rich set of peripherals with wireless Wi-Fi and Bluetooth communication capabilities making it ideal for Internet-of-Things, home automation and even industrial monitoring applications. With its low power consumption, high processing power, low cost, and extensive and highly supportive development ecosystem, ESP32 enables rapid prototyping and deployment of connected devices. Some main specifications of ESP32 are listed in the Table 3.

For high-speed applications like data logging, ESP32 makes an excellent candidate as it supports speeds up to 80 Mega bits per second (80 MHz) for Serial Peripheral Interface (SPI), 1 Mega bit per seconds (1 MHz) for Inter-Integrated Circuit protocol (I2C) and 5 Mega bits per second (5 MHz) for Universal Asynchronous Receiver-Transmitter (UART). Since the ADXL345 accelerometer outputs data at a rate of 3200 Hz or 3.2 kHz, the I2C protocol is speedy enough to log data without any hiccups.

Table 3. ESP32 specifications

Specification	**Value**
CPU	Dual-core Xtensa LX6, up to 240 MHz
Memory	520 KB SRAM
	4 MB Flash
Connectivity	Wi-Fi 802.11 b/g/n
	Bluetooth v4.2 + BLE
Interfaces	UART (up to 5 Mbps)
	SPI (up to 80 MHz)
	I2C (up to 1 Mbps)
	I2S, ADC, DAC, GPIO
Operating Voltage	2.2V to 3.6V
Operating Temperature	-40°C to +125°C
Power Consumption	Low-power modes available
Development Environment	Arduino IDE, ESP-IDF

A potentiometer is a variable resistor that is used for precise control of electrical resistance. It comprises of a resistive element and a slider, or wiper, which adjusts the output voltage based on the position of the slider. Potentiometers are broadly used in electronics for dimming, volume control, and even position sensing.

A 10 kilo ohm potentiometer is used to control the speed of the motor in this project.

The hardware setup is shown in Figure 8 below.

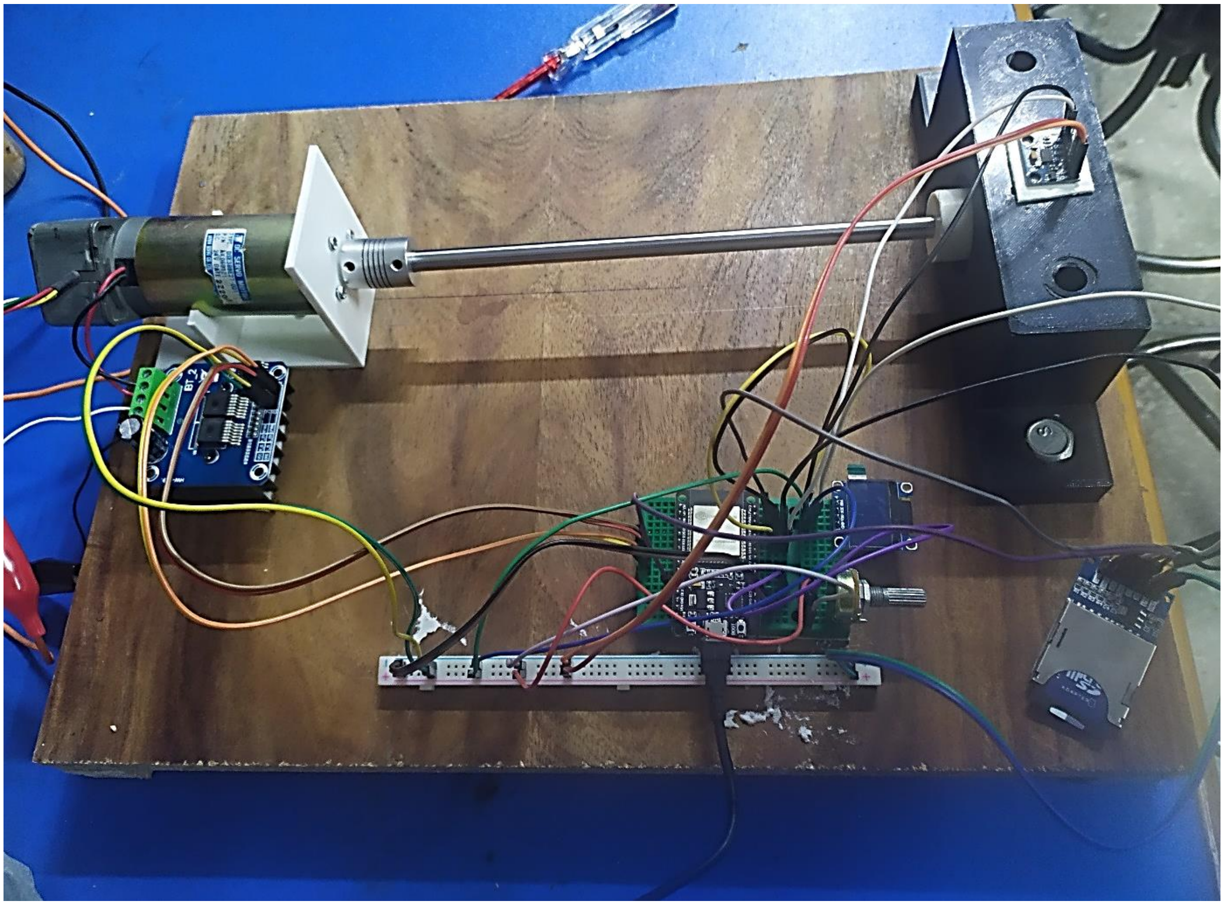

Figure 8. Test rig for data acquisition

4.2. Data Acquisition

The ADXL345 has a sampling rate of 400Hz which means every 2.5 milliseconds, the accelerometer outputs 400 data points. The data is logged for 30 seconds for each fault type. This effectively gives us about 12000 data points for each fault class. ESP32 reads this data from accelerometer and sends it to computer over serial port and a python script logs this data to a csv file. As compared to the KAIST dataset, we have faults of sizes 0.7 millimetres and 1 millimetre, on inner and outer races of the bearings. This gives us four faulty bearings and a total of 5 classes to classify. From here and onwards, the dataset acquired by our test rig would be referred as BearGuard dataset.

For feature extraction from BearGuard dataset, a window of 2 samples and a stride of 1 was used effectively giving us the same number of samples originally in the dataset.

4.3. Summary

Specifications of the test rig developed for the data acquisition are discussed in this chapter along with the datalogging conditions. In the upcoming chapter, we will discuss the results produced by the five machine learning algorithms and analysis will be conducted to comment on the performance of the algorithms to help choose the best of them.

Chapter 5: Results

Machine learning algorithms are implemented using Scikit-learn's python library. Data was manipulated using Pandas, Numpy, and Scipy and Seaborn was used for generating confusion matrices.

5.1. Performance Metrics

The five machine learning algorithms were trained and validated on the two datasets, and their performance was judged based on the accuracy of classification. Classification reports and confusion matrix of the respective models are also provided to give more meaningful insights into performance of the algorithms.

Classification report gives insights into the performance of the algorithm in terms of the correctly identified samples. These metrics are precision, recall and f1-score.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \qquad \text{(Eq. 6)}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad \text{(Eq. 7)}$$

$$\text{F1 Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision+Recall}} \qquad \text{(Eq. 8)}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \qquad \text{(Eq. 9)}$$

True positives are samples that are identified correctly, i.e. they belong to the predicted class. False positives are the samples that do not belong to the class they are predicted as. True negative is a sample that is not identified as sample of the class under consideration, which does not belong to the said class. False negative is a sample that does not belong to the class, but the algorithm classifies it as such.

Precision is the ratio of true positives to the total number of positive predictions, which might include incorrectly identified samples. Recall is the ratio of true positives to the total number of positive instances, true positives and false negatives. F1-score is the harmonic mean of precision and recall and is an indicator of the reliability of the algorithm.

Confusion matrix provides detailed information about the number of data points the model has managed to classify correctly. The algorithms are trained to classify the bearings into seven categories, six of them being faulty bearings and one being healthy. The fault types have been stated earlier in section 3.2.

5.2. KAIST Dataset Results

Table 4. Confusion matrix of the KAIST Random Forest classifier

	outer03	**outer10**	**outer30**	**inner03**	**inner10**	**inner30**	**normal**
outer03	11063	0	0	64	0	0	2
outer10	0	10884	0	0	0	0	0
outer30	0	0	11261	0	0	3	0

inner03	41	0	0	10737	0	276	0
inner10	0	1	0	0	11014	0	0
inner30	0	0	6	276	0	10761	0
normal	0	0	0	0	0	0	11009

Table 5. Classification report of KAIST Random Forest Classifier

Class	Precision	Recall	F1-Score	Support
Outer03	1.00	0.99	0.99	11187
Outer10	1.00	1.00	1.00	10823
Outer30	1.00	1.00	1.00	11071
Inner03	0.97	0.97	0.97	11145
Inner10	1.00	1.00	1.00	10992
Inner30	0.97	0.97	0.97	11088
Normal	1.00	1.00	1.00	11092

Random Forest classified the bearing faults with an accuracy of 99.07 percent and was the best performant model. A very useful feature included in scikit-learn's Random Forest implementation is that it calculates the feature importances by default which can be used to choose features that affect the performance of the model the most.

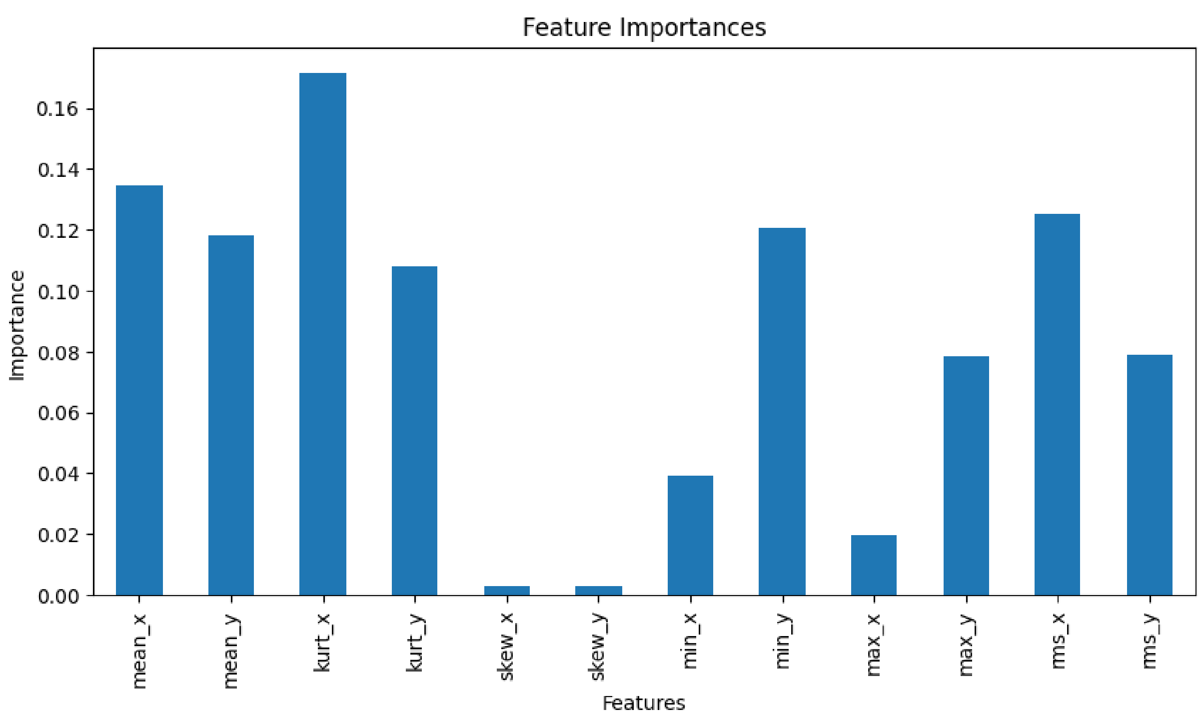

Figure 9. KAIST Random Forest feature importance scores

Table 6.KAIST Logistic Regression Confusion Matrix

	outer03	outer10	outer30	inner03	inner10	inner30	normal
outer03	10977	0	0	108	0	0	2
outer10	0	11037	0	0	4	0	0
outer30	0	0	11104	0	0	20	0
inner03	72	0	0	10389	4	513	0
inner10	0	4	1	0	11115	0	0
inner30	0	0	19	473	0	10528	0
normal	2	0	0	0	0	0	11026

Table 7. KAIST Logistic Regression classification report

Class	Precision	Recall	F1-Score	Support
Outer03	0.99	0.99	0.99	11187
Outer10	1.00	1.00	1.00	10823
Outer30	1.00	1.00	1.00	11071
Inner03	0.94	0.95	0.95	11145
Inner10	1.00	1.00	1.00	10992
Inner30	0.95	0.95	0.95	11088
Normal	1.00	1.00	1.00	11092

Logistic regressor managed to classify at a rate of 98.39 percent.

Table 8. KAIST KNN confusion matrix

	outer03	outer10	outer30	inner03	inner10	inner30	normal
outer03	10567	0	0	516	0	4	0
outer10	0	11038	0	0	3	0	0
outer30	0	0	11096	0	18	10	0
inner03	205	0	0	10112	0	661	0
inner10	0	6	30	0	11082	2	0
inner30	0	0	32	372	0	10616	0
normal	2	0	0	0	0	0	11026

Table 9. KAIST KNN classification report

Class	Precision	Recall	F1-Score	Support
Outer03	0.98	0.95	0.96	11187
Outer10	1.00	1.00	1.00	10823
Outer30	0.99	1.00	1.00	11071
Inner03	0.91	0.92	0.92	11145
Inner10	1.00	1.00	1.00	10992
Inner30	0.94	0.96	0.95	11088
Normal	1.00	1.00	1.00	11092

K-Nearest Neighbour algorithm managed to classify results faults with an accuracy of 97.41 percent.

Table 10. MLPC confusion matrix

	outer03	outer10	outer30	inner03	inner10	inner30	normal
outer03	11010	0	0	77	0	0	0
outer10	0	11041	0	0	0	0	0
outer30	0	0	11124	0	0	0	0
inner03	57	0	0	10360	0	561	0
inner10	0	0	0	0	11120	0	0
inner30	0	0	58	186	0	10776	0
normal	0	0	0	0	0	0	11028

Table 11. KAIST MLPC classification report

Class	Precision	Recall	F1-Score	Support
Outer03	1.00	0.99	0.99	11187
Outer10	1.00	1.00	1.00	10823
Outer30	1.00	1.00	1.00	11071
Inner03	0.98	0.93	0.95	11145
Inner10	1.00	1.00	1.00	10992
Inner30	0.94	0.99	0.96	11088
Normal	1.00	1.00	1.00	11092

MLPC classified at a rate of 98.81 percent.

Table 12. KAIST SVM classification report

	outer03	outer10	outer30	inner03	inner10	inner30	normal
outer03	10618	0	0	462	0	7	0
outer10	0	11038	0	0	3	0	0
outer30	0	0	11094	0	0	30	0
inner03	292	0	0	10091	0	595	0
inner10	0	6	0	0	11114	0	0
inner30	0	0	31	464	0	10525	0
normal	0	0	0	0	0	0	11028

Table 13. KAIST SVM classification report

Class	Precision	Recall	F1-Score	Support
Outer03	0.97	0.96	0.96	11187
Outer10	1.00	1.00	1.00	10823
Outer30	1.00	1.00	1.00	11071
Inner03	0.91	0.92	0.92	11145
Inner10	1.00	1.00	1.00	10992
Inner30	0.95	0.95	0.95	11088
Normal	1.00	1.00	1.00	11092

SVM predict faults with a mean accuracy of 97.37 percent which was the lowest performance out of all the five algorithms.

From the results, it can be concurred that Random Forest is the best performant model with an accuracy of 99 percent. The accuracy of all models is improved dramatically upon increasing the number of features from six, three each for one channel, to twelve.

It is observed that all algorithms tend to misclassify inner03, inner 30, outer03 and outer30 samples, confusing the classes with one another, as evident from precision and recall scores of respective algorithms. A good f1-score in high 0.90s suggest that the algorithms are reliable and can be deployed for classification tasks.

5.3. BearGuard Dataset Results

After logging the data from our setup, Random Forest Classifier was used to classify the faults and it achieved an accuracy of 97 percent.

Table 14. BearGuard Random Forest confusion matrix

	outer07	**outer10**	**inner07**	**inner10**	**normal**
outer07	2504	0	0	4	0
outer10	0	2593	70	2	0
inner07	0	21	2601	8	0
inner10	1	2	18	2374	150
normal	3	0	0	33	2547

Table 15. BearGuard Random Forest classification report

	Precision	**Recall**	**F1-score**	**Support**
0	1.00	1.00	1.00	2508
1	0.99	0.97	0.98	2665
2	0.97	0.99	0.98	2630
3	0.98	0.93	0.96	2545
4	0.94	0.99	0.96	2583

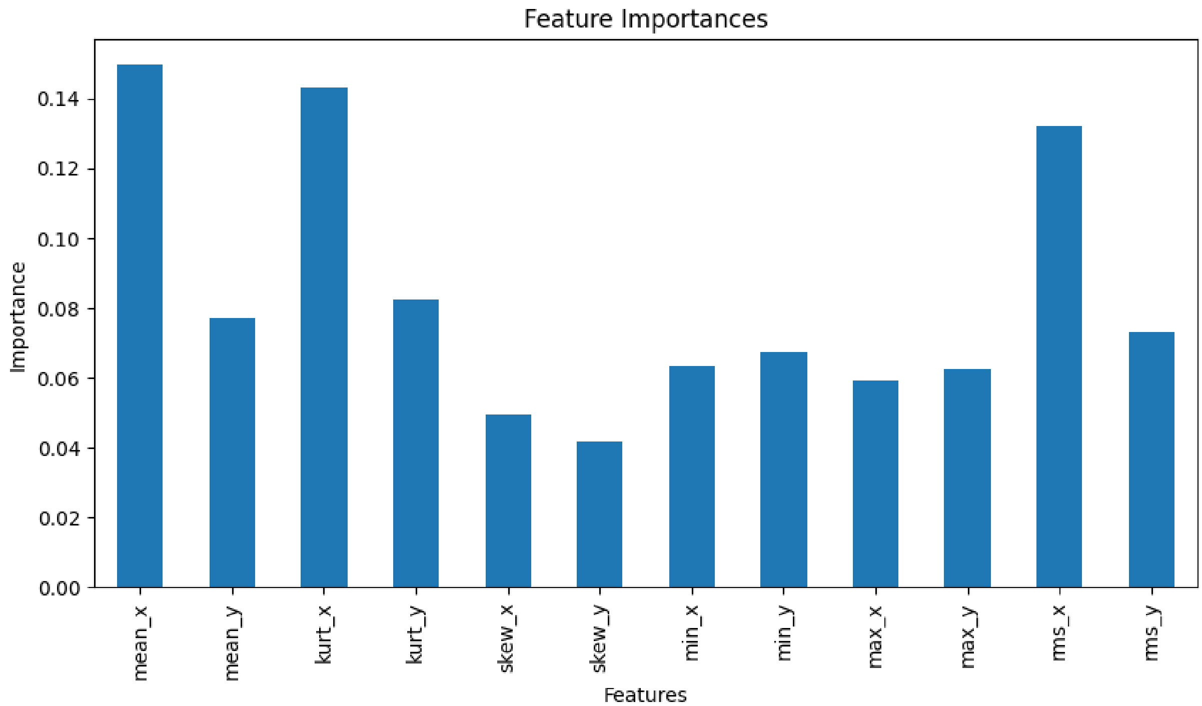

Figure 10. BearGuard Random Forest feature importance scores

Mean, kurtosis and rms are shown to have a big impact on classification as evident from the Figure 10.

Logistic regression performed with an accuracy of only 81.9 percent.

Table 16. BearGuard Logistic Regression confusion matrix

	outer07	**outer10**	**inner07**	**inner10**	**normal**
outer07	2408	3	0	78	19
outer10	0	2088	558	19	0
inner07	0	364	2123	127	16
inner10	82	18	141	1707	597
normal	18	0	0	301	2264

Table 17. BearGuard Logistic Regression classification report

	Precision	**Recall**	**F1-score**	**Support**
0	0.96	0.96	0.96	2508
1	0.84	0.78	0.81	2665
2	0.75	0.81	0.78	2630
3	0.76	0.67	0.71	2545
4	0.78	0.88	0.83	2583

KNN had the least accuracy of all the algorithms achieving just 80.9 percent.

Table 18. BearGuard KNN confusion matrix

	outer07	outer10	inner07	inner10	normal
outer07	2367	0	0	105	36
outer10	1	1985	631	48	0
inner07	1	351	2126	106	46
inner10	31	32	199	1555	728
normal	14	0	1	140	2428

Table 19. BearGuard KNN classification report

	Precision	Recall	F1-score	Support
0	0.98	0.94	0.96	2508
1	0.84	0.74	0.79	2665
2	0.72	0.81	0.76	2630
3	0.80	0.61	0.69	2545
4	0.75	0.94	0.83	2583

Multilayer perceptron achieved the second highest score of 87.8 percent.

Table 20. BearGuard MLPC confusion matrix

	outer07	outer10	inner07	inner10	normal
outer07	2494	0	0	13	1
outer10	0	2308	351	6	0
inner07	0	394	2196	39	1
inner10	4	36	76	1976	453
normal	0	0	1	202	2380

Table 21. BearGuard MLPC classification report

	Precision	Recall	F1-score	Support
0	1.00	0.99	1.00	2508
1	0.84	0.87	0.85	2665
2	0.84	0.83	0.84	2630
3	0.88	0.78	0.83	2545
4	0.84	0.92	0.88	2583

SVM classified at a rate of 82.2 percent.

Table 22. BearGuard SVM confusion matrix

	outer07	outer10	inner07	inner10	normal
outer07	2393	0	0	97	18
outer10	0	2111	549	5	0
inner07	0	397	2128	81	24
inner10	43	21	157	1619	705
normal	13	0	0	189	2381

Table 23. BearGuard SVM classification report

	Precision	Recall	F1-score	Support
0	0.98	0.95	0.97	2508
1	0.83	0.79	0.81	2665
2	0.75	0.81	0.78	2630
3	0.81	0.64	0.71	2545
4	0.76	0.92	0.83	2583

From classification reports of the algorithms trained on the BearGuard dataset, it can be observed that only Random Forest classifier managed to achieve an exceptional accuracy of 97 percent with a good f1-score in high 0.90s just like the algorithms trained on the KAIST dataset. However, rest of the four algorithms performed poorly on the BearGuard dataset barely reaching high 0.80s f1-score for faulty classes. The algorithms managed to classify only the normal bearings with a better accuracy than rest of the classes. This further strengthens the choice of Random Forest Classifier as the best algorithm for fault classification.

5.4. Summary

Although all the algorithms managed to classify the faults with an accuracy above 95 percent, however the algorithms have a problem confusing outer03 faulty bearings with outer30, and inner03 with the inner30 faulty bearings. This is because the features for these four classes are very similar to each other, and the algorithms are not capable to distinguish intricately among these faults. This problem stems from the nature of the algorithms and a more advanced learning algorithm could help overcome this issue.

Chapter 6: Conclusion

The objective of this project was to come up with a robust system for the classification of bearing faults using machine learning algorithms. The project managed to attain a high classification rate of up to approximately 99 percent through integration of hardware setup and advanced machine learning algorithms.

Our hardware setup proved to be a critical component in the whole experiment as in providing a high-fidelity data. It was observed that a higher number of features in the data improved the classification accuracy of the machine learning algorithms. But the dataset comprised only of two features initially, the vibration signals from the bearing in x and y directions.

Feature extraction helped in overcoming this hurdle and 12 features in time domain were extracted, six each for one direction. After extraction of the features, five machine learning algorithms were trained and validated, namely Random Forest classifier, Logistic Regression, Support Vector Machine, K-Nearest Neighbour, and Multi-Layer Perceptron.

The high accuracy achieved proved that artificial intelligence based maintenance methods are very effective in industrial setting, potentially reducing maintenance costs and downtimes.

There is room for improvement in this project as hardware could use an upgrade. Instead of a DC motor, an induction motor with a high-speed rating could be used. Shaft can be upgraded to a bigger diameter and length accommodating multiple bearings. Better accelerometer with technologies better than MEMS can be integrated.

The dataset can be expanded with other severities of bearing faults. The data logging system can be made wireless by saving the data directly to a computer which could save time by removing the element of having to remove the SD card and transferring data to computer.

Fault classification can be done, in the future, using advanced deep learning algorithms such as Convolutional Neural Networks and Recurrent Neural Networks. The Remaining Useful Life (RUL) of the healthy bearing could be predicted to establish a complete condition-based monitoring system.

A real-time diagnosis system can also be realized by making use of the concept of edge AI. The machine learning algorithms are trained and optimized to fit and run on a microcontroller.

REFERENCES

[1] S. M. Al-Ameri *et al.*, "Application of Frequency Response Analysis Method to Detect Short-Circuit Faults in Three-Phase Induction Motors," *Appl. Sci.*, vol. 12, no. 4, p. 2046, Feb. 2022, doi: 10.3390/app12042046.

[2] J. Dalzochio *et al.*, "Machine learning and reasoning for predictive maintenance in Industry 4.0: Current status and challenges," *Comput. Ind.*, vol. 123, p. 103298, Dec. 2020, doi: 10.1016/j.compind.2020.103298.

[3] M. Paolanti, L. Romeo, A. Felicetti, A. Mancini, E. Frontoni, and J. Loncarski, "Machine Learning approach for Predictive Maintenance in Industry 4.0," in *2018 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, Jul. 2018, pp. 1–6. doi: 10.1109/MESA.2018.8449150.

[4] A. Widodo and B.-S. Yang, "Support vector machine in machine condition monitoring and fault diagnosis," *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2560–2574, Aug. 2007, doi: 10.1016/j.ymssp.2006.12.007.

[5] A. Pelliccione, "Plant Engineering 2017 Maintenance Study," Plant Engineering. [Online]. Available: https://www.plantengineering.com/articles/plant-engineering-2017-maintenance-study/

[6] O. Surucu, S. A. Gadsden, and J. Yawney, "Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances," *Expert Syst. Appl.*, vol. 221, p. 119738, Jul. 2023, doi: 10.1016/j.eswa.2023.119738.

[7] P. Gangsar and R. Tiwari, "Signal based condition monitoring techniques for fault detection and diagnosis of induction motors: A state-of-the-art review," *Mech. Syst. Signal Process.*, vol. 144, p. 106908, Oct. 2020, doi: 10.1016/j.ymssp.2020.106908.

[8] H. Li *et al.*, "Design and experimental verification of an underwater ultrasonic drill for rock exploration," *Ocean Eng.*, vol. 294, p. 116824, Feb. 2024, doi: 10.1016/j.oceaneng.2024.116824.

[9] J. A. Jensen, "Medical ultrasound imaging," *Prog. Biophys. Mol. Biol.*, vol. 93, no. 1, pp. 153–165, Jan. 2007, doi: 10.1016/j.pbiomolbio.2006.07.025.

[10] I. Komura, T. Hirasawa, S. Nagai, J. Takabayashi, and K. Naruse, "Crack detection and sizing technique by ultrasonic and electromagnetic methods," *Nucl. Eng. Des.*, vol. 206, no. 2, pp. 351–362, Jun. 2001, doi: 10.1016/S0029-5493(00)00421-0.

[11] D. Mba, "The Use of Acoustic Emission for Estimation of Bearing Defect Size," *J. Fail. Anal. Prev.*, vol. 8, no. 2, pp. 188–192, Apr. 2008, doi: 10.1007/s11668-008-9119-8.

[12] A. Choudhary, T. Mian, and S. Fatima, "Convolutional neural network based bearing fault diagnosis of rotating machine using thermal images," *Measurement*, vol. 176, p. 109196, May 2021, doi: 10.1016/j.measurement.2021.109196.

[13] H. D. Nelson and J. M. McVaugh, "The Dynamics of Rotor-Bearing Systems Using Finite Elements," *J. Eng. Ind.*, vol. 98, no. 2, pp. 593–600, May 1976, doi: 10.1115/1.3438942.

[14] I. Atoui, H. Meradi, R. Boulkroune, R. Saidi, and A. Grid, "Fault detection and diagnosis in rotating machinery by vibration monitoring using FFT and Wavelet techniques," in *2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA)*, May 2013, pp. 401–406. doi: 10.1109/WoSSPA.2013.6602399.

[15] M. He and D. He, “Deep Learning Based Approach for Bearing Fault Diagnosis,” *IEEE Trans. Ind. Appl.*, vol. 53, no. 3, pp. 3057–3065, May 2017, doi: 10.1109/TIA.2017.2661250.

[16] W. Jung, S.-H. Kim, S.-H. Yun, J. Bae, and Y.-H. Park, “Vibration, acoustic, temperature, and motor current dataset of rotating machine under varying operating conditions for fault diagnosis,” *Data Brief*, vol. 48, p. 109049, Jun. 2023, doi: 10.1016/j.dib.2023.109049.

[17] M. C. Popescu and L. M. Sasu, “Feature extraction, feature selection and machine learning for image classification: A case study,” in *2014 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, May 2014, pp. 968–973. doi: 10.1109/OPTIM.2014.6850925.

[18] S. S. Udmale and S. K. Singh, “Application of Spectral Kurtosis and Improved Extreme Learning Machine for Bearing Fault Classification,” *IEEE Trans. Instrum. Meas.*, vol. 68, no. 11, pp. 4222–4233, Nov. 2019, doi: 10.1109/TIM.2018.2890329.

[19] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[20] D. A. Pisner and D. M. Schnyer, “Chapter 6 - Support vector machine,” in *Machine Learning*, A. Mechelli and S. Vieira, Eds., Academic Press, 2020, pp. 101–121. doi: 10.1016/B978-0-12-815739-8.00006-7.

[21] M. P. LaValley, “Logistic Regression,” *Circulation*, vol. 117, no. 18, pp. 2395–2399, May 2008, doi: 10.1161/CIRCULATIONAHA.106.682658.

[22] K. Taunk, S. De, S. Verma, and A. Swetapadma, “A Brief Review of Nearest Neighbor Algorithm for Learning and Classification,” in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, May 2019, pp. 1255–1260. doi: 10.1109/ICCS45141.2019.9065747.

[23] H. Ramchoun, M. Amine, J. Idrissi, Y. Ghanou, and M. Ettaouil, “Multilayer Perceptron: Architecture Optimization and Training,” *Int. J. Interact. Multimed. Artif. Intell.*, vol. 4, no. 1, p. 26, 2016, doi: 10.9781/ijimai.2016.415.