Ahsan Zahoor Khan

Master Thesis

Machine Learning-Based Motor Health Prediction for Enhanced Lifespan Management

School of Technology and Innovation.
Master'sThesis
Sustainable and Autonomous Systems.

Vaasa 2025

UNIVERSITY OF VAASA
School of Technology and Innovation.

Author:	Ahsan Zahoor Khan
Title of the thesis:	Master Thesis
Degree:	Master of Computer Science
Discipline:	Sustainable and Autonomous Systems
Supervisor:	Professor Petri Välisuo Professor Mohammed Elmusrati Principal Engineer Hauke Carstensen ABB Oy
Year:	2025 **Pages:** 71

ABSTRACT:

The aim of the research is to develop a model for predicting the condition of the motor (normal or faulty) by using the vibrational data recorded from the motor. The main question of the research is to find out, "Are these fault conditions predictable based on the vibrational data or Is vibrational data enough to predict different fault conditions of the motor using machine learning and deep learning?"

The objective of the research are to find the core differences of the machine learning and deep leaning strategies in managing the fault prediction of the motors and compares different performance parameters i.e. accuracy of finding the condition of the motor. The study progresses with highlighting the key differences between machine learning and deep learning techniques and the set of principles required to make the predictions accurately.

The research initiates with focusing on vibrational analysis of the motors and their key importance in detecting the faults of the motors. Most of the motor's faults such as imbalance, poor lubrication, bearing defect etc. are related to vibrations. Therefore, by extracting the meaningful pattern of the vibrational data and then training of the model on such pattern could leads in predicting the fault of the motor.

In this study, the vibrational dataset have been collected in controlled environment of ABB laboratories for varying speed-load conditions i.e. Normal or Faulty . The proposed methodology encompass both deep and machine learning algorithms in predicting the faults. Machine learning algorithms such as Support Vector Machine and the K-Nearest Means have been applied to the feature set fabricated from time, frequency, and time-frequency domain. In case of deep learning, Siamese Architecture with Feed Forward Neural Network have been applied to map between the feature vector and the desired output automatically.

The results of the employed algorithms have demonstrated that they can assist in the prediction of motor faults. The prediction of the motor condition directly correlated with the quality and the quantity of vibrational data of the motor. Furthermore, effective data processing, such as filtering, normalization, and transformation, could improve the accuracy of the developed models.

KEYWORDS: Machine Learning, Deep Learning, Support Vector Machine, K-Nearest Means, Siamese Architecture, Siamese Architecture, Feed Forward Neural Network

Contents

1 INTRODUCTION 8

1.1 BACKGROUND: 8

1.2 PROBLEM STATEMENT: 9

1.3 PURPOSE OF THIS STUDY: 12

1.4 SCOPE OF THE PROJECT: 13

1.5 IMPORTANCE OF THE STUDY: 14

1.6 RESEARCH QUESTIONS AND OBJECTIVES: 14

2 LITERATURE REVIEW 16

2.1 MACHINE LEARNING: 17

2.1.1 Machine Learning for fault diagnosis of the motor: 18

2.2 DEEP LEARNING: 24

2.2.1 Deep learning Model Development Process: 26

2.2.2 DL Techniques Used for Fault Diagnosis of Motor: 27

3 METHODOLOGY 30

3.1 RESEARCH DESIGN: 30

3.1.1 Experimental Setup: 30

3.1.2 Motor Specification: 30

3.1.3 Motor Placement: 31

3.1.4 Sensor Mounting: 32

3.2 DATA COLLECTED UNDER FAULT CONDITION: 32

3.3 DATA PREPROCESSING 35

3.3.1 Data Cleaning: 35

3.3.2 Feature Engineering: 35

3.3.3 Z- Score Normalization: 36

3.4 MODEL SELECTION: 36

3.4.1 Machine Learning & Deep Learning Models: 37

3.5 REASONING FOR THE SELECTION OF MODELS: 39

3.5.1 Complexities of Data: 40

3.5.2 Computational Efficiency: 40
3.5.3 Complex Structure: 40
3.6 PERFORMANCE EVALUATION PARAMETER: 40
4 RESEARCH RESULT AND ANALYSIS: 43
4.1 DATA CLEANING PROCESS 43
4.1.1 Sorting of Vibrational Data: 43
4.1.2 Analysis of Data: 44
4.2 FEATURE EXTRACTION: 47
4.2.1 Feature Extraction for ML: 47
4.2.2 Feature Extraction for DL: 51
4.3 EVALUATIONS AND RESULTS: 51
4.3.1 ML Performance: 52
4.3.2 DL Performance: 58
5 CONCLUSION 63
ACKNOWLEDGEMENT 65
REFERENCE 66

Figures:

Figure 1. Why motor fails? (Motor-EBook ABB). 10
Figure 2. Exploring the Diverse Branches of Artificial Intelligence. 16
Figure 3. Supervised Vs Unsupervised Learning. 17
Figure 4. Flow Chart of Supervised machine Learning. 20
Figure 5. Flowchart of Unsupervised Machine Leaning. 24
Figure 6. Schematic illustration of neuron (nerve cell) (Einevoll, 2006). 25
Figure 7. DL Model Development Process. 27
Figure 8.Motor Used for the Experimental Setup (ABB, n.d) 31
Figure 9. Sensor Mounting and Reference Coordinate. 32
Figure 10. From Data Collection to the Model Evaluation: Proposed Model. 34
Figure 11. Box Plot of RMS Values (Normal and Rotor Imbalance Operations) 36
Figure 12. Locality Preserving Projection After Dimensionality Reduction (Sugiyama, M. (2016)) 38
Figure 13. Siamese Network (Khandelwal, R. (2021). 39
Figure 14. Raw Recorded Data. 43
Figure 15. Normal Condition (Sorted OP's) 44
Figure 16. Rotor Imbalance Condition (Sorted OP's). 44
Figure 17. Plot of Vibrational Data for Three Different Conditions (Full Load) 45
Figure 18. Missing Values in the Data. 45
Figure 19. Duplicated Values in Each Row. 46
Figure 20. Box Plot of Normal Condition OP6. 47
Figure 21. Feature Vector. 51
Figure 22. PCA Plot (PC1 vs PC2) 52
Figure 23. LPP Plot (LPP1 vs LPP2) 53
Figure 24. FA Plot (Factor 1 vs Factor 2) 54
Figure 25. Confusion Matrix SVM. 55
Figure 26. Confusion Matrix KNN 56
Figure 27. Pipeline Diagram of ML 57

FIGURE 28. CONFUSION MATRIX (NORMAL VS ALL). 59
FIGURE 29. CONFUSION MATRIX SPECIFIC CLASS PAIR (LABEL 0-5). 60
FIGURE 30. CONFUSION MATRIX SPECIFIC CLASS PAIR (LABEL 0-4). 61
FIGURE 31. PIPELINE DIAGRAM OF DL 62

List of Tables

TABLE 1. THE MAJOR FAILURES OF MOTORS. 10
TABLE 2. NORMAL CONDITION OPERATIONAL POINTS. 33
TABLE 3. ROTOR IMBALANCE CONDITION OPERATIONAL POINTS. 33
TABLE 4. FEATURE SET OF TIME SERIES VIBRATIONAL DATA (ML) 50
TABLE 5. PERFORMANCE OF THE DL MODEL. 62

Abbreviations

AI	Artificial Intelligence
CRC	Corporate Research Centre
CF	Crest Factor
DNN	Deep Neural Network
FFT	Fast Fourier Transform
FC	Frequency Centre
FP	False Positive
FN	False Negative
IEA	International Energy Agency
IF	Impulse Factor
KNN	K-Nearest Neighbour
KF	Kurtosis Factor
KV	Kurtosis Value

LPP	Locality Preserving Projection
ML	Machine Learning
NN	Neural Network
OP	Operational Point
PCA	Principal Component Analysis
PPV	Peak to Peak Value
RMS	Root Mean Square
RMSF	Root Mean Square Frequency
RVF	Root Variance Frequency
RI	Rotor Imbalance
SV	Skewness Value
SF	Shape Factor
SVM	Support Vector Machine
SRA	Square Root Amplitude
SNN	Siamese Neural Network
SDG	Sustainable Development Goals
TP	True Positive
TN	True Negative
UN	United Nations
WPD	Wavelet Packet Decomposition
WPNE	Wavelet Packet Node Entropy

1 Introduction

1.1 Background:

” The motor will be the universal heart of industry, driving everything from factories to homes, making energy available to all.”

(Nikola Tesla)

Electric motors have been a pivotal point of the industry since its invention. Approximately 40% of electricity consumed by industry worldwide is consumed by the electric motors, commonly used in pumps, compressors, conveyors, and handling systems (ABB, n.d.). Global electricity demand projected to grow by more than 70 percent between 2015 and 2040 (International Energy Agency, 2015). With the increase of the global energy demand, the use of the electric motor in the industry will also be expanded. As the technology evolves, it is desirable to increase the efficiency of the motor but also ensures the sustainability development of the motors.

Induction motors are ubiquitous in the industrial sector and are the backbone of the many commercial and practical applications. Tough design, cost effectiveness, and low maintenance make induction motors the backbone of many industrial sectors (Sivaraju et al., 2012; Cheng et al., 2014). Practical applications range from conveyer belts to electric vehicles. Above all, the assurance of reliable operation and maintaining its operational efficiency illustrate its prevalence in the industry. These features allow the industry to avoid unwanted downtimes and achieve their targets without any hurdles.

High reliability allows for uninterrupted operation. On the other hand, the efficiency of the induction motor determines the conversion of the electrical power into mechanical power. With the groundbreaking invention of ABB cast iron Frame size 225, 45 kW IE4, 95.0% efficiency could be achieved ("Process performance IE4 cast iron motors," ABB, 2021). Another common type of motor is SynRm, offering even higher reliability such as efficiency greater than 95% (Efficiency Class IE5). The invention of the highly efficient motor not only contributes to the reduction of electricity, but it also has a great environmental impact such as low carbon emission. Considering the United Nations (UN), Sustainable Development Goal (SDG), it aligns with it to make the world a safer place to live.

1.2 Problem Statement:

Despite of the high operational efficiency and the reliability of induction motor, there are some other parameters which can cause an interruption in the industrial operations. This interruption in the operation of the industrial application leads to a significant cost. Moreover, using an unreliable motor can even cause a huge damage. These fault which arises frequently are the cause of skipping the scheduled maintenance. The main objective of the maintenance is mentioned below:

Objectives of Maintenance:

- Expand the life of the machine.
- Ensures its reliable operation.
- To avoid unwanted repairs.
- To find minor damages to avoid a severe one.

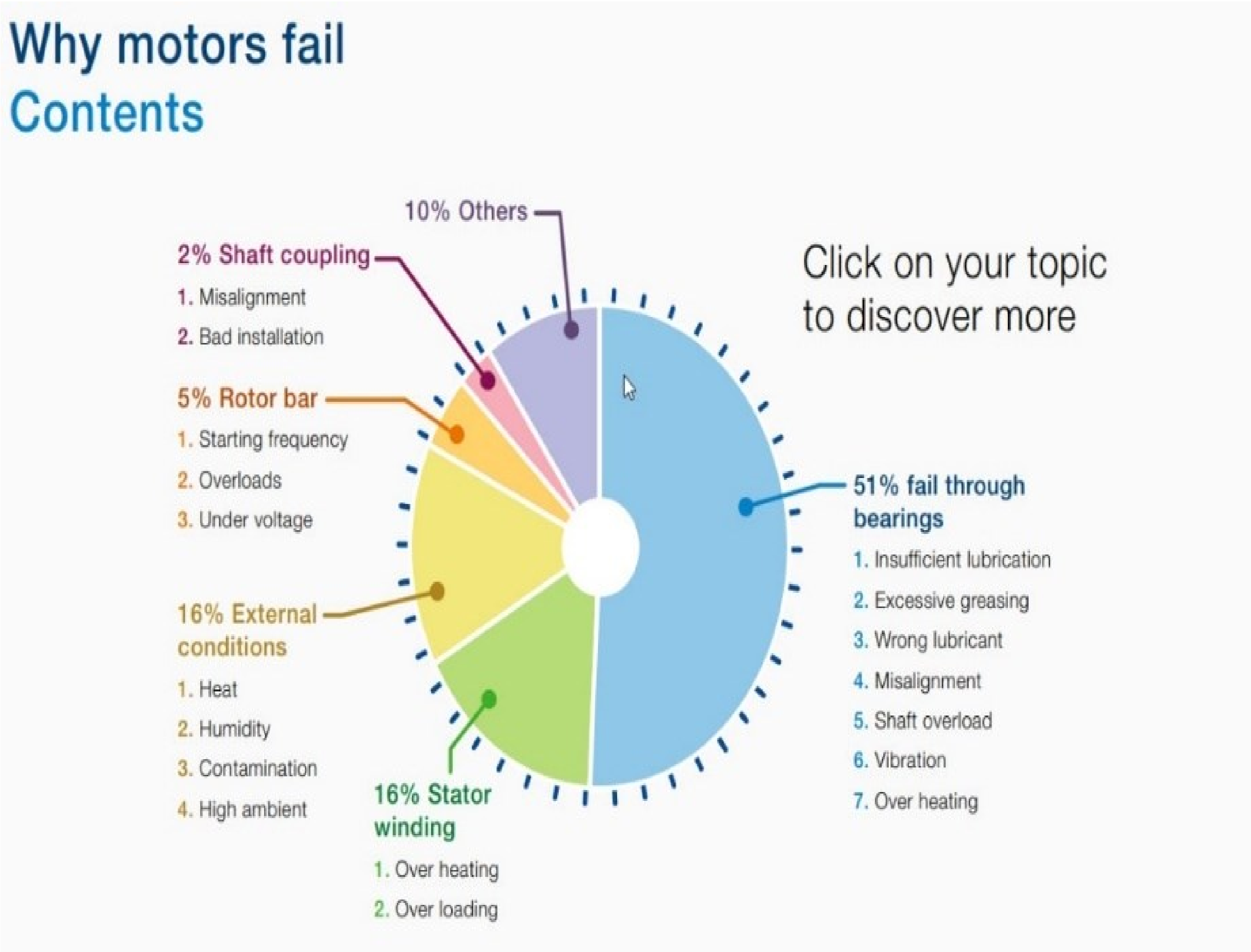

Figure 1. Why motor fails? (Motor-EBook ABB).

The major failures which arise in the motor are mentioned in Table 1:

Table 1. The major failures of motors.

No.	Fault Types	Failure % (IEA, n.d.)	Reasons	Warning Signs

1	**Bearing Failure**	51%	• Continuous operation and degradation. • Overloading. • Poor lubrication. • Overheating.	• Increased vibrational Level. • Excessive noise.
2	**External Condition**	26%	• Humidity in the environment. • Poor cooling. • High ambient temperature.	• Abnormal behavior of the motor. • Reduction in the efficiency.
3	**Stator Winding**	16%	• High starting Current. • Overloading.	• Unusual vibrational levels. • Overheating. • Drastic change in the efficiency of the motor.
4	**Rotor Bar**	5%	• Rotor experience stress at higher start-up frequency. • Temperature rise during locked rotor case.	• Increased vibrations. • Increased temperatures. •
5	**Shaft Coupling/ Misalignment**	2%	• Incorrect mounting of the shaft. • Incorrect alignment of the shaft	• Higher vibrations and noise. • Bearing failure. • Excessive heating.

6	**Bearing Current**	<2	• Electromagnetic effect induces the high frequency currents in the bearing. • Inappropriate grounding.	• Unwanted noise. • Bearing degradation.

1.3 Purpose of this Study:

The mechanical system is a system that converts the input force or motion into a desirable output. The oscillatory motion of the mechanical system transmits energy through the structure, cause a displacement, velocities or acceleration. The cause of the oscillation could be various factor like bearing defect, poor lubrication, misalignment, loosening, imbalance and degradation in mechanical system.

For induction motors, vibrations play a vital role in defining the performance parameters; The vibrations can affect the various aspects of the motor such as efficiency, speed stability, noise level and the temperature rise. Moreover, the excessive vibration could reduce the life of the motor, break the winding insulation, effect the shaft and could generate hindrance in the motion of the motor. If this vibration persists for longer period, this could result in the increase of the downtime and could ultimately result in the financial loss.

The high vibrations thus indicates that a fault diagnosis would be beneficial. In modern predictive maintenance, vibrational analysis is performed to optimize scheduling of periodic maintenance and to prevent faults of the motor. Thus, increases the life span of the motor and reduce the unwanted maintenance cost.

1.4 Scope of the Project:

The scope of the research is to study the comprehensive analysis of the vibrations generated by the induction motors under the influence of faults. The study mainly focuses on finding the specific fault based on the vibration using machine learning techniques. The research will include following steps:

- Acquiring data of the motor under various conditions.
- Pre-Processing of the data to acquire useful features.
- Training of the machine learning model.
- Find the best fitted features to identify the misalignment/shaft coupling fault in motors.

This study investigates the advanced machine learning and deep learning approaches to identify the abnormalities of induction motors. The scope of the study emphasizes on finding a robust approach for the feature extraction, data preprocessing and creation of a robust model to accurately identifying the type of the fault. The focus of the study will also emphasis on integrating the model with the real time monitoring system. Thus, facilitating the industrial application to work more precise and maintaining the accuracy, without giving unwanted downtimes.

The constraint of the study is the scarcity of the data availability. The data used in this study is taken under the controlled environment of ABB Oy, including various motors and various conditions. The limitation of study is defined to ensure the extends of the future work on the fault detection of the motor in broader perspective.

1.5 Importance of the study:

The ubiquitousness and the high dependencies of the induction motor in all industrial sector describes the importance and the need of the study. Integration of vibrational data obtained from motor with the advanced machine learning and deep learning techniques help the industries to deal with the complex challenges associated with the motors. such as early fault predictions, avoid unwanted downtimes and unplanned maintenance.

The importance of the machine learning methods is the inherent ability to detect the abnormalities and the specific pattern that is not possible to detect with the conventional methods. The main objective of the study to increase the efficiency of the motor and the increase the lifespan of the motor. Moreover, the early detection of the fault would be beneficial in avoiding unwanted damage and repair to the motors. This will help to optimize the performance of the motor. Thus, helping the industries to meet the SDGs and reducing the carbon footprints.

1.6 Research Questions and Objectives:

The research will evolve around the with following research questions and objectives:

Research Questions:

1. How to utilize a vibrational data to identify and classify the fault of an induction motor?
2. How to extract the best feature to classify the specific fault of an induction motor?
3. How to find the best possible machine learning/deep learning approach to get the maximum possible accuracy?
4. How can be a machine learning model integrates efficiently with the real-world monitoring system (Predictive Maintenance)?

Objective:

- To develop a reliable dataset of the vibrations (under various fault conditions) while maintain the integrity of the dataset.
- To find the best features out of the vibrational data by using advanced signal processing techniques.
- To implementation and compare the advanced machine learning/deep learning techniques to identify the faults with high accuracy.
- To integrate the machine learning model with the real time vibrational monitoring system, to improve the predictive maintainability of the system.

The thesis is organized in several chapters to provide comprehensive analysis of the research. The second chapter will emphasis on the literature review and to explore more on the approaches used for the fault diagnosis of the motor. The third chapter will emphasis on the methodology used in this thesis to find the best possible results. The fourth chapter will discuss more about the results and the finding of the research. The last chapter will be the conclusion of the thesis.

2 Literature Review

With the advent of the computer, it was primarily used for managing the data as an input and output device. Earlier, the computer always performs task based on the set of rules defines by the humans. Thus, the human interaction was always needed to implement a program to solve a new task by the computer. As the technology advances, the simple performing tasks transformed into the complex tasks but still with the instruction provided by the humans. As the data handling capabilities of the computer advances, this laid the foundation of the Artificial Intelligence (AI). The primary goal of AI was to replicate the human intelligence into machines to perform tasks such as reasoning, learning and decision making. Now a days, AI has become so advanced that it does not require any human involvement. It could make decision of its own and could learn new things with the experiences as human does. AI can now perform tasks like speech recognition, decision making, language translation, etc.

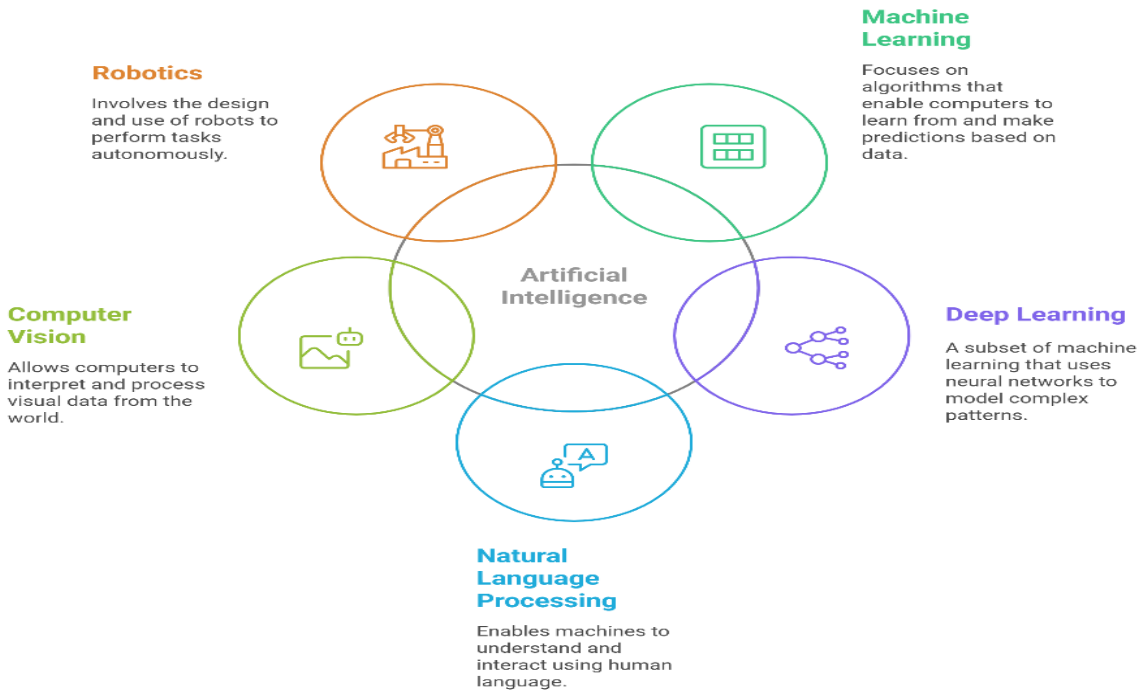

Figure 2. Exploring the Diverse Branches of Artificial Intelligence.

2.1 Machine learning:

Machine learning (ML) is the subbranch of AI, which allows the machine to learn from the patterns within the data and improve its decision-making abilities (Mian, Shawez, & Kaur, 2024). With the availability of the labelled data to make the prediction, such learning is called supervised machine learning and, in a case, if labelled data isn't available to train the model such learning is called unsupervised machine learning (Heller, S.2022; Naeem et al., 2023). The fundamental core concert of the machine learning models is to train it on a seen dataset and then to then gain insights on an unseen dataset. With the availability of large dataset, ML models are now performing better on many problems than the conventional set rule-based systems. Therefore, ML model is being used in many different sectors such as robotics, self-driving cars, power sector, etc.

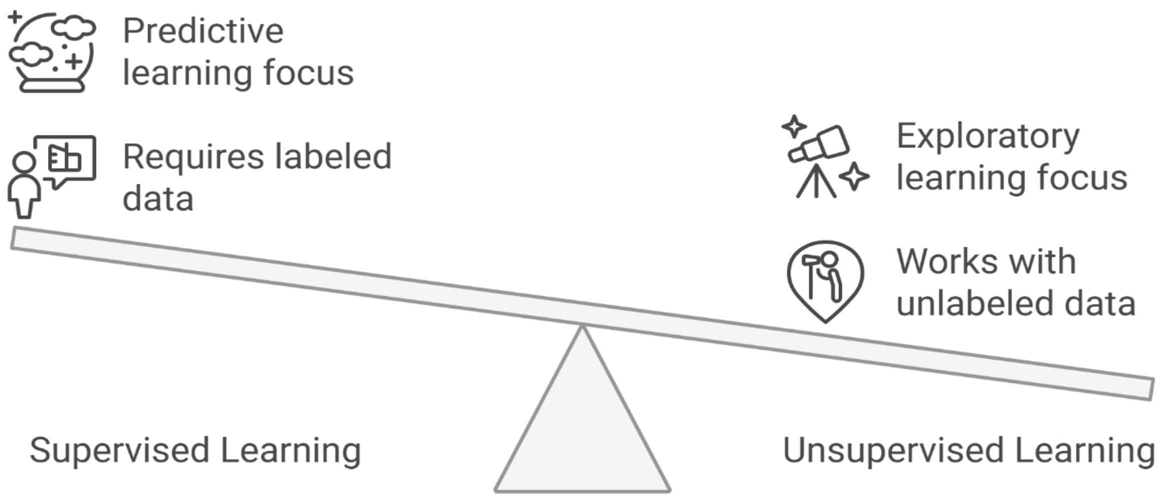

Figure 3. Supervised Vs Unsupervised Learning.

2.1.1 Machine Learning for fault diagnosis of the motor:

Machine learning has demonstrated its ability to diagnose the faults of the motors beforehand and helps in avoiding downtime. ML algorithms help us finding the solution of the complex issues by just utilizing the previous recorded data of the motors (during the faults) and enable us to notice and rectify the faults within the system (Treetrong, J. 2012).

The evolution in the field of ML has proven its abilities in all fields of life and made the system efficient enough of taking the decision by itself. ML algorithms combine many different fields of knowledge such as statistics, science and mathematics to enhance the decision-making ability of the machines. As human intelligence increases with the increase of the experience and learning new skills. Similarly, in context of 'experience of machines' refers as the previously recorded data available (labelled or unlabelled) to train the model. The decision-making capability of machine is largely dependent on the quality (integrity) and the quantity of the dataset, either it is applied on the fault detection or prediction of fault a head of time (Mohri et al. ,2016).

a. Supervised Learning for Fault Detection of Motors:

Supervised ML has become prominent due to its fault prominent diagnostic abilities in the field of electric motors. The abilities of the models to learn from the previous recorded data to train them and make effective predictive analysis. There have been many different approaches explored in the recent years to improve the predicting abilities of the model by focusing on different data sources and ML algorithms . In 2021, Rajapaksha et al. used the acoustic data collected from a three-phase motor to train the different classification models and found out that Support Vector Machine (SVM) performed the best (Rajapaksha et al., 2021). Nguyen et al. (2021) monitored current signal to predict the fault of the induction motor to make a cost-effective solution. In 2023, Mari et al. measured the vibration of the motor to predict the fault in the motor using different supervised machine learning techniques (Mari et al., 2023). These techniques enable the early fault detection, thus makes the reducing the maintenance cost and protecting the motors from a serious damage.

In case of the supervised learning, the main task is to learn a mapping function f from input features X to match the output labels **y,** based on the data provided in the training data, i.e.

$$\boldsymbol{y} = \boldsymbol{f}\,(\boldsymbol{X}) + \boldsymbol{\epsilon}$$

Where:
X in the input features.
y is the targeted value.
$f\ (X)$ is the main function to match the targeted value.
ϵ is the noise error.

In 2007, Kotsiantis stated that the main objective of the supervised machine learning model is to take set of features from the dataset as an input and then predict the label for unseen input (Kotsiantis, 2007).

Key Stages in Supervised machine learning Process:

1. Classification:

In a classification process, the model predicts the categorical labels at the output i.e. the email turns out to be a spam or not (refers as a binary classification) or like predicting a single object out of multiple categorical object (refers as multi-class classification.)

2. Regression:

In a regression process, the model predicts the continuous varying value bases on the previous recorded data i.e. the stock price of the market and estimating the price of the houses adds interoperability between labels.

3. Feature Engineering:

It is a process of finding the best parameters out of the dataset to train the model. This step is the crucial one for improving the prediction abilities of the model and reduces the complexity of the problem by formulating it in a simpler to solve.

4. Model Evaluation:

This step involves defining how accurate and efficiently the trained model is performing. The metrics to define the accuracy are accuracy, precision, recall etc.

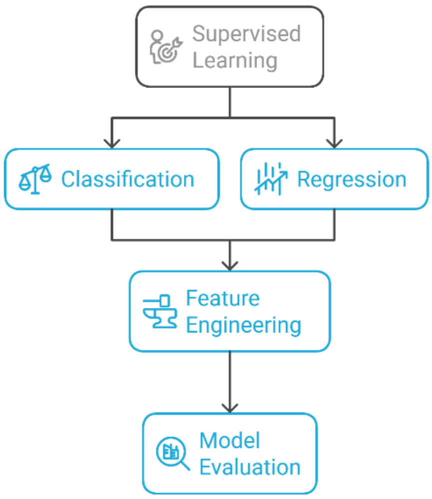

Figure 4. Flow Chart of Supervised machine Learning.

Key Algorithms in Supervised ML:

- **Linear and Logistic Regression:**

Linear and logistic regression are one of the supervised machine learning techniques used for predicting the continuous variable and to classify the discrete outcomes respectively (Shah, 2016; Li, 2019).

- **Support Vector Machine:**

It is one of the efficient algorithms used for the classification and regression problems. It locates the hyperplane to separate the different classes in feature domain. It is helpful in high dimensional datasets (Baser & Apaydın, 2015).

- **K Nearest Neighbour (KNN):**

KNN algorithm used in the classification problem. The data points of the same class are grouped together based on the closest proximity to one another. If a new data point comes, it grouped it to the same class which is in the smallest proximity range (Rani & Vashishtha, 2017).

- **Decision Tree:**

This supervised machine learning technique is useful in both regression and classification problems. It forms a tree like structure by asking questions about data and make groups based on the answers. Upon splitting these into different group, it predicts the answers (Fürnkranz, 2020).

- **Random Forest:**

This technique is an ensemble of decisions trees, where each tree makes an individual decision and based on the decision of the individual trees, the majority wins (Liu, Wang, & Zhang, 2012).

b. Unsupervised Machine Learing for fault detection of Motors:

Unlike the supervised machine learning, it focuses on the dataset that has not been labelled. The main objective of the unsupervised learning is to find the hidden pattern of the dataset without defining assistance of the specific output. There have been various algorithms that have been explored to improve the decision-making ability of the model. In 2019, Choi et al. applied K-Means clustering to cluster observations between a normal and a faulty motor and verified the results by replacing the sensors (Choi et al., 2019). To detect the faults early for the maintenance, unsupervised technique such as Principal Component Analysis (PCA), hierarchical clustering and fuzzy logic have been applied to the vibrational data of the induction motor to find the faults like broken bars and faulty bearing (Amruthnath & Gupta, 2018; Toscano & Carrera, 2020).

- **Key Algorithms in Unsupervised ML:**

a. K-means Clustering:

It is an unsupervised machine learning algorithm that splits the data points into distinct clusters based on the variance. It follows that each point in the cluster should have a minimal variance (to make the points fall under the cluster are like the same) (Vora & Oza, 2013). The objective function is given below, and it is minimised by optimising the locations of the centroids using Expectation Maximization algorithm (Na, Xumin, & Yong, 2010).

$$J = \sum_{k=1}^{K} \sum_{x \varepsilon C_k} ||x_i - \mu_k||^2$$

Here,

K = Number of clusters.

Ck = Grouped points in k-th cluster.

x_i = Single data point in cluster.

μ_k = Centroid (k-th cluster).

$||x_i - \mu_k||^2$ = Square Euclidean distance.

b. Principal Component Analysis (Dimensionality Reduction):

Dimensionality reduction is a common technique used in both supervised a well as in the unsupervised ML. The most common technique used for the dimensionality reduction is principal component analysis (PCA). PCA is used to transform data from a high dimensional space X to the low dimensional space by multiplying it with the top eigen vector of covariance matrix , thus forming λ_k (showing maximum variance) (Salem & Hussein, 2019). The generic equation of PCA is:

$$Z = X\,\lambda_k$$

here,

$Z =$ Data transformed into low dimensional space.

$X =$ High Dimensional Data.

$\lambda_k =$ Matrix consisting of Top Eigenvectors k.

c. Hierarchical Clustering:

Hierarchical Clustering is another unsupervised clustering technique. Unlike the K-mean, it does not require to specify the number of clusters. Instead, it produces a tree like structure which represents the merging of the cluster (dendrogram) (Godara & Kumar, 2012).

d. Isolation Forest:

It is an unsupervised technique used for finding the anomalies within the points. Isolation Forest as name suggests isolates the data points which seems to be different. This is accomplished by dividing the data points in smaller parts and if it finds that the point can be separated easily, then it is an anomaly (Buschjäger, Honysz, & Morik, 2020).

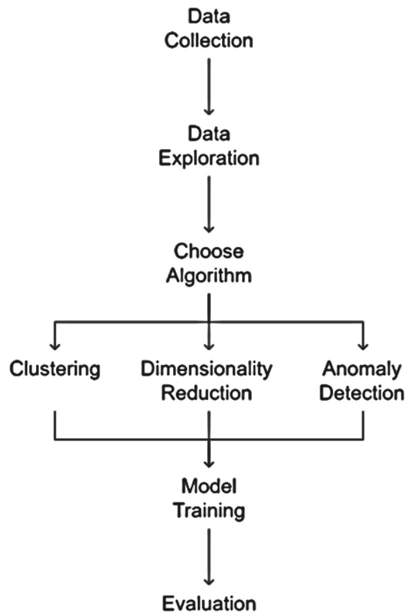

Figure 5. Flowchart of Unsupervised Machine Leaning.

2.2 Deep Learning:

Deep Learning (DL) has tremendously shown its promising results in fault diagnosis of the motor for recent years. Unlike the traditional methods, where crafted features were required to train the model, can learn the feature extraction too. It enables the model to acquire features and patterns from large amounts of data and make it a powerful tool in the field of AI (Kawaguchi, Kaelbling, & Bengio, 2017).

In recent years, DL techniques have been evolved in term of their architecture, computational power and the availability of data. Initially, the deep learning phenomenon was inspired by the human brain that it learns with the experience, thus specific feature engineering is not required. In human brain, the most important part of the nervous system is neuron. Neurons receives the inputs through dendrites, processes the received input, and the transmits the results through the axon to the other neuron. The connection between the neuron is facilitated by the synapse (connection strength between neuron). As the neuron receives the sufficient input which exceeds the certain threshold, it gets activated and transmitted the signal to the next neuron. These activations, results into the output (Einevoll, 2006).

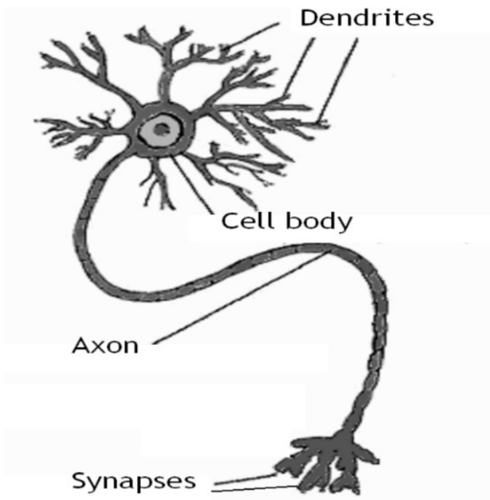

Figure 6. Schematic illustration of neuron (nerve cell) (Einevoll, 2006).

Due to its intrinsic behaviour of capturing complex patterns to train the model, it has revolutionized the traditional ML. As described above, neuron is the main component of nervous system, similarly, to increase the complexity of the system more layers of neurons are added in the neural network (NN). It processes the data using different advanced functions that allow it to pass through different layers, thus it recognizes the data more precisely. This features of DL like capturing complexity of the dataset, makes it more popular in the diagnosis of the motor.

2.2.1 Deep learning Model Development Process:

- **Data Collection:**

This process involves the collection of data. It could be a vibration, sound, pictures, etc.

- **Data Preprocessing:**

This process involves many crucial steps like normalization, cleaning, augmentation and reshaping of the data.

- **Model Selection:**

The selection of the model depends on the task being performed by the neural network. Most commonly architectures are:

- **Convolutional Neural Network:** Commonly used for images processing tasks.
- **Recurrent Neural Network (RNN) or Long Short-Term Memory (LSTM):** Commonly used for the processing of the text or sequential time-series.
- **Multilayer Perceptron or Fully Connected Neural Network:** Used for general purpose tasks.

- **Model Design and Configuration:**

This step involves the defining of the layers and their configuration. Moreover, the optimizer and loss function are also be selected here.

- **Model Training:**

The step involves the training of the model and involves the backpropagation and optimization of the algorithm.

- **Model Evaluation:**

This involves the evaluation of the modal based on the certain parameters e.g. accuracy, precision, and recall. Based on the evaluation, the model is then finetune. The model is then tested on the unseen data.

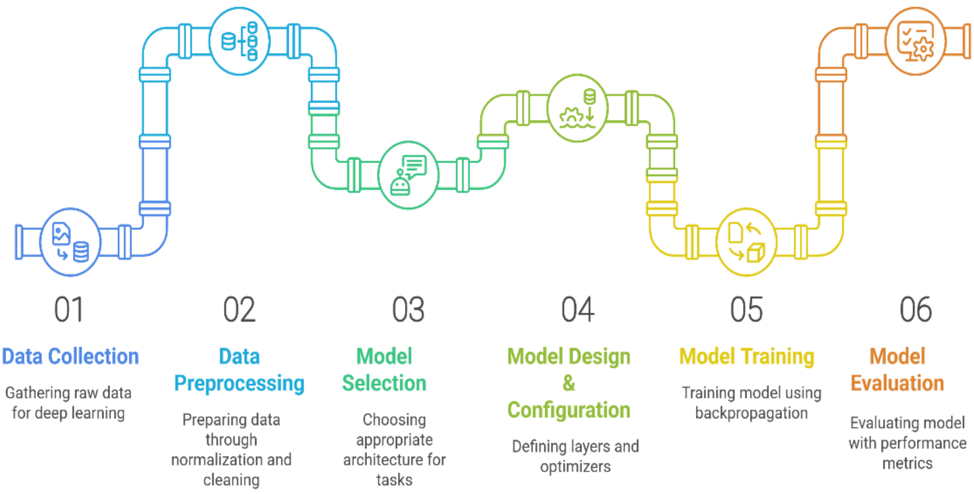

Figure 7. DL Model Development Process.

2.2.2 DL Techniques Used for Fault Diagnosis of Motor:

DL techniques have shown promising results over the past years in finding the early fault detection of the motors. Unlike the ML, where the feature engineering must be done by the data engineer and must highlight the important feature of the data set, DL utilized multiple layers of neural network, thus automatically extract the features from the raw data. This feature of DL makes it a powerful tool. Chattopadhyay et al. (2018) investigated the performance of the induction motor in term of early fault detection has been improved by using the semi-2D convolutional neural network (CNN) as compared to the conventional feature engineering methods. Singh et al. (2021) used the Deep Neural Network (DNN) to find the features from the vibration and the current data collected under different fault condition and find the optimal fault diagnosis technique. Iunusova et al. (2023) applied both knowledge-based approach and the data driven approach in diagnosing the fault of the motor. He and Jin (2021) utilized an autoencoder deep vibrational classifier to detect the faults of the motor.

A mathematical representation of a neural network is quite like the machine learning model, which is given below:

$$\boldsymbol{y} = \boldsymbol{f_{nn}}(\boldsymbol{x})$$

The series of interconnected function f_{nn}, represents the corresponding layers in the neural network. Thus,

$$\boldsymbol{y} = \boldsymbol{f_L}\{\boldsymbol{f_{L-1}}(\ldots(\boldsymbol{f_2}(\boldsymbol{f_1}(\boldsymbol{x})))\}$$

Where $f_L$represents the transformation of the K-th layer. The first layer would be:

$$\boldsymbol{f_1}(\boldsymbol{z}) = \boldsymbol{g_1}(\boldsymbol{W_1 z} + \boldsymbol{b_1})$$

Here,

W_1 = weight matrix of first layer.
b_1 = bias vector of first layer.
g_1 = Activation function (Relu, leaky Relu, etc.

Despite of huge research and advancement in ML and DL techniques the strive for the best possible solution has always been an ongoing challenge. According to the research and the advancement discussed above, we have seen many remarkable improvements in AI. But the field is still rapidly evolving in finding best techniques. Similarly, the performance of the ML and DL improves on large dataset. Technique like transfer learning the in the field of DL tries to alleviate us to deal with small dataset, but still, it is not universal. It is still desirable to solution which requires few data to be trained, less computational power and resolves the memory constrains of the devices.

3 Methodology

This chapter focuses on the research methodology and the equipment used to conduct the research. Furthermore, this chapter focuses on the procedures of the study conducted and explains each nuance of the study such as how data was collected, why DL and ML techniques were selected, preprocessing of the data and the data analysis tool.

3.1 Research Design:

This study is based on the experimental research design mainly focused on the collection of the data and then utilizes the data in finding the early fault detection of the induction motor. The experiment was conducted in the controlled environment of the Corporate Research Centre (CRC) of ABB AB. Moreover, the ML and DL approaches were applied to find the best possible accuracy in finding the faults of the motors.

3.1.1 Experimental Setup:

The experiment was performed in the CRC of ABB AB located in Västerås, Sweden. During the experiment, the primary objective was to collect the vibration data under different speed-load condition, store it in the cloud and then apply various techniques of ML and DL back in Finland.

3.1.2 Motor Specification:

The motor used for the testing purpose were manufactured by ABB and the product code is 3GBP162420, commonly used in the compressors, pumps, conveyers, etc. The specification of the motor is given below:

Motor Code	3GBP162420 (ABB)
Rated Power	15kW
Rated Speed	1500 rpm
Voltage	400/690V
Frequency	50
IE Class	4
Poles	4

Figure 8.Motor Used for the Experimental Setup (ABB, n.d)

3.1.3 Motor Placement:

The motor was mounted on the test bench, a common setup followed by the research centre. This setup is normally chosen to mount the motor properly and avoid the interference from the external sources. Moreover, the setup was appropriate in collecting the vibrational data. While ensuring the data collected was entirely due to the fault rather than the external noise.

3.1.4 Sensor Mounting:

The sensor used in the experiment was the 3-axis accelerometer, and it was mounted on top of the terminal box. It was intentionally place on the terminal box to record the maximum possible vibration under fault condition. The sensor mounting and reference coordinate system are shown below in the figure.

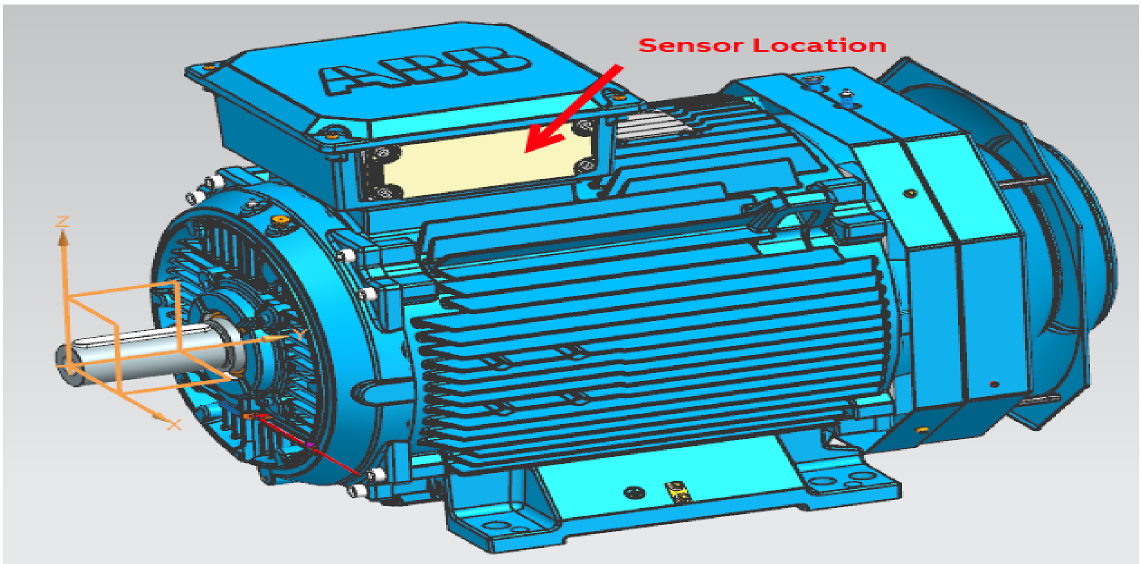

Figure 9. Sensor Mounting and Reference Coordinate.

3.2 Data Collected Under Fault Condition:

The experiment was conducted in a way that six (five) different Operational Points (OP) were introduced under two different condition such as:

1. **Normal.**
2. **Rotor Imbalance (RI).**

1. Normal Condition:

In normal condition, 6 different OPs were introduced (from no load - full load and 50% Speed to 100% Speed). Each OP was introduced for the 15 to 16 minutes. The number of samples collected were 2048 with the sampling frequency of 6.6kHz

Table 2. Normal Condition Operational Points.

Rounds	**Motor warmup [OP0] (Speed = 50%; Load = 0%)**	**OP1 (Speed = 50%; Load = 0%)**	**OP2 (Speed = 50%; Load = 25%)**	**OP3 (Speed = 50%; Load = 50%)**	**OP4 (Speed = 100%; Load = 0%)**	**OP5 (Speed = 100%; Load = 50%)**	**OP6 (Speed = 100%; Load = 100%)**
Normal	10:40:50-10:49:50	10:50:50 - 11:19:50	11:20:50-11:34:50	11:35:50-11:49:50	11:50:50-12:04:55	12:05:50-12:19:50	12:20:50-12:34:50

2. Rotor Imbalance Condition:

In Rotor Imbalance case, 5 different ops were used each for 15 to 16 minutes. The number of collected samples were 2048 with the sampling frequency of 6.6kHz.

Table 3. Rotor Imbalance Condition Operational Points.

Rounds	**Motor warmup [OP0] (Speed = 50%;**	**OP1 (Speed = 50%; Load = 0%)**	**OP2 (Speed = 50%; Load = 25%)**	**OP3 (Speed = 50%; Load = 50%)**	**OP4 (Speed = 100%; Load = 0%)**	**OP5 (Speed = 100%; Load = 50%)**	**OP6 (Speed = 100%; Load = 100%)**

	Load = 0%)						
Rotor Imbalance	N/A	12:57:49-13:25:49	13:26:49-13:40:49	13:41:49-13:55:49	N/A	15:03:49-15:17:49	15:18:49-15:32:49

The data collected for each op were recorded in the CSV files and different ML and DL algorithms were applied using python language package (3.12.1). The algorithms applied for the validation of the collected data, in case of ML, were the K- Nearest mean, the random forest and the dimensionality reduction technique used were the Principal Component Analysis (PCA), Factor Analysis (FA) and Locality Preserving Projection (LPP). Similarly, in case of the DL case, feed forward neural network with Siamese architecture.

These algorithms will be explained in the model selection, further in this chapter.

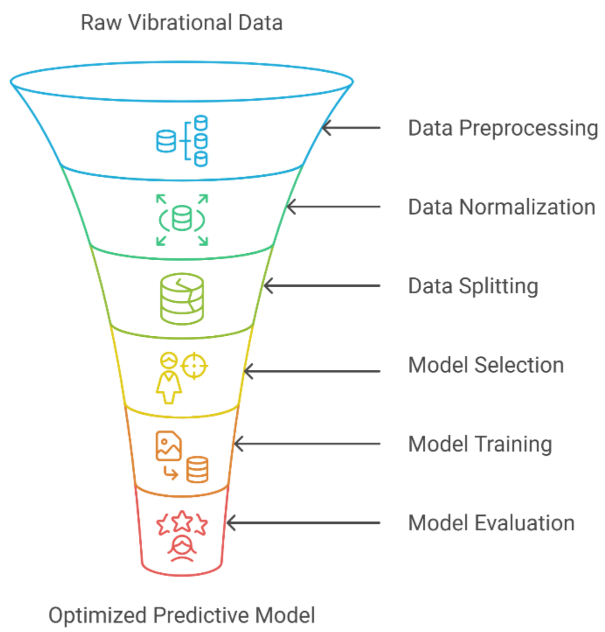

Figure 10. From Data Collection to the Model Evaluation: Proposed Model.

3.3 Data Preprocessing

According to the Gong et al. (2023), the accuracy and the performance of the model highly dependent on the quality of the data fed during training of the model. Therefore, the preprocessing of the collected vibrational data was the essential step to enhance the performance of the ML and DL models. To enhance the quality of the data, the following pre-processing steps were involved:

3.3.1 Data Cleaning:

Data cleaning is the crucial step in the process of the ML and DL. The data collected while the experiment was recorded in the single folder. The different ops were separated by using the time stamps recorded during the experiment and shown in Table. 1 and Table. 2. Many different cleaning techniques were applied without affecting the information preserved in the data.

3.3.2 Feature Engineering:

This step involves finding the best features to be fed into the ML model. To extract the features from a raw vibrational data, time domain, frequency domain and time-frequency domain features are extracted. In case of Time domain features such as Root Mean Square (RMS), Skewness, etc. were extracted. In case of frequency domain such as Root Variance Frequency (RVF), Frequency Centre (FC) and Root Mean Square Frequency (RMSF) were extracted. Similarly, the time-domain features were also extracted. In case of DL, the Fast Fourier Transform (FFT) is applied to all the raw vibrational data. Then, convert the signal into the vector of frequency domain signals. These vectors will be then transformed into the same class pair and different class pairs to feed the Siamese neural network.

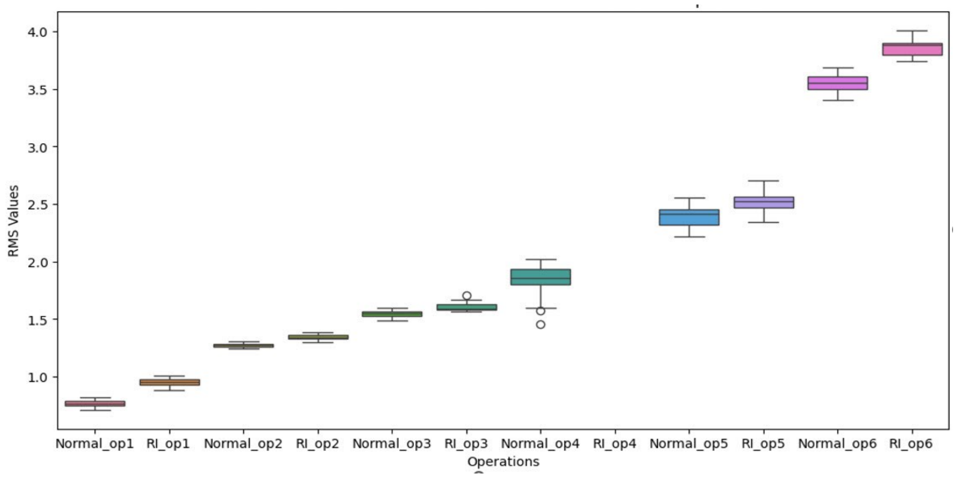

Figure 11. Box Plot of RMS Values (Normal and Rotor Imbalance Operations)

3.3.3 Z- Score Normalization:

The normalization technique used for the transformation of the data is Z-Score, which transformed the data with zero mean and the standard deviation of 1. This normalization is a good approach when dealing with the distance-based metrics. The formula for the Z-score transformation is:

$$z = \frac{x - \mu}{\sigma}$$

Here,

x = Data points (one by one),

μ = Mean,

σ = Standard deviation.

3.4 Model Selection:

The model selection for the classification of the fault is the most crucial part after the cleaning and the processing of the data. Because it is entirely dependent on the nature of the data and the desired outcomes. Both conventional ML and DL models were applied to find the best possible results. The following are the methods were applied during the research:

1. **Principal Component Analysis (PCA),**
2. **Factor Analysis (FA),**
3. **Locality Preserving Projection (LPP),**
4. **Support Vector Machine (SVM),**
5. **K-Nearest Neighbours (KNN),**
6. **Siamese Neural Network (SNN) (DL method).**

3.4.1 Machine Learning & Deep Learning Models:

Each model chosen for the study has its own ability to extract various aspects from the given data. These ML models described below are significant enough to locate the best features in identifying the fault. The technique i.e. SVM and KNN is explained in the chapter 2. The remaining model used while research will be explained below:

3.4.1.1 Principal Component Analysis:

PCA is applied to transform the feature domain from high dimensional space to the low dimensional space, while preserving most of the variance of the data. The research was conducted on the collected vibrational data, which usually have correlated features. Thus, making it difficult to analyse. Therefore, PCA was applied to reduce the noise and the outliner. PCA also improves the classifier performance and help to visualize the high dimensional data. It projects the data in any required number of principal component and make it easy to visualize. During the study, two principal components were used to visualize the data, and it helps to improve the performance of the SVM and K-NN classifier.

3.4.1.2 Factor Analysis (FA):

FA is also a dimensionality reduction technique. The goal of the PCA is to find the maximum variance in the data and transform into a new low dimensional feature domain. In case of FA, it focuses more on the underlying latent factor and find the correlation between the observed variable. It tries to explain the variance among the variable rather than maximizing it as in the case of PCA. Therefore, FA is essential to find the hidden relationships of the data, which is useful in the case of the vibrational data. FA also facilitates the classier model and safe them from overfitting.

3.4.1.3 Locality Preserving projection (LPP):

LPP is also a dimensionality reduction technique as of PCA and FA. While conducting the research, it was assumed that the vibrational data may contain the complex and nonlinear relationships. Therefore, the LPP was proposed because it is a nonlinear method to reduce the feature domain in low dimension space, but it preserves the local relationships, which is a remarkable trait. Thus, the points which were closed to each other in the high dimensional space will be somehow close to each other in the low dimensional space.

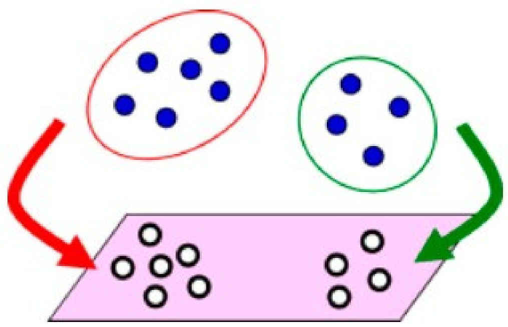

Figure 12. Locality Preserving Projection After Dimensionality Reduction (Sugiyama, M. (2016))

- **Siamese Neural Network (SNN):**

SNN is a widely used for the classification problem and the trait which makes it distinct from other models is its pairwise handling capacity. Unlike the traditional model, SNN are designed to capture the similarity and dissimilarity between two inputs and define it belongs to the same class or not.

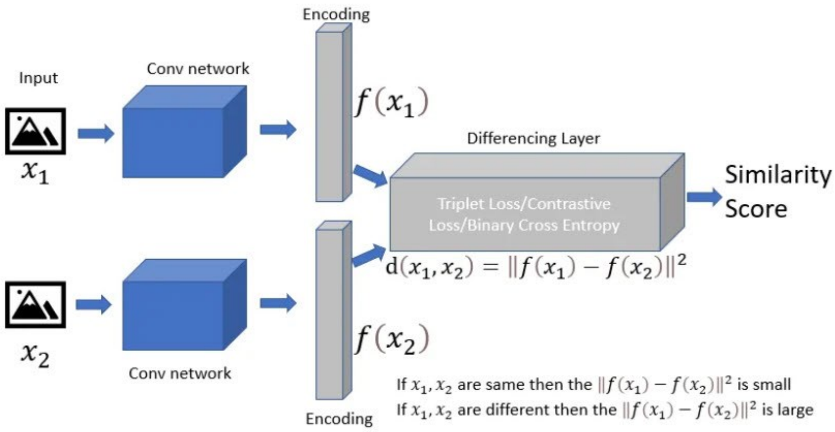

Figure 13. Siamese Network (Khandelwal, R. (2021).

SNN consists of two identical neural network that shares the same weight and are trained to compare the pairs of samples. When two inputs are compared on the same neural network (Feed Forward Neural Network), the resulting feature vector are compared based on their distance metric. If the distance is quite small and close to 0, this means the input pairs are of similar class, and if the distance is quite high or above the threshold level then the input pair was of dissimilar class.

3.5 Reasoning for the Selection of models:

The reason of selecting these models is justified below:

3.5.1 Complexities of Data:

As the raw vibrational data will be transformed into a feature domain, to address this complexity and preserving the features credibility such dimensionality reduction techniques (PCA, FA, and LPP) were selected.

3.5.2 Computational Efficiency:

To match the system with the computational capacity of the device, it is always desirable to select the computational efficient model. The model like SVM and K-means were selected because they are efficient, in term of computation.

3.5.3 Complex Structure:

The SNN were selected to enhance the fault-finding capabilities of the system. SNN can extract the hidden pattern, thus helps us finding the fault types more efficiently.

3.6 Performance Evaluation Parameter:

The performance evaluation of these model was based on these factors:

- **Accuracy Score:**

The accuracy score is a metric to determine the overall performance of the model. It is of range between 0 to 1 and it is defined by:

$$Accuracy\ Score\ = \frac{No.of\ Correct\ Prediction}{Total\ No.of\ Prediction}$$

- **Confusion Matrix:**

Unlike the accuracy score, the confusion matrix, it provides an intensive overview of the classifier in term of true class with respect to the predicted class. The confusion matrix is a good visualization of the model performance to find out where it is making an error. The key parameters are:

  - **True Positive (TP):**
    No. of correctly predicted positive instances of the targeted class.

  - **False Positive (FP):**
    No. of incorrectly predicted positive instances of the targeted class.

  - **True Negative (TN):**
    No. of correctly predicted negative instances of the targeted class.

  - **False Negative (FN):**
    No. of incorrectly predicted negative instances of the targeted class.

- **Classification Report:**

The classification report is also a comprehensive performance evaluation of the model compared to how well it was performing when dealing with multiple classes. The Key parameters are:

- **Precision:**

It helps in identifying the model's performance in finding the positive instances of the targeted class.

$$\boldsymbol{Precision} \ = \frac{\boldsymbol{TP}}{\boldsymbol{TP}+\boldsymbol{FP}}$$

- **Recall:**

It is the ratio of the TP to the sum of TP and FN. It helps in identifying all actual positive instances of the targeted class.

$$\boldsymbol{Recall} \ = \frac{\boldsymbol{TP}}{\boldsymbol{TP}+\boldsymbol{FN}}$$

- **F1- Score:**

This parameter is important in defining the performance of the model, in case we are concerned about both recall and the precision. This is important when the data is unbalanced. The formula is:

$$\boldsymbol{F1-Score} = \frac{\boldsymbol{2 * Precision * Recall}}{\boldsymbol{Precision + Recall}}$$

4 Research Result and Analysis:

This chapter will outline the steps involved in processing the data and the results obtained during the research conducted on the collected vibrational data. The purpose of this chapter is to highlight not just only the results but also answers the research question raised at the beginning of the study. The following chapter will provide the results of the research and contribute to understands the need of early fault prediction of the motors.

4.1 Data Cleaning Process

4.1.1 Sorting of Vibrational Data:

As described in Chapter 3, the raw data containing all the OPs were collected in a single folder. The data points were recorded in a CSV file. The sensor was recording the 2048 data points with the sampling frequency of 6.66kHz. Each CSV consist of file name, which is the name of the sensor identity and the time at which it was recorded e.g. acc_ID35_20250128_110142.

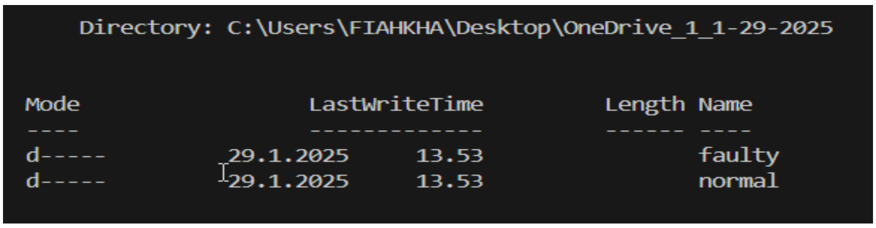

Figure 14. Raw Recorded Data.

There are 6 and 5 OP’s in case of Normal and Rotor Imbalance (RI) respectively, which were then separated based on the time stamps defined in the Table. 2 and Table. 3. The build-in sort () function of the python was used to sort the file based on the time stamp i.e. files.sort(key=lambda x: extract_time_from_filename(x)) and then stored in the respective folders.

```
    Directory: C:\Users\FIAHKHA\Desktop\Normal_Condition


Mode                 LastWriteTime         Length Name
----                 -------------         ------ ----
d-----          9.4.2025     21.33                Normal_op1
d-----          9.4.2025     21.33                Normal_op2
d-----          9.4.2025     21.33                Normal_op3
d-----          9.4.2025     21.33                Normal_op4
d-----          9.4.2025     21.33                Normal_op5
d-----          9.4.2025     21.33                Normal_op6
```

Figure 15. Normal Condition (Sorted OP's)

```
    Directory: C:\Users\FIAHKHA\Desktop\Rotor_Imbalance_Condition


Mode                 LastWriteTime         Length Name
----                 -------------         ------ ----
d-----          9.4.2025     21.41                Rotor_Imbalance_op1
d-----          9.4.2025     21.41                Rotor_Imbalance_oP2
d-----          9.4.2025     21.41                Rotor_Imbalance_oP3
d-----          9.4.2025     21.41                Rotor_Imbalance_op4
d-----          9.4.2025     21.41                Rotor_Imbalance_op5
```

Figure 16. Rotor Imbalance Condition (Sorted OP's).

4.1.2 Analysis of Data:

- **Plotting of the Data:**

After sorting of the data and completing the sifting of the data. The data was plotted to have a glimpse of the recorded vibration data to see its changing pattern over time. The condition 'Normal' refers as the normal state of the motor (under no influence of the fault), 'RI59' refers as the rotor imbalance condition approximately with 50% speed of the motor and ‘RI100’ refers as the rotor imbalance condition with 100% speed of the motor. The figure below shows the vibrational pattern for each condition.

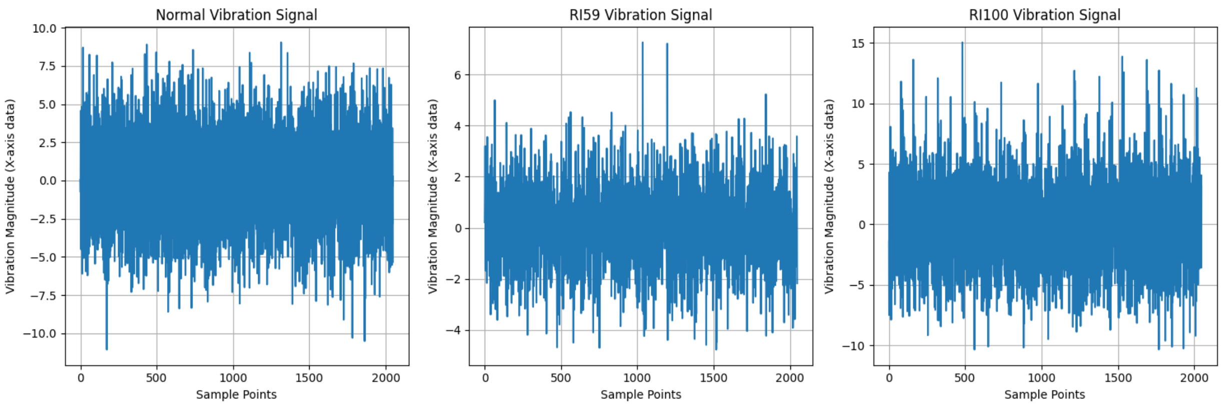

Figure 17. Plot of Vibrational Data for Three Different Conditions (Full Load)

- **Finding the Missing Value**

Once plotting the data, it was needed to find out if there’s any missing values in the recorded vibration data. The built-in function isna (). sum () of python was used such as df.isna().sum().

```
Missing values in each column:
0        0
1        0
2        0
3        0
4        0
        ..
2044     0
2045     0
2046     0
2047     0
label    0
Length: 2049, dtype: int64
```

Figure 18. Missing Values in the Data.

- **Duplicate Values:**

The built-in functions duplicated (). sum () was used such as df.duplicated().sum() to find out if there's any duplicated value in the rows of the measured vibrational data.

```
Duplicated values in each row:
0        0
1        0
2        0
3        0
4        0
        ..
2044     0
2045     0
2046     0
2047     0
label    0
Length: 2049, dtype: int64
```

Figure 19. Duplicated Values in Each Row.

- **Outliers:**

In case of the time series data set such as vibrational data, it is assumed that the dataset might contain outliers. The outliers are the values which drastically changes from the mean of the dataset. In this case seaborn library of the Python was used to plot the distributions of the values. A few outliners were found but these outliers were rectified in the further cleaning process i.e. by utilizing the dimensionality reduction techniques. The box plot below indicates 3.55 as the median value and the range define the spread of the values from median of OP6.

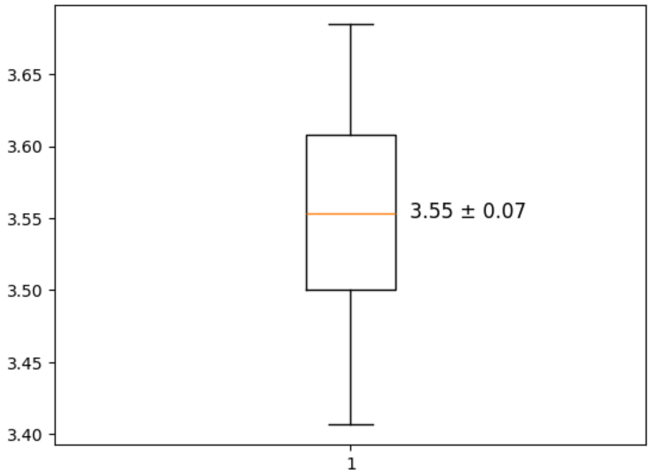

Figure 20. Box Plot of Normal Condition OP6.

4.2 Feature Extraction:

Feature extraction is one of the crucial parts in the development of the ML and DL models. It is a process in which the raw data is transformed into reasonable information and present them in a suitable format so that it can be used to train our ML and DL models. The aim of the feature extraction is not just to transform it from one form to another form but also to preserve the useful information.

4.2.1 Feature Extraction for ML:

According to Cai et al. (2018), feature selection (engineering) is the most important part in the machine learning to reduce the computational time and improve the accuracy. This highlights the importance of this feature selection process. The following are the steps performed to create a feature set to feed the machine learning model.

4.2.1.1 Time Domain Feature of Vibrational Data:

The data being investigated was a real time vibrational data. The time domain feature from the raw time domain signals can be extracted directly without changing it into a different domain. The following are the features obtained from the raw time domain data:

No.	Feature:	Advance Explanation
1	**Root Mean Square (RMS)**	RMS is the magnitude of the varying signal. (Finds the power of the signal)
2	**Square Root Amplitude (SRA)**	SRA is the average amplitude of the signal. (Finds the energy content of the signal)
3	**Kurtosis Value (KV)**	KV is useful for finding the outliers. (Finds the sudden changes)
4	**Skewness Value (SV)**	SV finds the asymmetry of probability distribution of the values.
5	**Peak to Peak Value (PPV)**	PPV defines the difference between the max and min values of the signal. (Finds the range how the signal is varying)
6	**Crest Factor (CF)**	It is the ratio of the Peak value to the RMS value of the signal. (Finds if there's a peak in the signal)
7	**Impulse Factor (IF)**	It is the ratio of the peak value to the mean value. (Finds the impulse in the signals)
8	**Shape Factor (SF)**	SF defines the shape of the signal. (Useful in finding the shape of the signal e.g. how close it to the known shape i.e. sinusoidal)
9	**Margin Factor (MF)**	MF defines the how close the values of the signals to the maximum values.
10	**Kurtosis Factor (KF)**	KF is also a useful factor in finding the outliner in the signal.

4.2.1.2 Frequency Domain Feature of the Vibrational Signal:

The frequency domain feature can be obtained from a raw time domain signal by transforming it to a frequency domain by utilizing a technique called Fast Fourier Transform (FFT). The following are the features obtained from the transformed signal:

No.	Feature	Advance Explanation
1	**Root Mean Square Frequency (RMSF)**	It is the square root of the average of the squared frequencies. (Finds the spread of the frequencies)
2	**Frequency Centre (FC)**	It is the ratio of the average of the frequency over the weighted Power Spectral Density (PSD). (Finds the central frequency)
3	**Root Variance Frequency (RVF)**	It finds the spread of the power of the frequency around the mean frequency.

4.2.1.3 Time-Frequency Domain Feature of the Vibrational Signal:

The Time-frequency domain feature can be obtained by transforming the raw time-domain signal. The Wavelet Packet Node Entropy (WPNE) which is the Time-Frequency feature can be obtained from the following step:

1. The time-domain signal is transformed into multiple frequency sub bands using wavelets packet decomposition (WPD). This transforms the signal into a treelike structure.
2. The next step is to calculate the energy of each node in the decomposition tree.
3. Then calculate the entropy of each node. The entropy is determined by the uncertainty of the signal at the specific frequency band.
4. In the final step, aggregate the entropy value of all the nodes to obtain the final WPNE value.

4.2.1.4 Feature Set:

A feature set is the collection of the various feature (characteristics) on which the ML model is trained. The performance of the ML is entirely dependent on the quality of the features extracted. The above extracted features were then merge into a single feature set, which is shown below:

Table 4. Feature Set of Time Series Vibrational Data (ML)

Fault Type	Normal	Normal	RI59	RI59	RI100	RI100
RMS	3.502126	3.499621	1.58899	1.572957	3.895793	3.797012
SRA	0.047192	0.037481	0.04202	0.041013	0.042563	-0.022221
KV	-0.686630	-0.648760	0.07570	-0.188767	-0.065339	0.072241
SV	3.501994	3.499008	1.58750	1.572897	3.895535	3.796806
PPV	20.22055	20.11885	10.68396	10.71386	28.9723	26.54837
CF	2.812251	2.584185	3.844486	3.444054	3.980981	3.312886
IF	-0.158228	0.067409	0.043267	-0.193455	0.168918	0.15517
MF	2.938141	2.925143	1.262891	1.264851	3.168806	3.055428
SF	1.191952	1.196392	1.258222	1.243590	1.229419	1.242710
KF	-0.68663	-0.64876	0.075704	-0.188767	-0.065339	0.072241
FC	754.5454	753.7314	329.6557	329.6557	754.5454	754.5454
RMSF	500.8187	512.7066	457.9430	460.0124	497.4328	491.3608
RV	2.961540	3.164679	3.844486	3.444054	3.980981	3.679023
WPNE(4,1)	272.2877	337.3427	1.587501	1.572897	3.895535	3.796804
WPNE(4,2)	386.3230	217.1783	592.1106	521.5179	592.1106	521.5179
WPNE(4,3)	592.1106	521.5179	427.5085	420.1316	752.4342	254.9858
WPNE(4,4)	427.5080	420.1316	592.1106	521.5179	592.1106	521.5179
WPNE(4,5)	517.8450	292.6155	593.1619	270.5692	13835.09	237.5894
WPNE(4,6)	592.1106	521.5179	427.5085	420.1316	592.1106	521.5179
WPNE(4,7)	1224.533	513.7413	592.1106	521.5179	859.7268	207.1488
WPNE(4,8)	595.7805	682.9746	592.1106	521.5179	13835.09	237.5894
WPNE(4,9)	855.2660	196.3810	859.7268	207.1488	2619.7551	193.7482
WPNE(4,10)	13835.09	237.5894	427.5085	420.1316	427.5085	420.1316
WPNE(4,11)	593.1619	270.5692	859.7268	207.1488	859.7268	207.1488
WPNE(4,12)	2619.755	193.7482	593.1619	270.5692	427.5085	420.1316
WPNE(4,13)	752.4342	254.985	13835.09	237.5894	595.7805	682.9746
WPNE(4,14)	1409.468	282.7796	855.2660	196.3810	595.7805	682.9746
WPNE(4,15)	859.7268	207.1488	595.78058	682.9746	2619.755	193.74821
WPNE(4,16)	629.4372	299.4388	2619.755	193.7482	13835.09	237.58949

4.2.2 Feature Extraction for DL:

The feature extraction is as important as in the case of DL as in ML. The measured vibrational time series data passed through the following steps to extract the features.

The steps are as follow:

1. The vibrational time series data will be transformed into a frequency domain signal using the FFT.
2. The obtained FFT will be in a shape of the FFT vector which is itself a feature domain for feeding the DL model.

```
First 5 FFT Feature Vectors:
[[ 94.97217     48.85322494  36.29402588 ...  72.37123409  88.98708284
   24.87176012]
 [134.15979     40.85295758  74.27247663 ...  23.2021941   44.99758753
   46.90933216]
 [104.43473    102.54523256  42.58578642 ...  30.89782809  44.48917581
   38.75077678]
 [ 43.6474      33.791118    29.94758686 ...  43.55807564  30.97842623
   51.43763724]
 [162.66168     90.16719107  24.16471037 ...  33.90693969  57.21506463
   57.28275803]]
```

Figure 21. Feature Vector.

Unlike the ML models, DL model does not require many features to map between the input and the output.

4.3 Evaluations and Results:

This section will present the most important part of the research, which are the results. The results of the ML and DL model will be presented separately along with the approaches used before feeding the features to train the model.

4.3.1 ML Performance:

In case of machine learning, the extracted feature set will be utilized to train the model, the model's performance can be improved by using the dimensionality reduction technique. The results of the dimensionality reduction are shown below:

4.3.1.1 PCA:

In case of the PCA, the feature set was transformed into four principal component and the covariance ratios were 61%, 21%, 4% and 3%. Once the dimensionality reduction is applied, the plot of the PC1 along with the PC2 had shown that the conditions can be grouped together, which means they are distinct.

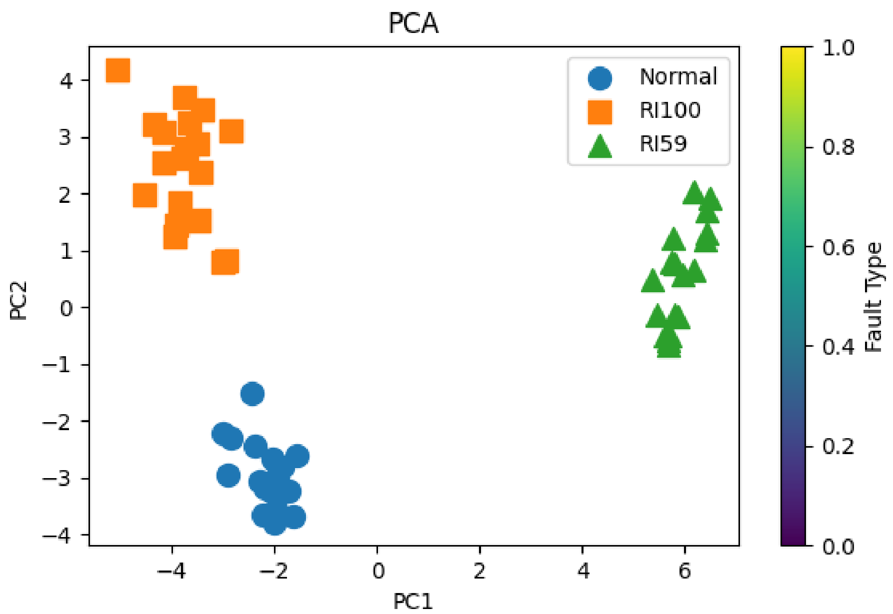

Figure 22. PCA Plot (PC1 vs PC2)

4.3.1.2 LPP:

In case of the LPP, the feature set was transformed into four n-components. Once the dimensionality reduction was applied, the two-dimensional plot of the LPP1 along with the LPP2, also shows the results of clustering where each condition is distinguishable.

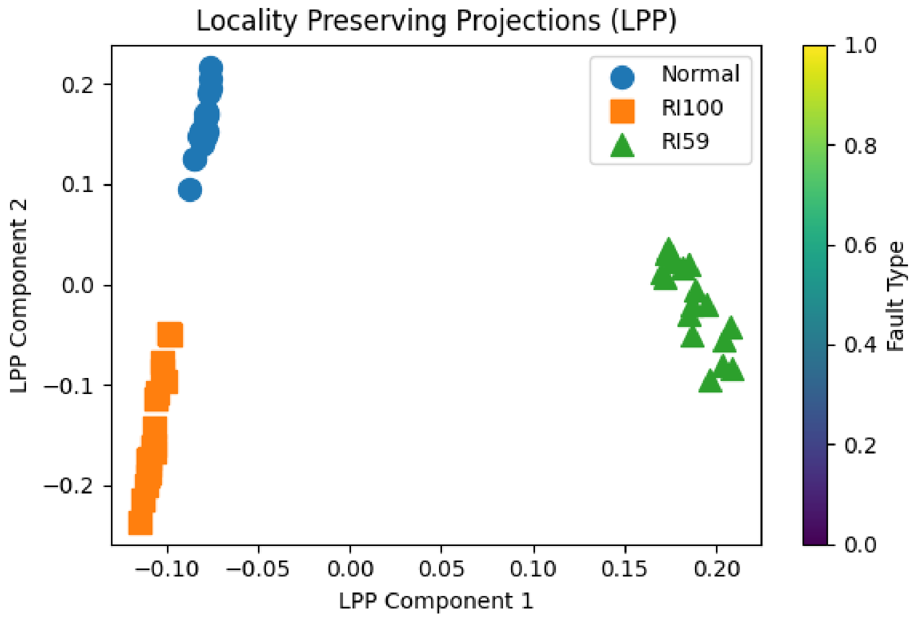

Figure 23. LPP Plot (LPP1 vs LPP2)

4.3.1.3 FA:

FA is also a dimensionality reduction technique; it was applied to the feature set with four n-component were selected. This also proves to be the good technique for the clustering the different conditions.

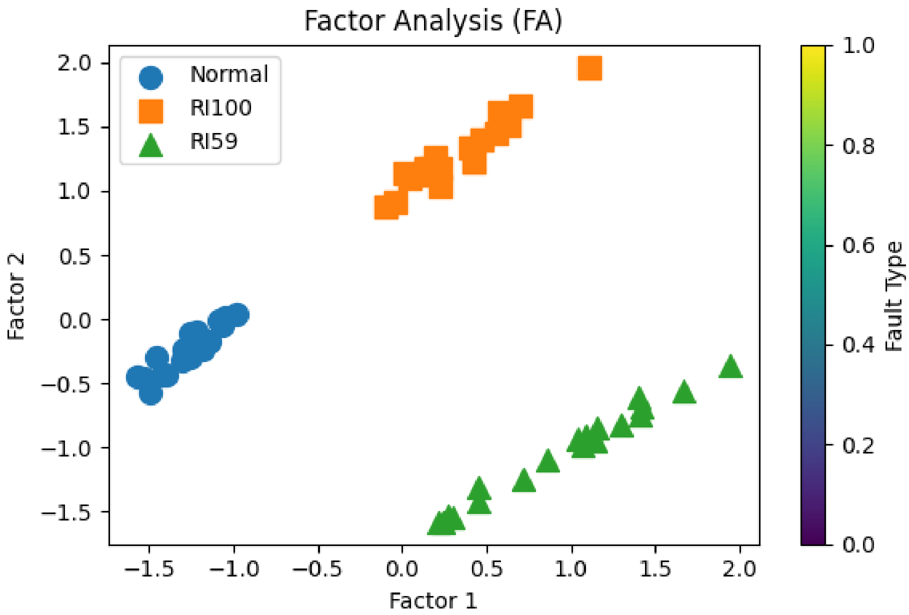

Figure 24. FA Plot (Factor 1 vs Factor 2)

4.3.1.4 SVM:

The results of the dimensionality reduction shows that these faults are distinguishable therefore the performance of the SVM was remarkable with a very high accuracy. The train split ratio was 70% training and 30% testing. The performance of the SVM can be seen with the predicted label in the confusion matrix shown below.

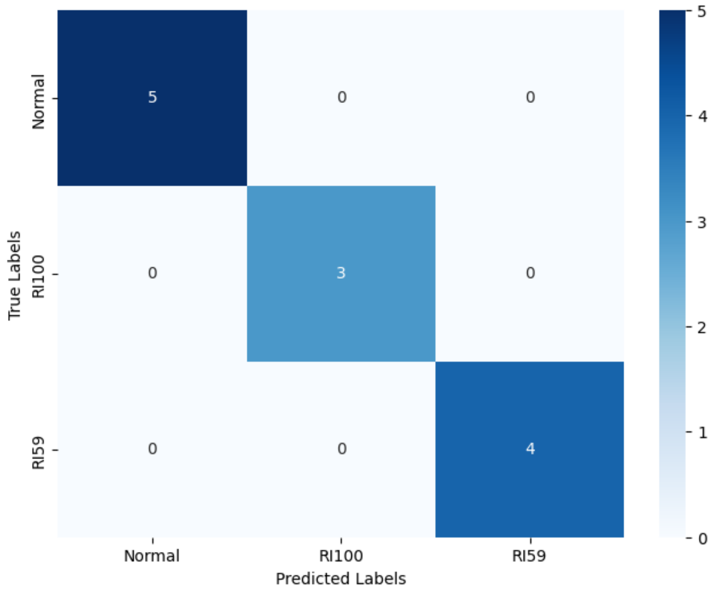

Figure 25. Confusion Matrix SVM.

4.3.1.5 KNN:

KNN also shows the best performance after the dimensionality reduction and the train split ratio was the same as in the case of the SVM. The confusion matrix of KNN is shown below:

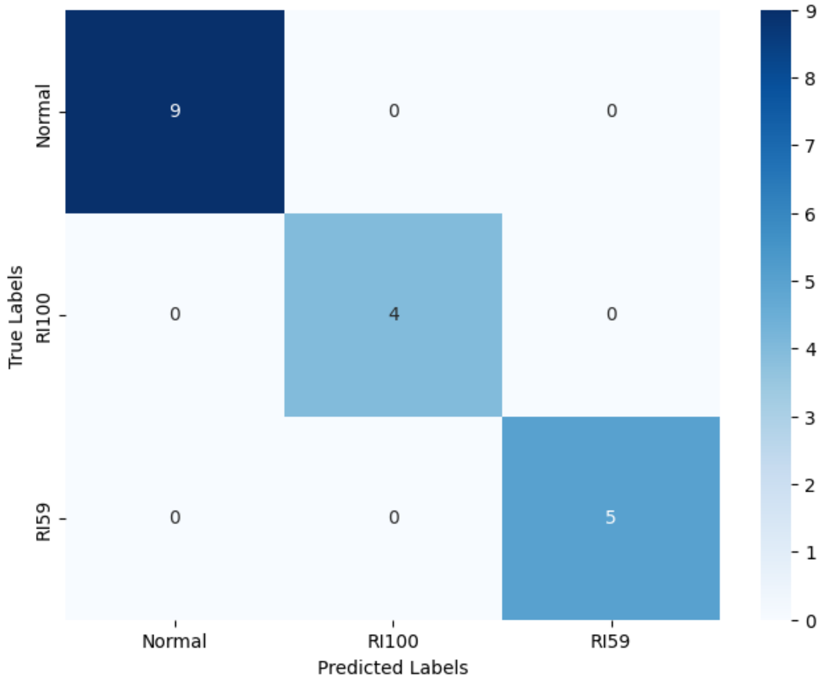

Figure 26. Confusion Matrix KNN

The pipeline diagram of the ML defining the whole process is shown below:

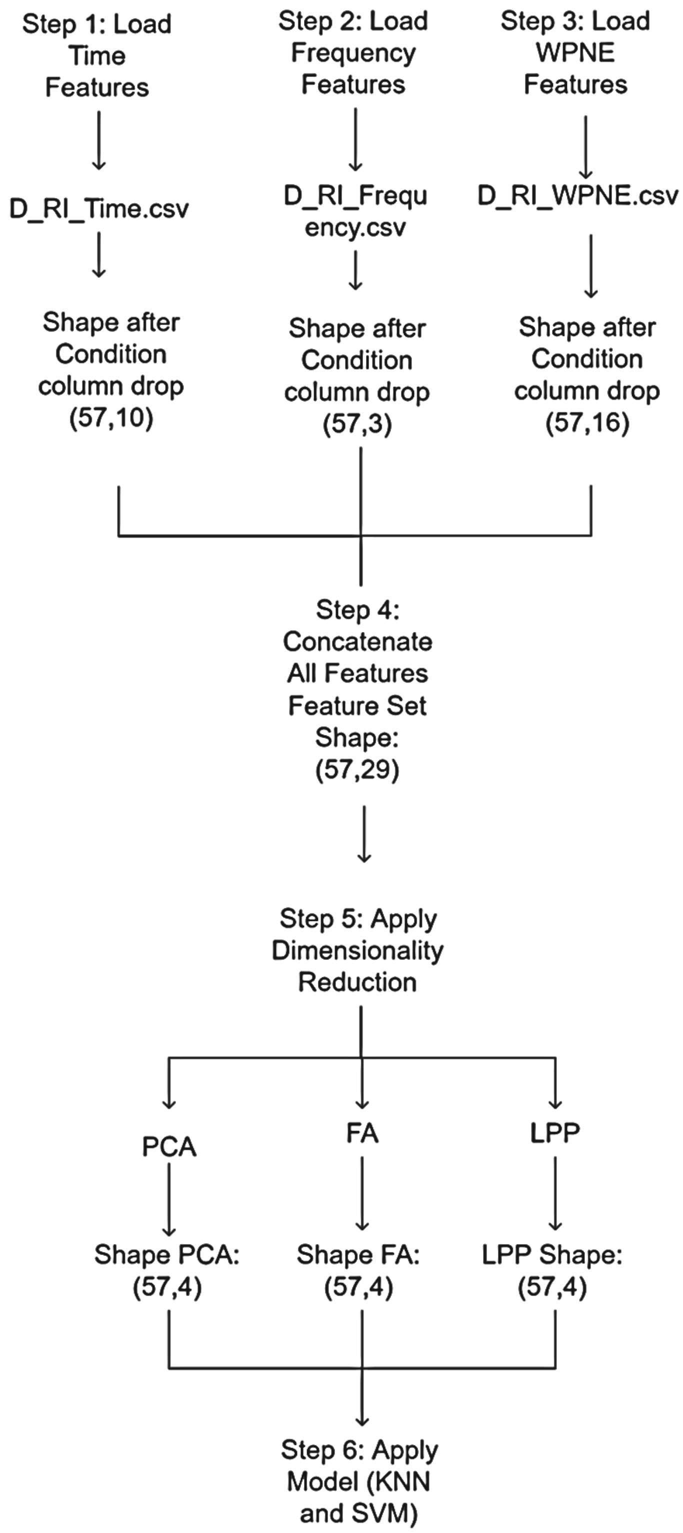

Figure 27. Pipeline Diagram of ML

4.3.2 DL Performance:

In case of Deep Learning, the feature vector extracted from the FFT of the feature set were used to create the pairs for the training and the testing of the SNN. As described in the previous chapter that the input to the SNN is the pair and the distance matric is used to identify the similarity and the dissimilarity. The Loss function employed was the 'Binary Cross Entropy' and the optimizer employed was the 'Adam. The batch size was 32 (usually adjusted to minimize the training time) and the number of epochs was 15.

4.3.2.1 Normal Vs ALL Condition:

In case of finding the fault, it is desirable that the motor always stays in the normal condition. Therefore, the model was trained to find the faulty condition other than the normal condition. If any variation comes over the threshold, that will be notified as a fault in the motor. The performance accuracy of finding the label was **98.76 %**. The SNN performed well in this case and the performance of the SNN can be seen on the confusion matrix below:

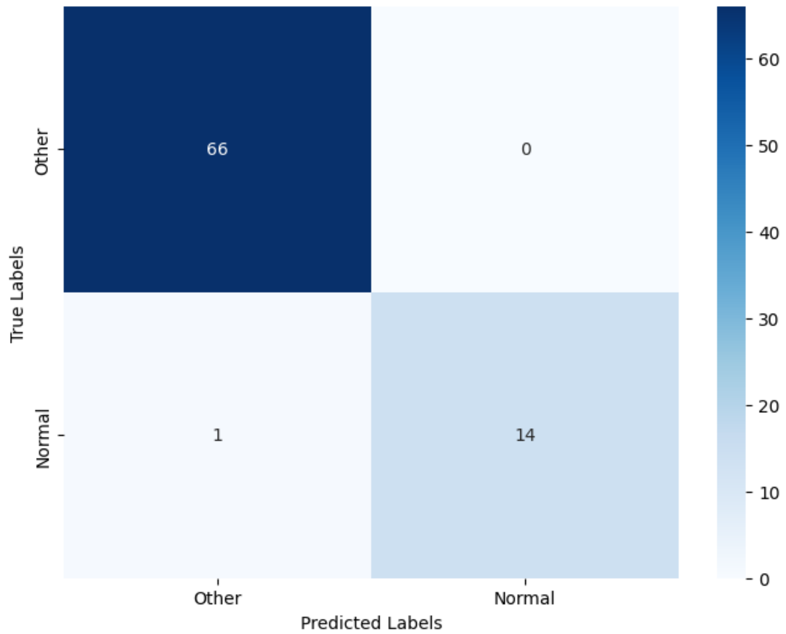

Figure 28. Confusion Matrix (Normal vs ALL).

4.3.2.2 Specific Class Pair Condition:

In this first experiment, the conditions were paired in specific ways from 0 to 5. The labelled pair with the number of trained pairs is shown below. The accuracy of the model was affected because of the Label pair No. 1, because the number of data point in that specific case was fewer than the other same condition pairs. The accuracy of predicting the condition was **83%.**

Label Map:

label no. 0; Pairs (**'Normal', 'Normal')** **:** 380

label no. 1; Pairs **('RI59', 'RI59')** : 110

label no. 2; Pairs **('RI100', 'RI100')** **:** 702

label no. 3; Pairs **('Normal', 'RI59')** : 96

label no. 4; Pairs **('Normal', 'RI100')** : 274

label no. 5; Pairs **('RI59', 'RI100')** **:** 172

The confusion matrix of the first experiment is shown below:

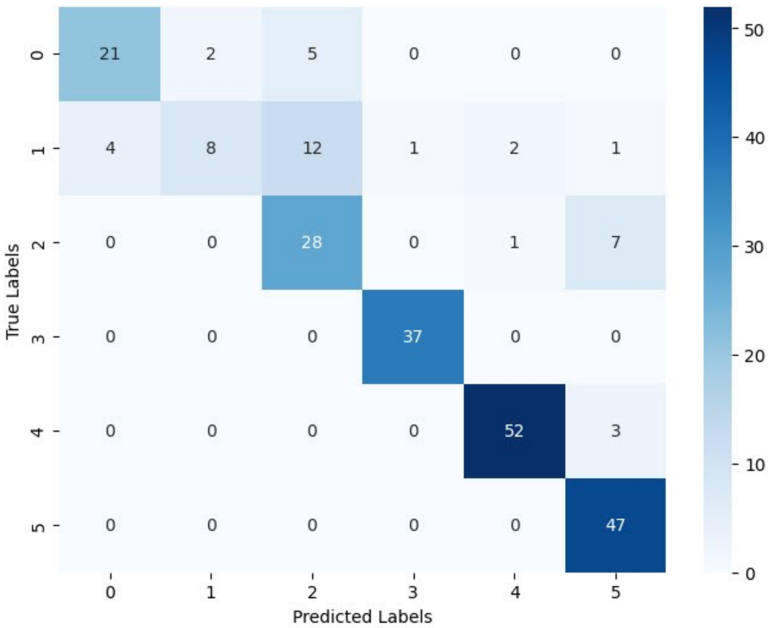

Figure 29. Confusion Matrix Specific Class Pair (Label 0-5).

In this second experiment, the conditions were label in specific ways from 0 to 4 and the condition which has the fewer number of pairs than the other in the first experiment were removed. The labelled pair with the number of trained pairs is shown below. The accuracy of the model was drastically increased because of because of the removal the Label pair No. 1 which had the fewer pairs. The accuracy of predicting the condition was **94 %.**

Label Map:

label no. 0; Pairs (**'Normal', 'Normal')** **:** 380

label no. 1; Pairs **('RI100', 'RI100')** **:** 702

label no. 2; Pairs **('Normal', 'RI59')** **:** 96

label no. 3; Pairs **('Normal', 'RI100')** : 274

label no. 04; Pairs **('RI59', 'RI100')** **:** 172

The confusion matrix of the second experiment is shown below:

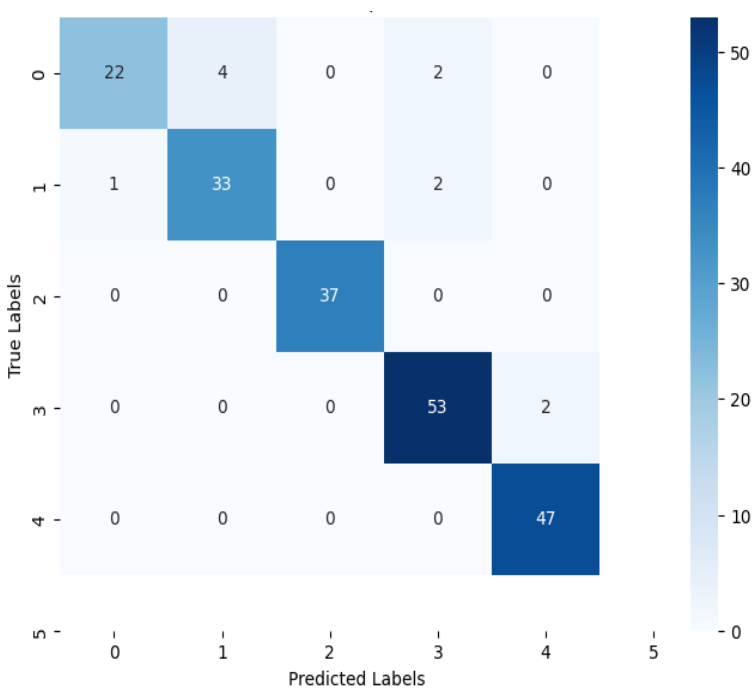

Figure 30. Confusion matrix Specific Class Pair (Label 0-4).

The table defines the performance of the DL experiment:

Table 5. Performance of the DL Model.

No.	Labelled Mapped	Accuracy
1	Normal Vs All	98.76%
2	0-5	83%
3	0-4	94%

The pipeline diagram of the DL defining the whole process is shown below:

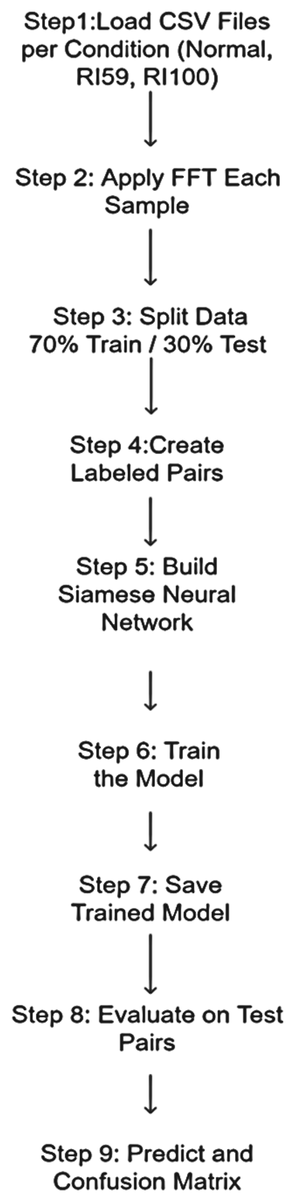

Figure 31. Pipeline Diagram of DL

5 Conclusion

The Study was conducted to distinguish the condition of the motor and find out either the motor working in a normal or a faulty condition. The importance of the study can be depicted with the fact that most of the fault in the motor are because of the poor lubrication, bearing problems and misalignment such as rotor imbalance. Each fault brings the change in the vibration of the motor. The primary objective of the research was to investigate that “Are these conditions such as normal and rotor imbalances are distinguishable based on the vibrational dataset?”.

Key Findings:

1. **Fault Prediction:**

Based on the research and results, it is verifiable that the different conditions of the motor can be distinguished. In case of the ML the results were quite good, but in DL, Siamese architecture provided inferior to ML but still good. If we acquire more data, then it might also outperform as ML.

2. **ML and DL Feature Extraction:**

Even though most of the time was utilized by the data cleaning and bringing it in a useful form, feature extraction does not consume much time.

3. **Memory Optimization**:

While looking for the models, one important thing which was under consideration was that the model should not utilize much memory. The model such as SVM is memory efficient model. Somehow, in case of DL, it requires a bit higher memory as compared to the ML model. DL model is still a good match because of its intrinsic nature of finding the complex pattern within the dataset.

4. **Computational Time:**

Comparing the performance of both DL and ML models based on the computational time, each model performs adequately (less than a minute for training and testing) based on the task being performed.

In conclusion, the performance of the ML and DL models are entirely dependent on the quality of the data availability for the training of the model. Moreover, the accuracy of the model depends on the size of the data set which can be split into training, testing and validation set. In case of ML, the model was predicting much better because the data preprocessing was done efficiently, therefore large amount of data was not needed. But, in case of DL, the quantity of data matters a lot. During the research, in case of DL, to predict the same class such as Normal vs Normal or RI59 vs RI59, for training the model it requires more data to create more pairs to feed the SNN. Both ML and DL has its limitations based on the task being performed. Additionally, in case of ML, the raw data needs to be passed through lots of preprocessing steps to relief the training of the model and best features need to be highlighted. While DL models have some advantages of capturing the hidden features of the raw data but here the quality of data holds a great significance. In summary of the research, both ML and DL works efficiently because the raw data was preprocessed quite a lot.

Acknowledgement

In the end, I would like to give gratitude to the R&D department of ABB Oy in Vaasa Finland and the CRC department of ABB AB in Västerås Sweden for allowing to conduct the research and working with their motors and equipment. I really admire the commitment of ABB Oy towards the sustainability development and making the motors reliable and efficient enough to make smart decisions and ensuring the investment of the customer stays safe. I hope more exploration in this sector of motors with the collaboration of different CRCs of ABB Oy.

Reference

ABB. (2021). *PPM IE4.* Retrieved from https://library.e.abb.com/public/833c42cc056142d7a5444041132b03a7/9AKK107424%20EN%2003-2021%20PPM%20IE4.pdf

ABB. (n.d.). *Driving down industrial energy consumption.* Retrieved from https://new.abb.com/drives/highlights-and-references/driving-down-industrial-energy-consumption

Amruthnath, N., & Gupta, T. (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. In *Proceedings of the Conference on Industrial Electronics and Applications.*

Başer, F., & Apaydin, A. (2015). Comparison between logistic regression and support vector machines for classification purposes. *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi - B Teorik Bilimler, 3*(2), 53–65. https://doi.org/10.20290/btdb.67263

Cai, J., Luo, J., Wang, S., & Yang, S. (2018). Feature selection in machine learning: A new perspective. *Neurocomputing*, *300*, 70–79. https://doi.org/10.1016/j.neucom.2017.11.077

Cauwenberghs, G., & Poggio, T. (2000). Incremental and decremental support vector machine. *In Advances in Neural Information Processing Systems* (Vol. 12, pp. 409–415). MIT Press.

Chattopadhyay, P., Saha, N., Delpha, C., & Sil, J. (2018). Deep learning in fault diagnosis of induction motor drives. In Proceedings of the 2018 Prognostics and System Health Management Conference (PHM-Chongqing) (pp. 1068–1073). IEEE. https://doi.org/10.1109/PHM-Chongqing.2018.00189

Choi, D.-J., Han, J.-H., Park, S.-U., & Hong, S.-K. (2019). Data preprocessing method in motor fault diagnosis using unsupervised learning. In *Proceedings of the International Conference on Control, Automation and Systems*.

Elidan, G. (2013). Copulas in machine learning. In *Lecture notes in statistics* (pp. 39–60). https://doi.org/10.1007/978-3-642-35407-6_3

Fürnkranz, J. (2020). Decision tree. In *Encyclopedia of machine learning and data mining* . Springer.

Garfinkle, J. (2024). Automation to approximate the Gestalt: applying machine learning to the general movement assessment. *Pediatric Research*. https://doi.org/10.1038/s41390-024-03558-w

Godara, S., & Kumar, P. (2012). An enhanced hierarchical clustering algorithm using WEKA tool. *ACADEMICIA: An International Multidisciplinary Research Journal*, *2*(7), 69–74.

Gong, Y., Liu, G., Xue, Y., Li, R., & Meng, L. (2023). A survey on dataset quality in machine learning. *Information & Software Technology*, *162*, 107268. https://doi.org/10.1016/j.infsof.2023.107268

He, X., & Niyogi, P. (2002). Locality preserving projections (LPP). *IEEE Transactions on Reliability, 16*, 153-160.

Heller, S. (2022). Supervised learning: From theory to applications. In *Microsoft* (Ed.), *Elsevier eBooks* (pp. 23–32). Elsevier.

International Energy Agency. (2015). *World Energy Outlook.* Retrieved from https://www.iea.org/reports/world-energy-outlook-2015

Inevoll, G. T. (2006). Mathematical modeling of neural activity. In A. T. Skjeltorp & A. V. Belushkin (Eds.), *Dynamics of complex interconnected systems: Networks and bioprocesses* (NATO Science Series II, Vol. 232, pp. 123–142). Springer. https://doi.org/10.1007/1-4020-5030-5_8

Iunusova, E., Gonzalez, M. K., Szipka, K., et al. (2024). Early fault diagnosis in rolling element bearings: Comparative analysis of a knowledge-based and a data-driven approach. *Journal of Intelligent Manufacturing, 35*, 2327–2347. https://doi.org/10.1007/s10845-023-02151-y

Kawaguchi, K., Bengio, Y., & Kaelbling, L. (2022). Generalization in deep learning. *Cambridge University Press*.

Khandelwal, R. (2021, January 28). One-shot learning with Siamese network. *The Startup*. https://medium.com/swlh/one-shot-learning-with-siamese-network-2c8d24c8902f

Kotsiantis, S. B. (2007). Supervised machine learning: A review of classification techniques. *Informatica, 31*, 249–268.

Lee, S., & El Naqa, I. (2015). Machine learning methodology. In *Machine learning in radiation oncology: Theory and applications* (pp. 21–39). Springer.

Li, J. (2019). Regression and classification in supervised learning. In *ICCBD '19: Proceedings of the 2nd International Conference on Computing and Big Data* (pp. xx–xx). ACM.

Liu, Y., Wang, Y., & Zhang, J. (2012). New machine learning algorithm: Random forest. In B. Liu, M. Ma, & J. Chang (Eds.), *Information computing and applications. ICICA 2012. Lecture Notes in Computer Science* (Vol. 7473, pp. 287–295). Springer. https://doi.org/10.1007/978-3-642-34062-8_32

Maier, A., Köstler, H., Heisig, M., Krauss, P., & Yang, S. H. (2022). Known operator learning and hybrid machine learning in medical imaging—a review of the past, the present, and the future. *Progress in Biomedical Engineering*, *4*(2), 022002. https://doi.org/10.1088/2516-1091/ac5b13

Mari, S., Bucci, G., Ciancetta, F., Fiorucci, E., & Fioravanti, A. (2023). Machine learning for anomaly detection in induction motors. In *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)* (pp. 1057–1062). IEEE. https://doi.org/10.1109/MetroXRAINE58569.2023.10405548

Mian, S. M., Khan, M. S., Shawez, M., & Kaur, A. (2024). Artificial intelligence (AI), machine learning (ML) & deep learning (DL): A comprehensive overview on techniques, applications and research directions. In *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)* (pp. 1404–1409). IEEE. https://doi.org/10.1109/ICSCSS60660.2024.10625198

Ming, C., Wang, Z., & Wong, Y. S. (2014). Induction motor drives. *In [Editor(s) name if applicable]* (Ed.), *Book title* (pp. 1-22). Wiley. https://doi.org/10.1002/9781118354179.AUTO210

Mari, S., Bucci, G., Ciancetta, F., Fiorucci, E., & Fioravanti, A. (2023). Machine learning for anomaly detection in induction motors. In *2023 IEEE International Conference on*

Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE) (pp. 1057–1062). IEEE. https://doi.org/10.1109/MetroXRAINE58569.2023.10405548

Mohri, M., Rostamizadeh, A., & Talwalkar, A. (2016). *Foundations of machine learning* (pp. 1–8). MIT Press.

Na, S., Xumin, L., & Yong, G. (2010). Research on k-means clustering algorithm: An improved k-means clustering algorithm. *2010 Third International Symposium on Intelligent Information Technology and Security Informatics*, 63–67. https://doi.org/10.1109/IITSI.2010.74

Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised Machine Learning Algorithms: Comprehensive Review. *International Journal of Computing and Digital Systems*, *13*(1), 911–921. https://doi.org/10.12785/ijcds/130172

Preeti, P. (2017). A review on Machine Learning Techniques. *International Journal of Advanced Research in Computer Science*, *8*(3), 778–782. https://doi.org/10.26483/ijarcs.v8i3.3096

Rajapaksha, N., Jayasinghe, S., Enshaei, H., & Jayarathne, N. (2021). Supervised machine learning algorithm selection for condition monitoring of induction motors. In *2021 IEEE Southern Power Electronics Conference (SPEC)* (pp. 1–10). IEEE. https://doi.org/10.1109/SPEC52827.2021.9709436

Rani, P., & Vashishtha, J. (2017). An appraise of KNN to the perfection. *International Journal of Computer Applications, 170*(2), 13–17. https://doi.org/10.5120/ijca2017914696

Sabuncu, M. R. (2015). A sparse Bayesian learning algorithm for longitudinal image data. *Lecture Notes in Computer Science*, 411–418. https://doi.org/10.1007/978-3-319-24574-4_49

Salem, N., & Hussein, S. (2019). Data dimensional reduction and principal components analysis. *Procedia Computer Science, 163*, 292–299. https://doi.org/10.1016/j.procs.2019.12.111

Shah, A. (2016). Prediction of malignant and benign tumor using machine learning. *International Journal of Computer Applications, 135*(5), 1–6.

Singh, V., Gangsar, P., Atulkar, A., & Porwal, R. (2021). Deep learning based optimum fault diagnosis of electrical and mechanical faults in induction motor. *IOP Conference Series: Materials Science and Engineering, 1136*, 012059. https://doi.org/10.1088/1757-899X/1136/1/012059

Sugiyama, M. (2016). Introduction to statistical machine learning. Springer.

Sørensen, N. L., Bemman, B., Jensen, M. B., Moeslund, T. B., & Thomsen, J. L. (2023). Machine learning in general practice: scoping review of administrative task support and automation. *BMC Primary Care*, *24*(1). https://doi.org/10.1186/s12875-023-01969-y

Sonnenburg, S., Braun, M. L., Ong, C. S., Bengio, S., Bottou, L., Holmes, G., LeCun, Y., Müller, K., Pereira, F., Rasmussen, C. E., Rätsch, G., Schölkopf, B., Smola, A., Vincent, P., Weston, J., & Williamson, R. (2007). The need for open source software in machine learning. *Journal of Machine Learning Research*, *8*(81), 2443–2466. https://doi.org/10.5555/1314498.1314577

Sivaraju, S., Devarajan, N., & Ferreira, J. T. E. (2012). Performance analysis of induction motor for optimum efficiency and power factor using soft computing technique. *I J E E E S, 4*(2), 77.

Sivaraju, S., Devarajan, N., Ferreira, J. T. E., & S. S. (2012). Performance analysis of induction motor for optimum efficiency and power factor using soft computing technique.

Toscano, D. S., & Carrera, E. V. (2020). Failure detection in induction motors using non-supervised machine learning algorithms. *ResearchGate*. https://doi.org/10.13140/RG.2.2.29964.54401

Treetrong, J. (2012). Fault detection of electric motors application using ML estimation Method. *Advanced Materials Research*, *591–593*, 1958–1961. https://doi.org/10.4028/www.scientific.net/amr.591-593.1958

United Nations. (n.d.). *Sustainable Development Goals.* Retrieved from https://sdgs.un.org/goals

Uschjäger, S., Honysz, P.-J., & Morik, K. (2020). Generalized isolation forest: Some theory and more applications extended abstract. In *Proceedings of the International Conference on Data Science and Advanced Analytics* (pp. 1–8).

Vora, P., & Oza, B. A. (2013). A survey on K-means clustering and particle swarm optimization. *International Journal of Advanced Research in Computer Science and Software Engineering, 1*(3), 24–26. https://doi.org/10.1109/C0150020213/2013

Zwanenburg, E., Nguyen, V.-D., Limmer, S., Luijben, W. J. M., Bäck, T. H. W., & Olhofer, M. (2021). A combination of Fourier transform and machine learning for fault detection and diagnosis of induction motors. *Honda Research Institute Europe*.