TRIBHUVAN UNIVERSITY
INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS

B-02-BAS-2018/2023

FAULT DIAGNOSIS OF A BRAOKEN ROTOR BAR IN AN INDUCTION MOTOR USING MOTOR CURRENT SIGNATURE ANALYSIS

By

Adarsa Acharya (075AER002)

Shashank Sharma Nepal (075AER040)

Subodh Pokhrel (075AER044)

A PROJECT REPORT
SUBMITTED TO THE DEPARTMENT OF MECHANICAL AND AEROSPACE
ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF BACHELOR OF AEROSPACE ENGINEERING

DEPARTMENT OF MECHANICAL AND AEROSPACE ENGINEERING
LALITPUR, NEPAL

MARCH, 2023

COPYRIGHT

TRIBHUWAN UNIVERSITY
INSTITUTE OF ENGINEERING
PULCHOWK CAMPUS
DEPARTMENT OF MECHANICAL AND AEROSPCAE ENGINEERING

The undersigned certify that they have read, and recommended to the Institute of Engineering for acceptance, a project report entitled **"Fault Diagnosis of a Broken Rotor Bar in an Induction Motor using Motor Current Signature Analysis"** Submitted by **Adarsa Acharya, Shashank Sharma Nepal** and **Subodh Pokhrel** for partial fulfillment of the requirements for the degree of Bachelor of Aerospace Engineering.

Supervisor: **Sanjaya Neupane**
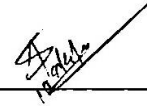Assistant Professor
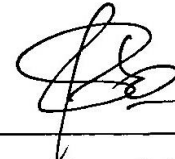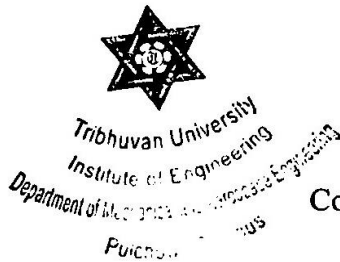Department of Mechanical and Aerospace Engineering
Pulchowk Campus, Institute of Engineering

External Examiner: **Anil Sapkota**
Junior Professor
Nepal Engineering College

Committee Chairperson: **Dr. Surya Prasad Adhikari**
Associate Professor
Department of Mechanical and Aerospace Engineering
Pulchowk Campus, Institute of Engineering

DATE: 16 March, 2023

ABSTRACT

The project report entitled "Fault Diagnosis of a Broken Rotor Bar in an Induction Motor using Motor Current Signature Analysis" reports about Motor Current Signature Analysis which is a non-invasive and cost-effective method for the fault diagnosis and classification of induction motors. This technique uses the current signals generated by the induction motor during its operation to diagnose faults such as rotor and stator winding faults, bearing faults, and misalignment faults. The process involves capturing the current waveform of the induction motor, processing it using signal processing algorithms, and analyzing it to identify the fault. The method is based on the principle that different faults generate distinctive current signatures. The current signals generated by the healthy and faulty induction motor, loaded with different loads, is captured using a current transformer, which is connected to a data acquisition system MyDAQ. The captured data is then processed using signal processing algorithms which is then processed to extract relevant features that are used to train machine learning models. The result indicates that Naïve Bayes algorithm was able to classify health condition of an induction motor with the accuracy of 94.4%. The algorithms like SVM and Decision tree also performed well with an accuracy of 88.9% and 91.7% respectively. The trained models can then be used to perform real-time fault classification on the induction motor, providing valuable information for predictive maintenance and condition monitoring.

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

N	Actual speed of rotor in rpm
A	Ampere
Ω	Angular velocity of rotor
$\in$	Belongs to
I	Current
ρ	Density of material
h	Depth of hole
F	Frequency
γ	Gamma (kernel parameter)
$\geq$	Greater than or equal to
Hz	Hertz
$\leq$	Less than or equal to
f_o	Line frequency
P	Number of Poles
π	Pi
r	Radius of drill bit
M	Reduced mass
f_b	Sideband frequency
s	Slip

Σ	Summation
N_s	Synchronous speed
Θ	Theta (angle)
T	Torque
F'	Unbalanced force
V	Volts

LIST OF ACRONYMS AND ABBREVIATIONS

AC   Alternating current

AI   Artificial Intelligence

ANN   Artificial Neural Network

BRB   Broken Rotor Bar

CNN   Convolutional Neural Network

DAQ   Data Acquisition

DBN   Deep Belief Neural Network

DL   Deep Learning

DTFT   Discrete Time Fourier Transform

EEMD   Ensemble Empirical Mode Decomposition

FEM   Finite Element Method

FFT   Fast Fourier Transform

FL   Federated Learning

FP   False Positive

FN   False Negative

K-NN   K-Nearest Neighbour

LabVIEW   Laboratory Virtual Instrument Engineering Workbench

LSB   Lower Side Band

MATLAB   Matrix Laboratory

MCSA	Motor Current Signature Analysis
ML	Machine Learning
MLP	Multilayer Perceptron Neural Network
OSELM	Online Sequential Extreme Learning Machine
RBF	Radial Basis Function
RNN	Recurrent Neural Network
RPM	Revolution Per Minute
RUL	Remaining Useful Life
PSA	Power Spectra Analysis
PSD	Power Spectral Density
SNR	Signal to Noise Ratio
STFT	Short Time Fourier Transform
SVM	Support Vector Machine
WT	Wavelet Transform

CHAPTER ONE: INTRODUCTION

1.1   Background

1.1.1 The Origin of Ideas

Data analytic and prediction techniques has always been our subject of interest. The genesis of an idea began from the engine health monitoring techniques that came across during the classes of fault monitoring and diagnosis. Initially, we thought of collecting aircrafts engine data from various airlines companies and forecast the Remaining Useful Life (RUL) of an aircraft through engine data. But due to private policies of an airline, we were not able to do so. Then rapid exploration was done for fault diagnosis of an Induction Motor (IM). This seemed to be the best idea upon consultation with our supervisor. As a result, the urge and desire to learn about predictive maintenance and fault diagnosis led to the idea of current signature analysis with the help of machine learning techniques enabling user to diagnose and find out faulty condition of broken rotor bar induced in an induction motor. These motors are also used in aviation industries and it was feasible for the project as it was technically and economically viable.

1.1.2 Attempts Made so Far by Others to Address the Issues

There are numerous attempts and has been successful in fault diagnosis of an IM with the help of motor current signature analysis. Various types of faults are present in induction motor. Various researchers and scholars have attempted to predict motor faults using different prediction, classification and data training techniques. Some of the crucial mechanical faults in IM are bearing fault, Broken Rotor Bar (BRB), eccentricity etc. Among these faults, bearing faults is the most common fault (Miljković & Dubravko, 2015). About (40-50%) of failure are due to bearing faults. As it is the common mode of failure, researchers have carried out extensive work for diagnosis of fault in an induction motor for bearing failure. For bearing failure, researchers had conducted research in abundant primitive techniques like thermal monitoring, torque monitoring, noise monitoring, and vibration monitoring where there are lots of use of sensors. On comparison to these techniques, electrical monitoring or current analysis requires minimum sensors and

is widely accepted all over the world. Researchers used motor current signature analysis techniques along with ANN algorithms for machine learning in order to diagnose bearing faults (Dhomad & A, 2020). Among various faults listed above, research and experimentation regarding fault diagnosis of a broken rotor bar of an induction motor was relatively less compared to other faults. Experimentation of researchers exhibited that MCSA was used to determine the frequency components connected to particular faults, specifically it was aimed to recognize the frequency components around the fundamental component (such as 50 or 60 Hz) that were connected to the BRB fault (Rodriguez & , 2020). Detailed steps of data acquisition, feature classification, feature extraction and interpretation with ML models were done in numerous research. Okwuosa presented MCSA at low load condition but taking three different faulty conditions, i.e. rotor misalignment, broken rotor bar & inter-turn short circuit winding. Researchers (Ejike Akpudo, et al., 2022) classified faults regarding broken rotor bar on the basis of loading conditions and was evaluated with the help of ML algorithms. A unified wavelet packet signature analysis was also used that identifies the signature of fault in the specific fault-oriented frequency bands by integrating the capabilities of both time-domain and frequency domain computational methodologies. Mehrjou identified that the fault severity during motor runtime is provided by the combination of wavelet analysis and a feed-forward neural network classifier. (Mehrjou, et al., 2017) Numerous studies has been carried out for feature extraction and classification. Various ML diagnostic techniques are implemented after fault classification. To decrease the dimension of characteristics feature and in order to take out the best features for the classification task, (Achmad & Widodo, 2009) used independent component analysis, principal component analysis, and their kernel to detect and diagnose faulty condition of broken rotor bar, phase unbalance. SVM, ANN and FL algorithms were used for vibration analysis. It was observed that SVM demonstrated 99.83% among other ML models. SVM is independent to many other parameters which influence the percentage of accurate detection. So, Silva cited that this was the reason for higher accuracy of SVM algorithms compared to others (Silva, et al., 2013). Various papers demonstrated detailed review of use of ML algorithms for fault detection. Researchers like Kumar and Keeman researched and explained the pros and cons of ML models like ANN, KNN, SVM, Decision tree, Deep learning etc. which enabled the

user to make correct choices according to the nature of problem and extent of accuracy provided by a particular model (Kumar & Shankar Hati, 2021; K, et al., 2018; Kecman & Vojislav, 2001).

1.1.3 Scope of Work

The main scope is to detect broken rotor bar faults in an induction motor using motor current signature analysis and machine learning techniques. Data extraction, characterization and fault classification of current signatures of an induction motor through Fast Fourier transform, wavelet transform technique is done. Then, with the help of ML algorithms, faults are diagnosed. Determining and creating different condition faults within the broken rotor bar failure condition is the work that is to be carried out which can possibly be beneficial during training as well as testing of data thus assisting in achieving higher accuracy.

1.1.4 Rational

Rapid increment in industrial and aviation sector is observed these days and will increase exponentially in coming future. In capitalist mindset world, saving billions of dollars will always be the priority of top industries and companies. Scientific maintenance technique like predictive maintenance has reached into multidimensional field. Researchers (Selcuk, 2017) explained advantages of Predictive maintenance techniques as compared to other techniques like corrective and preventive maintenance

Fault prediction mechanism of machinery components has become vital in today's context and deploying as well integrating these predictive machine learning models to real life systems will be advantageous as it saves both time and money (Choudhary, et al., 2019).The reasons or usefulness of fault diagnosis of an induction motor using ML techniques are as follows:

- It helps in early fault detection before catastrophic failure arises thus assisting user in minimizing downtime, repair cost and enhancing reliability of the system.

- It enables user to perform condition-based monitoring i.e. maintenance is carried out only when needed by observing the condition of the motor. Thus, time and money is saved.
- It improves safety by identifying potential faults that may occur before they cause safety hazards. It prevents accidents and boosts safety in the industry.

1.2   Problem Statement

Sudden failure of machinery equipment has been considered as serious problem in the world today. Among various kinds of faults that might occur in an IM, broken rotor bar fault is caused by high temperature, dynamic forces & severe current present in the rotor (R. A, et al., 1988). Garcia and his teams explained that these faults result in inefficiencies of working of the induction motor causing financial loss to the industries (Garcia, 2018). Their experiment suggested that motor becomes highly inefficient, i.e., decrement of efficiency to 20% as compared to healthy condition if severe faults are present. Also, if the problem is not diagnosed at low stage of severity, it can result to halt in processes and other machines attached at the same production line may be at higher risk as there is alteration in consumed current and production of new frequency components (Benbouzid, 2000). So, it is pivotal for any industry to detect broken rotor bar faults in the earliest stage possible. In the world of machine learning and artificial intelligence, trivialities as such should always be prevented in order to upscale the economics and business. Here, in the project, fault is artificially created in a rotor bar by drilling a hole. Current signals are acquired through Data acquisition card and fault is detected using machine learning algorithm.

1.3   Objective

1.3.1 Primary Objective

The main objective of this project is to diagnose and classify fault in an induction motor using motor current signature analysis along with supervised machine learning techniques.

1.3.2 Specific Objective

- To create experimental workbench for data acquisition from induction motor

4

- To generate datasets containing current signal from the healthy and faulty induction motor
- To test and train the Machine learning model for fault classification & diagnosis and select the best possible algorithm having highest accuracy

CHAPTER TWO: LITERATURE REVIEW

Motor Current signature analysis is a non-interfering technique done to recognize and isolate the mechanical and electrical faults and to indicate the failure of an induction motor. Paper of Thorsen and group suggested that unbalance supply of voltage and current, single phasing, earth fault, overload, inter-turn short circuit fault and crawling are electrical related faults whereas broken rotor bar, air gap eccentricity, bearing damage, rotor winding failure, stator winding failure are mechanical related faults (Motor reliability working group, 1985; Thorsen & M, 1995). The most commonly used 3 phase induction motor is squirrel cage type induction motor. The 3-phase squirrel cage motor is widely used motor in industries as it possesses features like versatility, reliability, low cost and high performance. Experiments performed by Babaa were able to validate that the inter-turn short circuits may alter or change particular harmonics in stator current which were thought to be as a default signature (Babaa, et al., 2021). Also, for the faulty condition of broken rotor bar in an induction motor, Didier explained that the presence of side bands around supply frequency was the decisive factor in detection of faults or cracks (Didier, et al., 2006). Various research conducted by Singh explained how IM are susceptible to different kinds of electrical and mechanical faults that can lead to unexpected failure and unscheduled down time (Singh, et al., 2016).

 MCSA is especially used for the early recognition and localization of electricals well as mechanical abnormal conditions present in an induction motor that may translate to failure of induction motors. It is explained that MCSA utilizes spectral analysis results of stator current for detecting faults like broken rotor bars, air gap eccentricity and bearing damage. Spectral analysis performed by Kliman explained that the spectral components of current will coincide with those caused by fault condition when rotor position varies with load torque (Benbouzid & Kliman, 2003). Liang explained that the presence of sideband in the fundamental motor current frequency is due to the broken rotor bars which is caused by distorted symmetry in the rotor circuit (Edomwandekhoe & Liang, 2018). The magnetic field anomaly is induced by broken rotor bars which can be identified by observing the current spectral components (Benbouzid, 2000).

The signal processing methods for detection of faulty condition fault in an induction motors is especially done in frequency domain using frequency domain features for broken rotor bars and bearing deformation. The major advantages and limitations of various techniques are further discussed which explains about time-frequency and time-scale transformations such as wavelets being optimum tool for detection of faulty induction motor rotors. Spectral analysis of operational process parameters such as current, temperature, pressure, etc. are performed according to the type of problem. Benbouzid suggested that time domain analysis and cestrum analysis are used as evaluation tools (Benbouzid, 2000). Fast Fourier Transform (FFT), wavelet transform, Hilbert transform, neural network and fuzzy inference-based techniques are used for signal processing for fault detection in induction motor. FFT is implemented traditionally for fault detections in an IM but processing of non-stationary signals cannot be done.SO, Short Time Fourier Transform (STFT) is deployed for the motor transient state with the time-frequency representation (Naha, et al., 2016). The pros of implementing wavelet techniques for fault monitoring of induction motors is growing as this method assist user to carry out stator current signal analysis during transients. This technique may be used for a localized analysis in the time-frequency or time-scale domain. Now, it can be regarded as important tool for performing condition monitoring and fault diagnosis (Benbouzid, 2000). STFT also contains its restrictions with characteristics of applying a fixed window to all signals, thus introducing the escape clause for data loss because of approximation and low resolution. Thompson Multitaper PSD estimation and Welch PSD estimation is proposed by (Edomwandekhoe & Liang, 2018) to eliminate the limitation of STFT Wavelet transform requires estimation of slip in order to localize the fault concerned frequency for diagnosis. The EEMD method is proposed to overcome the limitation. Wavelet Transform technique has shown enormous practicality in fault diagnosis of any machineries because of the benefits of multi resolution analysis. WT have been used in both discrete and continuous form for detection of fault on rotary machines where current can be analyzed during no load condition and condition during variation of load. (Wang, et al., 2011).

For other mechanical faults like air-gap flux variation Finite-element approach is used for analysis which is caused due to eccentricity in an induction motor. Using FEM based simulator, a 2D model is proposed to convert vibration into air gap change. (Hwang, et al.,

2005) explained that in case of static eccentricity, the rotor breaks down when vibration occurs without change in center of rotor and stator.

Fault diagnosis method involves process of data acquisition, feature extraction, feature selection and fault classification (Asghar, et al., 2016).FFT, STFT may not be desirable in all cases for processing of extracted data. So, process like wavelet packet analysis may be required. Wavelet packet analysis is a type of multi-resolution analysis. A bunch of orthogonal wavelet functions composed of subspace, creating a signal in various scale expansions are projected onto the signal. As a result, it keeps the signal characteristics at same scale in the time domain while extracting the signal characteristics in multiple frequency bands. These signals from a wavelet packet decomposition may be orthogonal, non-redundant, leak-free, and warrants that the high and low bands have the same high frequency resolution (Chao & , 2013).An induction motor defect detection and classification using machine learning model solely relies on the three phase voltages and current as input signals. As part of the experimental setting, (Kavana & Neethi, 2018) took 800 sets of the three phase voltages and currents of the defective motor were collected and sent as training data to a linear regression-based algorithm. Difficulties in MCSA was found in order to identify bearing defects due to uncertain features on the Fourier spectrum, error in reading the data manually without automation. So, an automatic system using machine learning based on motor current signature was proposed for induction motor fault diagnosis. A prediction model using support vector machine was developed which was able to identify defects up to 90% accuracy. SVM is a supervised learning model that analyses data for classification and regression analysis. It is concluded that the features calculated from wavelets coefficients from MCSA can be used as training data for Support Vector Machine (SVM) which can identify four conditions of an induction motor such as broken rotor bar, misalignment, bearing defect and healthy motor (Guo & Liu, 2018). Convolution Neural Network (CNN) is ML algorithm in detection of broken rotor bars considering different levels of severity. It focuses mainly on CNN approach of automatic image classification based on pattern recognition. Firstly, signal processing is done by STFT as a part of MCSA and the images obtained after this process were used as a training data for pattern recognition using CNN (Valtierra-Rodriguez, 2020). The major components of fault diagnosis are, (a) fault location identification, (b) faulty part

determination, (c) learning causes of failure (d) predicting fault patterns. Machine learning algorithms being powerful tools for classification of problems, are widely used for conditioning monitoring and fault diagnosis. Different algorithms like K- Nearest Neighbor (k-NN), SVM, ANN, Bayesian Classifier, Decision trees, random forest and deep learning are widely used. It has focused on application of deep learning as it has a major advantage on extraction of feature and selection that ultimately reduce the likelihood of occurrence of human error (Kumar, 2021).

Machine learning based approach for predicting failure of equipment in continuous production lines is initially assisted by MCSA. Suppose, one has to find faults in a gear box located in a place which is difficult to reach or inspect submersible pump situated underground that is run by an induction motor. The current drawn by an induction motor that drives the gear box or the submersible pump can be measured and this signature of current can be analyzed and fault can be detected using these machine learning models. Once these machine learning models are well trained and become robust, this ML models can be integrated with sensors that collect the data from the system. Real time monitoring is possible which results in early detection of fault thus ultimately reducing downtime. Integration of ML models with existing monitoring or control system is also possible which is beneficial for the company as it saves both time and money (Muller, et al., 2020). Another usefulness of Machine learning technique is remaining useful life (RUL) estimation which is done by the industries. Year wise analysis was done based on available data set in order to give detailed perspective of research trends in the field of RUL estimation. They also focused on existing maintenance method, predictive maintenance model, and RUL model. Concern for future and its challenges were also demonstrated. Countries like US, China and UK had tremendous research and investigation in field of RUL estimation with the help of machine learning as shown by the data base. (Sayyad, et al., 2021) also concluded that hybrid predictive maintenance could have tremendous possibility in the future Advanced and highly accurate Deep Learning is also used for RUL prediction in smart factories. CNN was used in Aero Propulsion System Dataset to estimate the RUL of engines. So, this gives us an idea and assurance of CNN algorithm and proper use and training of data can lead to result of high degree of accuracy (Jiang & Kuo, 2017). Four major deep learning techniques are extensively studies for the upcoming days. They

are: Auto-encoder, DBN, CNN and RNN. Deep learning assisted decision-making leads to new dimension into their operation. It also enables them to view real time performances using DL. A large data base is always needed and should be compared to observed data based on similarity approach in order to estimate RUL with high accuracy. Various researcher gave an insight about hybrid-based approach which could be the future of RUL estimation (Wang, et al., 2020).Brief explanation of RUL estimation and Deep learning is mentioned above as it holds deeper value in the future generation but is beyond the scope of this paper.

2.1 Induction Motor

One of most prominent type of motor used in industry is an induction motor. Pumps, fans, adjustable speed drives, with induction motors operate to translate about 60% of industrial electric energy into mechanical energy. (Kaur & Brar, 2016). Asynchronous or induction motors are AC electric motors that function in accordance with the faraday law of electromagnetic induction. Through electromagnetic induction from the revolving magnetic field of the stator winding, the electric current in the rotor required to generate torque is obtained. An induction motor's rotor can be categorized as either a wound type rotor or a squirrel cage rotor type. Single phase induction motors and three phase induction motors are the types of induction motor based on the input supply. Single phase induction motors are not self-starting, however three phase induction motors are. 3 phase induction motor is widely used in industries than single phase induction motor because of its simple and rugged construction. In terms of machine reliability, durability, efficiency, robustness, stable output voltage and torque, power factor and ripples, the three-phase motor outshines single phase (Kumar, 2018) .

An induction motor's operation is based on the electromagnetic induction phenomena. A rotating magnetic field is produced when an alternating current is supplied to the stator winding of the motor which induces an electromotive force in the rotor winding. It makes the current in the rotor flow. The rotor rotates as a result of the torque created by the magnetic field's interaction with the current in the rotor.

Synchronous Speed:

The synchronous speed of an induction motor is the speed at which the magnetic field produced by the stator rotates. It is given by the following equation:

$$N_s = (120*f) / P \tag{2.1}$$

Where N_s is the synchronous speed in revolutions per minute (RPM), f is the frequency of the power supply in hertz (Hz), and P is the number of poles in the motor.

Motor Slip:

The difference between the synchronous speed and the actual speed of the rotor is called slip. Slip (s) is given by the following equation:

$$s = (N_s - N) / N_s \tag{2.2}$$

Where N is the actual speed of the rotor in RPM (Alkadhim, 2020).

Torque:

The torque produced by an induction motor is proportional to the square of the current flowing through the rotor. The torque (T) produced by an induction motor is given by the following equation:

$$T = (3 * V^2 * R_2 * s) / (\omega_r * (R_1{}^2 + (R_2/s)^2)) \tag{2.3}$$

Where, T is the torque generated by the motor, V is the supply voltage to the motor, R_1 is the stator resistance of the motor, R_2 is the rotor resistance of the motor, s is the slip of the motor, ω_r is the synchronous speed of the motor (Ebadi, 2011).

Power:

The power output of a three-phase induction motor is given by:

$$P = (1.732 * V * I * \cos (\theta)) / 1000 \tag{2.4}$$

Where V is the voltage supplied to the motor in volts (V), I is the current drawn by the motor in amperes (A), $\cos (\theta)$ is the power factor of the motor (Soe, 2008).

2.2 Unbalanced Force due to Broken Rotor Bar
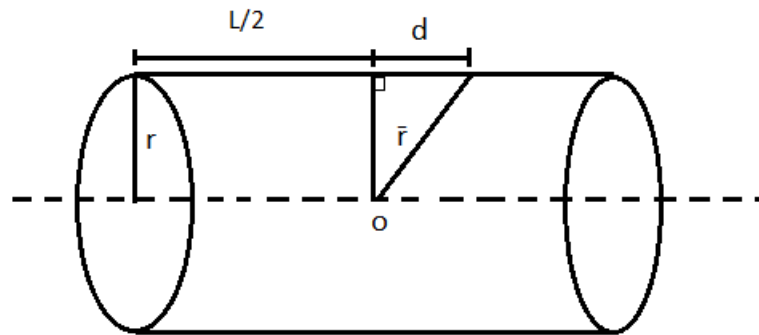
Figure 2.1 Cylindrical Rotor Bar of an IM

Let m be the reduced mass due to hole created in the cylindrical rotor bars.

$$m = V *\rho \qquad\qquad 2.5$$

And,

$$V = 1/2 * \pi * r^2 * h \qquad\qquad 2.6$$

Where, r is radius of drill bit, h is depth of hole, ρ is density of material

Now,

$$\text{Angular velocity (w)} = 2 \pi N / 60 \qquad\qquad 2.7$$

Where, N is the rpm

Finally,

$$\text{Unbalanced force (F')} = m\, w^2\, r' = m\, w^2\, (r/ \sin\theta) \qquad\qquad 2.8$$

2.3 Faults in Induction Motors

The common faults in induction motors can be classified into two groups,

2.3.1 Electrical Faults

The electrical faults are mainly classified into 4 types (Hammo & Rama, 2014). They are:

12

Overloading: It arises when mechanical torque crosses the threshold limit by applying mechanical load to the motor greater that rating at full load condition. This results in overheating, thus giving rise to the magnitude of phase current.

Short turn faults: A symmetrical three-phase AC machine with a short turn fault produces heat in the shorted turns by causing a large circulating current to flow. In accordance to a survey, stator winding insulation failures account for (35–40%) of IM failures.

Single phasing: Single phasing takes place when any one line among three lines are open. More current flows in remaining two line and excessive heat is generated in stator winding.

Phase reversal: Here, motors rotates in reverse direction as two phases are reversed from the normal sequence. For safety, in order to prevent phase reversal reverse phase relays and negative sequence relays are deployed.

2.3.2 Mechanical Faults

Mechanical faults are classified as;

Air gap eccentricity: It is a common rotor fault which produces problems such as vibration and noise in induction machines. The centre of rotation of rotor is perfectly aligned with the centre of stator bore in a healthy motor. When there is misalignment, it causes air gap eccentricity which can be the reason for failure of induction motor. Air gap eccentricity occur due to several causes like shaft deflection, bearing wear, wrong positioning of rotor with respect to stator, stator core movement etc.
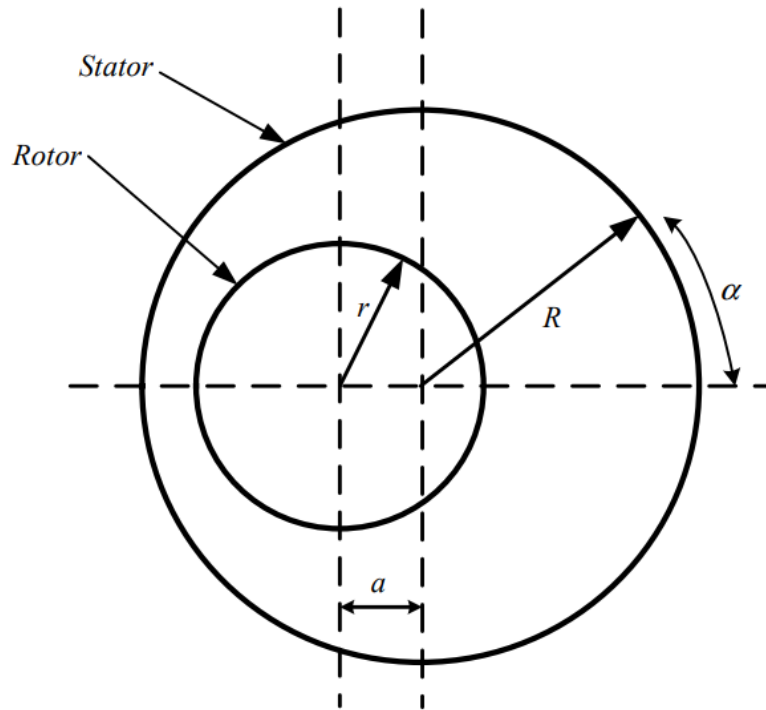
Figure 2.2 Schematic model of motor rotational under dynamic eccentricity (Zhang, et al., 2018)

Bearing faults: Bearings are common elements which are employed to permit rotary motion of the shafts. Continued stress on bearing can cause fatigue failures. Some sources of bearing faults are contamination, improper lubrication, improper installation, corrosion etc. (Riddle, 1955).
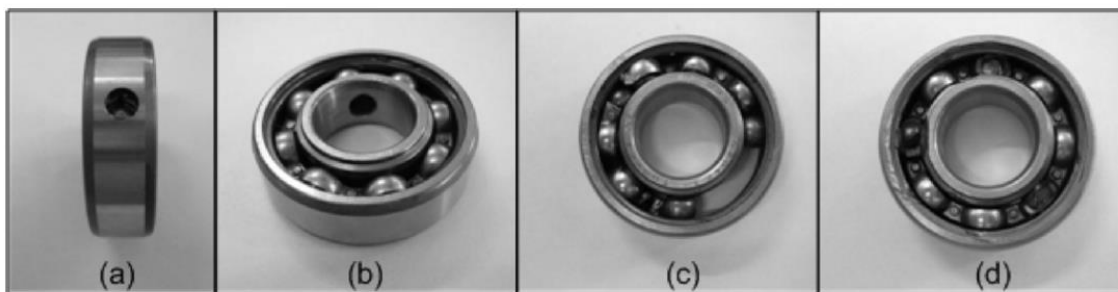
Figure 2.3 Bearing faults: a) outer race deterioration b) Inner race deterioration c) Cage deterioration d) Ball deterioration (Yilmaz & , 2008)

Load faults: Motors are sometimes connected to mechanical loads and gears which results to occurrence of faulty conditions. Misaligned couplings and defective gear systems that connect a load to the motor are a handful of examples (Dalpiaz & Meneghetti, 1991).

Broken rotor bar: An electrical fault such as broken rotor bar does not necessarily cause broken rotor bar to fail in the initial stage but can cause serious hazards over a period of time. It can cause mechanical damage when broken parts of rotor bar strikes to the rear winding of stator core with high velocity of high voltage motor (Bonnett & Soukup, 1986).
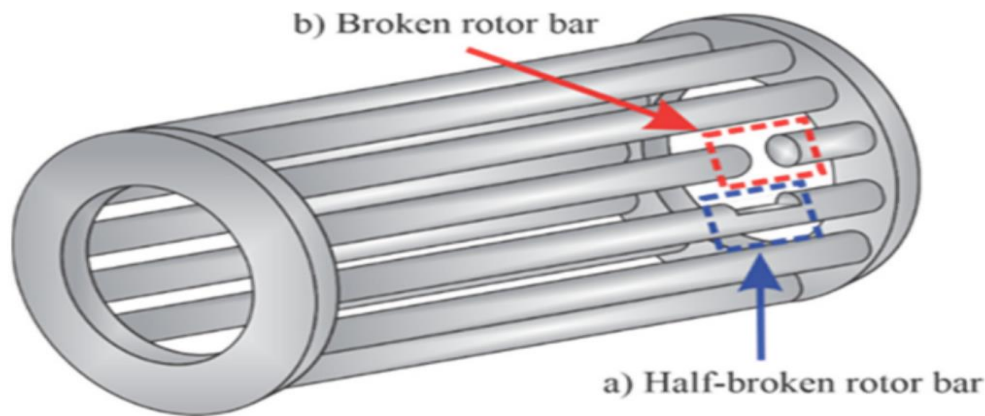
Figure 2.4 Broken rotor Bar: a) half- broken rotor bar   b) single broken rotor bar

2.4   Spectral Analysis:

When there is concentration of signal's energy for a given finite amount of time interval, power spectral density (power spectrum) can be computed. Power Spectral Analysis (PSA) is a technique used to analyze the power of frequency content of a signal taken for a finite time interval. The power spectrum diagram, also known as a periodograms, is a graphical representation of the PSD results and provides insight into the energy content in a frequency component of a signal. The idea behind PSA in case of this project is to find the power distribution in a specific range of frequency. Also, it gives variation in periodograms as the broken rotor bar faults is detected. PSD are based on square of absolute value of Fourier coefficients if FFT is deployed. But, (Romeral, et al., 2008) suggested that FFT may not be suitable for feature extraction process as transitory characteristics like trends, drift, abrupt changes which is most crucial part of signal are undetected. So, coefficient

15

from wavelet transform can also give PSD using wavelet transform. In this case, PSD helps in identifying faulty frequency bands and no single specific harmonic which was used previously. Without problem of slip variation, different bands can be analyzed using PSD along with wavelet transform. PSD along with wavelet transform is most widely used today because of its ability in fault diagnosis in any point during the operation. Another benefit is both constant and non-constant load torque condition results can be well analyzed with the help of PSD along with wavelet transform (Myung, et al., 2003).

The power spectrum diagram can be used for a variety of applications, including fault diagnosis, signal processing, and control systems design. For example, PSA can be used to analyze the current and voltage waveforms of an induction motor and detect any anomalies or faults in the system (Mitra, 2001). The power spectrum diagram can reveal the presence of harmonic components in the current and voltage waveforms, which can be used to diagnose faults such as broken rotor bars or stator windings. There are several methods for computing the power spectrum, including the fast Fourier transform (FFT) and the Welch method (Press, 2001).

2.4.1 Spectral Components for Broken Rotor Bar

The sideband components are examined around the provided current fundamental frequency i.e. line frequency, f_o, to find broken rotor bar faults.

$$f_b = (1 \pm 2s)\, f_o,$$

2.9

Where,

f_b are the sideband frequencies related with broken rotor bar,

s is per unit motor slip.

The slip s is defined as the relative mechanical speed of the motor, n_m with respect to the motor synchronous speed, n_s, as

$$s = (n_s - n_m)/n_s$$

2.10

The motor synchronous speed n_s is defined as:

16

$$n_s = (120f_o)/P \qquad\qquad 2.11$$

Where, P is the number of poles of the motor.

The lower sideband frequency component is specifically due to broken rotor bar

$$f_b = (1-2s)\, f_o \qquad\qquad 2.12$$

And, the upper sideband frequency component is due to consequent speed oscillation.

$$f_b = (1+2s)\, f_o \qquad\qquad 2.13$$

If there are numerous damaged rotor bar, the frequency components in the frequency spectrum of the stator current, in addition to the fundamental one, are provided by

$$f_b = (1+2ks)\, f_o \qquad\qquad 2.14$$

Where, $k = 1,2,3\ldots k_n$
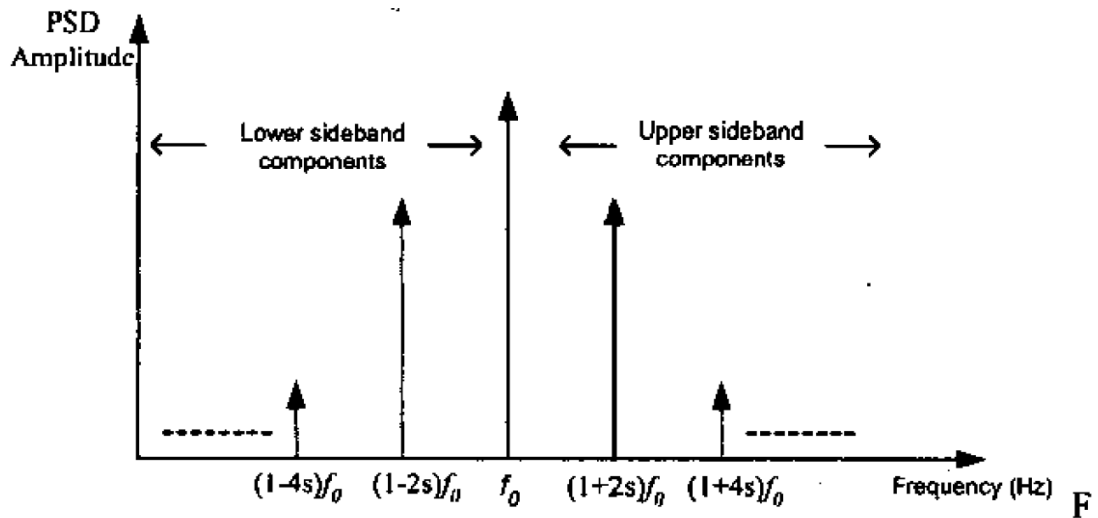
Figure 2.5 Ideal power spectrum

17

2.4.2 FFT

FFT is an algorithm that decomposes a signal into its constituent frequencies. It works by transforming the time domain representation of a signal into its frequency domain representation, which provides a representation of the signal in terms of its constituent frequencies. In the case of an induction motor, the FFT is applied to the current waveform to obtain the current spectrum. The algorithm works by dividing the input signal into smaller segments, and then performing a series of mathematical operations on each segment to obtain the frequency components of the signal. The results from each segment are then combined to obtain the final frequency domain representation of the signal. The FFT is a powerful tool for PSD estimation, but its results can be sensitive to the choice of window function and the length of the data segment used for the computation. In addition, the FFT provides a PSD estimate that can be biased by the presence of noise, leading to inaccurate results.

A discrete time series function's discrete Fourier Transform (DFT) can be calculated using the FFT algorithm with the least amount of computational work. Publication of Cooley and Tukey's paper received great response for FFT spectrum analysis. (James, et al., 1965)The DFT ofthe supplied time series is calculated by FFI's algorithms by breaking down the Npoint DFT computation into smaller computations. If these decompositions are done correctly, they significantly reduce the difficulties of computation. Consider a set of N evenly spaced samples of a finite discrete time series signal, x[n], specified for $0 < n < N - 1$ to revisit DTFT) of x[n], represents the sequence in terms of a complex exponential sequence $\{e^{-jwn}\}$, w is the real frequency the fundamentals of the DFT. The representation of the Discrete-Time Fourier Transform (variable and is depicted as $X(e^{jw})$. $X(e^{jw})$ is defined as: (Trussell, et al., 2003)

$$\sum_{n=0}^{N-1} x[n]\, e^{-jwn} \qquad\qquad 2.15$$

By sampling in uniform manner, the value of $X(e^{jw})$ on w- axis between $0 \le w \le 2\pi$ at $w_k = 2\pi k/N$, $0 \le k \le N-1$: we get relation between x[n] and it's DTFT,

At $w = 2\pi k/N$,

$$| X (e^{jw})| = \sum_{n=0}^{N-1} x[n] e^{-jwn} \qquad\qquad 2.16$$

Where, $0 \le k \le N-1$;

The discrete Fourier transform of x[n], X[k], will be finite sequence from

$$X[k] = |X (e^{jw}) | \text{ at } w = 2\pi k/N \qquad\qquad 2.17$$

Where $0 \le k \le N-1$

The price of computation gets significantly reduce by selecting composite number N in FFT algorithm. The PSD is also formulated from FFT function. FFT gives complex form two-sided spectrum. This needs to be converted to polar form for calculation of magnitude and phase. The relation for amplitude and phase vs frequency is given by FFT which is expressed as:

$$\text{Amplitude spectrum in quantity peak} = \frac{\text{Magnitude[FFT(B)]}}{N} \qquad\qquad 2.18$$

$$= \frac{\sqrt{\text{real[FFT(B)]}^2 + \text{imag[FFT(B)]}^2}}{N}$$

$$\text{Phase spectrum in radians} = \text{Phase [FFT (B)]} = \text{arctangent}\left[\frac{imag[FFT(B)]}{real[FFT(B)]}\right] \qquad\qquad 2.19$$

Where arctangent function gives value between $-\pi$ and $+\pi$. Power spectrum and amplitude spectrum are related to each other. When squaring single sided rms amplitude spectrum single sided PSD is obtained (G & Lyons, 1997). Amplitude spectrum in rms is calculated as:

$$\text{Amplitude spectrum in rms} = \sqrt{2} \cdot \frac{\text{Magnitude[FFT}(B)]}{N} \quad \text{for i} = 1 \text{ to N/2-1} \qquad\qquad 2.20$$

$$= \frac{\text{Magnitude[FFT}(B)]}{N} \qquad \text{for i} = 0 \qquad\qquad 2.21$$

Where i stands for frequency line number of Fast Fourier Transform of B.

2.4.3 Welch's Method

After data acquisition, FFT method illustrates irregularities in the frequency spectra during data processing. Thus, Welch method is used. Implementation of Welch method is a preprocessing step after which it is implemented in the ML model. The most important part of welch method is extract feature to compress data without deteriorating any of its information. It is also called low level feature extraction. In Welch's method, data is transformed to frequency domain from time domain. Though this seems similar to the task as performed by FFT but has significant difference. Welch's method assists user in subtracting associated noise components from the signal which is not possible via FFT method. Welch method acts as a combination of FFT & use of low pass filter as Welch method breaks signal into segments and averages the total segments thus acting like a low pass filter. Hence, power spectrum of these signals are estimated after preprocessing of data using Welch method is completed. It provides an estimate of the PSD that is more accurate and reliable than other techniques. This is because the Welch method uses overlapping data segments, which reduces the variance of the PSD estimate (Johannessen, et al., 2019). In addition, the Welch method provides a robust estimate of the PSD even in the presence of noise, making it an ideal choice for many applications, including induction motor analysis where signals may be noisy due to conditions like non-uniform loading condition or any other factors. (Vaidyanathan, 2006).

In the PSD estimation using Welch method, the data sequence x[n], x[0],x[1],…x[N-1] is first divided into k number of segments. The length of L samples &each segment can be overlapping on each other with (L-S) overlapping samples, where S is the number of points to shift between segments.

Segment 1: x[0],x[1]….x[L-1]

Segment 2: x[s],x[s+1],…..x[L+s-1]

Segment K: x[N-L],x[N-L+1],….x[N-1]

The weighted k^{th} segment will of the samples,

$$X^k[n]=w[n]x[n+ks], \qquad \text{for } 0<n<L-1, \, 0<k<K-1 \qquad\qquad 2.22$$

Where w[n] is the window function applied to the data at each segment before the computation of the segment periodograms. The sample spectrum of the weighted k^{th}segment is show in the frequency ranging as $-1/2T \leq f \leq 1/2T$ is given by

$$P^k \, (f) = \frac{1}{ULT} X^P \, (f) \, [X^P \, (f)] = \frac{1}{ULT} |X^P \, (f)|^2 \qquad\qquad 2.23$$

Where U is the discrete time window energy,

$$U = T \sum_{n=0}^{L-1} w^2[n] \qquad\qquad 2.24$$

And, $X^P \, (f)$ is the DTFT of the kth segment,

$$X^P \, (f) = T \sum_{n=0}^{L-1} x^{(k)} \, [n] e^{-j2\pi fnT} \qquad\qquad 2.25$$

Now, on averaging periodograms values of k segments, Welch PSD can be obtained. (Marple Jr, et al., 1989)

$$P_w \, (f) = \frac{1}{K} \sum_{k=0}^{K-1} P^k \, (f) \qquad\qquad 2.26$$

Thus, in this way, Welch PSD can be estimated.

2.5   Time Domain Features

Time domain features are features that describe the characteristics of signals in the time domain. They are often used in feature extraction in MATLAB to describe the signal's amplitude, frequency, and phase information. Some common time domain features include mean, standard deviation, root mean square, zero-crossing rate, energy, and power. These features can be extracted using built-in functions in MATLAB, such as mean(), std(), rms(), and others. The choice of features and the method of extraction depends on the type of signal being analyzed and the goals of the analysis.

Some of the time domain features that are used for feature extraction are;

2.5.1 Mean

The mean feature is a time domain feature that describes the average amplitude of a signal over a specified window of time. It is calculated by taking the sum of the signal's values

over the window and dividing it by the number of samples in the window. Mathematically, the mean is defined as

$$\text{mean} = 1/N * \sum(x) \qquad\qquad 2.27$$

,

where x is the signal, N is the number of samples in the window, and the sum is taken over the window.

2.5.2 Signal to Noise Ratio (SNR)

Signal-to-noise ratio (SNR) is a time domain feature that describes the ratio of the signal power to the noise power in a signal. It is calculated by dividing the power of the signal by the power of the noise. Mathematically, SNR is defined as:

$$\text{SNR} = 10 * \log_{10}(P \text{ signal} / P \text{ noise}) \qquad\qquad 2.28$$

,

Where P signal is the power of the signal and P noise is the power of the noise.

2.5.3 Root Mean Square (RMS)

Root mean square (RMS) is a time domain feature that describes the amplitude of a signal. It is calculated by taking the square root of the mean of the squares of the signal's values over a specified window of time. Mathematically, RMS is defined as:

$$\text{RMS} = \sqrt{\left(\frac{1}{N} * \sum(x^2)\right)}, \qquad\qquad 2.29$$

where x is the signal, N is the number of samples in the window, and the sum is taken over the window.

2.5.4 Peak Value

Peak value is a time domain feature that describes the maximum amplitude of a signal. It is a measure of the strength or intensity of the signal.

2.6 Frequency Domain Features

Frequency domain features are features that describe the frequency content of a signal. These characteristics are computed in the frequency domain, which is created by applying a Fourier transform to the signal to convert it from the time domain to the frequency domain.

2.6.1 Peak Frequency

Peak frequency is a frequency domain feature that describes the frequency with the highest amplitude in a signal. It is a measure of the dominant frequency content of the signal. To calculate peak frequency, the signal is typically transformed from the time domain to the frequency domain using a Fourier transform. The amplitude spectrum of the signal is then calculated and the frequency with the highest amplitude is identified as the peak frequency.

2.6.2 Peak Amplitude

Spectral peak amplitude is a frequency domain feature that describes the maximum amplitude of a specific frequency component in a signal. It is a measure of the strength or intensity of a particular frequency in the signal. To calculate spectral peak amplitude, the signal is typically transformed from the time domain to the frequency domain using a Fourier transform. The amplitude spectrum of the signal is then calculated and the maximum amplitude is identified as the spectral peak amplitude.

2.7 Machine Learning

The area of artificial intelligence known as machine learning is concerned with creating algorithms that can learn from data and predict the future. These algorithms build models from incoming data to make predictions or judgments, even though they are not particularly trained to do so. Machine learning algorithms are mathematical models and statistical methods used to create systems that can learn from data and make predictions or judgments. There are several types of machine learning algorithms including (Ayodele & Oladipupo, 2010);

23

- Supervised learning: The algorithm creates a function which maps inputs to get the required outputs. This algorithm learns from labeled training data to make predictions about unseen data.

- Unsupervised learning: The algorithms only models input data and learn from unlabeled data to identify patterns or groupings in the data.

- Reinforcement learning: The algorithms that learn a system in order to interact with an environment to maximize a reward signal.

- Semi-supervised learning: The algorithms that uses a mix element of supervised and unsupervised learning to work with partially labeled data generating suitable classifier.

- Deep learning: A subfield of machine learning which uses multi-layer neural networks to learn complex patterns in large amounts of data.

- Transduction: It does not readily construct any function as supervised learning but it forecast new outcomes which is based on training inputs, training output, and new input.

Among these ML algorithms, we used supervised learning as it fulfilled our requirement and was accurate for our purpose.

2.8   Fault Classification using Machine Learning

Machine learning techniques are used in fault classification to automatically classify faults or failures into several groups. A set of qualities or characteristics of the data gathered from sensors or other sources are used to determine the failure's primary cause. For instance, real-time temperature, vibration, and current data may be gathered by sensors in a manufacturing facility.

A machine learning system can forecast the possibility of a fault, such as a motor failure or bearing wear, by analyzing this data. The program can then categorize the defect as "mechanical failure," "electrical failure," or "thermal failure," among other categories. The

type of data that is available, the difficulty of the classification task, and the desired accuracy all influence the machine learning techniques that is chosen for fault classification. Several supervised learning methods, such decision trees, random forests, and support vector machines are used.

2.8.1 Supervised Machine Learning

Supervised learning is mostly used for classification type of problem. The problem is solved by the supervision of the user. These processes are dependent on knowledge given by pre-determined classification. The observations or output are found by the nature and magnitude of the input fed into the machine learning model. Here, is an example of how supervised learning is operated.
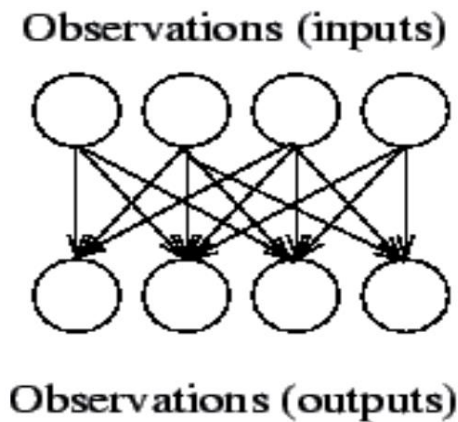
Figure 2.6 Examples of Supervised learning

Supervised learning algorithm mainly deals with classification. The types of these algorithms based on classification techniques are as follows:

- Linear classifier
    - ✓ Logical regression
    - ✓ Naïve Bayes Classifier
    - ✓ Perceptron
    - ✓ Support Vector Machine (SVM)
- Quadratic Classifier
- Decision Tree

- Neural networks
- Bayesian Networks

Here is the flowchart illustrating the process of supervised machine learning.
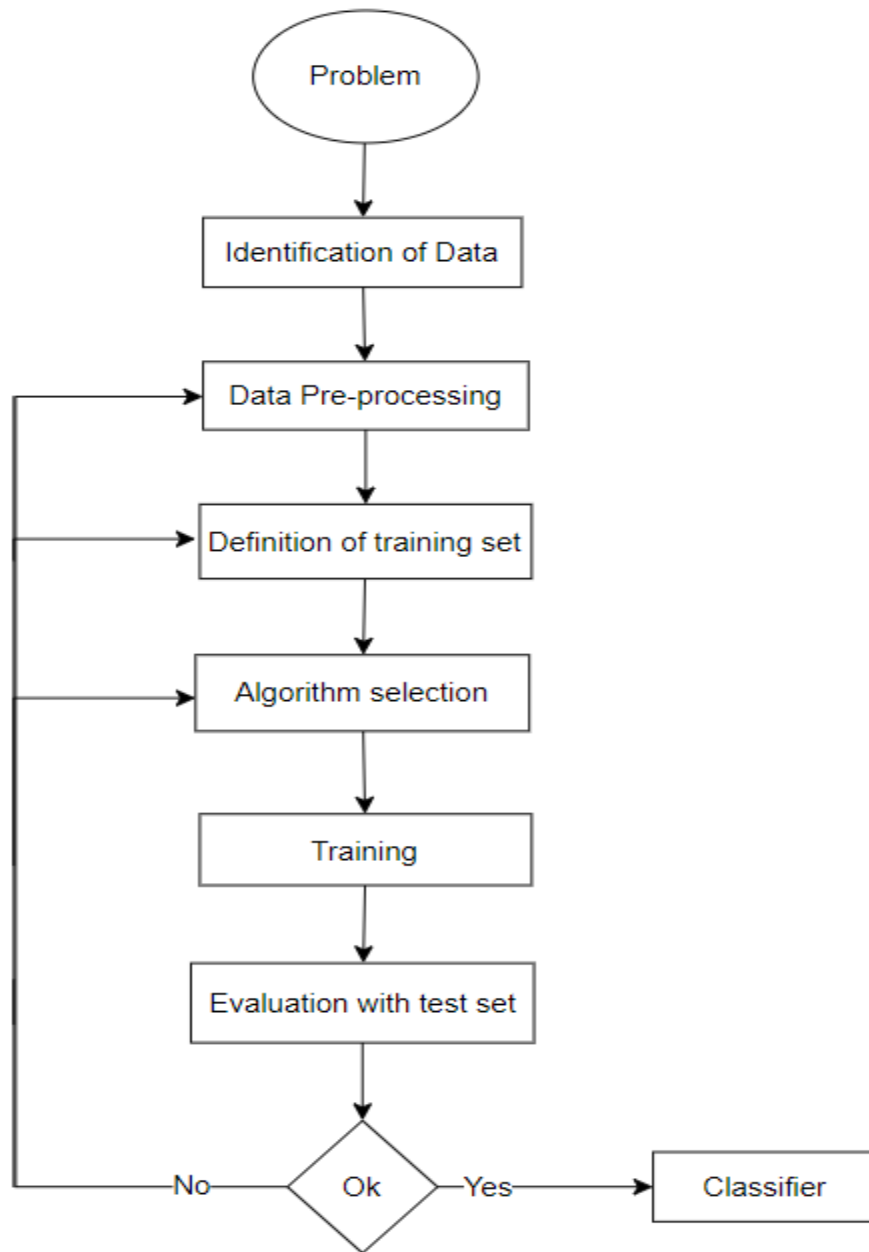
Figure 2.7 Flowchart depicting supervised ML

2.8.2   Support Vector Machine (SVM)

A supervised learning approach called Support Vector Machine (SVM) is employed for classification and regression analysis. Here, machine learning model learns from training data in order to find a particular class from variety of data. Support vector are pivotal in this model as they define the hyperplane. It is basically solution of problem optimization by selecting best hyperplane for a given datasets (Widodo & Achmad, 2007). It operates by locating a hyperplane that optimizes the distance between the nearest data points of two classes, or "support vectors." The support vectors are the data points that are closest to the hyperplane, and the margin is the separation between the two. Finding the hyperplane with the largest margin that best distinguishes the classes is the main goal of SVM. The side of the hyperplane that a new data point lies on determines the class in which it lies. SVM uses a technique called as the kernel trick to transform non-linearly separable data into a higher-dimensional space.

The steps illustrating how SVM works are:

Data Preparation: The first step in using the SVM algorithm is to prepare the data. In order to do this, the data must be cleaned, the features must be normalized, and the data must be divided into training and testing sets. The training set equation is given by:

$$(X_1, y_1)….(X_n, y_n), X_i \in R^n \tag{2.30}$$

Where, $y_i \in \{-1, 1\}$

Select a kernel function: SVM can transform data into a higher-dimensional space where it can be split by a hyperplane using either linear or non-linear kernels. Linear, polynomial, radial basis function (RBF), and sigmoid kernels are extensively applied (Lee , et al., 2005).

Some kernel function are explained as:

- Linear: $K(X_i, X_j) = (X_i)^T X_j$ 3.27
- Polynomial: $K(X_i, X_j) = (\gamma . (X_i)^T X_j + r)^d, \gamma > 0$ 3.28
- Sigmoid: $K(X_i, X_j) = \tanh(\gamma . (X_i)^T X_j + r)$ 2.31

27

Where r, d and γ are kernel parameters.

Model Training: The SVM algorithm discovers the hyperplane that optimizes the margin between the closest data points of two classes in order to train a model on the training data. By minimizing an objective function that penalizes points that are situated on the wrong side of the hyperplane, this is attained.

Hyperplane Selection: The SVM algorithm chooses the hyperplane with the largest margin once the model has been trained. The classes are effectively separated using this hyperplane. To distinguish multiple classes, "one vs one" and "one vs rest" various plans are implemented by researchers (Wang, et al., 2013).

Prediction: By determining which side of the hyperplane the data points fall, the SVM algorithm uses the trained model to forecast the class of fresh data points. One class is assigned to points on one side of the hyperplane, and another class is assigned to points on the opposite side.

2.8.3   Naïve Bayes Algorithm

Naïve Bayes algorithm based on Bayes theorem is a statistical classification algorithm that explains the probability of occurrence of an event based on the previous knowledge of conditions which may be in relation to the event. It is considered as "naïve" because its unique feature of assuming all the features in the given dataset are independent of one another. Its simplistic feature makes it highly effective in classification techniques.

Bayes theorem is based on Bayes rule which is explained mathematically as:

2.32
$$P (A/B) = P (B/A) * P (A) / P (B)$$

Where P(A/B) is the probability of event A occurring given that event B has occurred, P(B/A) is the probability of event B occurring given that event A has occurred, P(A) is the prior probability of event A and P(B) is the prior probability of event B.

It works on two phases: Training phase and prediction phase

During training, it identifies the likelihood of every feature occurring in class label. Here, prior probabilities at each class label is calculated which is also called probability of occurrence of each label in training dataset. After completion of training phase, naïve Bayes algorithm uses probabilities calculated during training period in order to classify new data. Finally, the evaluation of performance is done (Berrar, 2018).

2.8.4 Decision Tree Algorithm

It is a machine learning algorithm which builds a tree model of decision and forecast consequences. It is also supervised learning technique which fundamentally uses classification and regression analysis. It recursively splits the dataset into its subsets which looks branch like in order to make further decisions. There are three main components of decision tree algorithm. They are root node, decision node and leaf node.

Root node: It is the topmost node of tree representing entire dataset.
Decision node: It is used to split the data based on the values of input features. Each node has specific input features and this feature determines the specific input feature.
Leaf node: It represents the output of model. Each leaf node belongs to specific value of target variable.
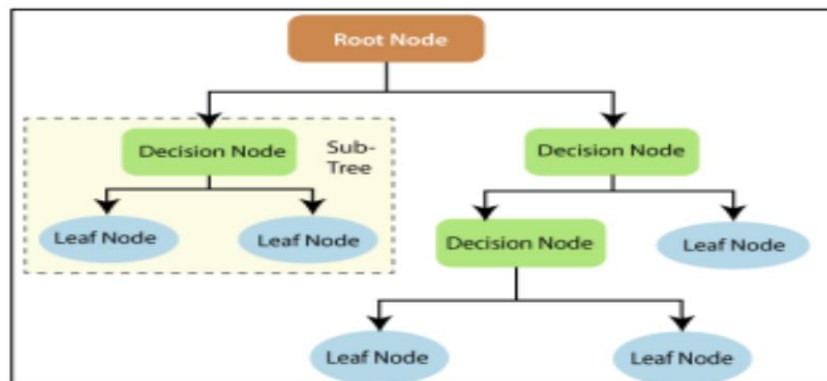
Figure 2.8 Decision tree algorithm (Janikow, 1998)

Firstly, it identifies input and output variables of entire dataset. Then it selects the best feature which is used to split data. The best feature is defined as those features having maximum information gain. It is a measure of decrement in entropy or uncertainty of data after split. Splitting of dataset into its subset is done based on chosen feature. Now, class

or value is assigned to each and every leaf node which based on majority or average value of target variable in is given subset. Now, new data points can be used to predict class (Jijo, 2021).

2.8.5 Neural Network

It is a machine learning model derived from the concept of human brain. It has a series of interconnected nodes which is called neurons, therefore performing complex tasks such as classification and regression. The components of neural network is explained below:

Input layer: It receives input data. Every input feature is indicated by a neuron in input layer.

Hidden layer: This is the middle layer between input and output layer. It contains multiple or single neuron for calculation.

Output layer: This produces output of the model. The nature of task indicates the number of neurons present in the output layer.

Figure 2.9 Neural network consisting of input, hidden and output layer

Weights and Biases: Every neuron in every layer is assigned respective weight and biases. Weight denotes the strength of connectivity between neurons and biases determine the output of the neuron for a given input.
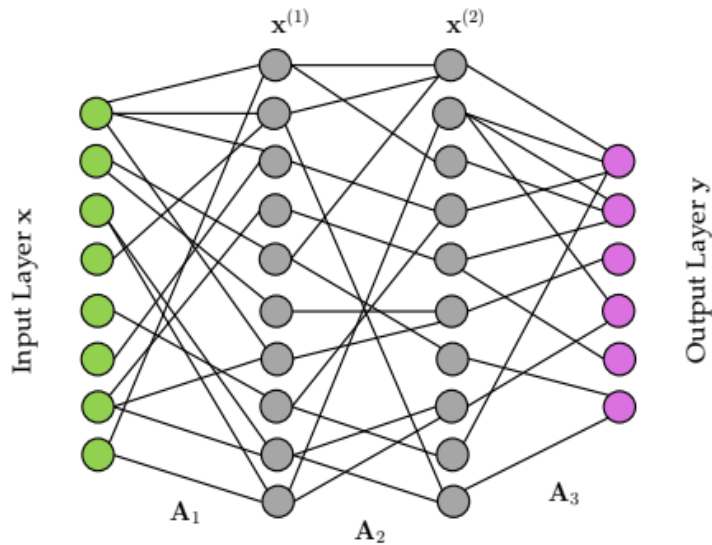
Initially, weights and biases are assigned randomly. The input data is fed into input layer. By the application of nonlinear activation function to the weighted sum of input, the input is transferred to hidden layers. Finally, output of the model is computed by propagating output of hidden layer into output layer. The resulted output is compared with true output and error is calculated. With the help of backpropagation of error, the error is minimized by changing the values of weights and biases. In this way, neural network algorithm works.

2.8.6 Confusion Matrix:

Confusion matrix also known as error matrix, is used for description of a classifier or classification model. It tells the user how well the classifier is performing on a given machine learning model. It depicts which is the confused classifier by illustrating confusion matrix. Here, summary of correct and incorrect prediction is given. It evaluates a set of data points' actual class labels with the predicted class labels the model produced.

The confusion matrix offers a number of metrics to evaluate the model's correctness, such as:

True Positives (TP): The number of times the model correctly predicted the positive class and that class really contained positive individuals.

False Positive (FP):The number of times the model predicted the positive class but the actual class was negative.

True Negatives (TN): The number of times the model correctly predicted the negative class when the actual class also fit the description.

False Negatives (FN): The number of instances where the model predicted the negative class but the actual class was positive.

The values in the confusion matrix can be used to construct a number of performance metrics, including accuracy, precision, and recall. These metrics offer a thorough analysis of the model's performance, enabling the observer to identify the model's benefits and drawbacks and decide whether to make any more adjustments or enhancements.

A confusion matrix can be used to calculate the accuracy of a machine learning model by comparing the actual class labels of a set of data points to the predicted class labels generated by the model. The accuracy of a model can be calculated as the proportion of correct predictions made by the model.

This can be expressed as:

Accuracy = (TP + TN) / Total Instances

Where:

True Positives (TP) are the number of instances where the model correctly predicted the positive class.

True Negatives (TN) are the number of instances where the model correctly predicted the negative class.

Total Instances is the number of instances in the dataset.

By using the values in the confusion matrix, the accuracy of the model can be calculated and used as a single measure of the model's performance.

Precision measures the model's accuracy in predicting instances that belong to the positive class. It is determined as the percentage of occasions out of all instances that were expected to be positive that really were positive.

Precision is calculated as follows:

Precision = True Positives / (True Positives + False Positives)

Where,

True Positives (TP) represent the number of times the model correctly predicted the positive class &

 False Positives (FP) represent the number of times the model correctly predicted the positive class but the actual class was negative.

CHAPTER THREE: MATERIALS AND METHODOLOGY

3.1   Materials Requirement

The materials required to setup the experimental workbench is described below:

3.1.1   Intangible Materials

- **MATLAB**

  A high-performance language for technical computing is called MATLAB. In a simple-to-use interface, it mixes computation, visualization, and programming while expressing issues and solutions using well-known mathematical language. Math and computing, algorithm creation, simulation, exploration, and visualization of scientific and engineering visuals, application development, including graphical interface construction, are some of its usual uses.

  MATLAB is a powerful tool for machine learning that provides a rich set of functions and tools for building, training, and evaluating machine learning models, as well as integrating with other tools and technologies. It is particularly well suited for tasks that involve data preprocessing, model selection, and performance evaluation (Moore, 2005).

- **LabVIEW**

  LabVIEW is a user-friendly visual programming language. It is a framework platform and building environment that was created to make it possible to create any type of system. It was created as a workbench for managing test instruments by National Instruments. It provides a comprehensive suite of tools for data acquisition, signal processing, and control, making it a powerful tool for applications in industry, academia, and research in graphical programming environment. Data acquisition device such as a data acquisition board or a data acquisition module is used to receive current signals from an IM in LabVIEW. It also provides a variety of functions and

tools for visualization, allowing you to view and analyze the signals in real-time (Higa, 2002).

3.1.2 Tangible Materials

- **MyDAQ**

Figure 3.1 MyDAQ

MyDAQ is a data acquisition device developed by National Instruments. Students may measure, produce, and manipulate signals in real-time with this small, portable equipment for electrical engineering and science education. It doesn't require extra power because it connects to a computer through USB and is powered by the computer. Students can carry out a variety of experiments and projects in fields including circuit design, control systems, and data acquisition due to the oscilloscope, multimeter, function generator, and digital I/O lines that are included into the device. The tool is compatible with LabVIEW, a graphical programming environment from National Instruments that makes it simple for students to create and carry out experiments. MyDAQ can be used for uses beyond education, including professional settings for data acquisition, instrument control, and automation tasks. MyDAQ is an economical, practical option for engineers and scientists that need to gather data or conduct tests outdoors thanks to its compact size and USB connectivity (Chesnutt, 2011).

- **Current Transformer**

Figure 3.2 Current transformer

A current transformer (CT) is a type of transformer used to increase or decrease an alternating current (AC). It generates a secondary current proportional to the primary current. The current ratio from primary to secondary is how CTs are identified. Typically, the rated secondary current is standardized at 1 or 5 amps. For instance, when the primary winding current is 1000 amperes, a secondary winding with a 1000:1 CT ratio will provide an output current of 1.

Table 3.1 Specification of Current Transformer

Model	ZMCT103C
Rated Input Current	5 A
Rated Output Current	5 mA
Turn Ratio	10000:1
Accuracy Class	2
Insulation Voltage	3000V
Operating Temperature	-40-70- Degrees

- **Induction Motor**

Figure 3.3 Three phase induction motor

A three-phase induction motor is a type of electric motor that uses a three-phase AC power source to run. It is made up of a stator and a rotor, and the three-phase power supply energizes the stator to create a revolving magnetic field. The rotor rotates as a result of the magnetic field's induction of a current, which results in mechanical output. Due to their effectiveness, dependability, and minimal maintenance requirements, three-phase induction motors are frequently utilized in industrial and commercial applications.

Table 3.2 Specification of an Induction Motor

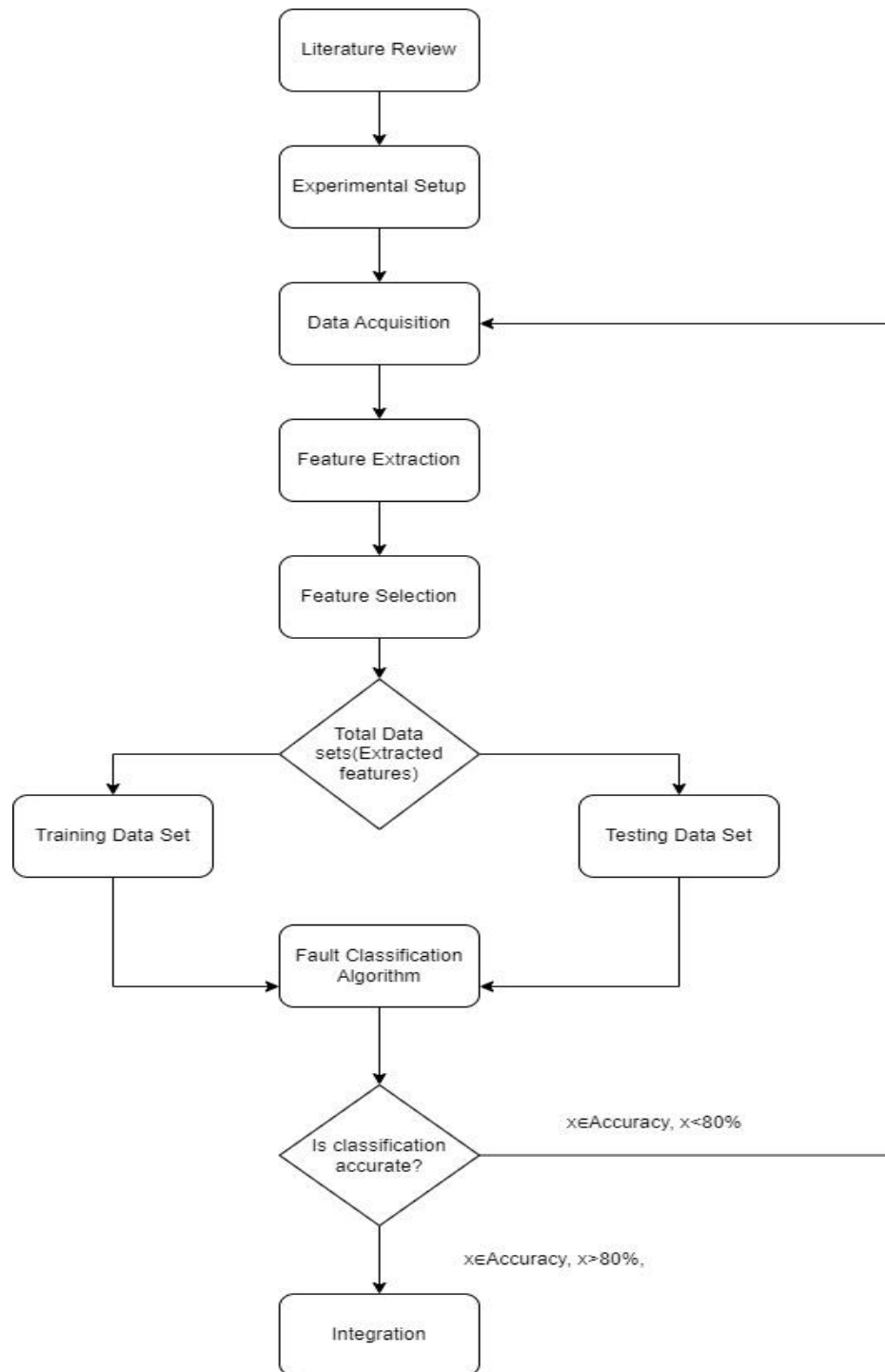| Power	1 hp
Frequency	50 hz
Phase	3 Phase
Speed	1390 r.p.m
Voltage	380 V
Current	2.01 Amp
No. of rotor slots	22
No of pole pair	4

3.2   Methodology

Figure 3.4 Flowchart representing research methodology

3.2.1 Data Acquisition
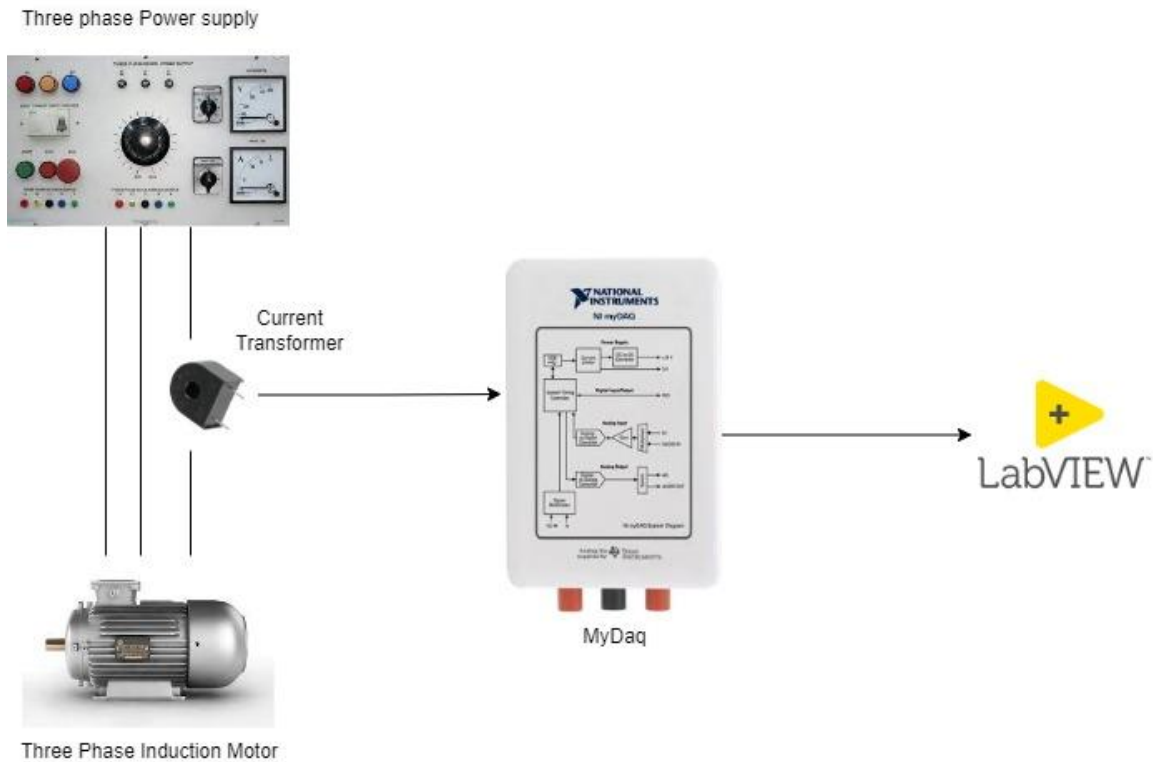
Figure 3.5 Block diagram repersenting experimental setup

The 3-phase induction motor is employed for testing the offered technique to detect the broken rotor bar fault. Various loading conditions were used in order to observe the variation in stator current. Stator current from single phase of induction motor is measured using current transformer using NI data acquisition card MyDAQ. To simulate the failure of broken rotors bars in the squirrel cage rotor of the three-phase induction motor it is necessary to drill the rotor. A rotor without a hole is to be tested first, that is, a healthy rotor, and then it will be successively replaced in order to obtain a database of monitored variables. The received signal is then fed to LabVIEW program for the detection and diagnosis of failures for healthy machines and machines with rotors consisting 2 and 4 bars broken adjacent to one another.

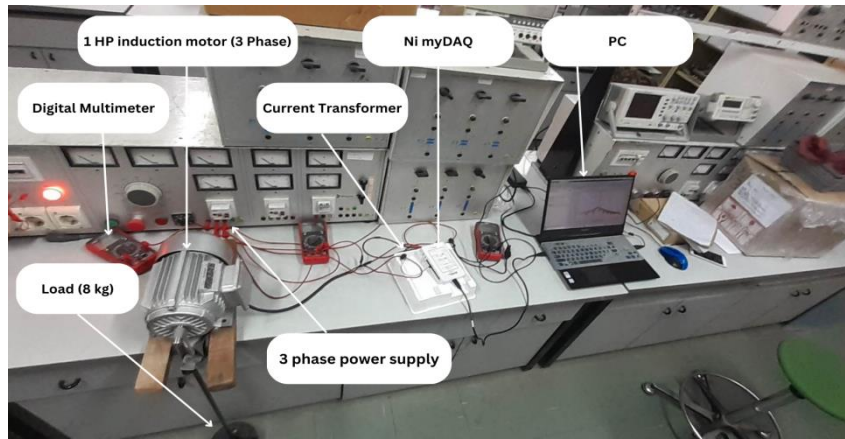### 3.2.2 Experimental Setup

Figure 3.6 Experimental setup for fault diagnosis of broken rotor bar in IM

The experiment was conducted in order to detect faults in a broken rotor bar in a 3-phase induction motor. The induction motor (Type: Y-802-4 & model: 202111170) consists of 22 rotor bars, has a nominal power of 1 hp, and supply frequency is 50 Hz at 220V. The power factor of IM was 0.76 with efficiency of 73%. Various loading condition were given to an induction motor in order to analyse the current signatures. Load of 2, 5 and 8 kg was given to an induction motor. Digital multimeters were used in order to check the supply current and voltage. Current transformer of turn ratio 1: 1000 was used in order to decrease the magnitude of a current so that it becomes feasible to feed the measured current in the data acquisition card. MyDAQ was used as an interface between measured current drawn by the motor and PC user. The data acquisition system based on MyDAQ from National instruments was used. With the help of LabVIEW software, the DAQ was configured to sampling frequency of 10000 samples/ sec. The data was taken for a duration of 10 seconds. So, the total sample was 100000 samples. These samples were taken for no load condition as well as loading condition. The fault was artificially induced in the rotor bar by drilling the rotor bar with a drill bit having diameter of 6mm. The first fault condition was taken as 2 broken rotor bars. All current samples for loading and unloading condition was taken. Again, the condition for 4 broken rotor bar was analysed by drilling holes in 4 rotor bars. All the current data obtained in a LabVIEW software was exported to Matlab for further analysis.

3.2.3   Labview Progarmming

The block diagram contians current indicator, current waveform genartor, DAQ assistant, write to measurement and while loop. The DAQ assiatnt is an automated feature in labview which is going to ease the process of collecting current measurements from our instrument. The write to measurement function allows us to save data file and allows us to configue how the data file is saved. The data will flow though the system only once. Continous measurement can be done by adding while loop. The stop condition is added so that while loop does not run continously. The real time current vlaue and current wavefrom can be seen in front panel.
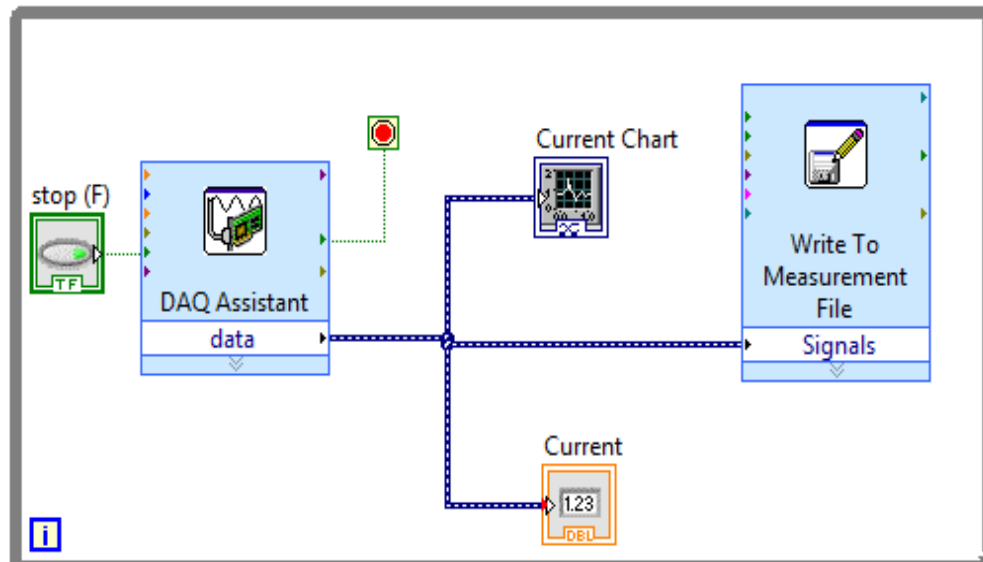

Figure 3.7 Block diagram to read the current VI
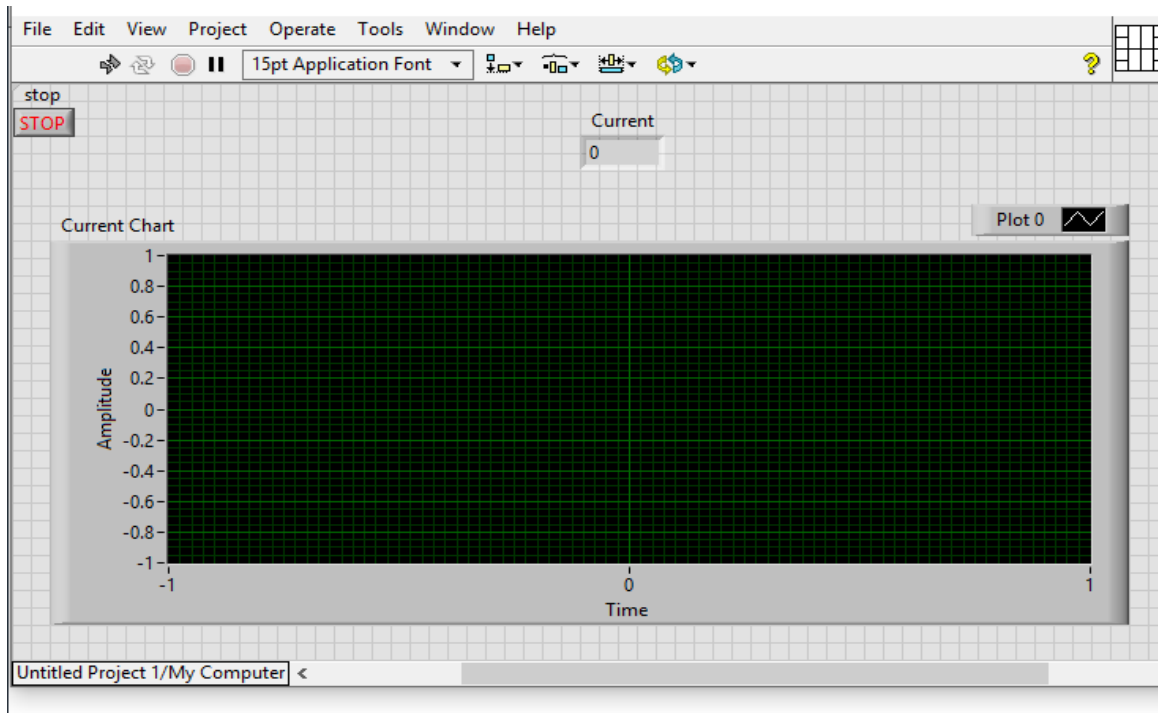
Figure 3.8 Front panel window

3.3   System Description

3.3.1   Signal Processing

The power spectrogram, or the power of a signal's frequency content over time, gives details about the signal's spectral energy content and is frequently used in a variety of applications, including signal processing. To quickly calculate the power spectrogram of a signal in MATLAB, use the spectrogram function in the Signal Processing Toolbox. Along with the corresponding time and frequency vectors, the function also returns the signal's power-frequency representation. The window function, window length, and window overlap parameters can all be set using the spectrogram function.

3.3.2   Feature Extraction

Feature extraction is a key step in the process of building a machine learning model. It is the process of transforming raw data into a set of meaningful and informative features that can be used as inputs to the model. The goal of feature extraction is to capture the most

important information in the data while reducing the dimensionality and removing any redundant or irrelevant information. Welch method is a widely used technique for feature extraction from spectral data in MATLAB. It provides a robust estimate of the frequency content of a signal and is particularly useful for signals with low SNR. MATLAB provides functions for the Welch method in the Signal Processing Toolbox. The pwelch function in MATLAB provides a convenient way to perform the Welch method and provides options for specifying the window function, the length of the segments, and the overlap between the segments. The function returns the estimate of the PSD, as well as the corresponding frequencies.

3.3.3   Feature Selection

Choosing relevant features from a wide set of attributes is a crucial step in the machine learning process that. The feature selection aims to strengthen the performance of the model by reducing the dimensionality of the data, eliminating unnecessary features, and enhancing the model's interpretability. The feature selection can be carried out in MATLAB using a subset of the features that are most important for creating the predictive model after the power spectrogram has been computed. The T-test and One Way Anova are two feature selection techniques that can be applied in this situation.

3.3.4   Machine Learning Model

Building a Machine Learning Model involves selecting an appropriate algorithm, such as linear regression, logistic regression, decision trees, or neural networks, and training the model on the data. MATLAB provides a number of functions and tools for building machine learning models. MATLAB also provides a number of functions and tools to evaluate its performance such as confusion matrix.

CHAPTER FOUR: RESULTS AND DISCUSSION

4.1 Output

The experimental workbench was setup to acquire the current signal from induction motor. The displaced material of the rotor bar caused due to fault injection was calculated theoretically. Then, the features were extracted from faulty and healthy cases. Finally, machine learning model was developed for computation of accuracy.

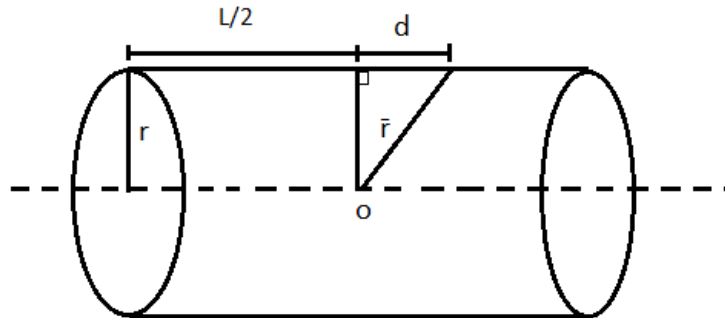### 4.1.1 Calculation of Unbalanced Force due to Broken Rotor Bar

Figure 4.1 Cylindrical Rotor Bar of an IM

Consider squirrel cage rotor as cylindrical object having radius r and length L.

Radius of drill bit (a) = 6 mm = 0.006 m

Depth of hole (h) = 5 mm = 0.05 m

Now,

$$\text{Volume of reduced material (V)} = 0.5\pi\, a^2\, h$$

4.1

$$= 2.82 * 10^{-7} \text{ m}^3$$

Density of aluminum (ρ) = 2710 kg/m^3

So,

Mass of reduced material (m) = V * ρ

4.2

$$= 7.66 * 10^{-4} \text{ kg}$$

The speed of motor after drilling (N) = 900 rpm

Then,

Angular velocity (w) = $2 \pi N / 60 = 94.2$ rad/s

4.3

We have,
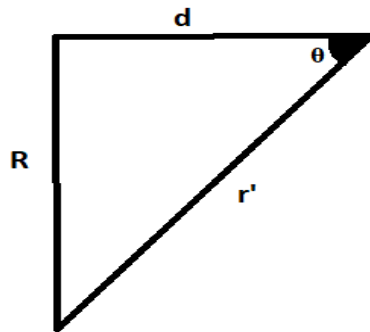
Radius of rotor (R) = 4 cm = 0.04m

Distance of hole from center (d) = 0.023 m

Then,

$\Theta = \tan^{-1}(R / d) = 60°$

Now,

Unbalanced force due to broken rotor bar

$F' = m w^2 r'$

4.4

$= m w^2 (R / \sin \Theta) = 0.313$ N.

Therefore, unbalanced force due to two and four broken bar is calculated as 0.627N and 1.253 N respectively.

Now, how these magnitude of mechanical unbalanced force results into variation in current signal and power spectrum is demonstrated below.

4.1.2 Feature Extraction

The entire signal processing is done on MATLAB. For this, the acquired experimental data is stored in an ensemble matrix form using a MATLAB code. The data is stored in order by assigning them with condition variables like health conditions and load conditions. Each sample of data contains 100000 values of current signals. The data were obtained for three different health conditions like healthy, 2 broken bar and 4 broken bars respectively and four load conditions like no load, 2 kg, 5 kg and 8 kg respectively. These nonlinear loads were used to simulate the real-life conditions for training ML algorithm. For each case 3 different samples were taken resulting 36 different samples in total.

1 Ia	2 Health	3 Load
100000x1 double	"healthy"	"8 kg"
100000x1 double	"healthy"	"5 kg"
100000x1 double	"healthy"	"2 kg"
100000x1 double	"healthy"	"no load"

1 Ia	2 Health	3 Load
100000x1 double	"2 broken bar"	"8 kg"
100000x1 double	"2 broken bar"	"5 kg"
100000x1 double	"2 broken bar"	"2 kg"
100000x1 double	"2 broken bar"	"no load"

1 Ia	2 Health	3 Load
100000x1 double	"4 broken bar"	"8 kg"
100000x1 double	"4 broken bar"	"5 kg"
100000x1 double	"4 broken bar"	"2 kg"
100000x1 double	"4 broken bar"	"no load"

Figure 4.2 Ensemble matrix containing acquired data
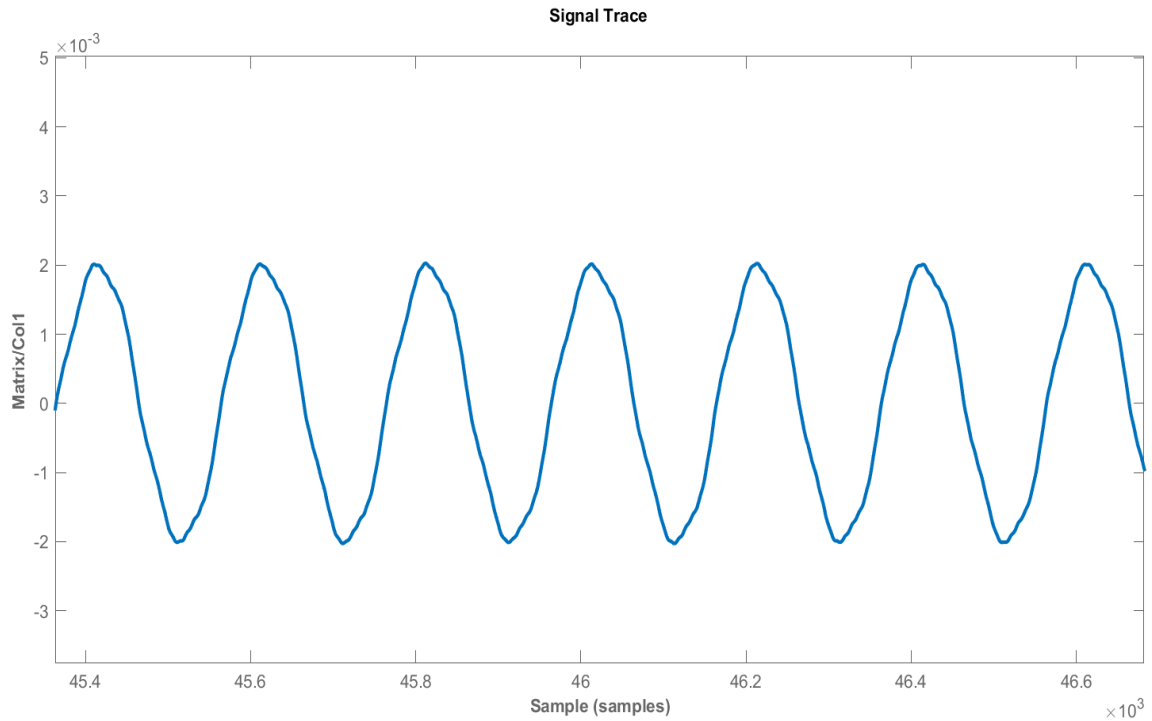
Figure 4.3 Signal tracing in time domain

Signal processing is done using diagnostic feature designer toolbox in MATLAB. For this, the time domain signal is converted into frequency domain signal using welch method to plot the power spectrum of the given data. Welch method is used because it is based on fast Fourier transformation and converts time domain signal into frequency domain signal. Also, the graphical view of this signal is quite large with a lot of noise containing frequency from 0.3 Hz to 5000 Hz. So, it is trimmed within a range of 10 to 90 Hz because the fundamental frequencies are present between this range.
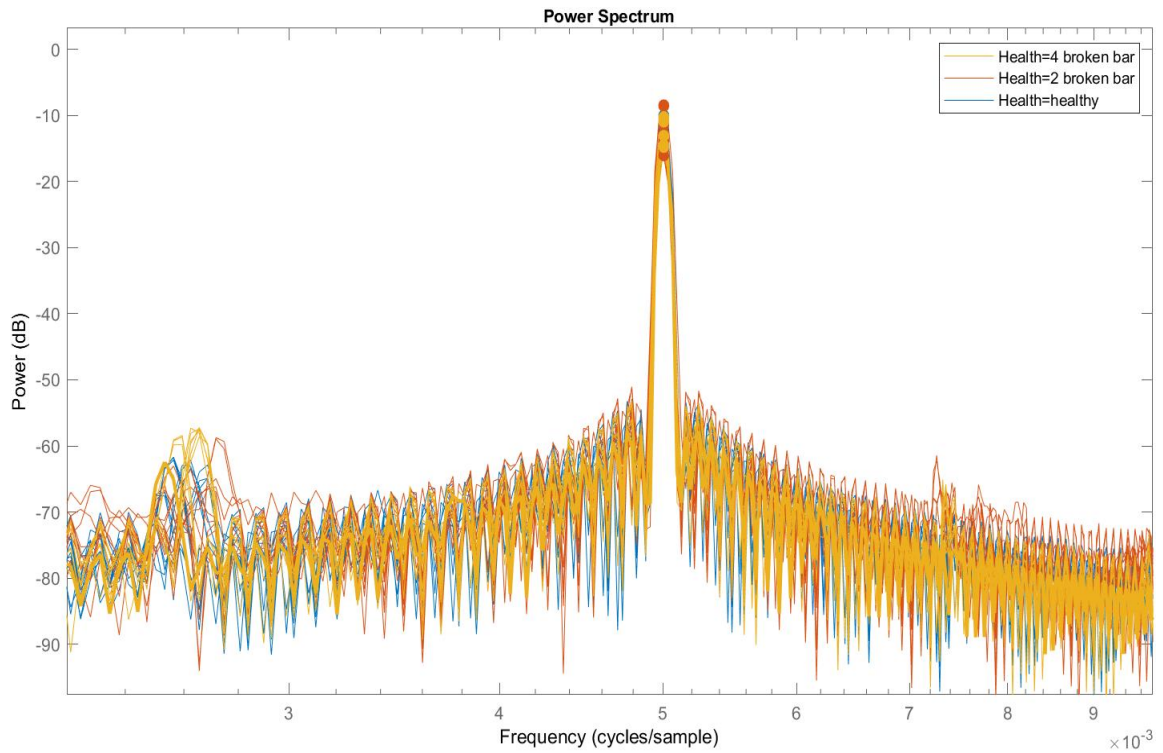
Figure 4.4 Power Spectrum

Peak frequency, peak amplitudes and band power are selected as frequency domain features because they carry characteristic information of the signal.

4.1.3   Calculation of Expected Frequencies of Sidebands.

The synchronous speed of rotor can be defined as

$$n1 = 120*f1/p$$

4.5

Where, pole (p)= 4 and f1= 50 Hz

$$n1= 1500 \text{ rpm}$$

Also, nominal speed of the motor (n)=1315 rpm

$$slip(s)= (n1-n) /n1$$

4.6

$$s = 0.123$$

48

This is the slip at maximum load condition.

When k=1,

$$f_b = (1 \pm 2ks)$$

<div align="right">4.7</div>

,

f_1 =37.7Hz and 62.3 Hz are the frequencies for first harmonics.

when k=2,

f_1 = 25.4Hz and 74.6Hz are the frequencies for second harmonics.

And, when k=3,

f_1 = 13.1Hz and 86.9 Hz the frequencies for third harmonics.

Table 4.1 Expected frequencies of sidebands for different harmonics

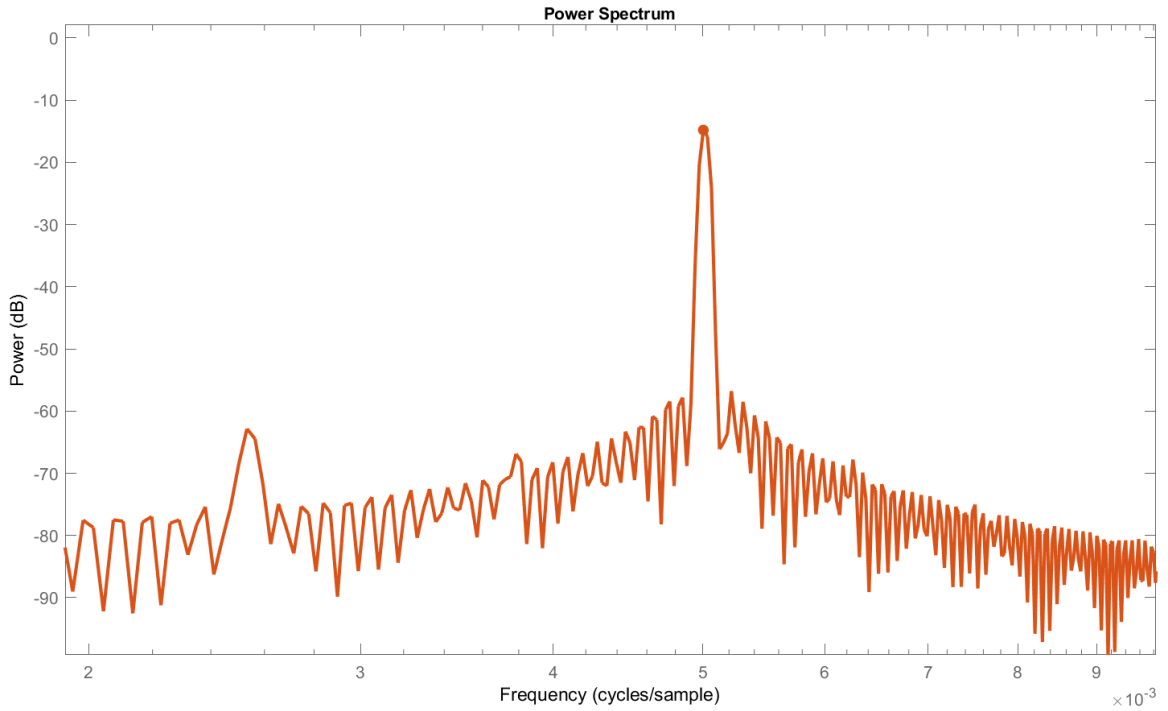| Load conditions	Speed (rpm)	slip	K=1		K=2		K=3	
			LSB (Hz)	USB (Hz)	LSB (Hz)	USB (Hz)	LSB (Hz)	USB (Hz)
No load	1315	0.123	37.7	62.3	25.4	74.6	13.1	86.9

Figure 4.5 Current spectrum at 5kg load

Table 4.2 Obtained frequencies of sidebands for different harmonics

Load conditions	Speed(rpm)	slip	K=1		K=2	
			LSB(Hz)	USB(Hz)	LSB(Hz)	USB(Hz)
No load	1315	0.123	37	63	26	73

The expected and obtained values were very close within an error tolerance of <5%.

4.1.4 Current Signal at Different Load Conditions
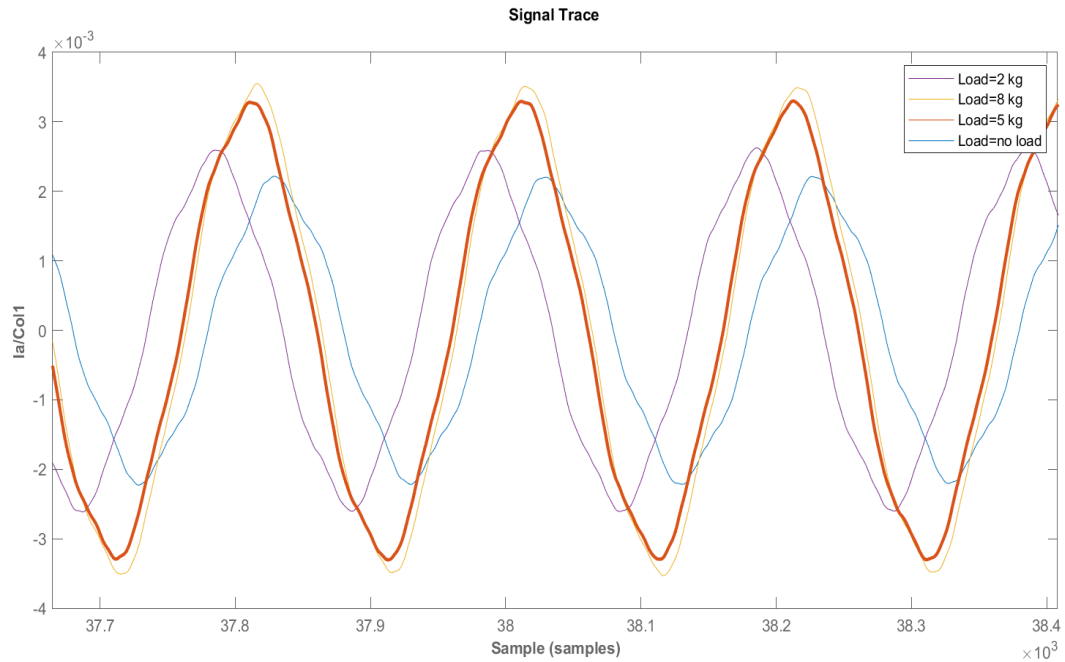
Figure 4.6 Current signals at different load condition

The graph above represents the signal tracing of current values at different load conditions indicated by different color codes. We can see that the peak values range from 2mA and 3.5mA for no load condition up to 8 kg load. So, the increase in peak values is directly directly proportional to the increase in load condition. It is because when the mechanical loads are applied to the shaft of the motor, speed of the motor decreases which reduces the back EMF. So additional current is drawn from the source to carry the increased load at a reduced speed.

4.1.5 Current Signal at Different Health Conditions
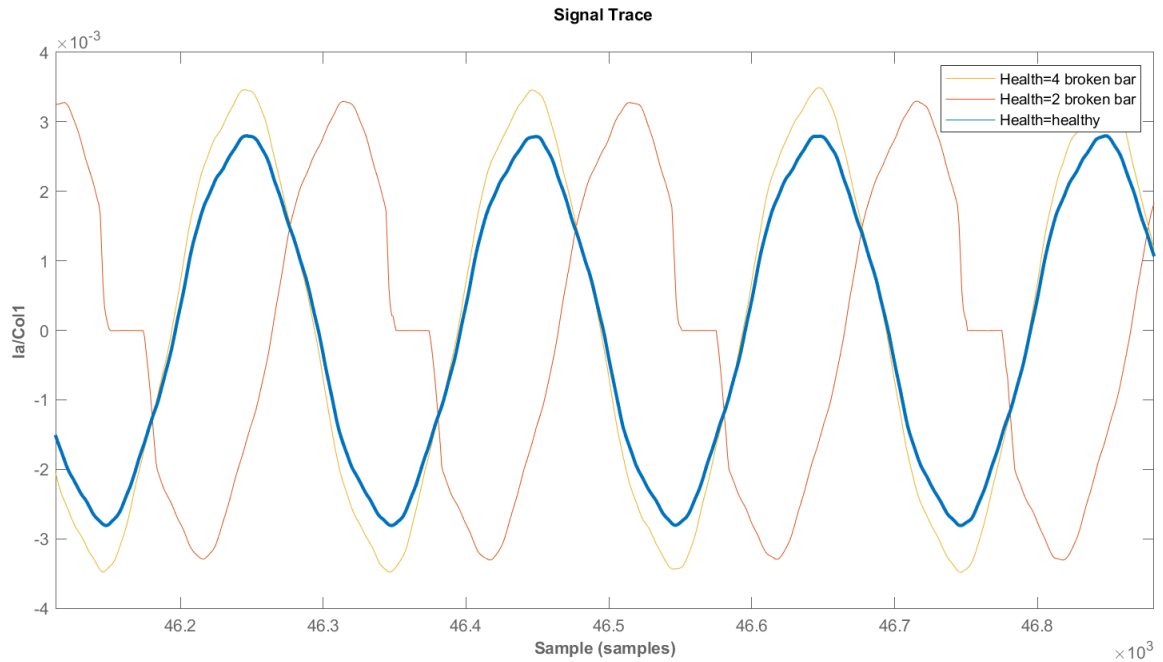
Figure 4.7 Current signals at different health conditions

The graph above represents the current signal in time domain at different health conditions indicated by different color codes. All of these signals are taken at 5 kg loading condition. We can observe that with increase in broken bar of the rotor, the peak values of current increases.

4.1.6 Power Spectrum at Same Health Condition but Different Load Conditions

Figure 4.8 Spectrum at same health but different load conditions

The graph above shows the power spectrum of different samples which are of same health condition but different load conditions that are represented by different color code. We can see that the frequency of sidebands increases when loads are increased. We can also see that the amplitudes are relatively equal when loads are applied whereas the amplitude is lower at the no load condition. It is because the rpm of motor decreases with increasing load which also increases slip resulting in increased frequency.

4.1.7 Power Spectrum at Same Load Conditions but Different Health Condition
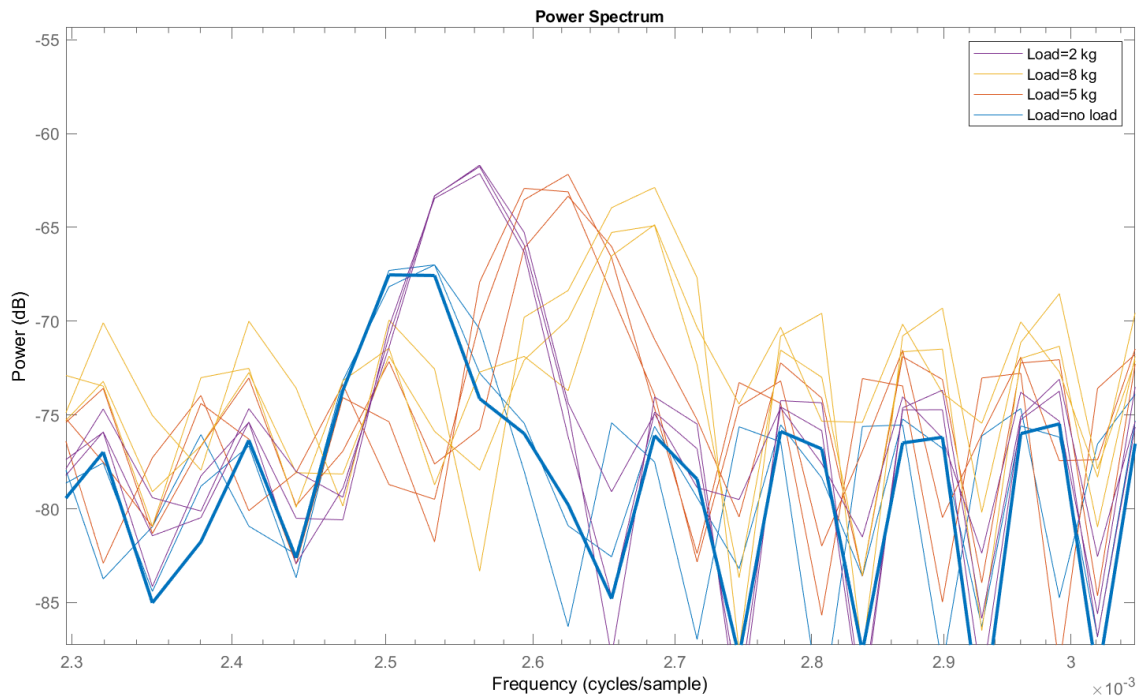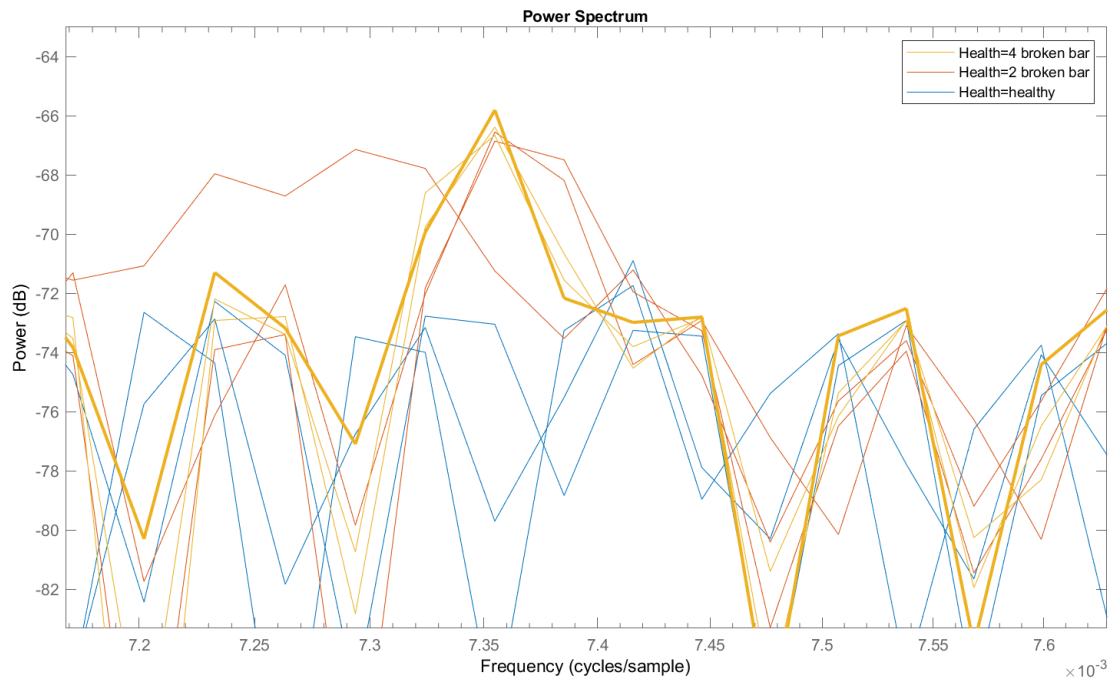

Figure 4.9 Spectrum at same load but different health conditions

The graph above shows the power spectrum of different samples which are of different health condition but same load condition that are represented by different color code. We can see that the amplitudes of sidebands increased when broken rotor bars are increased with frequency remaining relatively similar.

4.2   Feature Selection

For this, all the possible features of frequency domain like band power, peak frequency and peak amplitudes are calculated and shown in histogram as probability diagram below:
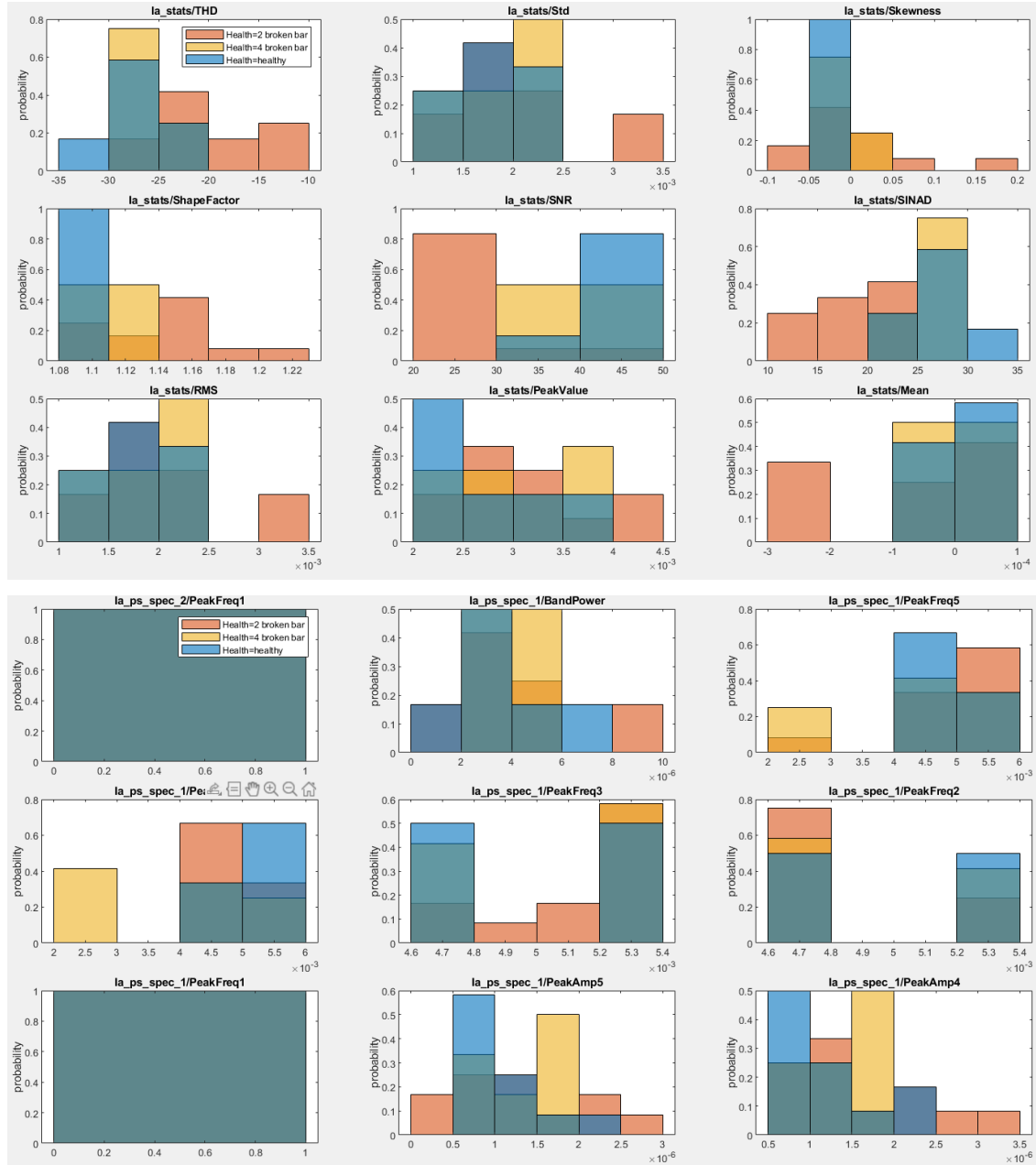
Figure 4.10 Histogram of all features with different health condition

The histogram shown above contains all the time domain features like mean, SNR, peak value etc. and frequency domain features like peak frequency, peak amplitude and band power. We can visually inspect that mean, SNR (time domain) and band power, peakamplitude5 (frequency domain) have separate value ranges for healthy and faulty condition which has two distinguished color code.

4.3 Feature Ranking

One-way ANOVA test is used for feature ranking.

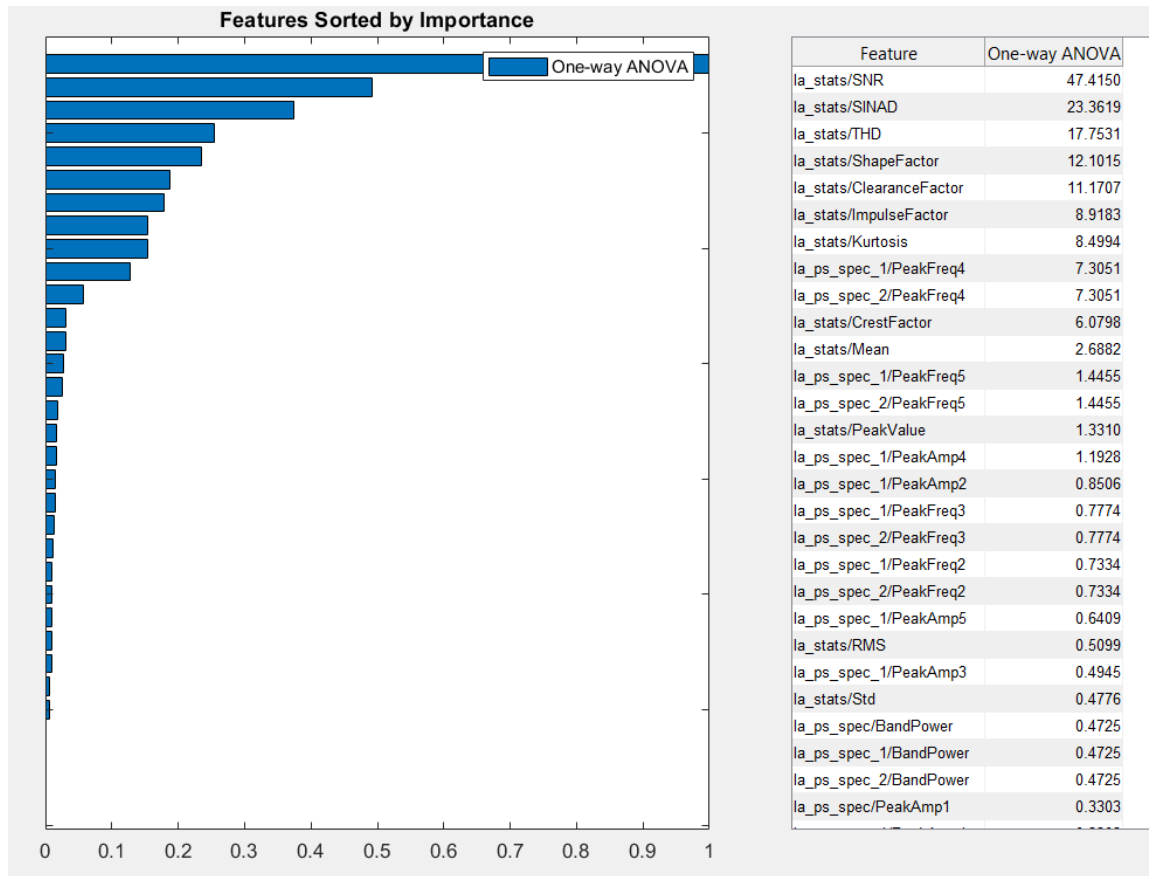| Feature	One-way ANOVA
la_stats/SNR	47.4150
la_stats/SINAD	23.3619
la_stats/THD	17.7531
la_stats/ShapeFactor	12.1015
la_stats/ClearanceFactor	11.1707
la_stats/ImpulseFactor	8.9183
la_stats/Kurtosis	8.4994
la_ps_spec_1/PeakFreq4	7.3051
la_ps_spec_2/PeakFreq4	7.3051
la_stats/CrestFactor	6.0798
la_stats/Mean	2.6882
la_ps_spec_1/PeakFreq5	1.4455
la_ps_spec_2/PeakFreq5	1.4455
la_stats/PeakValue	1.3310
la_ps_spec_1/PeakAmp4	1.1928
la_ps_spec_1/PeakAmp2	0.8506
la_ps_spec_1/PeakFreq3	0.7774
la_ps_spec_2/PeakFreq3	0.7774
la_ps_spec_1/PeakFreq2	0.7334
la_ps_spec_2/PeakFreq2	0.7334
la_ps_spec_1/PeakAmp5	0.6409
la_stats/RMS	0.5099
la_ps_spec_1/PeakAmp3	0.4945
la_stats/Std	0.4776
la_ps_spec/BandPower	0.4725
la_ps_spec_1/BandPower	0.4725
la_ps_spec_2/BandPower	0.4725
la_ps_spec/PeakAmp1	0.3303

Figure 4.11 Features ranking using One-way Anova

From the above feature ranking table, we can see that the time domain features like SNR and SINAD has highest scores of 47.41 and 23.36 respectively followed by other time domain features like THD, shape factor, clearance factor etc. Also, we have frequency domain features like peak amplitude and peak frequency with scores 1.19 and 0.777

respectively. One-way ANOVA test is done because it is a versatile method to demonstrate the extent of comparative differences between more than two groups.

Out of all these features, top ten features were then selected and extracted into numerical values for classification training using ML algorithm.

The exported features as numerical value is shown as:

	Health	Load	la_ps_spec/PeakAmp1	la_ps_spec/PeakFreq1	la_ps_spec/BandPower	la_ps_spec_1/PeakAmp1	la_ps_spec_1/PeakAmp2	la_ps_spec_1/PeakAmp3	la_ps_spec_1/Peak
1	healthy	no load	0.0321	0.0050	1.9825e-06	0.0321	1.5441e-06	1.2613e-06	6.66
2	healthy	5 kg	0.0574	0.0050	3.7793e-06	0.0574	3.2267e-06	2.3983e-06	1.40
3	healthy	8 kg	0.0916	0.0050	5.8742e-06	0.0916	5.1127e-06	4.0460e-06	2.11
4	healthy	no load	0.0316	0.0050	1.9797e-06	0.0316	1.5456e-06	1.4412e-06	6.27
5	healthy	5 kg	0.0582	0.0050	3.6094e-06	0.0582	2.8111e-06	2.4154e-06	1.25
6	healthy	8 kg	0.0952	0.0050	6.0088e-06	0.0952	4.5171e-06	4.4844e-06	1.73
7	healthy	2 kg	0.0425	0.0050	2.6180e-06	0.0425	2.0925e-06	1.8099e-06	9.25
8	healthy	2 kg	0.0405	0.0050	2.5736e-06	0.0405	2.1312e-06	1.8337e-06	8.54
9	healthy	no load	0.0335	0.0050	2.1401e-06	0.0335	1.7594e-06	1.5210e-06	6.94
10	healthy	5 kg	0.0708	0.0050	4.3962e-06	0.0708	3.4091e-06	2.9302e-06	1.44
11	healthy	8 kg	0.1003	0.0050	6.1563e-06	0.1003	4.7327e-06	4.4071e-06	2.14
12	healthy	2 kg	0.0406	0.0050	2.5552e-06	0.0406	1.9752e-06	1.9071e-06	7.60
13	2 broken bar	no load	0.0426	0.0050	2.8607e-06	0.0426	2.7025e-06	1.5477e-06	1.21
14	2 broken bar	5 kg	0.0708	0.0050	4.8423e-06	0.0708	4.2622e-06	3.1644e-06	2.00
15	2 broken bar	8 kg	0.1434	0.0050	9.1715e-06	0.1434	7.3655e-06	6.8127e-06	2.81
16	2 broken bar	no load	0.0433	0.0050	2.6561e-06	0.0433	2.0006e-06	1.9337e-06	9.47
17	2 broken bar	5 kg	0.0772	0.0050	4.9716e-06	0.0772	4.2749e-06	3.3826e-06	1.80
18	2 broken bar	8 kg	0.1379	0.0050	9.0169e-06	0.1379	7.8623e-06	5.9764e-06	3.17
19	2 broken bar	2 kg	0.0520	0.0050	3.5063e-06	0.0520	2.8827e-06	2.3770e-06	1.41
20	2 broken bar	2 kg	0.0542	0.0050	3.5137e-06	0.0542	3.3029e-06	2.4597e-06	1.32
21	2 broken bar	no load	0.0247	0.0050	1.5495e-06	0.0247	1.2116e-06	1.1702e-06	5.07
22	2 broken bar	5 kg	0.0496	0.0050	3.0528e-06	0.0496	2.4821e-06	2.1361e-06	1.02
23	2 broken bar	8 kg	0.0815	0.0050	5.7554e-06	0.0815	5.0763e-06	4.4300e-06	2.49
24	2 broken bar	2 kg	0.0303	0.0050	1.9659e-06	0.0303	1.7235e-06	1.3382e-06	7.15
25	4 broken bar	no load	0.0330	0.0050	2.1995e-06	0.0330	2.1013e-06	1.4298e-06	8.54
26	4 broken bar	5 kg	0.0916	0.0050	5.6779e-06	0.0916	4.5252e-06	4.0686e-06	1.91
27	4 broken bar	8 kg	0.0926	0.0050	5.7876e-06	0.0926	4.3498e-06	4.3296e-06	1.79
28	4 broken bar	2 kg	0.0491	0.0050	3.0695e-06	0.0491	2.5260e-06	2.4369e-06	1.17
29	4 broken bar	no load	0.0356	0.0050	2.1890e-06	0.0356	1.6912e-06	1.5849e-06	7.32

Figure 4.12 Exported features

The same process was repeated for selecting features when the sample was grouped into different load conditions as shown below,
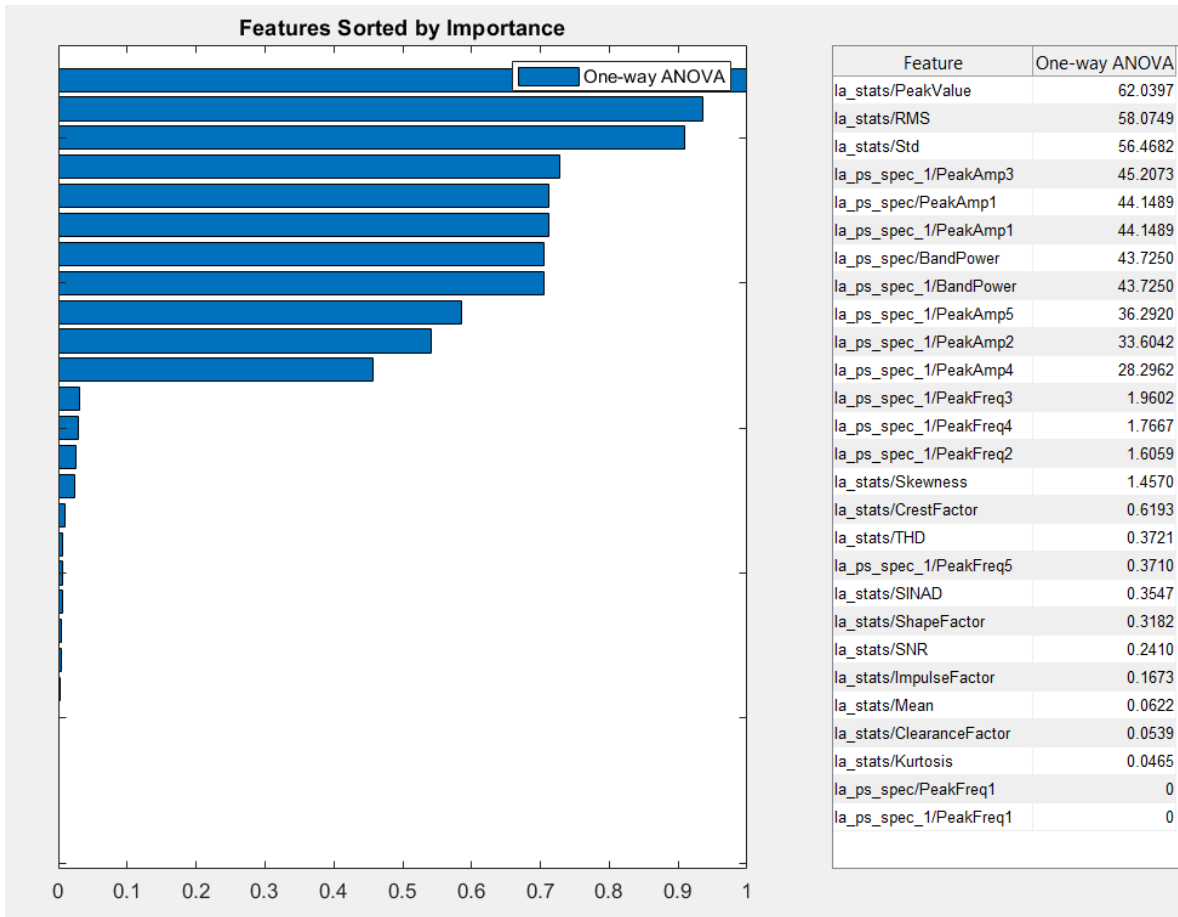

Figure 4.13 Features for different load conditions

From the feature ranking table, we can see that Peak value and RMS of time domain features have high scores for One-way ANOVA test with scores 62.03 and 58.07 respectively followed by peak amplitude and band power with scores 45.20 and 43.72 respectively which are frequency domain features. All these features are ranked taking load as condition variable.

Similarly, top 10 features are selected and exported as numerical values which will be used on for training ML algorithm.

4.4 Machine Learning

All the exported features were then used in classification learner app which is an inbuilt toolbox in MATLAB for training. For the training process all the available algorithms like SVM, KNN, naïve baise, tree, neural network etc. were used and based on the accuracy the algorithm will be selected as a trained model.

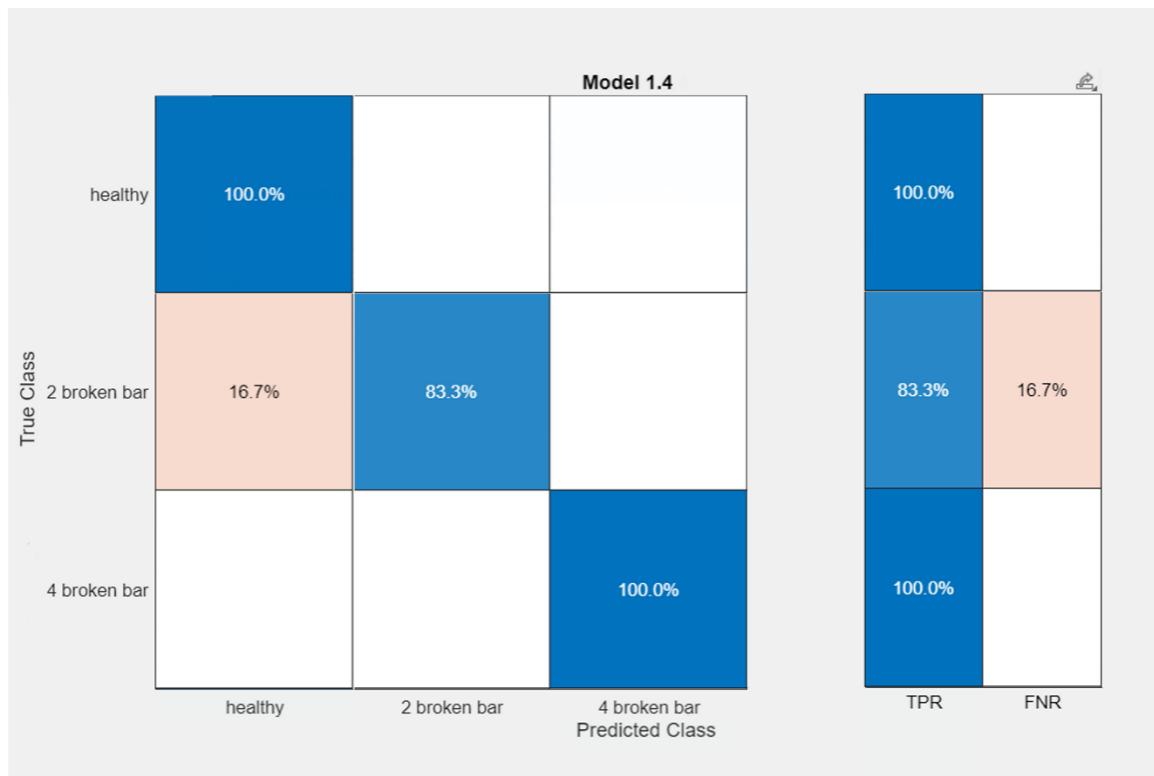After the training process is completed, a confusion matrix is obtained as shown below:

Figure 4.14 Confusion matrix for different health conditions and 5kg load

The above figure shows the confusion matrix obtained after training data for different health conditions with various ML algorithms. The observations can be discussed as following;

1) The overall classification accuracy turned out to be 94.4% accurate. The algorithm used for training with this accuracy was Naïve Bayes. Similarly, other algorithms

like Tree and neural network also classified the faults with accuracy more than 90% which is considered excellent for any machine learning model (Oviously.ai, 2022).

2) We can see that the predicted class were 100% accurate for healthy and 4 broken bar case whereas the accuracy reduced to 83.3% for 2 broken bar case.

Various similar studies have demonstrated accuracy of 80-99% and is considered as suitable machine learning models. (Esakimuthu Pandarakone, 2019). So, it also gives us an indication that this project is suitable for integrating machine learning model into real industries. The precision and recall value of this model is 0.8756 and 0.861. Also, the F-score of this model is 0.8683 which indicates it averts the possibility of false positive and false negative cases to higher extent which is beneficial for our problem (El Menzhi, 2022).

(i)   (ii)

Figure 4.15 Accuracy results for health conditions (i) and load condition (ii)

In the same way, the extracted features for a different number of load conditions were trained using all the available ML algorithms and the confusion matrix obtained is shown.
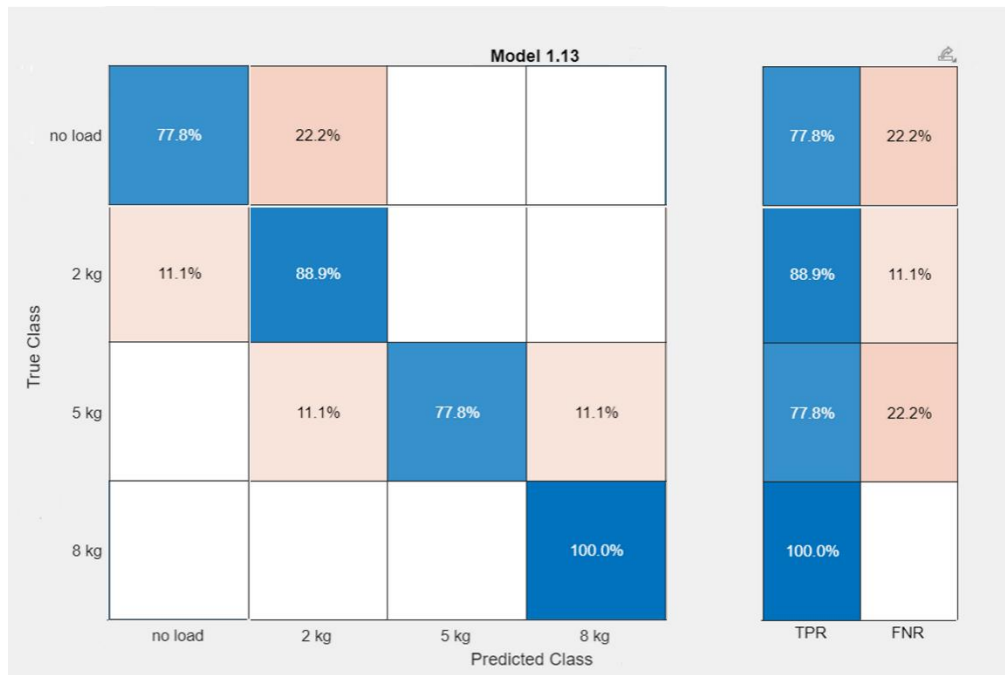
Figure 4.16 Confusion matrix for different load conditions and 4 broken bar

Table 4.3 Accuracy of all algorithms for classification of health & load conditions

S.no	Classification Method	Sub-groups	Classification accuracy for health condition (%)	Classification accuracy for load condition (%)
1	Naïve Bayes	Kernel	**94.4**	50
		Gaussian	-	-
2	SVM	Linear	88.9	63.9
		Quadratic	83.3	61.1
		Cubic	75	66.7
		Medium Gaussian	75	63.9
		Fine Gaussian	61.1	61.1
		Coarse Gaussian	47.2	41.7
3	Decision tree	Fine	91.7	77.8
		Medium	91.7	77.8
		Coarse	91.7	77.8
4	Ensemble	Bagged trees	88.9	**86.1**
		RUSBoosted Trees	41.7	30.6
		Boosted Trees	27.8	16.7
5	Neural network	Medium	86.1	69.4
		Wide	86.1	69.4
		Bilayered	86.1	80.6
		Narrow	83.3	75
		Trilayered	75	63.9

The above figure shows the confusion matrix obtained after training data for different load conditions with various ML algorithms. The observations can be discussed as following;

1) The overall classification accuracy turned out to be 86.1% accurate and the algorithm used for training is Ensemble.

2) For 8 kg load, model was able to predict the fault with no error. It is because at higher load conditions the sidebands are clearer with greater amplitude and there is less noise.

3) On the other hand, the classification accuracy decreased to 77.8% for no load condition. It is because at no load condition, the sidebands are very small (almost

negligible) as the current in the rotor bars is minimal which makes them harder to detect for user by observation.

4.5   Limitations

- The other IM faults like air gap eccentricity, bearing faults etc. were not created in IM during data acquisition.
- Higher loading conditions (greater than 10 kg) in IM were not explored.

4.6   Problem Encountered

The problem faced during the task are enlisted below:
- Three phase supply for the induction motor was difficult to find.
- Difficulties in finding the current transformer as it is not easily available.
- Extensive iterations were required during machine learning process as desired accuracy was not easily achievable.

4.7   Budget Analysis

Table 4.4 Budget Division

S. N	Particular	Amount (NRs)	Remarks
1	Induction motor	9,000	
2	Current transducer	500	
3	Documentation	1500	
	Grand total	**11000**	

CHAPTER SIX: CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The experimental workbench was successfully setup to acquire stator current signal from three phase induction motor. The data acquisition system MyDAQ was used to acquire current signals from single phase of the motor using current transformer. The current transformer of turn ratio 1000:1 was clamped to one of the supply line wires of induction motor. The motor was loaded with 2kg, 5kg and 8 kg load. The LabVIEW software was used to configure the MyDAQ to sampling frequency of 10000 samples/sec. The data was taken for a duration of 10 seconds. So, the total sample was 100000 samples. The samples were taken for no load condition as well as loading condition. The fault was artificially induced in the rotor bar by drilling the rotor bar 5mm deep with a drill bit having diameter of 6mm. The first fault condition was taken as 2 broken rotor bar and later the condition of 4 broken bar was taken into consideration. All current samples for loading and unloading condition was taken.

During the experiment, the healthy and faulty condition data were collected and preprocessing of data was done. The data were obtained for three different health conditions i.e., healthy, 2 broken bar and 4 broken bars respectively. Four load conditions were taken i.e., no load, 2 kg, 5 kg and 8 kg respectively. For each case 3 different samples were taken resulting 36 different samples in total.

Finally, various supervised machine learning algorithms like SVM, Naïve Bayes, Tree, and Ensemble were used for fault classification of broken rotor bar in an induction motor. The result indicated that Naïve Bayes algorithm was the most suitable among other algorithms with an accuracy of 94.4%. Similarly, other algorithms like SVM and Decision tree also performed well with an accuracy of 88.9% and 91.7%. Hence, it can be concluded that Naïve Bayes ML model can possibly be reliable in order to integrate into real life systems for the prediction of faults and condition monitoring of the induction motor in industries.

5.2 Recommendations

5.2.1 Scope for Future Enhancement

- The advanced machine learning algorithm techniques like deep learning can be deployed for the fault detection of induction motor in case of broken rotor bar.
- All the cases of electrical and mechanical defects can be studied and robust machine learning model can be developed which can be integrated to real life industries.

REFERENCES

Alkadhim, S. A. S., 2020. Three-phase Induction Motor: Types and Structure. *Available at SSRN ,* p. 3647425.

Ate., 2018. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Transactions of the Institute of Measurement and Control,* Volume 4, pp. 2681-2693.

Asghar, F., Talha & Ho Kim, S., 2016. A Matlab and Simulink based three-phase inverter fault diagnosis method using three-dimensional features. *International Journal of Fuzzy Logic and Intelligent Systems,* 16(25), pp. 173-180.

A. & W., 2009. Intelligent fault diagnosis system of induction motor based on transient current signal. *Mechatronics,* 19(6), pp. 680-689.

Ayodele & Oladipupo, 2010. Types of Machine Learning Algorithms. *New advances in machine learning,* 3(53), pp. 19-48.

Babaa, F., Bennis, O., Kratz, F. & Hadri-Hamida, A., 2021. Combined Electrical Faults Detection and Diagnosis Using Current Signature Analysis.. *Power Electronics and Drives (SDEMPED),* 1(1), pp. 52-57.

Benbouzid, M. E. H., 2000. A review of induction motors signature analysis as a medium for faults detection. *IEEE transactions on industrial electronics,* 47(12), pp. 984-993.

Benbouzid, M. E. H., 2000. A review of induction motors signature analysis as a medium for faults detection.. *IEEE transactions on industrial electronics,* 47(5), pp. 984-993.

Benbouzid, M. & Kliman, G., 2003. What stator current processing-based technique to use for induction motor rotor faults diagnosis?. *IEEE Transactions on energy Conversion,* 2(3), pp. 238-244.

Berrar, D., 2018. Bayes' theorem and naive Bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics,* Volume 403, p. 412.

Bonnett, A. H. & Soukup, G. C., 1986. Rotor failures in squirrel cage induction motors.. *IEEE transactions on industry applications,* 6(21), pp. 1165-1173.

Brar, R. K. a. G., 2016. Induction Motor Drives-A Literature Review.

Chao, C. & C., 2013. Realizing Method of MATLAB Software for Fault Diagnosis of Freight Bearings. *Applied Mechanics and Materials,* 273(26), pp. 245-249.

Chesnutt, C. M. C. B., 2011. *Incorporation of NI myDAQ exercises in electric circuits.* Washington DC: s.n.

Choudhary, A., Goyal, D., Akul, A. & Shimi, S., 2019. Condition monitoring and fault diagnosis of induction motors: A review. *Archives of Computational Methods in Engineering,* 26(11), pp. 1221-1298.

Cortes & Vapnik, V., 1995. Support Vector networks. *Machine learning,* 20(55), pp. 273-297.

Dalpiaz, G. & Meneghetti, U., 1991. Monitoring fatigue cracks in gears.. *NDT & E International,* 6(20), pp. 303-306.

Dhomad & A, J., 2020. Bearing Fault Diagnosis Using Motor Current Signature Analysis andthe Artificial Neural Network. *International Journal on advanced scince Engineering Information Technology,* 10(2).

Didier, G., Ternisien , E., Caspary, O. & Razik , H., 2006. Fault detection of broken rotor bars in induction motor using a global fault index. *IEEE Transactions on industry applications,* 42(17), pp. 79-88.

Ebadi, A. M. M., 2011. Torque analysis of three-phase induction motor under voltage unbalance using 2D FEM. *International Journal of Engineering Science and Technology,* Volume 3, pp. 871-876.

Edomwandekhoe, K. & Liang, X., 2018. Current spectral analysis of broken rotor bar faults for induction motors.. *IEEE Canadian Conference on Electrical & Computer Engineering (CCECE),* Issue 4, pp. 1-5.

El Menzhi, L. a. X. C., 2022. Supervised Classification Of induction Motors faults. *E3S Web of Conferences,* Volume 336.

Esakimuthu Pandarakone, S. Y. M., 2019. A comparative study between machine learning algorithm and artificial intelligence neural network in detecting minor bearing fault of induction motors. *Energies,* Volume 12, p. 2105.

Garcia, P. A.-D. G., 2018. Efficiency assessment of induction motors operating under different faulty conditions. *IEEE Transactions on Industrial Electronics,* Volume 66, pp. 8072-8081.

G, R. & Lyons, 1997. *Understanding digital signal processing.* India: 49.

Guo, H. & Liu, M. K., 2018. Induction motor faults diagnosis using support vector machine to the motor current signature.. *2018 IEEE Industrial Cyber-Physical Systems (ICPS),* Issue 10, pp. 417-421.

Higa, M. L., 2002. An introduction to LabVIEW exercise for an electronics class. *32nd Annual Frontiers in Education,* Volume 1.

H. & R., 2014. Faults identification in three-phase induction motors using support vector machines. Issue 37.

Hwang, D. H. et al., 2005. Analysis of a three phase induction motor under eccentricity condition.. *In 31st Annual Conference of IEEE Industrial Electronics Society,* Issue 8, pp. 5-pp.

James, W., Cooley & Tuckey, J. W., 1965. An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation,* 19(47), pp. 297-301.

Janikow, C., 1998. Fuzzy decision trees: issues and methods. *IEEE Transactions on Systems, Man, and Cybernetics,* Volume 28, pp. 1-14.

Jiang, J. R. & Kuo, C. K., 2017. Enhancing convolutional neural network deep learning for remaining useful life estimation in smart factory applications.. *2017 International*

Conference on Information, Communication and Engineering (ICICE), Issue 15, pp. 120-123.

Jijo, B. T., 2021. Classification based on decision tree algorithm for machine learning. *evaluation,* Volume 6, p. 7.

Johannessen, A., Robbersmyr, K. G. & Van Khang, H., 2019. Data-driven fault diagnosis of induction motors using a stacked autoencoder network. *International Conference on Electrical Machines and Systems,* Issue 50, pp. 1-6.

Kang, Z., Catal, C. & Tekinerdogan, B., 2021. Remaining useful life (RUL) prediction of equipment in production lines using artificial neural networks.. *Sensors 21,* 3(13), p. 932.

Kaur, R. & Brar, G. S., 2016. Induction motor drives-A literature review.. *Control Theory and Technology,* 16(17), pp. 8071-8081.

Kavana, V. & Neethi, M., 2018. Fault analysis and predictive maintenance of induction motor using machine learning.. *In 2018 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT),* Issue 9, pp. 963-966.

K., R. A, K. & . M. M., 1988. Noninvasive detection of broken rotor bars in operating induction motors. *IEEE Transactions on Energy Conversion,* 3(12), pp. 873-879.

Kumar, D., 2018. Performance analysis of three-phase induction motor with AC direct and VFD.. *IOP Conference Series: Materials Science and Engineering,* 331(18), p. 012025.

Kumar, P., 2021. Review on machine learning algorithm based fault detection in induction motors.. *Archives of Computational Methods in Engineering,* 28(12), pp. 1929-1940.

Kumar, . P. & Shankar Hati, . A., 2021. Review on machine learning algorithm based fault detection in induction motors. *Archives of Computational Methods in Engineering,* 28(8), pp. 1929-1940.

K. & V., 2001. *Learning and soft computing: support vector machines, neural networks, and fuzzy logic models.* s.l.:MIT press.

Lee , Y. C., J.H & Min, 2005. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert systems with applications,* 28(56), pp. 603-314.

Marple Jr, Lawrence, S. & Carey, W. M., 1989. Digital spectral analysis with applications. Issue 52, p. 2043.

Mehrjou, . R. et al., 2017. Broken rotor bar fault detection and classification using wavelet packet signature analysis based on fourier transform and multi-layer perceptron neural network. *Applied Sciences,* 8(5), p. 25.

Miljković, D. & Dubravko, 2015. BRIEF REVIEW OF MOTOR CURRENT. *HDKBR Info magazin,* 5(1), pp. 14-26.

Mitra, S. K., 2001. *Digital signal processing: a computer-based approach.* s.l.:McGraw-Hill Higher Education.

Moore, H., 2005. *MATLAB for engineers.* s.l.:14.

M. r. w. g., 1985. Report of large motor reliability survey of industrial and commercial installations. *IEEE Trans,* 21(14), pp. 853-872.

Muller, D., Baranski, M., Thul, S. & Bode, 2020. Real-world application of machine-learning-based fault detection trained with experimental data. *Energy,* 198(31), p. 117323.

Myung, H. S., A. & Trussell, H. J., 2003. A case study on the comparison of non-parametric spectrum methods for broken rotor bar fault detection. *ECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society,* 3(44), pp. 2835-2840.

Naha, A., Samanta, A. K., Routray, A. & Deb, A. K., 2016. A method for detecting half-broken rotor bar in lightly loaded induction motors using current.. *IEEE Transactions on Instrumentation and Measurement,* 65(7)(6), pp. 1614-1625.

nmn, n.d. *bmb,* p. 4.

O., Ejike Akpudo, U. & Nwabufo, C., 2022. "A Cost-Efficient MCSA-Based Fault Diagnostic Framework for SCIM at Low-Load Conditions. *Algorithms,* 15(4), p. 212.

Oviously.ai, 2022. *How to know if ML mode has good performance,* s.l.: s.n.

Press, W. H., 2001. *Numerical recipes 3rd edition: The art of scientific computing.* Cambridge: Cambridge university press..

Proakis, J. G., 2001. *Digital signal processing: principles, algorithms, and applications.* s.l.:Pearson Education India.

Riddle, J., 1955. *Ball bearing maintenance: a survey of design features and maintenance principles which govern performance and service life with a comprehensive study of service failures.* Oklahoma: University of Oklahoma Press.

Rodriguez, V. & M., 2020. Convolutional Neural Network and Motor Current Signature Analysis during the Transient State for Detection of Broken Rotor Bars in Induction Motors. *Sensors,* 13(3), p. 3721.

Romeral, L., Jordi & Ortega, J. A., 2008. Fault detection in induction machines using power spectral density in wavelet decomposition. *IEEE Transactions on Industrial Electronics,* 55(43), pp. 633-643.

Sayyad, S. et al., 2021. Estimating remaining useful life in machines using artificial intelligence: A scoping review.. *A scoping review,* Issue 14, pp. 1-26.

Selcuk, S., 2017. Predictive maintenance, its implementation and latest trends.. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* Volume 231, pp. 1670-1679.

Silva, AD, V. & Pederiva, . R., 2013. Fault detection in induction motors based on artificial intelligence. *Proceedings of the International Conference on Surveillance,* 7(7).

Singh, A. et al., 2016. A review of induction motor fault modeling.. *Electric Power Systems Research,* 133(2), pp. 191-197.

Soe, N. N. T. T. H. Y., 2008. Dynamic modeling and simulation of three-phase small power induction motor. *World Academy of Science, Engineering and Technology,* Volume 42, pp. 421-424.

Thorsen, O. & M, D., 1995. A survey of faults on induction motors in offshore oil industry,. *IEEE Trans,* 31(15), pp. 1186-1196.

Trussell, H. J., Chow, M.-Y., Ayhan & Bulent, 2003. A case study on the comparison of non-parametric spectrum methods for broken rotor bar fault detection. *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society,* 3(48), pp. 2835-2840.

Vaidyanathan, P. P., 2006. *Multirate systems and filter banks.* s.l.:Pearson Education India.

Valtierra-Rodriguez, M., 2020. Convolutional neural network and motor current signature analysis during the transient state for detection of broken rotor bars in induction motors.. *Sensors 20,* 13(11), p. 4721.

W. & A., 2007. Support vector machine in machine condition and fault diagnosis. *Mechanical systems and signal processing,* 21(54), pp. 2560-2574.

Wang, J., Gao, R. X. & Yan, R., 2011. Broken-rotor-bar diagnosis for induction motors.. *In Journal of Physics: Conference Series,* 305(7).

Wang, M., Liu, F. & Wu, 2013. ReRecognition of Control Chart Patterns Using a Support Vector Machine Based Classifier with Features Extracted from Clustering. *Metalurgia International,* 18(57), p. 5.

Wang, Y., Zhao, Y. & Addepalli, S., 2020. Remaining useful life prediction using deep learning approaches: A review.. *Procedia manufacturing,* 49(16), pp. 81-88.

W., . J., K, Y. & Z., 2018. Comparison of random forest, artificial neural networks and support vector machine for intelligent diagnosis of rotating machinery. *Transactions of the Institute of Measurement and Control,* 40(9), pp. 2681-2693.

Yilmaz, I. & O., 2008. Induction motor bearing failure detection and diagnosis: Park and concordia transform approaches comparative study. *IEEE/ASME Transactions on mechatronics,* 13(39), pp. 257-262.

Zhang, T., Zhang, W. & T., 2018. Induction motors dynamic eccentricity fault diagnosis based on the combined use of WPD and EMD-simulation study. *Applied Sciences,* 8(37), p. 1709.

APPENDIX

Code: To store data in ensemble matrix in MATLAB

```
Fs_elec = 10000; % Sampling frequency of electrical signals in Hz.


folder = 'data_files';
if~exist(folder, 'dir')
mkdir(folder);
end


% Iterate over the number of broken rotor bars.
fori = 1:numel(health)
fprintf('Processing data file %s\n', files(i))


% Load the original data set stored as a struct.
  S = load(files(i));
fields = fieldnames(S);
dataset = S.(fields{1});


loadLevels = fieldnames(dataset);
% Iterate over load (torque) levels in each data set.
forj = 1:numel(loadLevels)
experiments = dataset.(loadLevels{j});
data = struct;


% Iterate over the given number of experiments for each load level.
fork = 1:numExperiments
signalNames = fieldnames(experiments(k));
% Iterate over the signals in each experimental data set.
forl = 1:numel(signalNames)
```

```matlab
% Experimental (electrical) data
    data.(signalNames{l}) = experiments(k).(signalNames{l});
end


% Operating conditions
data.Health = health(i);
data.Load = string(loadLevels{j});


% Constant parameters
data.Fs_elec = Fs_elec;


% Save memberwise data.
name = sprintf('rotor%db_%s_experiment%02d', i-1, loadLevels{j}, k);
fprintf('\tCreating the member data file %s.mat\n', name)
filename = fullfile(pwd, folder, name);
save(filename, '-v7.3', '-struct', 'data'); % Save fields as individual variables.
end
end
end


location = fullfile(pwd, folder);
ens = fileEnsembleDatastore(location,'.mat');


ens.ReadFcn = @readData;
ens.WriteToMemberFcn = @writeData;


ens.DataVariables = [ "Ia"];
ens.ConditionVariables = ["Health"; "Load"];
ens.SelectedVariables = ["Ia"; "Health"; "Load"];


ens
```

```
T = readall(ens)

cm = confusionmat(actual, predicted);
cmt = cm'

diagonal = diag(cmt)
sum_of_rows = sum(cmt, 2)

precision = diagonal ./ sum_of_rows
overall_precision = mean(precision)

sum_of_columns = sum(cmt, 1)

recall = diagonal ./ sum_of_columns'
overall_recall = mean(recall)

f1_score = 2*((overall_precision*overall_recall)/(overall_precision+overall_recall))
```

Figures:

Figure: Broken Rotor Bar