

1-1-2010

An Application Of A Hybrid Wavelet-SVM Based Approach For Induction Motor Fault Detection

Shermineh Ghasemi
Ryerson University

Follow this and additional works at: <http://digitalcommons.ryerson.ca/dissertations>

 Part of the [Electrical and Computer Engineering Commons](#)

Recommended Citation

Ghasemi, Shermineh, "An Application Of A Hybrid Wavelet-SVM Based Approach For Induction Motor Fault Detection" (2010).
Theses and dissertations. Paper 1363.

This Thesis is brought to you for free and open access by Digital Commons @ Ryerson. It has been accepted for inclusion in Theses and dissertations by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.

**AN APPLICATION OF A HYBRID WAVELET-SVM
BASED APPROACH FOR INDUCTION MOTOR FAULT
DETECTION**

by

Shermineh Ghasemi

Bachelor of Electrical and Computer Engineering, Azad Tehran University ,2000

A thesis

Presented to Ryerson University

in partial fulfillments of the degree of

Master of Science

In the program of

Electrical and Computer Engineering

Toronto, Ontario, Canada, 2010

© S. Ghasemi 2010

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Name:.....

Date:.....

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, at the request of other institutions or individuals for the purpose of scholarly research.

Name:.....

Date:.....

BOWRROWER'S PAGE

Ryerson University requires the signatures of all persons using or photocopying this thesis.

Please sign blow, and give address and date.

APPLICATION OF HYBRID WAVELET-SVM FOR INDUCTION MOTOR FAULT IDENTIFICATION

Shermineh Ghasemi

Master of Science

Department of Electrical and Computer Engineering

Ryerson University, Toronto, Ontario, Canada, 2010

ABSTRACT

Induction motors have been widely used in the industries due to their simple and rugged construction. Failures of this electrical machinery may cause considerable losses. Therefore, adapting an efficient method to diagnose a fault at a very early stage would prevent any further consequences of this deficiency. The major concern is related to the mechanical failures, normally caused by the inner component deficiencies. Application of intelligent methods have attracted interest in recent years. Support Vector Machine is a supervised learning method, based on the statistical learning theory. This thesis presents three different SVM algorithms: SVM, KPCA-SVM and ROC-SVM, applicable for broken rotor bars detection. SVM proved to be reliable method for classification. While application of KPCA-SVM, shows nonlinear feature extraction can improve the performance of classifier with respect to reduce the number of overlapping samples. Furthermore, ROC-SVM has improved the accuracy by selecting a decision threshold for the classifier.

ACKNOWLEDGEMENTS

I would like to thank my supervisor Dr. Alireza Sadeghian for his encouragement, unlimited support and providing a pleasant working environment in the department. I'm also grateful to Dr. Maryam Davoudpoor for her guidance through the work and moral support. Also I would like to express my appreciation to my committee members.

Furthermore, I want to thank Dr. Bin Wu for providing us the required hardware to collect the data which has been used in this work.

In addition I want to thank my dear friend Shawn Maharaj for helping me editing the thesis and for his unexpected visits.

Finally, I want to dedicate this thesis to the dearest persons in my life, my mother and my brother for their inspirable support and unconditional love. My lovely father for encouraging me to follow my education and pursuing this degree was a part of his dream which I tried to fulfill.

Shermineh

September 2010-Ryerson University

TABLE OF CONTENTS

TITLE PAGE.....	i
AUTHOR’S DECLARATION.....	ii
BORROWER’S PAGE.....	iii
ABSTRACT.....	iv
ACKNOWLEDGMENTS.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABELS.....	ix
LIST OF FIGURES.....	x
NOMENCLATURE.....	xi

CHAPER 1: INTRODUCTION

1.1 AC induction motors.....	1
1.2 Broken rotor bars identification issues	2
1.3 Existing approaches	3
1.4 Support Vector Machines	5
1.5 Proposed approach	6
1.6 Methodology.....	7
1.7 Thesis outline.....	8

CHAPER 2: BACKGROUND

2.1 Structure and principal of induction motor’s operation	10
2.2 Induction motor faults.....	11
2.3 Condition monitoring of the induction motor.....	12
2.3.1 Motor current signature analysis (MCSA).....	13
2.4 Artificial intelligent methods for induction motor fault detection	16
2.4.1 Analytical-based models	16

2.4.2 Knowledge-based models	18
2.4.3 Data-based models	20

CHAPTER 3: METHODOLOGY

3.1 Data acquisition method	30
3.2 Feature-extraction method	33
3.2.1 Wavelet Packet Decomposition mathematical analysis.....	35
3.2.2 WPD Feature Coefficients.....	37
3.2.3 Load torque.....	39
3.3 Support Vector Machine	40
3.3.1 Statistical learning theory.....	40
3.3.2 Δ -Margin separating hyperplane.....	43
3.3.3 Linear classification.....	44
3.3.4 Non-linear classification.....	48
3.3.5 Tuning the meta-parameter of the support vector machine.....	51
3.3.6 KPCA-SVM.....	52
3.3.7 ROC-SVM.....	55

CHAPTER 4 :EXPRIMENTAL VERIFICATION

4.1. Experiment's Layout.....	60
4.2 Calculating the wavelet coefficients of monitored signal.....	63
4.3 Forming a Feature Vectors of the classifier.....	67
4.4 Application of a binary class -support vector machine.....	68
4.5 Fault identification based on KPCA-SVM	75
4.6 Using ROC-SVM to identify broken bars	79

CHAPTER 5: CONCLUSION AND FUTUREWORK

4.1 Conclusion.....82

4.2 Future work.....85

REFERENCES.....87

LIST OF TABELS

4.1 Mean values of the feature coefficients for faulty induction motor with two broken bars and healthy motor with no broken bar.....	70
4.2 Performance of the SVM under various loading conditions.....	73
4.3 Performance of the standard SVM for induction motor fault detection.....	74
4.4 Performance of the KPCA-SVM for induction motor broken bars diagnosis.....	77
4.5 Performance comparison of the KPCA-SVM and standard SVM.....	78
4.6 Comparing the accuracy rate of the ROC-SVM, SVM and KPCA-SVM in identifying broken rotor bars	80

LIST OF FIGURES

3.1 The fault detection proposed algorithm.....	30
3.2 Comparison of the Stator current waveform for faulty induction motor with two broken rotor and a healthy motor with no broken	32
3.3 Wavelet packet filter bank decomposition.....	37
3.4 Δ -Margin separating hyperplane.....	43
3.5 Linear-separating hyperplanes for the separable case.....	46
3.6 Non-linear classification by mapping data into a feature space.....	51
3.7 The ROC-SVM curve.....	55
4.1 Induction motor-generator Setup bench.....	60
4.2 Faulty induction motor with a fabricated hole on the rotor bar.....	61
4.3 Stator current waveform for faulty induction motor with two broken rotor bars at depth 10 under full load and 50% of full load.....	64
4.4 Feature coefficients at Depth 10 Nodes 60, 61, 62 and 63 for Faulty and healthy motors.....	65,66
4.5 The feature coefficients at node 62 of depth 10 for both healthy and faulty motors under a heavy loading	67
4.6 KPCA Features distribution.....	76

NOMENCLATURE

Mathematical notations

w_s	the synchronous speed of the revolving field
w_r	the slip speed
w_m	the rotor speed
P	the number of poles
f	the frequency of the current in the stator
s	the slip speed of the motor
f_{brb}	the frequency of the broken rotor bar
f_s	the excitation current frequency
$c_{a,b}$	the wavelet coefficients
$\psi_{a,b}$	the mother wavelet
h	the low-pass filter
g	the high-pass filter
n_w	the nyquisite frequency
f_m	the rotational speed frequency
$p(x)$	the unconditional probability for occurrence of x
$f(x)$	the learner function
C	the error penalty for SVM training
$R[f]$	the structural risk
$R_{emp}[f]$	the empirical risk
VC	the Vapnik Chervonenkis dimension
w_i	the weight vector

b	the bias value
y_i	the label for a sample vector x_i
x^+	the sample belonging to positive classes
x^-	the sample belonging to negative classes
x^0	the sample on a separating hyperplane
γ	the geometrical margin between positive and negative class
L	the Lagrange function
α	the Lagrange multipliers for all samples
$\#SV$	the number of support vectors
ξ_i	the slack variable introduced to the i th inequality condition $\xi_i > 0$
$\varphi(x)$	the nonlinear vector function that maps the input x into the feature space
$K(x_i, x_j)$	the Kernel function that calculates inner products of feature space in original data space
C_y	the covariance matrix
λ	the eigenvalue
V	the eigenvector
ρ	the regularization parameter
H	the Hilbert space

CHAPTER 1: INTRODUCTION

1.1 AC induction motors

Since the introduction of induction motors by Galileo Ferraris in 1885 and further improvement made by Nikola Tesla in 1888, these motors have been widely applied in the industry [1]. These motors are probably the simplest and most rugged of all electrical motors and often provide core capabilities essential to many industrial operations. Therefore monitoring induction motors to insure a high degree of reliability has always been an important issue. Induction motors are either single phase or poly-phase. However, the three-phase motors are the most common in the industry [1,2].

An AC induction motor consists of two main parts: stationary part or stator, and rotating component or rotor. The stator winding is connected to the AC source and the rotor is either short-circuited or closed through external resistances. The induction motors, based on their rotor type, are known either as squirrel cage or wounded rotor. Noticeably, an induction motor requires no direct connections to the rotating parts. Since, the energy transformation from the stationary part to the rotating part is carried out by the mean of the electromagnetic induction [1]. This electromagnetically transformation appears across the air gap between the stator and the rotor. Thus, the air gap should be quite small to offer minimum reluctance [2].

Fault occurrence in induction motors might be due to several distinguishable reasons such as internal failing, which inherent to the machine itself, or external faults, generated by the operating conditions [3,4]. Regardless of the nature of the fault, any fault in induction motor demands an emergency maintenance to avoid excessive down times and the lost of profits. Therefore, adapting an efficient method to diagnose a fault at a very early stage is highly desired.

The major faults of an induction motor are categorized as stator faults, broken rotor bars or cracked rotor end-rings, air-gap eccentricity, bent shaft and bearing failures.

The mechanical failures are the major concerns of the industry. They are normally caused by the inner component deficiencies such as winding faults, unbalanced stator, broken rotor bars, air-gap eccentricity or bearing faults [3,4].

Traditionally, the fault detection task is performed by a human expert on fixed time scheduled maintenance or by a reaction to the plant failure as the fault happens. Although the traditional methods are still the dominant methods in identifying a fault but the manufacturers and users of these drives are now showing more interest toward using intelligent diagnostic methods in order to insure the machine's stability and reliability. In the last decade the advancement in pattern recognition has delivered more effective methods to detect the faults by applying soft computing methods such as neural networks, expert systems and support vector machines to observe the machine's behavior constantly and react to the abnormal conditions. In this thesis, we mainly focus on broken rotor bar (BR) detection, which is responsible for almost 5% of faults in the squirrel cage induction motors [4]. The required information is obtained by monitoring the stator current under specific conditions.

1.2 Broken rotor bars detection issues

In spite of all the progress made to improve the fault detection methodology, the unexpected system failures are still inevitable.

Generally to identify a broken rotor bar, any distorted harmonic in the stator current should be identified as a sign of the failure. After monitoring the current signal with one of the prevalent techniques, the fault related measurements such as the amplitude of the side band harmonics,

should be extracted. Respectively, acquired measurements should be transformed into the time-domain, frequency-domain or time-frequency domain, to be useful [5,6,7]. Fast Fourier Transformers (FFT) traditionally applied for this purpose. However; the non-stationary nature of the current signal makes the FFT spectrum analysis unreliable. In fact, the FFT can reach the desirable accuracy as long as the signal is stationary. Besides, the machine should be sufficiently loaded to be able to separate the sideband frequencies from the fundamental frequencies of the motor. However, piling up a machine with full load, specifically in a stationary condition would reduce its lifetime. As an alternative in some cases, the FFT could be replaced by Short Time Fourier Transformers (STFT). The STFT analyzes the transient signals using a time-frequency representation by applying a fixed-sized window for all frequencies, which leads to frequency resolution. However, the deficiency still exists due to its inability to provide multiple frequencies and temporal resolution [8,9].

Furthermore, the environmental effects on the monitored signal should also be considered due to their noisy contribution to the harmonic spectrum, which generates more inaccuracy in detecting the abnormal harmonics.

1.3 Existing approaches

In recent years various Artificial Intelligence techniques such as expert systems, artificial neural networks (ANN), fuzzy logic, fuzzy neural networks and genetic algorithms, have been adopted to identify the induction motor faults; specifically the broken rotor bars.

In many research works, neural networks (NN) have been utilized mostly due to their general nonlinear function approximation and their adaptive learning abilities. Moreover, neural networks does not require a prior knowledge about the diagnostic model .The NNs are capable to

produce the best results if appropriate network structures are selected and proper weights can be assigned.

However, the neural networks have some drawbacks due to their numerically oriented black-box nature. This phenomenon results in the impossibility of utilizing the qualitative and linguistic information directly from the operator of the machine. Furthermore, the input and output mapping of a trained neural network cannot be interpreted into the meaningful fault diagnosis rules. Also, since the exact architecture of a NN is not known in advance and normally obtained by a trial-and-error procedure, selecting an inappropriate structure might lead to over and under fitting problems and slow convergence speed, affecting the overall performance accordingly [10].

Fuzzy logic systems are expert rule-based systems, but they also could be considered as a type of nonlinear function approximations. It should be pointed out that Fuzzy logic lacks an effective learning capability. In contrast to NN, they give a clear physical description of the approximated function by defining the appropriate rules.

In spite of the fact that Fuzzy logic-based methods have the advantages of linguistic knowledge in contrast to NN, but the rule definition depends on the intuitive experience acquired from an expert operator. Hence, the fuzzy membership functions and fuzzy rules cannot be guaranteed to be optimal in any sense [11].

To overcome the disadvantages of neural networks and fuzzy logic systems, a hybrid arrangement is usually used. In this approach, the neural network would only deal with numerical approximation while fuzzy logic systems define the rules and the linguistic model.

Even though, this hybrid neural-fuzzy system has been employed successfully in many industrial applications, there exist some cases such as an ANFIS system [12].

Genetic algorithm (GA) is a stochastic optimization method, capable of parallel search and global optimization. GA can be combined by NN or Fuzzy logic methods in order to optimize their parameters. Adding GA to any classification method for motor fault diagnosis could improve their overall performance. This is a direct result of optimization with GA often involving with heavy computation, which prolongs the CPU timing [13].

1.4 Support Vector Machines

Support vector machine (SVM) is a supervised learning method, which can be applied for both classification and regression. The implementation of SVM and kernel methods in various fields such as bioinformatics and data mining has produced significant results. Support vector machines are based on the statistical learning theory introduced by Vapnik [14]. According to which, the statistical learning theory describes the properties of learning machines, which allow them to produce reliable predictions on the problems.

SVM models, to some extent, are similar to multilayer perceptron neural network. Since, they both deal with solving a quadratic programming problem with linear constraints.

In short, the key features of SVM are: independency of dimensionality due to its kernels based nature, the absence of local minima, the sparseness of the solution and the capacity control obtained by optimizing the margin, the flexibility, scalability, and speed [16].

In Binary SVM classification, the classifier discriminates data points of two categories by applying a separating hyperplane. In order to achieve maximum separation between the two

classes, SVM picks the hyperplane, which introduces the largest margin, called an optimal hyperplane. The margin is in fact the summation of the shortest distance from the separating hyperplane to the nearest data point of both categories. Such a hyperplane is most likely to generalize better by correctly classifying unseen or testing data points.

SVM uses a kernel function to map the data from input space to feature space when there is no exact formulation of a mapping function to overcome the curse of dimensionality. In such a case, finding a solution to the linear classification in the feature space is equivalent to solving a nonlinear classification in the original space, both end up in constructing a maximal separating hyperplane [15,16].

1.5 Proposed approach

This thesis proposes a method to detect broken bars in AC induction motors by applying an intelligent statistical method called support vector machine. In this approach a wavelet based support vector machine has been used to detect the abnormal harmonics related to broken rotor bars. The support vector machine has been selected as a classifier to separate the faulty samples from the healthy ones. In the next step, to optimize the performance and accuracy rate of the classifier, a kernel based Principal Component Analysis (KPCA) has been combined with the SVM to introduce a new algorithm called KPCA-SVM [16]. By applying KPCA, which normally leads to the input dimensionality decrement, we expect an increment in the convergence speed and accuracy rate of the standard SVM. At the end, one more step is taken toward optimizing the accuracy rate of the SVM, by utilizing a new support vector machine algorithm called (Receiver Operation Curve) ROC-SVM. To our best knowledge, this algorithm has never been applied before for industrial applications.

1.6 Methodology

In this thesis we have followed Sadeghian et al. [17] work to some extent. In their proposed algorithm, wavelet packet decomposition (WPD) and neural networks have been applied for online detection of rotor bar breakage in induction motors. Respectively in our work, the WPD would be applied to the monitored signal to extract features, representing the multiple frequency resolutions for faulty modes in order to accurately differentiate between healthy and faulty machines. In the original work, the extracted features with different frequency resolutions, together with the slip speed, formed the input vector to the classifier.

However, we will make some modification to the proposed algorithm [17] by replacing the slip speed with the load torque. Knowing that whenever the load oscillations are getting damped due to a high inertia, an adequate detection can be accomplished by monitoring fault frequencies at sufficiently high harmonics [18]. This criterion would be used to form the input vector in our approach.

Furthermore, a statistical classification technique named support vector machines would substitute the neural network. Different aspect of SVM classification would be discussed from generated results.

Also, the feasibility of using the nonlinear feature extraction method known as KPCA in co-operation with SVM leading to the (KPCA-SVM) algorithm would be studied in this research. The main objective of implementing this algorithm is to study the effects of dimensionality reduction on the overall performance of SVM in this specific case study.

Finally, to obtain an optimal solution of the support vector machine for broken rotor bars detection, application of another SVM algorithm called ROC-SVM also has been experimented to study the capability of this method in increasing the accuracy rate of the applied SVM.

1.7 Thesis outline

Section 2 provides the required background knowledge for the rest of the thesis. Section 2.1 briefly describes the mechanical structure of an AC induction motor in particular also; the principal of the induction motor's operation has been discussed. Section 2.2 provide a short introduction to the various faulty condition an AC induction motor may encounter while operating. Section 2.3 discusses a condition monitoring techniques available for extracting the fault related features for diagnostic purposes while subsection 2.3.1 mainly discuss the Motor Current Signature Analysis (MCSA) technique [17]. Respectively section 2.4 gives some insight into the various artificial intelligent methods, which have been used for industrial applications, and specifically for induction motors fault detection. The available methods have been divided into three major groups as analytical-based models, knowledge-based models, and data-based models which have been presented in sections 2.4.1, 2.4.2 and 2.4.3 accordingly.

Chapter 3 presents the proposed methodology for the intelligent fault identification of an induction motor according to which the diagnostic process consists of four steps: data acquisition 3.1, preprocessing or feature extraction 3.2, fault detection using support vector machine and post processing 3.3.

The experimental procedure of this work has been presented in chapter 4, starting with experimental layout in 4.1. continuing with sections 4.2 and 4.3 which introduces the feature selection methodology, approached in this thesis. Section 4.4, 4.5 and 4.6 are representing three

different algorithms, named standard SVM, KPCA-SVM, and ROC-SVM respectively, which have been employed in this thesis for broken bars identification.

In chapter 5 the main results of our work and the future objectives have been summarized.

CHAPTER 2: BACKGROUND

In this chapter the performance of the induction motor has been studied 2.1. In section 2.2, the major fault of the induction motor has been reviewed. Also the most common monitoring methods for detecting the induction motor faults have been discussed in 2.3, while the artificial fault detection methods have been presented in 2.4. Including analytical-based models 2.4.1, knowledge-based model 2.4.1 and data-based model 2.4.2 such as neural networks and support vector machines

2.1 Structure and principal of induction motor operation

An AC induction motor is a rotating electrical machine designed to operate by a three-phase source of alternating voltages .The induction motor could be single phase or poly-phase. However, the poly-phase induction motor is more common for industrial application. Considering the mechanical structure, the poly-phase induction motor consists of five essential parts: laminated stator core, laminated rotor core, stiff shaft, frame and a cooling system. The stator core has been made up of laminations about 3-6 mm thick, which are insulated from each other. The rotor also has the same build [1,2].

Generally the induction motor is categorized into two groups: wounded rotor and squirrel-cage motor. The AC motors used in domestic and light industries are mainly squirrel-cage motors [1].

Normally, when the stator winding gets excited from a balanced ploy-phase source, produces a magnetic filed in the air gap which rotates at a synchronous speed determined by the number of poles and applied frequency. The speed, at which the AC motors magnetic field rotates, is the synchronous speed of the AC motor. The frequency of the current induced in the stator has been

shown by f and P shows the number of the poles [1]. The synchronous speed of the revolving magnetic field is:

$$w_s = \frac{4\pi f}{P} = \frac{120f}{P} \quad (2.1)$$

This rotating field produces the Electro Motive Force (EMF) in the motor and the rotor windings form closed loops as a result of this induced EMF in each coil, which arise the induced current in that coil [1]. When a current-carrying coil is in the magnetic field, it experiences a force that tends to rotate it. The rotor that receives it, powers by induction only when there is motion between the rotor speed and the revolving field. Since the rotor rotates at the speed slower than the revolving field this kind of machine is called an asynchronous machine [1].

The difference between the speed of the rotating flux and the speed of the rotor is called slip speed; and the ratio of the slip speed to synchronous speed is known as slip, which is expressed as:

$$s = \frac{w_s - w_m}{w_s} \quad (2.2)$$

w_s is the synchronous speed and w_m is the actual speed of the motor.

2.2 Induction motor faults

The rotating machinery is a subject to various deteriorating stresses such as thermal, electrical, mechanical and environmental [3]. Consequently, the fault existence in the AC machines could be a result of the internal deficiency, which inherent to the machine itself such as mechanical failure or it can be external such as insulation failing as a result of the bad environmental

condition. According to the statistics, the pure insulation system failures are responsible for less than 5% of all electrical machine failures [4]. The insulation failures are mainly caused by bad environmental conditions such as moisture or over-temperature. However, broken rotor bars and cracked end rings on cage machines could be a participant to insulation fault as well. There also exist various mechanical faults in induction motors such as rotor rubbing, stator and rotor fatigue, bent shaft as a consequence of the physical damage like misalignment or eccentricity [4].

Due to the presence of these failures in three main components of the induction motors stator, rotor and bearings, the internal faults are mainly categorized by referring them to their location such as stator or rotor faults.

The stator faults are mostly the result of the degraded insulation in stator windings, which leads to inter-turn, phase to phase or phase to ground short circuit. The rotor faults could be divided into rotor eccentricity related faults, and physical damage of the rotor such as the broken rotor bar and broken rotor ring [3].

The existence of the mechanical faults of induction machines could be linked with presence of the particular harmonics in the current or vibration spectrum of the stator. Literally, each fault is associated with the presence of a specific harmonic in the spectrum [5].

2.3 Condition monitoring of the induction motor

Due to the vulnerability of the rotating machinery to the internal faults resulting in considerable economical and performance loss, intelligent fault identification methodology has received a considerable attention in recent years. This is considering the importance of an early fault diagnosis, which can increase performance, reduce consequential damage, and prolong machine's lifetime [19]. Different variables could be measured as a sign of a fault occurrence

such as magnetic flux due to the stator deficiency as a result of the abnormality in the air gap flux or stator vibration produced by supply voltage unbalanced, air gap torque to identify flux leakage [3]. Depending on the measured variable, there exist different monitoring methods such as Vibration analysis, partial discharge measurement, and temperature monitoring or MCSA [20]. It should be pointed out, that all these methods requires the user to have some degree of expertise in order to be capable of distinguishing a healthy operating condition from a potential failure mode. The reason behind this is the co-existence of the samples related to both the design and construction of the motor and the variation in the load torque, which could be misleading in most cases.

2.3.1 Motor Current Signature Analysis (MCSA)

The development of a reliable condition monitoring technique has always been a controversial issue in the fault identification procedure. Although vibration monitoring is a prevalent technique for induction motor condition monitoring, in most researches, the priority has been given to the current signal monitoring. Since the stator current spectrum has been proved to be a reliable source to provide the fault related information due to its sensitivity to almost all of the motor faults. Furthermore, it is not necessary to have a direct access to the motor due to non-invasive nature of the current monitoring technique.

In practice, an advanced signal processing method should be accompanied by the monitoring technique to carry out the monitored signal into time-domain, frequency-domain or time-frequency domain, where the significant features of the signal could be identified.

In time-domain analysis the statistical features are normally calculated from time waveform signals such as mean, peak, standard deviation skewness or kurtosis [21].

The Frequency-domain analysis, on the other hand is more explicit to identify the specific fault related side band frequencies. Although Fast Fourier Transform (FFT) analysis is traditionally used in this stage but transforming a signal from time-domain to the frequency-domain leads to information loss. This problem is considered to be one of the main disadvantages of the frequency-domain transition [21].

Finally, the time-frequency domain analysis represents the power spectrum of waveform signals in two-dimensional functions of both time and frequency to better reveal fault patterns. Wavelet transforms are the dominant feature extractor in this field [21].

When a fault occurs in stator, rotor or bearing, it generates a specific harmonic in the stator current. For instance, the rotor asymmetry on the current spectrum is detected by the existence of the following side band harmonics [20]:

$$f_{brb} = f[k(\frac{1-s}{P}) \pm s] \quad (2.3)$$

$k=1,2,3,\dots$, s is per unit slip speed, P is number of the pole pairs and f stands for electrical supply frequency.

Moreover, the amplitude of the left sideband frequency component is proportional to the number of broken rotor bars. In case, the number of the broken bars needs to be calculated. Therefore the amplitude I_{brb} of frequency component $f_s(1-2s)$ evaluated from [20]:

$$\frac{I_{brb}}{I_s} \cong \frac{\sin \alpha}{2P(2\pi - \alpha)} \quad (2.4)$$

In most researchers, the priority has been given to MCSA over other available monitoring techniques.

Generally, intelligent fault identification in rotating machinery using an MCSA analysis compromise of three major steps: data acquisition, feature extraction and classification. The data acquisition has been discussed thoroughly however the feature extraction can be carried out either in frequency, time, or time-frequency domain.

The strength of FFT, which is traditionally used in frequency-domain, lies beneath its ability to analyze a signal in the time-domain by its frequency content. FFT is the main tool for signal spectral decomposition, which can describe a stationary signal profoundly. However when it comes to non-stationary signals, it starts getting sensitive to the presence of the short periodic signals. Thus, the drawbacks of applying FFT are loss of information because of the transition from time-domain to frequency-domain plus its incapability of time localization of the signal.

The FFT could be replaced by The Short-Time Fourier Transform (STFT) to solve the time localization problem. In fact, STFT is a Fourier- related transform used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time by using a sliding window. Deficiency of using the STFT is related to selecting the size of the window. By applying a wide window, a good frequency resolution and a bad time resolution would be resulted on the other hand by reducing the window size, the frequency resolution decreases while the time resolution increases. Consequently, by utilizing STFT it is impossible to acquire the time-frequency information simultaneously. Thus, STFS could be replaced by a method of time–frequency localization that is scale independent, such as wavelet analysis. The wavelet transform can be used to analyze time series that contain non-stationary power at many

different frequencies [20,22,23].

Similar to a Fourier analysis which breaks up the original signal into sine waves of various frequencies, the wavelet analysis is the breaking up of a signal into shifted and scaled versions of the original which is called mother wavelet. However, the priority of the wavelet over FFT is due to its capability of space localization. Furthermore, it can achieve a good frequency and time resolution at the same time by obtaining high frequency resolution at lower frequencies and high time resolution from the higher frequencies [24].

2.4 Artificial intelligent methods for induction motor fault detection

In contrast to hard computing methods, which tend to achieve precision and certainty in solving real-world problems, soft computing demonstrates the ability of the human mind to deal with uncertainty, imprecision, and prediction. The lack of precise data and a human expert in real-world problems, increase the demand for replacing the traditional methods by more intelligent techniques. Knowing that, the intelligent methods are capable of providing us with the unique features of adaptation, flexibility, and embedded linguistic knowledge over conventional methods. In this chapter the analytical methods (2.4.1) have been compared to knowledge based models (2.4.2) and data-based models (2.4.3).

2.4.1 Analytical-based models

Generally, a diagnostic scheme in analytical-model is based on comparing the actual system behavior against the system model. Thus, existence of any inconsistency in the measured data (input /output) is a sign of fault occurrence, which could be used for fault classification. However generating the diagnostic residuals requires a complex mathematical model. Besides, the analytical approaches only can handle linear models, means the system need linear

approximation of a nonlinear model to be able to deal with it. Consequently, the efficiency of this method expected to be low. Furthermore, if a fault were not specifically modeled in the system, it wouldn't be detected [25].

A. Knight et al. In [26] presents a method of identifying a mechanical fault by analyzing a combination of permeance and magneto-motive force (MMF) harmonics. These harmonics are appearing as a result of irregularities in an induction motor. In general, the statistic and dynamic eccentricity can be detected by monitoring the magnitudes of a specific current harmonic such as $f_n + f_r$. However, the presence of both static and dynamic eccentricity would increases in the magnitude simultaneously. Hence, differentiating the sideband harmonics would be difficult. To overcome this problem, in this approach, the side band harmonics has been reformulated as $f_n \pm (N+1)f_r$ to be distinguishable. Where f_n is the stator current n^{th} harmonic and f_r is the rotor mechanical speed. Accordingly, the first part shows the harmonics affected by static eccentricity and the second one indicates the harmonics caused by dynamic eccentricity. The test has been run under different loading situations. The results have proved the efficiency of the applied method.

Combastel et al. [27] used a Park vector to model electrical variables of an induction motor in order to inspect specific system faults such as stator winding failures or broken rotor bars. Two different methods have been applied: the first one uses the analytical redundancy method and the second one uses the wavelet transform. In the first test, residual gradient inconsistency has been used to discriminate between normal and abnormal motor conditions. It has been observed that these residuals are more sensitive to stator short-circuits than to rotor broken bars .In the second examination, the wavelet transform is applied directly to the residuals not to the raw stator

currents. The acquired wavelet coefficients have been used to identify the healthy from faulty machine. Both methods tested on different machines with different types of faults such as stator short-circuits and broken rotor bars. Consequently, their capacity to detect the faults affecting the induction machine in realistic situation has been verified.

2.4.2 Knowledge-based models

A knowledge-based model like Expert system is capable to apply a human-like knowledge to diagnose the fault. The main advantages in the development of expert systems for diagnostic problems are: ease of development, transparent reasoning, the ability of reasoning and explaining. However, the role of the expert system in the fault identification process is only limited to identifying the specific harmonic components and filtering the large amount of spectral information to provide more informative data. In fact, the expert system eliminates unrelated components of the data spectrum according to the machine history of the various faulty conditions. Besides, the expert system can also be a threshold handler to decide about considering or ignoring a particular failure component. The threshold should be computed by using simplified models or formulas in advance.

Liu et al. [28] presents a fuzzy expert system, which is PC-based, menu driven and user-friendly, defined by specific knowledge of various aspects about bearing monitoring such as diagnostic methods, defect frequencies, feature selection and fuzzy bearing classification. This expert system shell consists of seven modules as of:

- 1) Short-Cut-References,
- 2) How-To-Use-Bearing-Expert,

3) General–Information,

4) Diagnose–Methods,

5) Defect–Frequency,

6) Feature–Selection,

7) Fuzzy–Bearing–Class

The system requires the user's feedback to select one of available modules; when the engine receives the user's selection, it would start searching the knowledge base to find the match and fire the rule. If the conditions haven't been met, the system would ask for more specification. If the fuzzy logic module is being chosen, the fuzzy logic would be used to reach conclusions and the process would be ended. By applying this method, 100% reliability in finding a bearing fault has been attained.

Schoen et al. in [22] presents a fault identification algorithm, which utilizes a combination of a rule-based expert system, and a neural network to identify a fault in an induction machine. In fact the proposed methodology provides a rule-based model to determine the importance level of the spectral components in condition monitoring. The neural network then would detect the changes in the frequency components, which have been designated by the expert system.

This algorithm contains the five processing sections. First, the sampler and preprocessor convert the current signal from time-domain into a frequency-domain spectrum. The current signature monition analysis (MSCA) has been applied to the input signal. Then the monitored signal has been transformed into the frequency–domain by applying the FFT. Second, a rule-based (expert system) frequency filter provide the system with four levels of frequency to determine which

level contains the fault related information and should be monitored by the neural network. Finally, the neural network and postprocessor determine the sufficiency of abnormality in the current spectrum to indicate a possible fault condition in the monitored machine. This method proved to be successful in identifying the faults.

2.4.3 Data-based models

In contrast to the model-based approaches, which require a priori expert knowledge to make a proper model. A data based model is only dependent on the availability of large amounts of historical process data [25]. In fact the goal of a data based model is, to categorize the data into pre-determined classes. As one category of data based models, statistical methods use knowledge of a priori class distributions to perform classification [29]. A supervised (artificial neural network) ANN as a data-based model is a function approximation, which can synthesize the relationships between the different input and output variables.

Nejjari et al. in [30] presents a fault identification method based on a combination of the stator current Park's vector with an artificial neural network (ANN). Two diagnostic patterns tested in this work are: classical and decentralized methods. First, stator current Park's vector has been used to form the input data vector of a multilayer feed forward NN (FFNN) with a back propagation-training algorithm. The used NN consisted of 42 neurons in the input layer, a variable neuron number in the hidden layer for accuracy improvement tests and one neuron in the output layer. Two methods discussed here are: the classical approach in which healthy motor condition according corresponds to "0" output. The outputs "0.5" and "1" will correspond to the faulty motor condition. The second part, the decentralize method applied according to which, the healthy motor condition achieved through current sensors corresponds to "0" output.

Consequently, the output “1” will correspond to the condition faulty motor. From results it has been concluded that the discussed NN has been successfully trained on the above data by producing 100% correct prediction.

Sadeghian et al. in [17] proposed a novel algorithm for online detection of broken rotor bars in induction motors by adapting the Wavelet Packet Decomposition (WPD) analysis and neural networks. This method is involved with the stator current signature analysis (MCSA), the application of multilayer perceptron (MLP) networks and WPD. In this four step algorithm including data acquisition, preprocessing, fault detection, and post processing, the WPD coefficients of the monitored signal at certain Depths and Nodes have been calculated to provide the neural networks with the input vector of the significant features. The extracted features with different frequency resolutions in combination with the slip speed have been fed into a MLP network. Finally in the post processing stage, the fault related information obtained by the classifier has been stored, the faults which are in a database to be used as a reference adaptively. This algorithm reached up to 96% accuracy rate.

Ghate et al. In [31] presents a noninvasive NN based fault detection scheme for small and medium sized induction motors. In overall thirteen statistical parameters have been calculated and are used to define the significant features of the monitored current including the Root Mean Square (RMS) of a zero mean signal, the maximum, and minimum values, the skewness coefficient and kurtosis coefficients. Generalized Feed Forward Neural Network (GFFD-NN) and Support Vector Machine-Neural Network (SVM-NN) models are designed and verified for optimal performance in fault identification. Output layer has been defined by four conditions of motor namely Healthy, Inter-turn fault, Eccentricity and Both faults. In second test Principal Component Analysis (PCA) has accompanied the SVM. In fact PCA has been introduced to

reduce the number of statistical input feature or the dimensionality of the features. Consequently, reduced support vector machine (SVM-DR) based classifier works as an elegant multi classifier for fault diagnosis of three-phase induction motors. The average classification accuracy rates on testing as well as cross validation instances are obtained as 99.61% for neural network and 98.72% for SVM.

Zidani et al. in [32] demonstrates a fuzzy approach based on the stator current Concordia patterns (park vector). Here, the stator current Concordia patterns for both healthy and faulty motor formed the input data for the classifier and the membership function has been determined by three fuzzy if-then rules such as: (a) when the error between a healthy pattern and a faulty one is negative and if this error increases, the output will apply a high severity index, or (b) When the error between a healthy pattern and a faulty one is positive and if this error decreases, the output will apply a medium severity index. The fault severity is less significant than in the above situation and (c) both patterns are equal but the error may increase or decrease. The error in this method is defined by the pattern deformation. These results clearly confirm the reliability of this method.

Chabu et al. in [33] presents a fuzzy logic approach to detect a broken rotor bar based on a mathematical model of magnetic flux density. In fact, this mathematic equation of the flux spectrum has been plotted by MATLAB .It has been stated that the area under the positive curve shows the gravity of the fault and the number of the broken bars. Consequently, the fuzzy rules defined accordingly. This method shows to be acceptable.

Nejjari et al. in [34] proposed another a fuzzy logic method for diagnostic purposes .in this method, fuzzy logic rules applied to the stator current amplitudes .In fact, the required Fuzzy

rules and membership functions are constructed by observing the data set. The membership functions have been generated as zero, small, medium, and big. Also, for the measurement related to the stator condition, it is only necessary to know if the stator condition is good, damaged, or seriously damaged. Totally 12 heuristic IF–THEN fuzzy inference rules applied to detect the faults. According to the author an excellent diagnosis has been achieved for this test.

Widdodo et al. in [35,36,37] presents a method for an induction motor fault diagnosis based on a transient signal using Principal Component Analysis (PCA) and SVM. The start-up transient current signal is selected as features source for fault diagnosis. The Motor current signature analysis has been developed to perform condition monitoring. The Preprocessing of transient current signal is performed by applying discrete wavelet transform (WDT). In this approach induction motor fault diagnosis were acquired from faulty motors with broken rotor bars, bowed rotor, faulty bearing, rotor unbalance, eccentricity and phase unbalance and the motors in normal condition were considered as a benchmark to be compared with the faulty one. The related features have been calculated statistically from time waveform, frequency domain. Furthermore, the application of PCA, independent component analysis (ICA) and their kernel (KPCA and KICA) for feature extraction and reduction has been studied. A multiclass Support vector machine has been employed for fault classification due to the excellent generalization performance. In this work, the accuracy rate achieved by applying PCA for feature extraction and reduction, are 100% and 78.57% for training and testing, respectively. However, it has been noted that any increment in the number of principal components could raise the training accuracy rate up to 100%. In the second test, where the classification has been done by using SVM and ICA as a feature extractor, the accuracy rates of 100% and 80.95% in training and testing has

been achieved. Finally, it can be concluded that the application of component analysis and SVM as a multi-class classifier, has the potential to be used for intelligent fault diagnosis.

Pyhonen et al. in [38,39] Applied Different coupling strategies to reconstruct a multi-class Support Vector Machine (SVM) based classifier for fault diagnostics of a cage induction motor. A Powers spectrum density estimate of circulating current signal in parallel branches of the motor has been calculated with Welch's method then SVM has been trained to distinguish a healthy motor from a faulty one. Four different approaches to this problem including: coupling pair wise SVM, majority voting schemes with soft and rough reconstructions, a mixture matrix which linearly combines outputs of previously stated coupling pair wise SVM classifiers have been studied and compared. The measurements acquired from motor in healthy situation, faulty with a broken rotor bar and in a severe fault situation with broken end ring in addition to three broken rotor bars. In a case where the pair wise SVM has been trained for each load situation independently, all coupling techniques resulted in a 100% correct classification rate. By applying mixture matrix coupling pair wise SVM, the fault detection rate decreased to 88.6%. Majority voting with rough and soft reconstruction resulted in 91.7% and 80% correct classification rates correspondingly. The weak success of the mixture matrix coupling is understandable since the mixture matrix is built to improve the results only in a no-load situation.

As has been discussed before, the fuzzy logic model is capable of applying an IF-THEN type of expert rules and linguistic variables to capture the qualitative aspects of the problem. However, there are some difficulties in defining the membership function and tuning the rules accordingly. On the contrary, an NN-based technique could find an accurate solution to a particular fault problem, without having an accurate knowledge of the faulty system. However, the exact architecture of the NN which is not clear due to its black box nature. Therefore, by combining

NN and the rules of the fuzzy techniques, it's possible to solve the problem by a neuro-fuzzy (fuzzy-neural) method without any limitation. There is variety of combination for Fuzzy logic and NN such as adaptive-network-based fuzzy inference system (ANFIS), which is an adaptive neural network where input–output function has been defined by a fuzzy inference. Furthermore, using the genetic algorithm (GA) would optimize the parameters and structures of neural networks and fuzzy logic systems: knowing that The GA is mainly a stochastic optimization algorithm not a classifier itself [29].

Ye et al. In [13] proposed a novel online diagnostic algorithm for mechanical faults of electrical machines using multiple Adaptive Neuro-Fuzzy Inference Systems (ANFIS). In this method, Wavelet Packet Decomposition (WPD) has been used to extract features from the stator current over a wide range of speed. The parameters of the network are the mean and standard deviation of the membership functions and the coefficients of the output linear functions, which would be obtained by the ANFIS learning algorithm. This learning algorithm is a hybrid algorithm consisting of the gradient descent and the least-squares estimate. First, A Fuzzy Inference System has been formed by the means and deviations of the feature coefficients of the observe data from a healthy machine. Then the samples from two fault conditions have been inserted to the ANFIS subsystem. Finally, the errors between the expected and real output of the network has been calculated and sent back to tune the parameters in the fuzzification and defuzzification layers. The accuracy of the proposed diagnostics system is about 91.7%. According to author both the number of training iterations and the CPU processing time are significantly in comparison to the traditional neural network applied to the same data under the same condition.

Tran et al. In [40] presents an adaptive neuro-fuzzy inference system (ANFIS) in combination with decision trees Classifier (CART). In fact, the decision tree has been applied to select

relevant features from data set. Then the crisp rules would be converted to fuzzy if-then rules to identify the structure of ANFIS classifier. The membership function of each input parameter was divided into three regions, namely small, medium and large and a hybrid of back-propagation and least squares algorithm is utilized to tune the parameters of the membership functions. Here, the vibration signals and current signals both have been monitored and the most significant features have been calculated from them statistically. The system parameters and the chosen membership functions have been automatically adjusted during the learning process. The convergence of the Root Mean Squared (RMS) error has been utilized to evaluate the learning process. The results indicate that the CART–ANFIS model has potential for fault diagnosis by achieving an overall classification accuracy rate around 91.11% for the vibration signal and 76.67% for current signal.

Altug et al. in [41] demonstrates the application of two popular methods by NN/FZ structures as of Fuzzy Adaptive Learning Control/Decision Network FALCON-Based Fault Detector (FFD), and the Adaptive-Network-Based Fuzzy Inference System ANFIS-Based Fault Detector (AFD). The FFD architecture is a five-layered feed forward network-based fuzzy inference system. The fault detector is a fuzzy inference system implemented on an adaptive network structure. The stator current I rotor speed s and load torque τ of the motor are the inputs of the fault detector, and are used to estimate the motor friction f . Three fuzzy sets are defined on each of the input spaces, corresponding to low, medium, and high for each variable. Fuzzy rules of the motor fault detection/diagnosis problem and membership functions used in these rules are implemented in the hidden-layer nodes. It has the same architecture and membership function as FFD. However, in second approach, the AFD can construct an input–output mapping based on either human knowledge (in the form of Takagi–Sugeno-type if–then rules) or observed input–output data.

Takagi–Sugeno-type if–then rules used in this work have fuzzy antecedent but with a crisp consequence, which is a linear combination of the input values. Both structures can provide good fault detection/diagnosis under varying load torque, while the results of the AFD being slightly more accurate. It has been proposed to make a hybrid of the neural/fuzzy fault detector architecture with the rule extraction power of the FFD and the speed and accuracy of the AFD.

Hong et al. in [42] presents a fuzzy neural network based on-line fault detection method to detect the stator winding turn faults. A B-spline membership fuzzy neural network has been trained with the data acquired from voltage and the positive-sequence current. The target output is the negative-sequence current. The selection of the weighting factors, the knot positions and the control points of the B-spline membership fuzzy-neural networks are crucial to obtaining good approximation for complex nonlinear systems, therefore a genetic algorithm with an efficient search strategy is developed to optimize network parameters. Based on it, experiments have been carried out on a special wound induction motor. The results show that a fuzzy neural network based diagnosis model determines the shorted turns exactly, and is more effective than the parameters estimation method.

Betta et al. in [13] proposed methods to use genetic algorithm (GA) to optimize a neural network-based induction motor fault diagnosis scheme. Actually, two types of GA have been applied for the design and training of the neural network: designer GA and trainer GA. For more clarification, the designer GA optimizes the design parameters of the neural network such as number of hidden layers, as well as characteristics of the neuron transfer function. The connecting weights inside the neural network have been optimized with the trainer GA. The application of these two types of GA in the tuning procedure leads to the optimal neural network

structure and parameters for motor fault diagnosis. The percentage of correct classification rate is higher than 98%.

Gao et al. in [43] discussed a method in which Elman neural network–based motor fault detection algorithm has been applied due to its time series prediction capabilities. A genetic algorithm also has been added to this algorithm to improve the accuracy. Elman neural network is a globally feed forward locally recurrent network model. Which consist of four layers: input, hidden, context and output layers. The Elman neural network has a powerful time series prediction capability because of its memory nodes and local recurrent connections. In this work, it has been applied to give one-step-ahead prediction of the motor feature signals. A GA-based training strategy is further introduced to optimize the initial outputs of the context nodes inside Elman neural network in order to improve its prediction accuracy, and thus achieve better detection performance.

Considering all available intelligent methods and the progress made so far in this field, we decided to choose Support Vector Machine as our classifier due to its limited implementation in for industrial applications considering its significant results in other research areas such as bioinformatics.

CHAPETR 3: METHODOLOGY

In this chapter, the proposed diagnostic, which is a combination of advanced signal processing and pattern recognition techniques, has been presented. The proposed algorithm is a modified version of the algorithm presented by Sadeghian et al. [17] which is originally a four-step process including data acquisition, preprocessing, fault detection, and post processing. In this approach, the current signal is monitored and low-pass filtered to discard the frequency above the Nyquist frequency, which is a bandwidth of a sampled signal equal to half of the sampling frequency before being digitalized by an analog-to-digital converter. In the preprocessing unit, the converted signal is transmitted to time-frequency domain using a Wavelet Packet Decomposition (WPD) analysis. The obtained coefficients from WPD are used to form the classifier's input vector. In the next step the Support Vector Machine is employed to identify the broken rotor bars. Finally, in the post-processing unit, the fault related information is stored in a database as a reference to improve the overall performance of the classifier. The proposed algorithm has been shown in Figure (3.1).

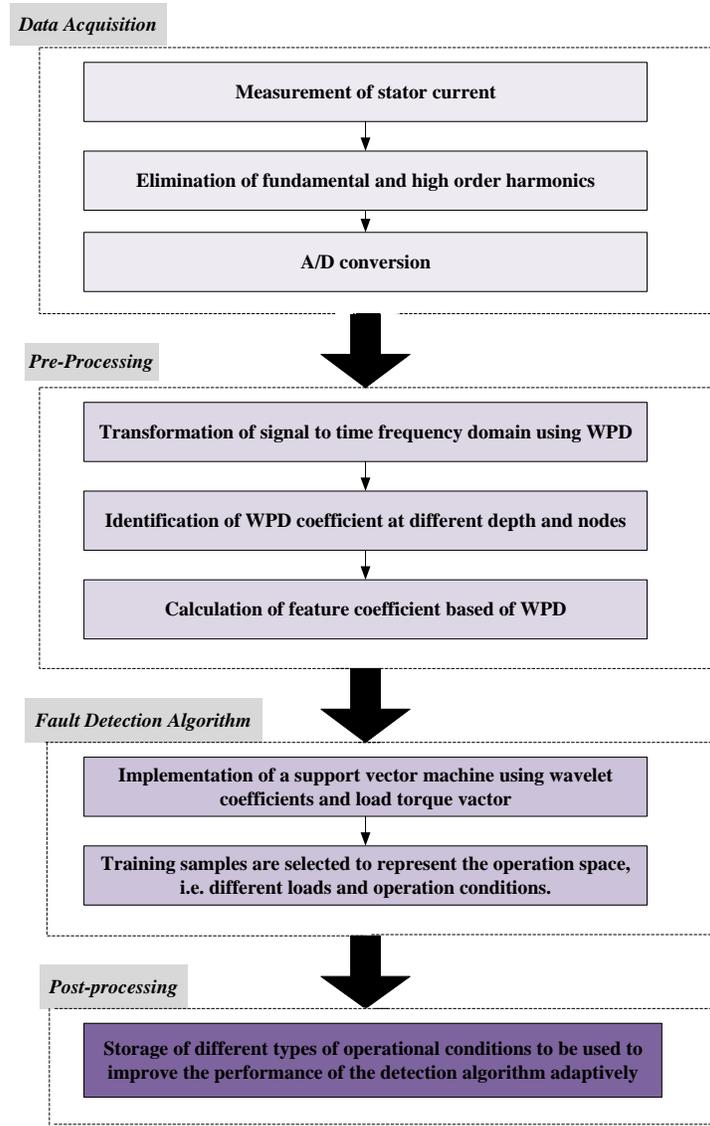


Figure (3.1) the fault detection proposed algorithm

3.1 Data acquisition method

Under the healthy condition, the two-pole induction motor is assumed to be rotating at $(1-s)f_s$, where s is the per unit motor slip and f_s is the excitation current frequency. However, in the presence of fault due to the broken rotor bars, the resulted asymmetry leads to the occurrence of

the positive and negative side bands around the fundamental frequency, which can be formulated as [20,22]:

$$f_b = (1 \pm 2ks)f_s \quad \text{Where } k = 1, 2, 3, \dots \quad (3.1)$$

Consequently, the broken rotor bar can be detected by identifying the fluctuating amplitude around the fundamental frequency, on the right and left sidebands.

The Motor current signature analysis (MCSA) is a well-known monitoring technique, applicable in advanced signal processing to attain the characteristic of the harmonic components in the frequency spectrum from the monitored stator current. The MCSA has a priority over other monitoring methods such as vibration monitoring due to the noninvasive and simplicity of its implementation [20].

Figures (3.2) depict the stator current waveform and spectrum around the fundamental frequency for a three-phase induction motor with two broken rotor bars. Apparently, the low sideband harmonic component around 60 Hz, caused by the rotor bar breakage is comparatively small in a healthy motor while the slip speed is constant. However while studying the stator current the impact of the various loads on the frequency spectrum should be noted

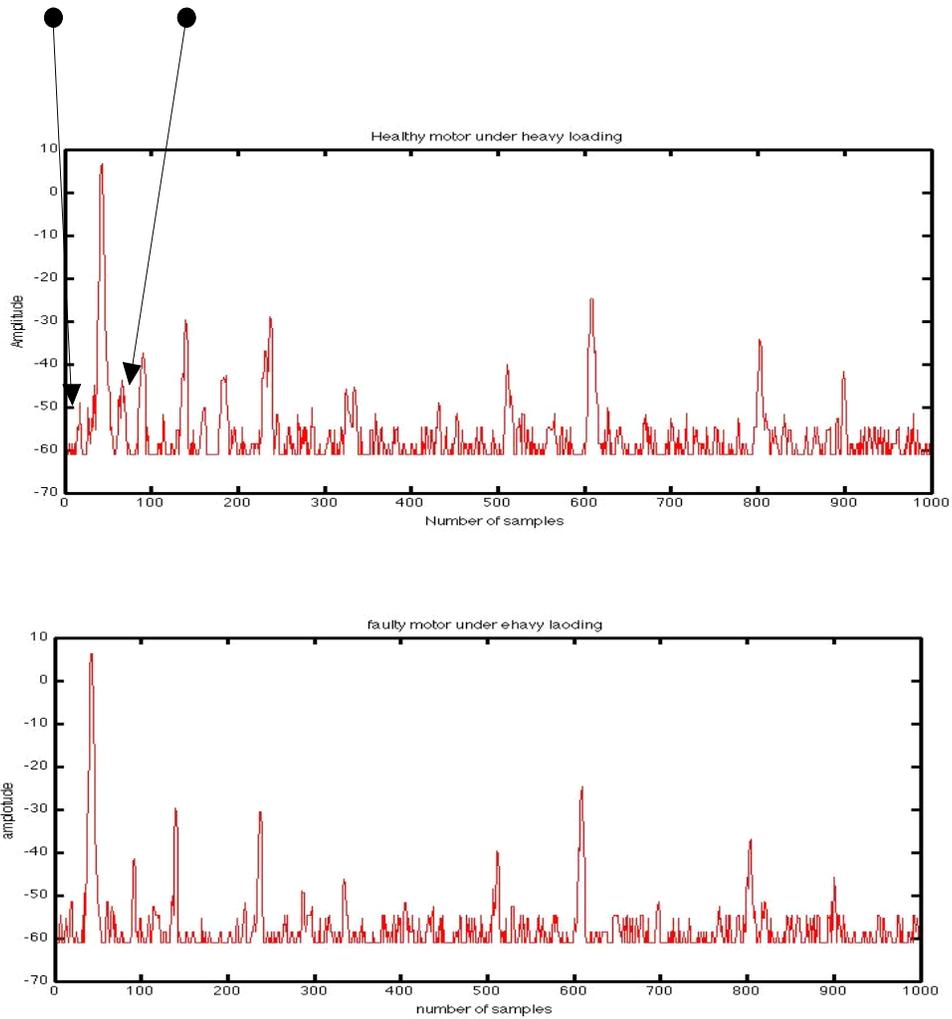


Figure (3.2) comparison of the Stator current waveform for faulty induction motor with two broken rotor and a healthy motor with no broken bar.

Generally, running the machine under steady-state speed is almost impossible in real cases. On the other hand, under non-stationary speed condition, the Fluctuations around fundamental frequency resulted from machine operation in association with broken rotor bars is hardly differentiable. As a result, accurate fault detection under transient conditions cannot be easily accomplished. Therefore a proper method to obtain the required information should be applied to the current signal [44].

3.2 Feature-extraction method

Even though, the current signal has been monitored and oscillation around the fundamental frequency has been detected, extracting the desired features is still complicated. Knowing from pervious discussion, the transient current signal has non-stationary fundamental frequencies as well as non-stationary frequencies associated with the rotor bars, which should be accurately separated. Fast Fourier transformation (FFT) is traditional approach to this problem however this algorithm lacks accuracy due to the side lobe leakage as well as its dependency on the stationary condition of the startup signal.

Therefore in this thesis the wavelet packet decomposition (WPD) analysis has been preferred over the FFT to extract the required features for stator current signal.

Practically, the original signals only would be appropriated for the fault diagnostic purposes only if it changes to a discrete form of a signal sequences by an analog to digital converter. The advantage of the discrete wavelet transformation (DWT) over other methods such as FFT is the capability of extracting more peculiar information from observed signals in practical applications.

The subsample sequences of the discrete signal attained by passing the signal through a digital low-pass filter. The results in a sequence of wavelet coefficients representing the degree of similarity between investigated subsample sequences and the feature wavelet. Recursively performing this process for signals in narrower frequency bands result in producing sets of wavelet coefficients representing the similarities of examined subsamples with the selected feature wavelet in those bands [57]. The ordinary wavelet transformation only analyzes the approximation portion of subsamples of each decomposed level; the detail portion of sub-

samples in those levels may carry some critical information about the motor conditions. Hence, applying the WPD will make it possible to obtain all useful information from the discrete signal sequence by extracting both the detail and approximated coefficients which reveal the necessary and sufficient information for fault diagnosis.

In contrast to WDT, the WPD analysis uses both a low-pass filter and a high-pass filter to decompose a signal string. Normally, the half-frequency of the band used to determine the approximation (from the low-pass filter) and the other half to acquire the detail (from the high-pass filter) coefficients [24].

Since, the two filters are related to each other, they are known as the quadrature mirror filters. It is important to mention the possibility of repeating this decomposition as many times as needed to increase the frequency resolution. Notably, since the current signal is being monitored constantly, the length of the data sets is infinite. However To implement a proper diagnostic method, the length of the data should be limited.

The WPD also breaks the signal down into finite sets, which are the smaller components named packet, and only contains small frequency bands of the original signal. This decomposition simplifies the inspection of the signal bands carrying specific feature related to the machine's healthy or faulty condition. More importantly, the length of the sequence should be sufficient enough to support adequate loops of wavelet analyses to determine wavelet coefficients of the subsample carried in the packet [48].

After the detail and approximation strings being separated by the WPD technique, regular wavelet analysis should be applied to the resulting data sets in order to extract diagnostic wavelet

coefficients. In fact, this multilayer decomposition technique provides a powerful tool for extracting available information in various bands of signal strings to identify the faults.

Moreover, the level of the Wavelet Packet Decomposition should be specified depending on the frequency bands, carrying the signature signals; since the sensitivity of the wavelet relies on choosing the appropriate level of decomposition. In fact, by applying the WPD analysis, the original signal is decomposed into 2^m different signals segments. By the same token, the frequency bandwidth of decomposed signals can also be determined by dividing the sampling frequency by 2^m where m is the decomposition levels [45-49].

In the following sections the assets of feature extraction technique has been defined. Starting with wavelet packet decomposition mathematical analysis (3.2.1), continued by WPD Feature Coefficients (3.2.2) and Load torque in section (3.2.3)

3.2.1 Wavelet packet decomposition mathematical analysis

Mathematically, the wavelet transformer (WT) considered as the inner product of the signal $x(t) \in L^2(R)$ and formulated as [45]:

$$\begin{aligned} c_{a,b} &= \langle x(t), \psi(t) \rangle = |a|^{-1/2} \int_{-\infty}^{+\infty} x(t) \psi(t-b) dt \\ c_{a,b} &= \sum x(t) \psi_{a,b}(t) \end{aligned} \quad (3.2)$$

Where $\psi_{a,b}(t)$ called a mother wavelet function at scale level (a) and location (b). The signal $x(t)$ can be reconstructed from the wavelet coefficients $c_{a,b}$ without error, only if the generating wavelet $\psi_{a,b}$ satisfy the admissibility condition [45]:

$$\int_{-\infty}^{+\infty} \psi_{a,b} t dt = 0 \quad (3.3)$$

Hence, the original signal can be reconstructed by using the wavelet functions and the wavelet coefficients as:

$$x(t) = \sum_a \sum_b c(a,b) \psi_{a,b}(t) \quad (3.4)$$

To ensure existence of the inverse wavelet transformer the mother wavelet should satisfy an admissibility condition such as:

$$c_w = \int \left| \frac{\psi(w)}{w} \right| < +\infty \quad (3.5)$$

As states in previous chapter, by employing a quadrature mirror filter a (discrete) two-channel mutilated filter bank convolves a signal a_0 with a low-pass filter $h_1[2n]$ and a high-pass filter $g_1[2n]$ as of:

$$\begin{aligned} a_1[n] &= a_0 * h_1[2n] \\ d_1[n] &= a_0 * g_1[2n] \end{aligned} \quad (3.6)$$

A reconstructed signal a_2 can be obtained by filtering the zero expanded signals with a dual low-pass filter h_2 and a dual high-pass filter g_2 . If $z(x)$ denotes the signal obtained from x by inserting a zero between every sample which can be written as:

$$a_2[n] = z(a_1) * h_2[n] + z(d_1) * g_2[n] \quad (3.7)$$

The following Figure (3.3) illustrates the decomposition and reconstruction process [45].

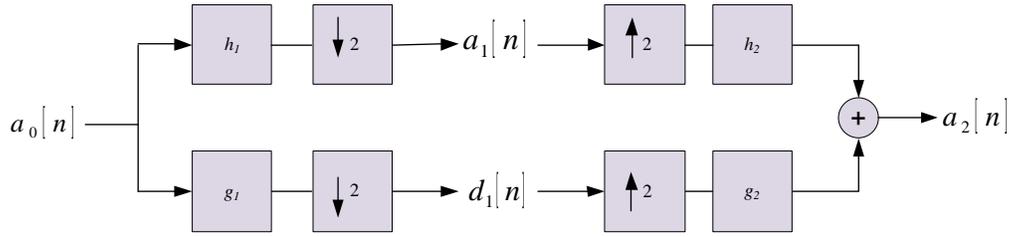


Figure (3.3) Wavelet packet filter bank decomposition where h_1 : low-pass filter; g_1 : high-pass filter. Down sampling \downarrow operator.

The filter bank is said to be a perfect reconstruction filter bank when $a_2 = a_0$. Moreover, if $h_1 = h_2$ and $g_1 = g_2$, the filters are called conjugate mirror filters [45].

3.2.2 WPD Feature Coefficients

The applicability of using wavelet transform analysis for motor health monitoring and fault diagnosis is based on the existence of some sensitive bands in the monitored signals which reflects the machine's healthy or faulty status. The wavelet coefficient measures the similarity between sequential patterns and the probed signal and the feature signal strings (the feature wavelet). In fact, wavelet coefficient residuals differences are used to detect the deficiency in the motor [47]. By applying WPD analysis to the probed signal, a time–frequency spectrum containing the wavelet packet coefficients for all the Depths and Nodes would be generated. For a signal with m samples, the maximum number of coefficients for any given Depth is m , and the total number of Depths is $(1 + \log_2^m)$ [4]. Accordingly, at each Depth, the WPD spectrum has been arranged in a way to locate those low frequency components at small Node values and the high-frequency components at large Node values.

It should be pointed out that the harmonic frequencies only exist in a limited number of Nodes of the WPD spectrum Due to the limited frequency and time support of the wavelet, at each Depth. Consequently, The feature coefficient $w(j,k)$ for Depth j and Node k is [13,17].

$$w(j,k) = \sqrt{\sum_{j,k} x_{j,k}^2(n)/N} \quad (3.8)$$

N is the number of coefficients and determined by the Depth level M and the length of the analyzed data m :

$$N = M2^m \quad (3.9)$$

Even though, the multi resolution ability of the WPD results in different frequency resolutions at different Depth levels but the product of the frequency and time resolutions remains the same Due to the limited frequency and time support of the wavelet, at each Depth, so as the number N . For a given harmonic component, the corresponding feature coefficients are localized at a number of Nodes for a given Depth level. The center of these Nodes is determined in terms of Nyquist frequency and Depth as [45]

$$n_w = 2^m w/w_N \quad (3.10)$$

In practice, due to the frequency overlap between successive filters, some harmonics may cover more than one pocket of the time–frequency spectrum. Therefore, at the same Depth, some harmonics can be presented by more than one feature coefficient. Furthermore, it might not be necessary to use all the coefficients of one node of the WPD spectrum to calculate the feature coefficients. In some cases, computing the feature coefficients with a few WPD coefficients might induce a better result specifically for fault localization.

3.2.3 Load torque

The analysis of the torque load is relatively new subject in induction motor condition monitoring. However, the efficiency of it has been verified.

In theory, the occurrence of a broken rotor bar, hamper the current flow. Consequently the torque would be reduced and the slip speed would be introduced. This increment in the voltage should be compensated by the current in the good rotor bars by producing the needed torque. This circulation continues until the torque of the motor is restored to equal the torque of the load [18].

Considering an induction motor in an ideal condition where the stator flux linkage is sinusoidal, any minor change in the load torque results in current changes at specific harmonic such as [18]:

$$f_{load} = f_e \pm mf_m = f_e [1 \pm (\frac{1-s}{P/2})] , m = 1, 2, 3, \dots \quad (3.11)$$

f_m is the rotational speed frequency,

f_e is the fundamental frequency,

P is the number of poles, and

s is the sleep frequency.

Generally, the lowest harmonic component produced by a fault consider as a fault identifier since it's the largest harmonic and easiest to be measured. The Human expert knowledge of the torque load fluctuation also can contribute to this task. The load torque is known to contain only specific harmonic components. Unlike the normal routine in this approach the current harmonics that are not affected by the load should be monitored as a sign of the broken rotor bar existence.

In fact, if a faulty condition in the stator current harmonics corresponds to the odd values of m while the load torque corresponds to the $2(m/2)$ harmonics. In fact, these harmonics could not co-exist at the same time [18].

Notably, the same frequencies apply for the broken bar detection:

$$f_{brb} = f_e [s \pm k(\frac{1-s}{P/2})] \quad (3.12)$$

Considering the similarity, the torque oscillation frequencies can easily be confused by the frequencies related to the fault existence. Therefore, an adequate discrimination only can be accomplished by monitoring fault frequencies at sufficiently high harmonics, which is the lowest one as it has been stated before.

3.3 Support Vector Machine

In this section the principal of SVM has been explained thoroughly, starts with statistical learning theory (3.3.1), which SVM has been defined, based on it. Following by Δ -Margin separating hyperplane in (3.3.2). Linear classification (3.3.3), non-linear classification (3.3.4) and tuning the meta-parameter of the support vector machine (3.3.5). Furthermore, two other SVM based algorithm has been discussed as of Kernel based Principal Component Analysis – Support Vector Machine (KPCA-SVM) (3.3.6) and Receiver Operation Curve-Support Vector Machine (ROC-SVM) in section (3.3.7) respectively.

3.3.1 Statistical learning theory

Statistical learning theory provide a frame work to study the problem of learning by acquiring knowledge about the supervisor's responses to obtain the best possible prediction and decision

on a particular problem. Accordingly, after defining the Phenomenon, a model should be constructed and finally the decision would be made based on the proposed model.

In a learning model, a random vector x_i is being selected as an input to the classifier independently from an unknown distribution $p(x_i)$. A supervisor learner defined by the function $f(x_i)$ which returns an output vector y_i for every input vector x_i , under the conditional distribution function $p(y_i|x_i)$. At the end, to evaluate the classifier's performance the differences between input x_i and the output y_i called loss function can be measured as[14,15]:

$$C(x, y, f(x)) = \frac{1}{2} |f(x) - y| \quad (3.13)$$

$C = 0$ means the data is classified correctly and $C = 1$ interpreted as a misclassification. The loss value called structural risk (SR). Hence, the learning machine's objective can be specified as minimization of the structural risk by selecting an appropriate decision function $f(x_i)$ where the probability $P(x_i, y_i)$ is unknown. Hence the structural risk is:

$$R[f] = \int_{X \times Y} L(y, f(x_i)) P(x_i, y_i) dx_i dy_i \quad (3.14)$$

Structural risk and the empirical risk $R_{emp}(\alpha)$ could be equally minimized if:

$$\lim_{l \rightarrow \infty} \min_{f \in H_n} R_{emp}[f] = \min_{f \in H_n} R[f] \quad (3.15)$$

H_n is a Hilbert, a separable and infinite-dimensional Euclidean space [16]. It is necessary to restrict the set of function from which f is chosen from in accordance to the Vapnik-chervonenkis (VC) theory. The VC theory provides band on the test error to provide a suitable

capacity for available training data. Consequently, There should be a h in the VC dimension of the classifier which can predict a probabilistic upper bound on the test error of a classification model.

According to the explained VC dimension, each function can separates the pattern in a certain way by a labeling them as ± 1 . Therefore, there should be at most 2^m different labeling for m patterns.

If $h < m$ then the VC dimension of the class of the functions that the learning machine implements, is independent from the underlying distribution $p(x_i)$. Which means, the generated data, has at least minimum probability h over the selected training data [16].

$$R[f] \leq R_{emp}[f] + \sqrt{\frac{h \ln(\frac{2m}{h} + 1) - \ln(\frac{\delta}{4})}{m}} \quad (3.16)$$

h is the VC dimension of the classifier and the equation on the right hand side is the risk bound. Independency of the equation from $P(x_i|y_i)$ is noticeable and it is easy to be calculated only if h is known. If the δ be chosen sufficiently small, the classifier that minimizes the right hand side could be easily obtained in accordance to the structural risk minimization theory.

Support Vector Machine should map the input vectors into a very high-dimensional feature space by constructing an optimal separating hyperplane named Δ -margin separating hyperplane.

Subsequently, if vectors $x_i \in X$ belongs to a sphere of radius R , then the Δ -margin separating hyperplanes has the VC dimension h bounded by the inequality [15]:

$$h \leq \min([\frac{R^2}{\Delta^2}], n) + 1 \quad (3.17)$$

A good generalization obtained by constructing an optimal separating hyperplane, which maximize the margin [15,50].

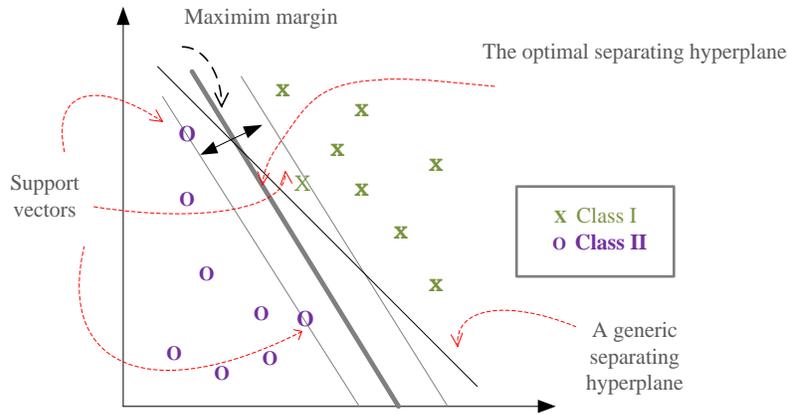


Figure (3.4) Δ -Margin separating hyperplane

3.3.2 Δ -Margin separating hyperplane

If there exist a separable data set such as: $x_i = (x_1, x_2, \dots, x_m)^T, i = 1, 2, \dots, m, m \in R^m$ the input vector x_i can be labeled as $y_i \in -1, +1$. The positive and negative labeled data can be separated by a Δ -margin separating hyperplane into two differentiable classes. Where the decision function is:

$$f(x_i) = wx_i + b = \sum_{i=1}^m wx_i + b = 0 \quad (3.18)$$

$w_i \in R^m$ is the weight vector and scalar b represents bias value. These elements are defining the position of the separating hyperplane.

The separating hyperplane, should satisfy the following conditions as [15,16]:

$$y = \begin{cases} 1 & \text{if } wx_i + b \geq \Delta \\ -1 & \text{if } wx_i + b \leq -\Delta \end{cases} \quad (3.19)$$

$$\alpha_i \geq 0 \quad i = 1, 2, \dots, m$$

3.3.3 Linear classification

In order to formulate a linear SVM classifier if a separating hyperplane, assumed to be:

$$\gamma = \frac{1}{\|w\|} \quad (3.20)$$

Then the decision function could be defined as:

$$\begin{aligned} y_i &= +1 & \text{if } f(x_i) \geq +1 \\ y_i &= -1 & \text{if } f(x_i) \leq -1 \end{aligned} \quad (3.21)$$

Combination of these two equations from (3.21) with decision function as

$f(x_i) = wx_i + b = \sum_{i=1}^m wx_i + b = 0$, results in the following equality:

$$y_i f(x_i) = y_i (wx_i + b) \geq 1 \quad \text{where } i = 1, 2, \dots, m \quad (3.22)$$

To find the optimal margin geometrically according to Figure (3.5), supposing that x^+ belongs to the hyperplane $h_1: wx^+ + b = 1$ and x^- belongs to class $h_2: wx^- + b = -1$ and x^o lies on the separating hyperplane $h: wx^o + b = 0$ then geometrical margin could be found by the projection of the distance between these two data points from different classes on the direction

perpendicular to the hyperplane ,consequently the distance between them could be measured as[15] :

$$\left[\frac{w^T}{\|w\|} (x^+ - x^-) \right] = \frac{2}{\|w\|} \quad (3.23)$$

Putting it back to (3.21) results in:

$$\gamma = \frac{1}{2} \left(\left(\frac{-w^T (x^0 - x^+)}{\|w\|_2} \right) - \left(\frac{-w^T (x^0 - x^-)}{\|w\|_2} \right) \right) = \frac{1}{2\|w\|_2} ((w^T x^+) - (w^T x^-)) \quad (3.24)$$

$$\gamma = \frac{1}{\|w\|_2} \quad (3.25)$$

The solution for the convex optimization problem to maximize the margin should satisfy the following conditions:

$$\text{Minimize} \quad \frac{1}{2} \|w\|^2 \quad (3.26)$$

$$\text{Subject to} \quad y_i (wx_i + b) \geq 1 \quad (3.27)$$

Those training points located on one of the hyperplanes h_1 or h_2 whose removal, could change the obtained solution, are support vectors (SVs). In Figure (3.5) support vectors have been depicted by empty circles.

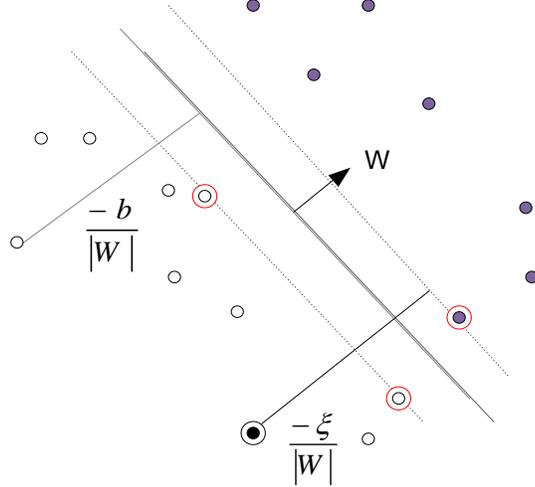


Figure (3.5) Linear-separating hyperplanes for the separable case. The support vectors are circled.

To calculate the dot product between vectors for the large data sets, the Lagrange multiplier applies to decision function to simplify the process as [14,15]:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \alpha_i y_i ((w_i^T x_i) + b) - 1 + \sum_{i=1}^m \alpha_i \quad (3.28)$$

Knowing that $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_m)^T$ is the Lagrange multiplier string, the dual problem should be solved with regard to the following conditions:

$$\text{Minimize } L(w, b, \alpha) \text{ to } \alpha_i \geq 0 \text{ for } x_i \quad (3.29)$$

Hence, to find a proper solution for the convex problems (if the regularity condition holds), the (Karush-Kuhn-Tucker theorem) KKT conditions should be satisfied. More specifically, solving the SVM problem is equivalent to finding a solution to the KKT conditions [50] :

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^m \alpha_i y_i w = 0 \quad (3.30)$$

$$\frac{\partial L}{\partial b} = -\sum_{i=1}^m \alpha_i y_i = 0 \quad (3.31)$$

$$y_i(w \cdot x_i + b) - 1 \geq 0 \quad \alpha_i \geq 0 \quad \forall i \quad (3.32)$$

$$\alpha_i(y_i(w \cdot x_i + b) - 1) = 0 \quad \alpha_i \geq 0 \quad \forall i \quad (3.33)$$

From (3.30), (3.31), dual presentation of the Lagrange multiplier is:

$$\text{Maximize} \quad L_D = \sum_{i=1}^m \alpha_i - \frac{1}{2} \sum_{i=1}^m \alpha_i \alpha_k y_i y_k x_i^T x_k \quad \text{subject to} \quad \sum_{i=1}^m \alpha_i y_i = 0 \quad (3.34)$$

$$\alpha_i \geq 0 \quad i = 1, 2, \dots, m$$

L and L_D arise from the same objective function but with different constraints. Thus, the solution could be determined either by minimizing L or maximizing L_D [50].

Explicitly, if there is any Lagrange multiplier α_i associated to the training data (x_i, y_i) and $\alpha_i \geq 0$ then, x_i is a support vector.

Finally, the generalization ability of the classifier on the test error could be calculated where m is the number of the training samples:

$$E[P_{error}] \leq \frac{\#SVs}{m} \quad (3.35)$$

3.3.4 Non-linear classification

So far only the cases where the training data is linearly separable have been discussed. In many cases that data could not linearly be separated, some modification should be made to equations.

First, to relax the constraint, a positive slack variable ξ_i has been introduced to equations [14], [15]:

$$x_i \cdot w + b \geq +1 - \xi_i \quad \text{for } y_i = +1 \quad (3.36)$$

$$x_i \cdot w + b \leq -1 - \xi_i \quad \text{for } y_i = -1 \quad (3.37)$$

$$\xi_i \geq 0 \quad \forall i$$

$\sum \xi_i$ Represents the upper bound on the training errors. Respectively, the primal Lagrange multiplier equality changes to:

$$L = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_i \alpha_i [y_i(x_i \cdot w + b) - 1 + \xi_i] - \sum_i \mu_i \xi_i \quad (3.38)$$

In this equation, μ_i is a new Lagrangian multiplier which is introduced to enforce the positivity of the ξ_i [50].

Consequently by introducing the ξ_i , the KKT condition also is converted to:

$$\frac{\partial L}{\partial w} = w - \sum_i \alpha_i y_i x_i = 0 \quad (3.39)$$

$$\frac{\partial L}{\partial b} = -\sum_i \alpha_i y_i = 0 \quad (3.40)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \quad (3.41)$$

$$\xi_i \geq 0, \alpha_i \geq 0, \mu_i \geq 0$$

$$\alpha_i \{y_i(x_i \cdot w + b) - 1 + \xi_i\} = 0 \quad (3.42)$$

$$\mu_i \xi_i = 0$$

If $\xi_i = 0$ then the boundaries would be limited to $0 \leq \alpha_i \leq C$.

To map the n-dimensional non-linear input data into the high dimensional feature space a nonlinear vector function $\varphi(x) = (\varphi_1(x), \dots, \varphi_l(x))^T$ should be utilized. Accordingly, the decision function changes to:

$$f(x) = \sum_{SVs} \alpha_i^* y_i \varphi(x_i)^T \varphi(x) + b^* \quad (3.43)$$

In non-linear cases the dot product of the input data needs a kernel function K to form a dot product of the mapped data $\varphi(x_i) \cdot \varphi(x_j)$ as:

$$K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j) \quad (3.44)$$

Some renowned kernel functions are:

$$\begin{aligned} x_i \cdot x_j & \quad \text{linear} \\ (\gamma x_i x_j + 1)^d & \quad \text{polynomial} \\ \exp(-\gamma |x_i - x_j|^2) & \quad \text{RBF} \end{aligned}$$

However, the kernel function should satisfy the mercer condition in order to map the input data into the feature space. According to mercer condition $K(x_i, x_j) = \sum_i \varphi(x_i) \cdot \varphi(x_j)$,if only

$$\int g(x)^2 dx \text{ is finite} \quad (3.45)$$

Then

$$\int K(x_i, x_j) g(x_i) g(x_j) dx_i dx_j \geq 0 \quad (3.46)$$

The decision function would be:

$$f(x) = \sum_{i=1}^m \alpha_i^* y_i K(x_i, x) + b^* \quad (3.47)$$

The unknown sample x_i belongs to:

$$x_i = \begin{cases} \text{class I} & \text{if } f(x_i) > 0 \\ \text{class II} & \text{otherwise} \end{cases} \quad (3.48)$$

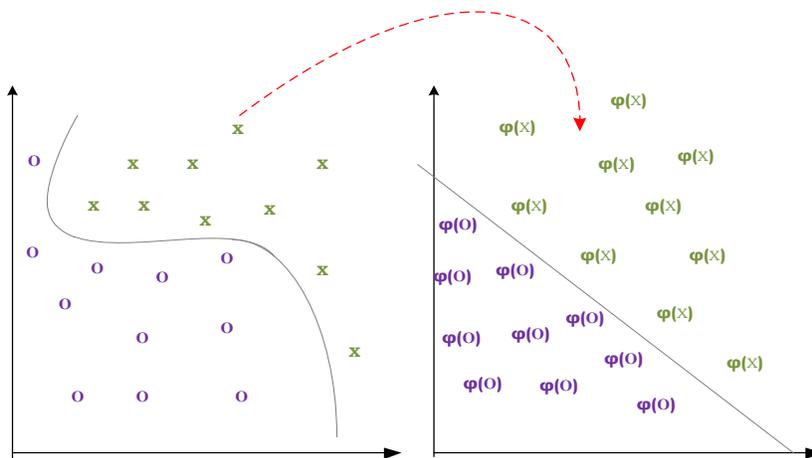


Figure (3.6) Non-linear classification by mapping data into a feature space

3.3.5 Tuning the meta-parameters of Support Vector Machine

Selecting appropriate parameters for SVM directly affect the performance of the classifier. The generalization ability of the SVM controlled by two factors: the capacity of the learning defined by the upper bound c and the error rate on the training data. In fact, the upper band controls the tradeoff between margin maximization and error minimization. Applying a high error penalty forces the SVM to avoid classification error. Therefore, a very high error penalty would be expected to lead to a perfect classification. However, it does not guarantee a good performance due to the boundary irregularity. More importantly c also controls the number of the SVs, which affects the generalization ability of the SVM, which means by increasing the number of SVs the generalization ability of the classifier would be degraded. Moreover, in the non-linear cases, the kernel parameters should be tuned as well. Normally, only limited types of kernels have been used in researches so far due to the difficulty of the tuning the required parameters. Selecting a proper kernel is a key factor in SVM performance. According to literature review, the radial basis function (RBF) kernel is the mostly preferred over other ones due to its resistance to the outliers.

3.3.6 KPCA-SVM

In this section, another SVM based algorithm which has been built by combining a standard Support Vector Machine and Kernel-Principal Component Analysis (KPCA) has been introduced. Normally in classification, by applying a Principal Component Analysis (PCA) the eigenvectors of the covariance matrix of the original inputs are calculated and linearly transforms the high-dimensional input vector with correlated components into a low-dimensional one whose components are uncorrelated [21].

In a similar way, KPCA maps the original inputs into a high-dimensional feature space using the kernel method before applying the PCA on it.

Application of the PCA on the input vector x_i results in:

$$y_i = u^T x_i = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ y_m \end{bmatrix} \quad x_1 \quad \cdot \quad \cdot \quad x_m \quad (3.49)$$

$i = 1, \dots, m$ And u^T is a matrix transforming x_i to y_i . Which means, all coefficient of y_i are the results of the dot product of x_i and the corresponding row in u .

The covariance matrix C_y as a base of the PCA analysis is [21]:

$$C_y = \frac{1}{m} \sum_{i=1}^m y_i y_i^T \quad (3.50)$$

In fact C_y is the value of the covariance between all possible pairs of measurements and reflects their noise, and redundancy values.

Choosing a proper eigenvalues $\lambda > 0$ and eigenvectors V to meet the following condition, forms the basis of PCA application:

$$\lambda C = C u \quad (3.51)$$

To calculate the values for KPCA mathematically, equations (3.50) combined with (3.51) as:

$$C_y = \frac{1}{m} \sum_{i=1}^m y_i y_i^T = \frac{1}{m} \sum_{i=1}^m (u_i x_i)(u_i x_i)^T = u_i \left(\frac{1}{m} \sum_{i=1}^m x_i x_i^T \right) u_i^T \quad (3.52)$$

$$C_y = u_i C_{x_i} u_i^T \quad (3.53)$$

Thus, the Principal Components of x_i are the eigenvectors of $x_i x_i^T$ or the rows of u_i . To expand the PCA to the KPCA, the linear PCA should be transformed to a none-linear case using a kernel function. In fact; the original input x_i is mapped into a high-dimensional feature space as $\varphi(x_i)$ first before applying the linear PCA. Consequently:

$$C_y = \frac{1}{m} \sum_{i=1}^m \varphi(y_i) \varphi(y_i)^T = \frac{1}{m} \sum_{i=1}^m \frac{1}{m} \sum_{i=1}^m (u_i \varphi(x_i)) (u_i \varphi(x_i))^T \quad (3.54)$$

Similarly the eigenvalue $\lambda \geq 0$ and the eigenvector V could be found by transforming the original data into a high dimensional feature space as:

$$\lambda(\varphi(x_i)V) = (\varphi(x_i)CV), \forall i \in 1, \dots, m \quad (3.55)$$

If there exists some co-efficient as $\alpha^k = \{\alpha_1, \dots, \alpha_l\}$

Where
$$V^k = \sum_{i=1}^m \alpha_i \varphi(x_i) \quad (3.56)$$

Putting it into (3.55) and (3.56) into (3.54):

$$K(x_i, x_j) = (\varphi(x_i), \varphi(x_j)) \quad (3.57)$$

$$l\lambda K\alpha^k = K^2\alpha^k \quad (3.58)$$

The solution α^k should be normalized by $(V^k \cdot V^k) = 1$. Consequently From (3.56), (3.57) and (3.58):

$$1 = \sum_{i,j=1}^l \alpha_i^k \alpha_j^k (\varphi(x_i), \varphi(x_j)) = (\alpha^k . K \alpha^k) = \lambda_k(\alpha^k, \alpha^k) \quad (3.59)$$

The Principal Component is extracted by projecting the image of the test point $\varphi(x_i)$ to the eigenvector V^K in the F as:

$$(V^K . \varphi(x_i)) = \sum_{i=1}^l \alpha_i^k (\varphi(x_i), \varphi(x_j)) \quad (3.60)$$

Finally

$$y_i = u^T \varphi(x_i) = \sum_{i,j=1}^m \alpha_i k(x_i, x_j) \quad \forall i \in 1, \dots, m \quad (3.61)$$

Application of the KPCA in combination to SVM is expected to reduce the dimensionality of the input data and increase the performance and accuracy rate consecutively [21,51].

3.3.7 ROC-SVM

Generally, the performance of the classifier is evaluated by estimating its generalization error. However, in many cases even when a classifier gains a high detection rate, the accuracy rate is still poor. In advanced signal processing, the Receiver Operating Characteristics (ROC) curve which is a two-dimensional measurement representing the performance rate of the classifier. In fact, the ROC curve defines the probability of correctly classified samples against the misclassified ones. In such a case the decision rule would be set by selecting a proper threshold, capable of separating the positive classes from the negative ones. Literally, ROC curve is determined by this threshold value. In practice, the performance rate is calculated by measuring the Area Under the ROC Curve, called AUC [52]. Figure (3.7) represents the ROC-SVM curve.

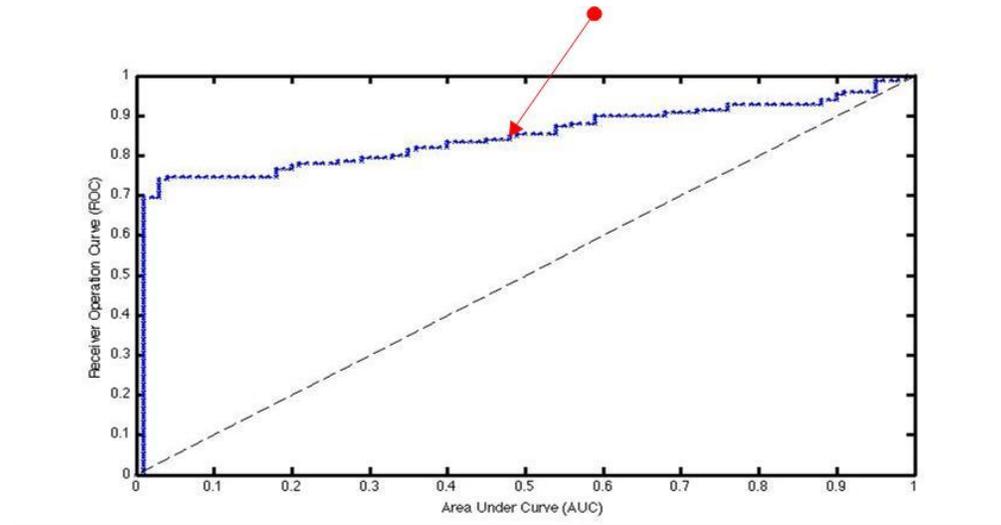


Figure (3.7) ROC curve, representing the ROC-SVM CURVE

Respectively an algorithm has been proposed based on AUC optimization called ROC-SVM, which expected to produce a higher accuracy rate in comparison to the standard SVM. In this approach, the decision function has been defined by a decision threshold separating the positive classes from negative ones. For discrete cases the AUC threshold equation is [52]:

$$AUC = \frac{\sum_i^{m^+} \sum_j^{m^-} f(x_{i,j})}{m^+ m^-} \quad (3.62)$$

x^+ and x^- are the positive and negative samples and m^+ and m^- shows the number of samples by the same token. Thus, the AUC achieves the maximum value only if $f(x_i^+) > f(x_j^-)$, $\forall i \in 1, \dots, m^+$ and $\forall j \in 1, \dots, m^-$.

In linear cases where the input of the classifier has is $\{x_i, y_i\}_{i=1}^m$, $\forall i \in 1, \dots, m$, similar to standard SVM, the linear decision function is $f(x) = wx_i + b$ and the problem of maximizing AUC becomes:

$$\max \quad AUC = \frac{\sum_i^{m^+} \sum_j^{m^-} f(x_{i,j})}{m^+ m^-} \quad (3.63)$$

$$\text{with} \quad \xi_{ij} = f(x_i^+) - f(x_j^-) \quad (3.64)$$

To find a linear hyperplane capable of separating all data points AUC equation has been modified as:

$$\min \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{m^+} \sum_{j=1}^{m^-} \xi_{ij} \quad (3.65)$$

$$f(x_i^+) \geq f(x_j^-) + \rho - \xi_{ij} \quad \xi_{ij} \geq 0 \quad \rho > 0 \quad (3.66)$$

$$\begin{aligned} 1 \leq i \leq m^+ \quad 1 \leq j \leq m^- \\ 1 \leq i \leq m^+ \quad 1 \leq j \leq m^- \end{aligned}$$

$\rho > 0$ is a square-norm regularization term according to Tikhonov based-regularization rules [52] and c defines the trade-off between the regularization term and the constraint violation ξ_{ij} .

In practice to link the linear ROC-SVM equations with the hard margin SVM, a regularization parameter $\rho > 0$ applied which links the ROC-SVM with standard SVM such as [52]:

$$w_{ROC} = \beta w_{SVM} \quad (3.67)$$

$\beta = \frac{\rho}{2}$ And ρ is a regularization parameter.

Consequently, the Lagrange multipliers function changes to:

$$L(w, \xi_{ij}) = \frac{1}{2} \|w\|^2 + c \sum_{i,j=1}^{m^+, m^-} \xi_{ij} - \sum_{i,j=1}^{m^+, m^-} \alpha_{ij} (\langle w, x_i^+ - x_j^- \rangle - \rho + \xi_{i,j}) - \sum_{i,j=1}^{m^+, m^-} \gamma_{i,j} \xi_{i,j} \quad (3.68)$$

Applying the KKT condition results in:

$$\frac{\partial L}{\partial w} = w - \sum_{i,j} \alpha_{i,j} (x_i^+ - x_j^-) = 0 \quad (3.69)$$

$$\frac{\partial L}{\partial \xi_{i,j}} = c - \alpha_{i,j} - \gamma_{i,j} = 0 \quad (3.70)$$

$$\forall i \in 1, \dots, m^+$$

$$\forall j \in 1, \dots, m^-$$

Putting them back into the lagrangian equation (3.68) leads to the following quadratic optimization problem:

$$\text{Maximize } -\frac{1}{2} \sum_{i,j=1}^{m^+,m^-} \sum_{u,v=1}^{m^+,m^-} \alpha_{i,j} \alpha_{u,v} \langle x_i^+ - x_j^-, x_u^+ - x_v^- \rangle + \rho \sum_{i,j=1}^{m^+,m^-} \alpha_{i,j} \quad (3.71)$$

$$\text{With } C \geq \alpha_{i,j} \geq 0$$

The decision function is:

$$f(x) = \sum_{i,j=1}^{m^+,m^-} \alpha_{i,j}^* (x_i^+ - x_j^-, x) + b \quad (3.72)$$

α^* is optimal solution to the dual problem called support vector [52].

For all pairs of positive and negative examples optimized with regards to AUC, the $f(x)$ is ranked as:

$$\begin{cases} f(x_i^+) \geq f(x_j^-) + \rho & \text{if } \alpha_{i,j} = 0 \\ f(x_i^+) \leq f(x_j^-) + \rho & \text{if } \alpha_{i,j} = C \\ f(x_i^+) = f(x_j^-) + \rho & \text{if } 0 < \alpha_{i,j} < C \end{cases} \quad (3.73)$$

To speed up the process instead of ranking the positive example above the negative ones, only a subset of a positive-negative samples are taking into account. In fact, if N^+ be a set of all positive samples there would be m-nearest positive neighbors of each negative samples.

In non-separable cases, like original SVM the data being mapped into the feature space before getting separated, using a kernel function $k(x_i, x_j)$.

The decision function for the ROC-SVM is:

$$f(x) = \langle w, \varphi(x) \rangle + b = \sum_{i,j=1}^l \alpha_i^* y_i k(x_i, x_j) + b^* \quad (3.74)$$

To obtain the optimal separating hyperplane bellow conditions should be satisfied:

$$\min_{w,b,\xi_i} \quad \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i^2 \quad (3.75)$$

$$\text{With} \quad y_i f(x_i) \geq 1 - \xi_i$$

Applying the KKT conditions to lagrangian equation in order to find the support vector α_i^* results in:

$$\text{Maximize} \quad -\sum_{i,j}^l \alpha_i \alpha_j y_i y_j (k(x_i, x_j) + \frac{1}{C} \delta_{x_i}(x_j)) + \sum_i \alpha_i \quad (3.76)$$

$$\text{With} \quad \sum_i \alpha_i y_i = 0 \quad \alpha \geq 0$$

Literally, the kernel function in a new Hilbert space H' where the data is linearly separable changes to:

$$k'(x_i, x_j) = k(x_i, x_j) + \frac{1}{C} \delta(x_j) \quad (3.77)$$

By evaluating the decision function in H' instead of H to preserving the AUC-maximal property produces a decision function:

$$f(x) = \sum_{i,j=1}^l \alpha_i^* y_i (k(x_i, x_j) + \frac{1}{c} \delta(x_j)) + b^* = \sum_{i,j=1}^l \alpha_i^* y_i k'(x_i, x_j) + b^* \quad (3.78)$$

CHAPTER 4: EXPERIMENTAL VERIFICATION

In this chapter the experimental procedure of our work has been explained completely in six sections as: experiment's Layout (4.1), calculating the wavelet coefficients of monitored signal (4.2), forming a Feature Vectors of the classifier (4.3), application of a binary class-support vector machine (4.4), KPCA-SVM (4.5) and ROC-SVM in section (4.6).

4.1 Experiment's Layout

Our Experiments performed on a three-phase 208-V, 60-Hz, 5-HP motor known as a squirrel-cage rotor with 28 rotor bars two of which are broken. The rated speed of the induction motor is 1750 r/min. The bench test has been set up, consisting of two identical squirrel cage induction motors connected end to end. One of them is acting, as a driving motor, and the other one is an induction generator, which provides various loading conditions for the motor under test. The leading phase current has been provided by Delta-connected capacitors, attached to the output terminals of the generator. The complete layout of the circuit for this experiment is depicted in Fig (4.1).



Figure (4.1) Induction motor–generator setup.

Practically, to achieve the real faulty samples from squirrel cage induction motor with two broken bars, a faulty condition has been fabricated by drilling a hole in the middle of the rotor bar, which deteriorates the whole electrical bar. However the impact of the mechanical asymmetry we made is still not clearly verifiable comparing to the dominant effect of the corresponding electrical asymmetry. Thus, to amplify the failure condition and satisfy the training requirements, two physically successive rotor bars have been impaired to induce the actual faulty condition. In reality, the rotor bars close to the broken bar are more prone to breakage than the other ones [17].



Figure (4.2) Faulty induction motor with a fabricated hole on the rotor bar

Measurements carried out, once with the healthy motor then with the faulty one separately. The sampling frequency is 60HZ and the number of samples in each data set is 25k in length. Four different loadings have been applied to the drive motor as of: 25%, 50% and 100% of the full load under both healthy and faulty conditions to study the effects of putting various loads in fault identification procedure. The samples have been collected during 100ms. In all experiment, the above-mentioned loading conditions would be named as light, medium and heavy loadings. In this work half of the samples from both healthy and fault motors have been used for training and the other half used for testing.

Furthermore, to keep down the starting current while maintaining adequate starting torque, a speed controller has been placed in the circuit, the controller consist of a diode bridge, a dc link and an IGBT based PWM controlled inverter. This controller provides variable frequency voltage for the inductor motor for speed adjustment.

In theory, the perfect results could be obtained from monitoring the steady state current of the induction motor, which is not applicable, specifically when the machine is fully loaded. Subjecting a machine to heavy loads could be impractical due to the overheating problem, which could reduce the operating lifetime of the machine. Besides, the steady state monitoring methods are not accurate enough while the running machine has been lightly loaded or is in the transient conditions. Since, under the light loading the presence of fault related harmonics is not easily differentiable from the harmonics around the fundamental frequencies. Also, the motor current spectral harmonics produced by the load could overlap the harmonics caused by the fault conditions, if the load torque varies with the rotational speed [53].

In fact, the main challenge of transient current analysis is differentiating the non-stationary fundamental frequency from non-stationary frequencies associated with the broken rotor bars existing in a transient signal.

Consequently, the main objective of this study could be summarized as: finding a proper method to detect faulty rotor bars in an induction motor by monitoring the startup transient signal instead of the steady state signal which requires a prior knowledge about the extremely unsteady and relatively short duration of the startup signal. Second, to be able detect and correctly classify the fault related samples consecutively.

4.2 Calculating the wavelet coefficients of monitored signal

To obtain the fault related features, a fundamental feature extraction algorithm should be applied to the transient start up signal. Consequently, the algorithm proposed in [17] has been followed according to which the WPD analysis which is the fourth-order biorthogonal wavelet Coiflets has been applied to the monitored signals in order to find the specific harmonic related to the broken rotor bars.

The Coiflets are one family of compactly supported wavelets with highest number of vanishing moments for a given support width [54]. For an N th order Coiflet, the support width is $6N - 1$ and the vanishing moment for the wavelet function is $2N - 1$ and the scaling function is $2N$ respectively [17,48,49].

Respectively, up to ten detailed scales (called Depth 10 or D10) of the Discrete Wavelet Transform have been performed. Since, no distinction could be made between signals monitored under healthy condition and under faulty condition with two broken rotor bars up to level 7. However, examining scales 8-10 represents the existence of some differences in the wavelet coefficients.

Furthermore, the wavelet coefficients of the signals under different loading situation have been studied. The results confirm that there is not enough information in the monitored signal from the stator current working under 25 % of the full loading; since, the sideband components overlap the supply frequency component. Therefore the diagnosis of the fault would be almost impossible or it may lead to misclassification.

The results obtained from applying the level D10 of the Wavelet Packet Decompositions to the current signal by 100 % and 50 % of full loads, has been pictured in Figure (4.3). Accordingly, the increment in the amplitude of the observed harmonics is undeniable.

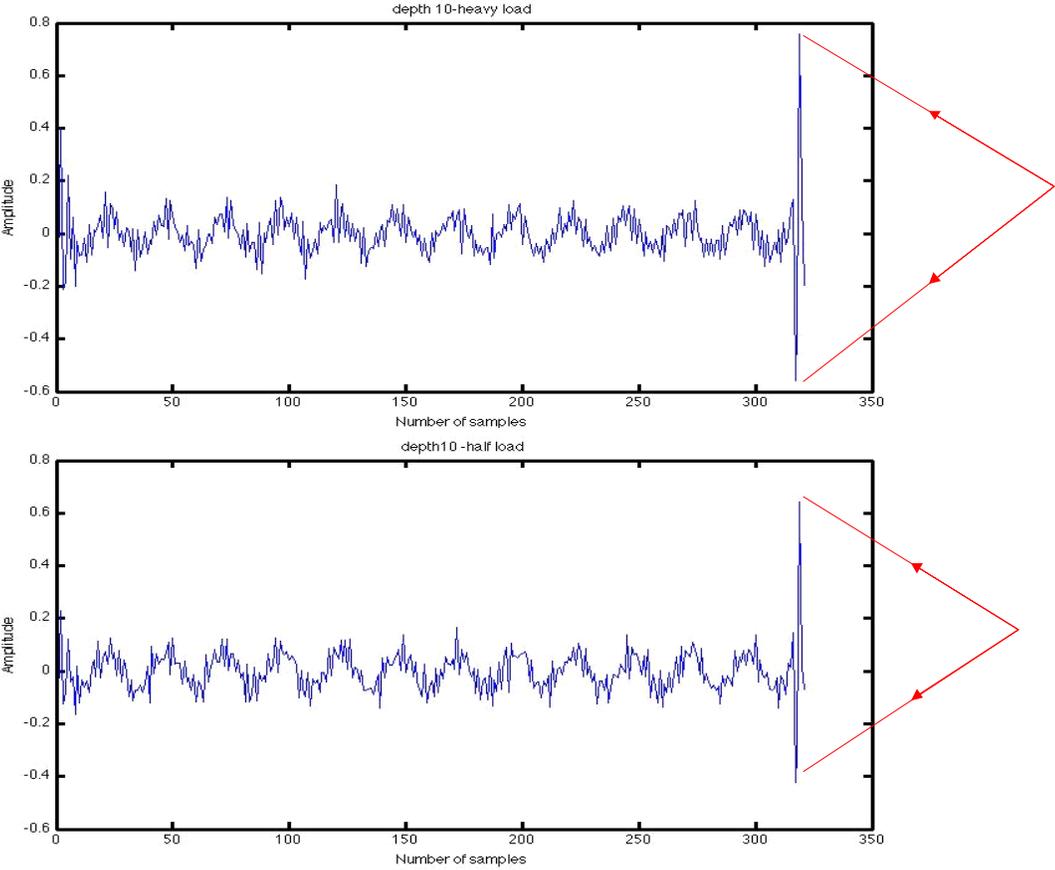
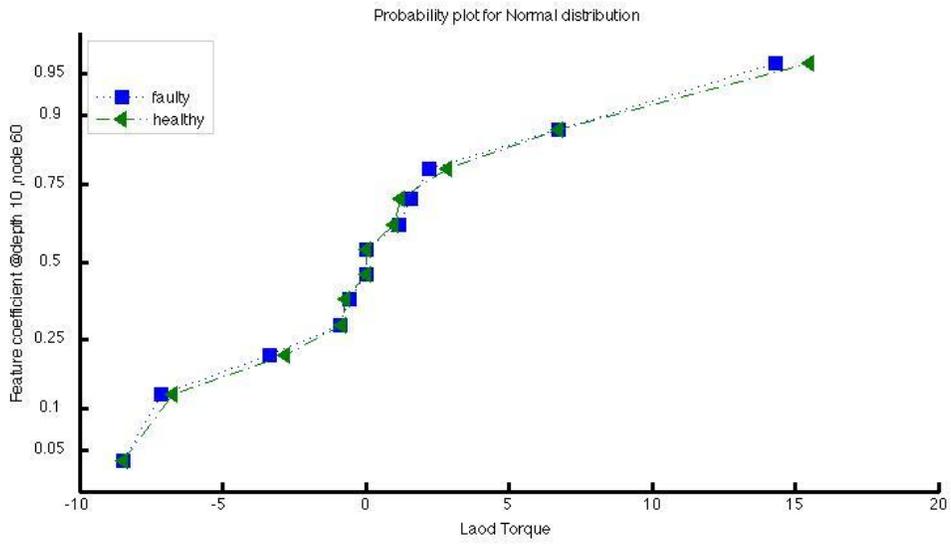
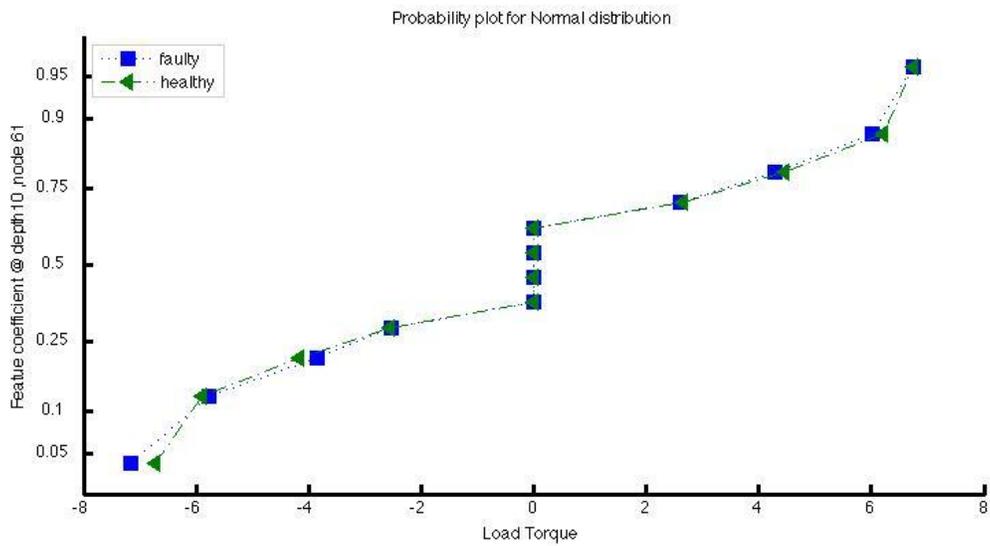


Figure (4.3) Stator current waveform for faulty induction motor with two broken rotor bars at depth 10 under full load and 50% of full load.

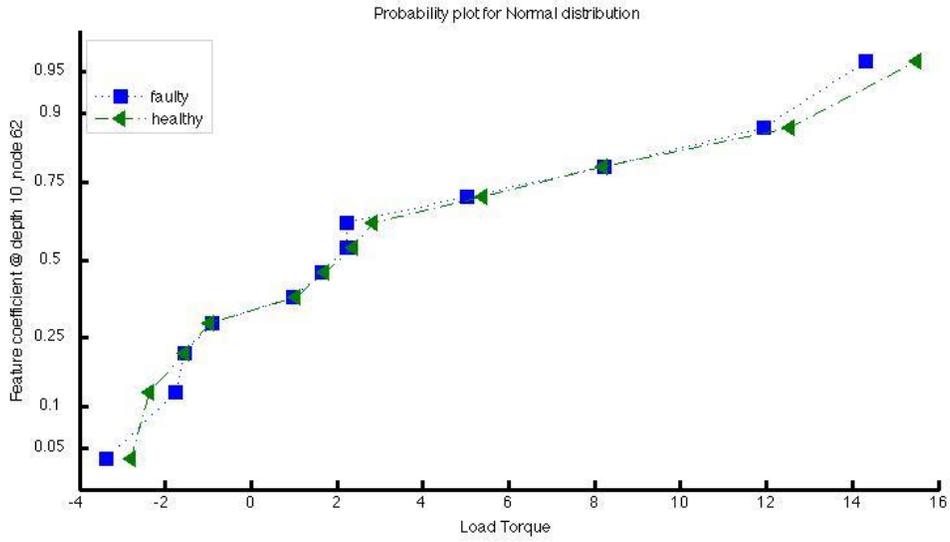
Finally, the distribution of the wavelet coefficients in the selected Depth and Nodes has been inspected. Figure (4.4) presents the spreading pattern of the wavelet coefficients from Nodes 60, 61, 62, and 63 of Depth 10 under different loadings.



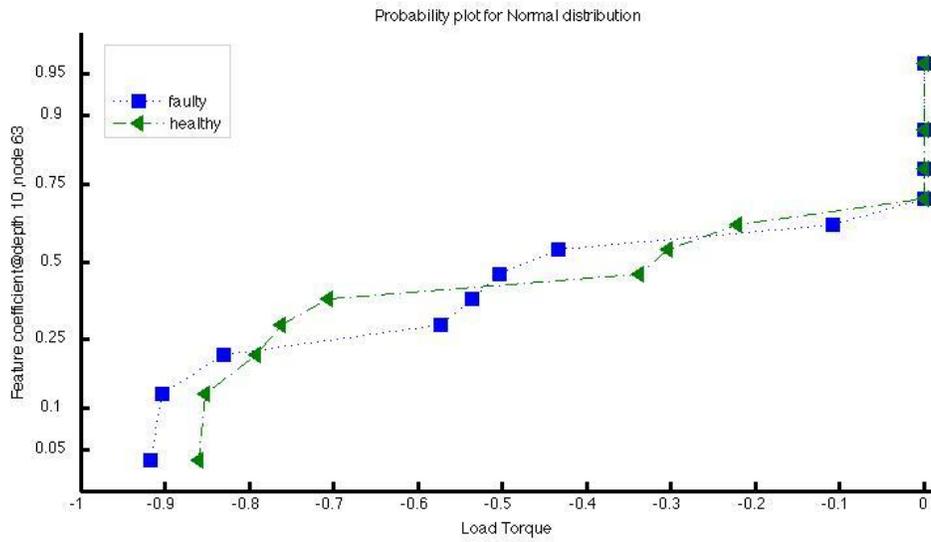
(a)



(b)



(c)



(d)

Figure (4.4) Feature coefficients at (a) Depth 10 Nodes 60 ,(b) depth 10 node 61,(c) depth 10 node 62 and (d) depth 10 node 63 for Faulty and healthy motors.

Figure (4.5) represents the feature coefficients at Node 62 of Depth 10 for both healthy and faulty motors extracted from the signal under heavy loading to accentuate the differences in the amplitude of the healthy and faulty coefficients. According to this figure, the value of the coefficient for the faulty motor is always greater than the healthy one.

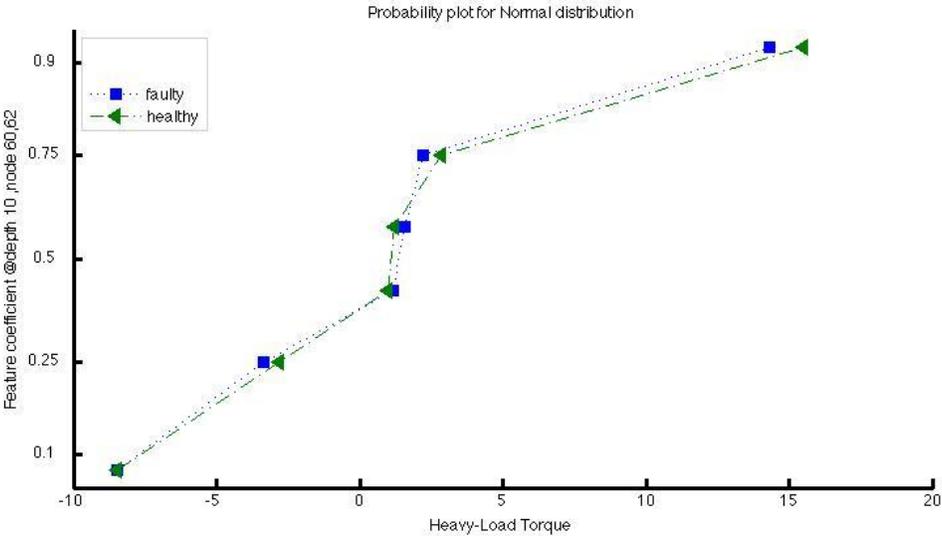


Figure (4.5) the feature coefficients at node 62 of depth 10 for both healthy and faulty motors under a heavy loading

4.3 Forming a Feature Vector of the classifier

Choosing an appropriate feature to form the feature vector in pattern recognition would increase the detection rate of the classifier. Normally, instead of using the rough measured data directly, a feature vector of fixed dimension could be used as input sets for the classifier.

We have chosen to make a feature vector consist of wavelet coefficients obtained extracted from different depth and nodes in addition to the load torque. The results of the classification verified the impeccability of this selection.

Accordingly, the wavelet coefficients have been chosen from different Depths and any two coefficients from variant Depths such as $d9$ or $d9$ selected to build a vector [17,48,49]:

$$F^P = [W(8, n_1), \dots, W(9, n_3), \dots, W(10, n_z), \dots] \quad (3.80)$$

Where P representing the node number of the features, and P denotes the number sequence of the samples. These features vector together with the load torque T_L of the motor used to build the input vector used for training:

$$I_p = [F^P, T_L] \quad (3.81)$$

Where $T_L \in T_{L100\%}, T_{L50\%}, T_{L30\%}, T_{Lzero}$.

Accordingly, the entire input and output-training spaces would be represented as:

$$\begin{aligned} & [I_1, \dots, I_p, \dots, I_P] \\ & [T_1, \dots, T_p, \dots, T_P] \quad \text{for } p = 1, 2, \dots, P \end{aligned} \quad (3.82)$$

The decision function a of the trained support vector machine can then be expressed as

$$T_p = f(I_p) \quad (3.83)$$

Where I_p is the input vector and T_p denote the target data vector for each given vector in the input feature sets, which reflects the actual condition of the machine.

4.4 Application of a binary class -Support Vector Machine

In this part, the application of the binary Support Vector Machine for detecting broken bar has been studied. Considering the noisy nature of the data acquired from the stator's current signal, a

soft-margin binary SVM classifier has been chosen to discriminate between the healthy motor with no-broken bar and a faulty one with two broken bars. The loading contribution in fault identification process also has been investigated.

The input to the classifier is the two dimensional feature vectors consisting of the wavelet coefficient and the load torque values, which has been clearly defined before. The SVM-KM toolbox in MATLAB has been used in this work.

Generally in classification, a proper subset of data samples should be selected to improve the discriminative ability of a classifier. By considering both the quality of the solution and the computational expense, we have divided the collected data into two subsets to be used as:

Dataset # 1) all acquired samples has been treated as one whole training data set, where

$$\text{Dataset 1} = \{\text{Training Set}\} = \{\text{Test Set}\}.$$

Dataset # 2) data samples has been separated into two subsets where

$$\text{Dataset 2} = \{\text{Training Set}\} \cup \{\text{Test Set}\} \text{ while } \{\text{Training Set}\} \cap \{\text{Test Set}\} = \emptyset.$$

In the first classification experiment, the means of the wavelet coefficients at depth 8,9,10 with various nodes for both healthy and faulty condition under various load conditions have been calculated separately [17,48,49]

$$mean = \frac{1}{n} \sum_{i=1}^n x_i \quad (3.84)$$

The training data shaped by 20 samples from both healthy and faulty motor under light, medium and heavy loads which results in (3x20 = 60) samples in total. The testing data set consists of

100 samples for each load 300 in total for both healthy and faulty machine. Table (3.1) displays the statistics mean of the feature coefficients located at different Nodes and Depths for a healthy motor with no broken bar and the fault machine with two broken bars under light, medium and heavy loadings.

Table (4.1) Mean values of the feature coefficients for faulty induction motor with two broken bars and a healthy motor with no broken bar.

(Depth#, Node#)	Light load		Medium load		Heavy load	
	# Broken bars		# Broken bars		# Broken bars	
	Non	2	Non	2	Non	2
	Mean		Mean		Mean	
(8,14)	*2.5496	*2.9896	*1.6958	* 2.1112	2.0473	2.2146
(8,15)	-0.7438	-0.7824	-0.6337	-0.6338	-0.7260	-0.6453
(9,29)	-0.6084	-0.6481	-0.3998	-0.4178	*-0.3485	*-0.5987
(9,30)	-0.2013	-0.2026	-0.0179	-0.0124	0.0615	-0.0590
(9,31)	-0.3091	-0.3632	-0.4554	-0.3216	-0.3168	-0.2445
(10,60)	0.1144	0.2553	*0.3619	*2.0793	*0.5574	*1.6522
(10,61)	-1.2809	-1.1431	-1.2773	-1.1289	-1.5794	-1.5408
(10,62)	0.3099	0.3685	*0.1675	* -0.9471	*0.4328	*0.9918
(10,63)	0.1212	0.1745	*0.1494	*0.7967	0.2206	0.2408

The table (4.1) reveals that the feature coefficients at Depth 10 and Nodes 60, 62 and Depth9 node 29 are the best selections for detecting a fault under heavy-load conditions knowing that feature coefficients would be changed by load increment. The features at Depth 10, Nodes 61 62 and 63 for medium and node 14 of Depth 8 for the light loadings are the best choices. Selecting an appropriate kernel is crucial in SVM classification.

Generally, the RBF and polynomial kernels are the basic kernels:

$$\text{RBF kernel: } k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

$$\text{Polynomial kernel: } k(x_i, x_j) = (1 + \gamma x_i \cdot x_j)^d$$

There are few parameters related to these kernels that should be tuned in advance: c and γ for Gaussian and c, γ and d For the RBF kernel. The parameter "gamma" controls the width of the kernel and the regularization parameter c or error penalty determines the tradeoff between minimizing the training error and minimizing model complexity. Parameter d in polynomial kernel controls the degree of the polynomial by setting it to 1 for the linear kernel, 2 for the quadratic kernel and 3 for the cubic kernel. Therefore, improper selection of parameters c, γ and d can cause over fitting or under fitting problem.

In fact, choosing an appropriate upper band (error penalty) c modifies the generalization ability of the classifier in addition to other kernel parameters. It has been noticed that, the RBF kernel is the most favorable choice of a kernel in similar works, which has been carried out so far. However, we have applied both the RBF and polynomial kernels to verify it.

Practically, we could not find a promising method to select the soft margin parameters other than trial and error.

Thus, we decided to follow the routine in variety of SVM based researches and use cross validation to find the proper kernel parameter (c, γ) although it's not cost efficient. We have tried the k-fold cross validation in which the training set is divided equally into k subsets. Then, k iterations of training and validation are performed. In each iteration, a different fold of the data is held-out for validation while the remaining k-folds are used for learning. The 10-fold cross-validation (k = 10) which has been applied in this work, is the most common cross-validation in machine learning. Furthermore, to avoid over-fitting, an independent test set is preferred in which the data is to split into two non-overlapping parts: one part for training and the other one for testing. The test data is held out and not contributed in the training procedure. This approach is called holdout cross validation which normally results in more accuracy considering the generalization ability of the classifier. Accordingly, a 10-fold hold-out cross-validation has been applied on data set #1 to find the proper kernel parameter c, d, c , and γ . In practice, all the pairs of c, γ for RBF kernel and d, c, γ for polynomial kernel obtained from cross-validation have been examined. Based on the acquired results considering both detection and accuracy rate, the best values have been selected to be used in classification. For RBF kernel, the range of parameters has been limited to $c = \{10^0, \dots, 10^3\}$ and $\gamma = \{0.3, 0.5, 1\}$. The same parameters have been also applied for polynomial kernel plus the kernel Option as $d = \{1, 2, 3\}$ as discussed before. Then after, classification tests has been carried out using data set # 2 with different loadings and performance and accuracy rate of the classifier has been inspected. The value of the kernel

parameters has been chosen randomly from the applicable ones obtained in the previous stage the overall performance results has summarized in the table (4.2).

Table (4.2) Performance of the SVM under various loading conditions

Load	Set	# All samples	#Samples - Healthy motor	#Samples- faulty motor	Training accuracy rate	Testing accuracy rate
25%	Training	20	10	10	99%	61%
	Testing	100	50	50		
50%	Training	20	10	10	99%	76%
	Testing	100	50	50		
100%	Training	20	10	10	99%	99%
	Testing	100	50	50		

It should be pointed out that in this test only the dependency of the detection rate on the loadings of the motor has been examined.

According to this table, by decreasing the load values, the accuracy rate of the classifier has been degraded which means, detecting the broken rotor bars is almost impossible when the motor has been loaded under 30% of the full load.

Based on the achieved results, further experiments only have been carried out on the induction motor, which has been heavily loaded to achieve maximum possible accuracy rate.

In the second test, besides the classifier performance rate, the number of the SVs and CPU timing (cost efficiency) has also been inspected. The result of this study summarized in Tables (4.3).

Table (4.3) performance of the standard SVM for induction motor fault detection

Kernel	Parameter	Training accuracy rate	Testing Accuracy rate	Training time	Number of SVs
RBF(c, γ)	$(10^2, 1)$	97%	97%	0.0317	10
	$(10^3, 1)$	97%	96%	0.0283	11
	$(10^3, 0.1)$	97%	97%	0.0346	7
Polynomial (d, c, γ)	$(3, 10, 0.05)$	99%	99%	0.0274	9
	$(3, 10^2, 0.05)$	99%	99%	0.0262	11

Although all possible kernel option have been examined, in this table only the kernel functions and parameters, which produced the best results considering the number of support vector, training time, training and testing accuracy rate has been presented. In overall, an excellent accuracy rate (99%) has been obtained by applying SVM.

The highest performance rate belongs to the polynomial kernel (3,10 ,0.05) corresponding to the third order polynomial with highest accuracy rate and minimum number of SVs. Considering the impact of the quantity of the SVs on the generalization ability of the classifier, which means, the generalization ability increase if the number of SVs decreases.

4.5 Fault identification based on KPCA-SVM

In the following experiment, a feature selection method has been combined with the standard SVM and the overall performance of the classifier has been tested. The main objective of applying a feature extraction algorithm such as KPCA is to reduce the dimensionality of the input feature vector in order to increase performance accuracy rate. Normally the PCA is a well-known method, mostly used in image analysis. However, high dependency of the SVM on the kernels, made us to try the kernel of the PCA in conjunction to the original SVM. Literally, by applying KPCA, the data has been mapped into the feature space twice. Once the first kernel has been used to map the input vector into the Hilbert space in order to select proper features then the second kernel has been applied to separate healthy samples from the faulty ones.

The importance of feature extraction methods such as KPCA lies beneath the overlapping nature of the input data in most cases. Due to the limitation of the available data in this work, just the overall performance of the KPCA-SVM has been observed however in more complicated cases the more significant results would be expected. Figure (4.6) depicts the features distribution acquired by KPCA.

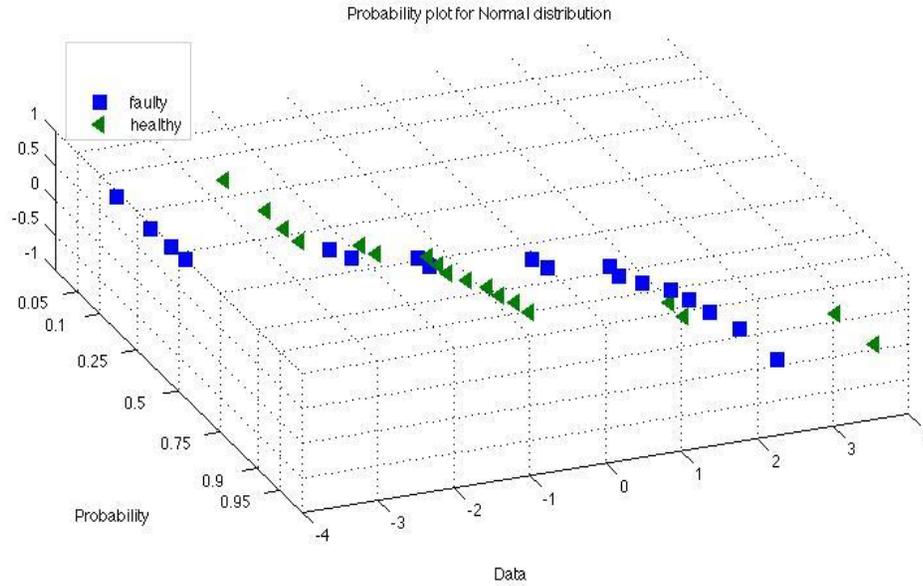


Figure (4.6) KPCA Features distribution

The appropriate SVM parameters have been selected according to the previous experiment. However the KPCA parameters have been selected based on test and trial. Consequently, five kernels showing the best performances such as RBF($10^2,1$), ($10^3,1$), ($10^3,0.1$) polynomial ($3,10,0.05$), ($3,10^2,0.05$) has been used in combination with SVM. The results of this experiment have been summarized in Table (4.4). According to which the overall performance rate achieved by KPCA-SVM is 99 %. Furthermore, it should be pointed out that the highest misclassification rate obtained by applying the RBF kernel for both KPCA and SVM. In short, the KPCA-SVM has successfully clustered the faulty samples and represented promising classification skills.

Table (4.4) Performance of the KPCA-SVM for induction motor broken bars diagnosis

Upper bound (C)	KPCA-SVM		Training accuracy rate	Testing accuracy rate
	KPCA	SVM		
C=10	RBF d=1	Poly d=1	97%	97%
	Poly d=2	Poly d=1	97%	97%
	Poly d=3	Poly d=1	99%	99%
	Poly d=2	RBF d=3	97%	97%
C=100	Poly d=3	Poly d=1	99%	99%
	Poly d=2	Poly d=1	97%	97%
	RBF d=1	Poly d=3	95%	95%
C=100	RBF d=1	Poly d=2	99%	99%

In the third part, the results obtained by KPCA-SVM have been compared with the ones gained from the standard SVM. The SVM parameters are similar to the SVM parameters in KPCA-SVM in order to be comparable. The results summarized in Table (4.5).

Table (4.5) performance comparison of the KPCA-SVM and standard SVM

Upper bound (C)	KPCA-SVM		Accuracy Rate	
	KPCA	SVM	Binary SVM	KPCA-SVM
C=10	RBF d=1	Poly d=1	89%	97%
	Poly d=2	Poly d=1	89%	97%
	Poly d=3	Poly d=1	89%	99%
	Poly d=2	RBF d=3	82%	97%
C=100	Poly d=3	Poly d=1	89%	99%
	Poly d=2	Poly d=1	89%	97%
	RBF d=1	Poly d=3	79%	95%
C=100	RBF d=1	Poly d=2	81%	99%

According to the table, the accuracy rate has been improved up to 10 %. Particularly in cases that standard SVM obtained weak results. Consequently, KPCA-SVM could be good replacement for the SVM in particular cases. However, KPCA-SVM involves with more mathematical terms, which would decrease the cost efficiency of this method.

4.6 Using ROC-SVM to identify broken bars

In the very last experiment, the performance of the ROC-SVM algorithm in broken bar identification has been investigated. To the best of our knowledge, there is no similar work available using this algorithm for industrial application and more specifically for broken rotor bars detection.

In the first part, the impact of changing the kernels parameters and tuning the upper bound penalty c for the ROC-SVM classifier has been studied to see the changes on the ROC curve.

Generally, the ROC and SVM parameters should be defined separately. However, since in this work, the focus is mainly on the SVM performance so the ROC parameters consist of the regularization parameter ρ and the size of neighborhood m have been respectively fixed to one and INF intentionally to cover up all samples and accentuates the rule of the SVM parameters in optimizing the area under the curve.

Furthermore, to be able to compare the results, the SVM parameters such as the upper bound c , the degree of polynomial d plus the width of the Gaussian (γ) for RBF kernel, selected similar to the standard SVM (summarized in Table (3.3)).

The numerical accuracy rate of the ROC-SVM classifier achieved by both polynomial and RBF kernels in five cases are presented in Table (4.6) and has been compared with the results acquired with previous algorithms, original SVM and KPCA-SVM, using the same parameters.

Table (4.6) Comparing the accuracy rate of the ROC-SVM, SVM and KPCA-SVM in identifying broken rotor bars.

Kernel	Parameters	Accuracy rate		
		SVM	ROC-SVM	KPCA-SVM
RBF (c, γ)	$(10^2, 1)$	97%	99.78%	100%
	$(10^3, 1)$	96%	99.35%	100%
	$(10^3, 0.1)$	97%	99.40%	100%
Poly (d, c, γ)	$(3, 10, 0.05)$	99%	100%	100%
	$(3, 10^2, 0.05)$	99%	100%	100%

According to the Table (4.6), accuracy rate of ROC-SVM shows has been improved in comparison to the standard SVM For all cases with both kernels. For instance for the RBF kernel $(10^3, 1)$, the accuracy rate has jumped from 96% to 99.35% which means the performance gained a 3% improvement. In conclusion, applying the exact same parameters for ROC-SVM and the original SVM, the ROC- SVM outperformed the standard SVM by decreasing the accuracy rate around 1% to 3 %, as we expected.

Moreover, KPCA-SVM shows the best performance by 100% accuracy rate ,following by the ROC-SVM and standrad SVM around 99-100 % and 96-99% respectively.

In short ,all three algorithms provide excellent accuracy rates ranging from 96 % to 100% .However ,KPCA-SVM involves with considerable mathematical complexity wich decrease the cost efficy of this method also ROC-SVM needs more tuning compared to the standard SVM.

CHAPTER 5: CONCLUSION

5.1 Conclusion

The importance of the induction motors role in the industry brings their fault diagnostic and condition monitoring into more attention. Besides, the traditional monitoring methods, which are still in effect, many companies are still faced with unexpected system failures and reduced motor lifetime. Consequently, significant research activities have been performed on replacing the traditional methods with more intelligent diagnostic methods such as neural network to be able identify the fault at early stages or even replace them. Comparing to the traditional methods, intelligent methods and specifically, data-based models are more reliable and less Time consuming.

Support Vector Machines method is a new type of the data-based models which has been recently introduced to pattern recognition. Unfortunately, despite the reliability of the SVM, the application of it in industries is still limited.

Unlike most available works based on simulated data, in this study real data has been collected from test bench; we specifically made in our laboratory in order to produce more realistic and accurate results. Knowing that there exist significant differences between the simulated data, which are measured under perfect condition and the real ones influenced by all possible environmental deficiency such as a noise or mechanical failure. The induction motor fault related data has been acquired by the use of start-up transient current signal. Motor Current Signature Analysis, has been employed to monitor the stator current. This method is a powerful monitoring tool for industrial applications, which is capable of providing nonintrusive measurements.

After monitoring the signal, advances signal processing method have been applied to extract the required information out of the monitored signal. Although, the Fast Fourier Transformers are traditionally used for this purpose but considering its flaws such as side lobe leakage and dependency on the stationary signal, made us to apply an alternative method.

Hence, to obtain more accurate results we applied a novel algorithm [17], using Wavelet Pocket Decomposing analysis to extract the features coefficients related to the existence of the fault related harmonics, which could clearly being differentiated from fundamental harmonics. Then consequently, the accuracy reached by this method is significantly higher than the ones achieved by applying the FFT or any other similar methods. Furthermore a non-stationary signal can be easily used for this purpose in contrast to the FFT.

In the first experiment, we chose the standard binary SVM as a model-based fault identifier for several reasons. First of all, there is no dimensionality limitation. Since the SVM calculations are mostly applied in high dimensional space. Second, in contrast to other available methods relying on Empirical Risk Minimization Principal the SVM implements the Structural Risk Minimization Principal by which tries to minimize an upper bound of the generalization error rather than minimizing the training error. Hence, SVM shows better generalization ability. Third, SVM seeks global optimal solution to the problem unlike other methods, which are looking for local minima. Moreover, the SVM has a sparse solution since only support vectors, which are a subset of training data points, determine the solution. Finally, the outlier can be removed from training data set while applying the optimal Lagrange multiplier in SVM training. Finally, the time elapse of the SVM is pretty short.

The first test provides promising results by achieving almost 96-99% accuracy rate. In comparison to other intelligent methods such as neural network, the SVM showing more accuracy in less time since it can be trained and tested with small number of samples.

In second test, to show the importance of feature extraction and kernel parameters selection, we trained the SVM on the data input with and without applying a feature extraction algorithm, also the criteria for choosing proper kernel parameters have been discussed.

To reduce the dimensionality of the input data, the KPCA-SVM which is a combination of SVM with a feature extraction and reduction method called KPCA has been used. In this approach the Kernel of the Principal Component Analysis, which highlights the influence of the number of component analysis towards the classification accuracies, has been studied thoroughly. Normally PCA principal component analysis is a known technique, used to reduce the dimensionality. In a similar way, the KPCA is a nonlinear PCA developed by using the kernel method. In fact, KPCA maps the original inputs into a high-dimensional feature space using the kernel method and then calculates PCA on them. Considering the fact that more number of principal components could also be extracted by applying the KPCA, a better generalization ability of the classifier is expected in comparison to the standard SVM. As we expected, application of the KPCA-SVM improved the general accuracy rate up 10 %, in cases that SVM provided poor results, however the cost efficiency of KPCA-SVM is high due to its mathematical complexity.

Finally a new algorithm called ROC-SVM has been applied. This method was supposed to improve the accuracy rate of the classifier by optimizing the area under the ROC curve called AUC. Since, this algorithm has not rarely used before, there is not enough information about tuning its parameters. Hence, we decided to just focus on its overall performance for this case

without approaching the Meta-parameters tuning criteria. Consequently, the ROC parameters selected to be fixed while the SMV parameters have been changed. The obtained result confirms that the ROC-SVM is capable of improving the accuracy rate of the SVM up to 4%. Therefore, it might be possible to reach even better results by tuning the ROC and SVM parameters at the same time.

In spite of all the promising results obtained by applying a model-based SVM classifier, some deficiency also has been observed. The quality of SVM models depend on a proper setting of SVM meta-parameters so the main issue for practitioners trying to use SVM for classification, is how to set these parameter values. One of advantages of SVM over other classification methods is its generalization ability. However, there is no proof about SVM minimizing the structural risk in high dimensional space since the SVM parameter is being selected by the cross fold validation based on examining all the available possibility. Hence, there is no promising approach to choose the SVM parameter such as the upper bound C .

Although some approaches have been made by vavnik [55] but the foreseen future is still in doubt. SVM performance absolutely depends on applying a proper kernel function, which needs a prior knowledge about the problem to select the proper kernel. Knowing that the choice of kernel function is data dependent and there are no definite rules on choosing it to get a satisfactory performance.

5.2 Future work

Our next objective would be focused on localization of the broken rotor bars, and estimating the severity, which would make it possible to predict the occurrence of the fault in advance in order to be able to replace the faulty component, or maintain it at very early stages. Moreover, further

studies on ROC-SVM application and parameter tuning could be performed since it seems to have a high potential of improving the SMV performance specifically in industrial application. Therefore, development of a theoretical framework of generalization bound on AUC and a more extensive comparison of SVM to other algorithms and on a larger amount of data sets could be set as an objective. Finally, more investigation mainly focused on providing a proper method to define the SVM meta-parameters such as upper bound is a further objective to be carried out.

REFERENCES

- [1] M.S.Sarma, *Electrical machines Steady-state theory and dynamic performance*, second edition: Thomson learning, 1994.
- [2] C.I.Huber, *Electric machines theory, operation, application, adjustment and control*, Second edition, New Jersey: Prentice hall, 2002
- [3] S.Nandi, H.A.Toliyat and L. Xiaodong,” Condition monitoring and fault diagnosis of electrical motors-a review,” *IEEE Transaction on energy conversion*, vol.20, no.4, pp.719-729, Dec.2005.
- [4] Y.Zhongming and B.Wu,” Induction Motor Mechanical Fault Simulation and Stator Current Signature Analysis,” *IEEE Transaction on Power System Technology* ,vol.2, no.2, pp.789-794, Aug 2002.
- [5] M. Haji, and H.A. Toliyat, “Broken rotor bar detection in induction machines with transient Operating Speeds,” *IEEE Transactions on Energy Conversion*, vol.16, no.4, pp. 312-317, Dec. 2001.
- [6] J. Penman, M. N. Dey, A. J. Tait, and W. E. Bryan, “Condition monitoring of electrical drives,” *IEEE Transaction on Electrical Power Technology*, vol. 133, no.3, pp. 142-148, Nov.2008.
- [7] R. Supangat,E.Nesimi,L.S.Wen,A.G.Douglas,H.Colin and G.Jason ,” Broken Rotor Bar Fault Detection in Induction Motors Using Starting Current Analysis, “ *IEEE European Power Electronics and Applications conference* , Dresden , Germany , 2005 .
- [8] H.Douglas, P.Pillay and A.K.Ziarani, “Pattern Recognition – A Technique for Induction Machines Rotor Broken Bar Detections,” *IEEE Transactions on Energy Conversion*, vol. 16, no. 4, pp. 312-317, Dec. 2001.
- [9] B.Ayhan, M. Chow and M. Song, Member “ Multiple Signature Processing-Based Fault Detection Schemes for Broken Rotor Bar in Induction Motors, ” *IEEE Transaction on Energy Conversion* , vol. 20 , no. 2 , pp. 336-343 , May. 2005.

- [10] F. Filippetti, G. Franceschini, and C. Tassoni, “Neural networks aided on-line diagnostics of induction motor faults,” *In Proceeding IEEE Industry Applications Society Annual Meeting Conference*, vol. 1, pp. 316-323, Aug.2002.
- [11] F. Filippetti, G. Franceschini, C. Tassoni and P. Vas, “A fuzzy logic approach to on-line induction motor diagnostics based on stator current monitoring,” *in Proceedings of Stockholm Powertek*, Stockholm, Sweden, pp. 150–161, June 1995.
- [12] Z. Ye, A. Sadeghian and B. Wu, “Mechanical fault diagnostics for induction motor with variable speed drives using Adaptive Neuro-fuzzy Inference System,” *IEEE Transaction on Electric Power research*, Amsterdam, vol.76, no.9-10, pp.742-752, June.2006.
- [13] G. Betta, C. Liguori and A. Pietrosanto, “The use of genetic algorithms for advanced instrument fault detection and isolation schemes,” *In Proceedings of IEEE conference on Instrumentation and Measurement Technology Conference*, Brussels, Belgium, vol.2, pp. 1129–1134, Aug.2002.
- [14] V.N.Vapnik, *The Nature of Statistical Learning Theory*, NewYork: John Willey, 2000.
- [15] N.Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*: Cambridge University Press, 2000.
- [16] B. Scholkopf and A.J. Smola, *Learning With Kernels*, The MIT Press, Cambridge, 2001.
- [17] Sadeghian, Zhongming Ye, and Bin Wu, ‘Online Detection of Broken Rotor Bars in Induction Motors by Wavelet Packet Decomposition and Artificial Neural Networks’, *IEEE Trans. on Instrumentation and Measurement*, vol.58, no.7, pp2253-2263, July 2009.
- [18] R. R. Schoen and T. G. Habetler, “Effects of Time-Varying Loads on Rotor Fault Detection in Induction Machines,” *IEEE Transaction on Industrial Applications*, vol.31, no.4, pp.900–906, Aug.2002.
- [19] F.Filipetti, G.Franceschini, C.Tassoni, and P.Vas, “Recent developments of induction motor drives fault diagnosis using AI techniques,” *IEEE Transaction on Industrial Electronics*, vol.47, no.5, pp. 994-1004,2000,.

- [20] M. E H. Benbouzid, “A Review of Induction Motors Signature Analysis as a Medium for Faults Detection,” *In 24th IEEE conference on Industrial Electronics*, vol. 47, no. 5, pp.1950-1955, Oct. 2000.
- [21] S.V.Vaseghi, *advanced digital signal processing and noise reduction* ,4th edition: john Wiley, London,UK,2008.
- [22] R.R. Schoen, B.K. Lin,T.G.Habetler, J.H. Schlag, and S.Farag, “An unsupervised, on-line system for induction motor fault detection using stator current monitoring ,” *IEEE Transaction on Industry Application* ,vol.31,no.6,pp.1280-1286 ,Aug.2002.
- [23] A.M. Knight and S.P. Bertani,” Mechanical Fault Detection in a Medium-Sized Induction Motor Using Stator Current Monitoring,”*IEEE Transaction on Energy Conversion*, vol.20, no.4, pp753-760, Dec. 2005.
- [24] M. J. Mohlenkamp and M. C. Pereyra , *Wavelets, Their Friends, and What They Can Do for You*,Zimmermann Freiburg ,germany,2008.
- [25] V.Venkatasubramanian , R.Rengaswamy and S.Kavuri. “A review of process fault detection and diagnosis: Part I: Quantitative model-based methods,” *Computer and Chemical Engineering*, vol.27, no.3, pp.313–326, March2003.
- [26] A.M. Knight and S.P. Bertani,” Mechanical Fault Detection in a Medium-Sized Induction Motor Using Stator Current Monitoring,” *IEEE Transaction on Energy Conversion*, vol.20, no.4, pp753-760, Dec. 2005.
- [27] C. Combastel, S. Lesecq, S. Petropol, and S. Gentil, “Model-based and wavelet approaches to induction motor on-line fault detection,” *Control Engineering* , vol. 10, no. 5, pp. 493–509 , May 2002.
- [28] T.I. Liu, J.H. Singonahalli and N.R. Iyer, “Detection of roller bearing defects using expert system and fuzzy logic,” *Mechanical Systems and Signal Processing*, vol.10, pp.595–614, 1996.

- [29] V.Venkatasubramanian , R.Rengaswamy and S.Kavuri. “A review of process fault detection and diagnosis: Part III: Process history based methods,” *Computer and Chemical Engineering*, vol. 27, no.3, pp.327–346, March2003.
- [30] H.Nejjari and M.H.Benbouzid,” Monitoring and Diagnosis of Induction Motors Electrical Faults Using a Current Park’s Vector Pattern Learning Approach,” *IEEE Transaction on industry applications*, vol. 36, no. 3, pp730-735, Aug. 2003.
- [31] V. N. Ghate and S. V. Dudul ,” Fault Diagnosis of Three Phase Induction Motor Using Neural Network Techniques ,” *2d international IEEE conference on Emerging Trends in Engineering and Technology* , Nagpur ,India, pp. 922-928 , Jan 2010.
- [32] F. Zidani, M. E. H. Benbouzid, D. Diallo, and M. S. Nait-Said, "Induction motor stator faults diagnosis by a current Concordia pattern-based fuzzy decision system,” *IEEE Transaction on Energy Conversion*, vol. 18, no. 4, pp. 469-475 ,Nov.2003.
- [33] I.E.Chabu and C.G.Dias ,“ A fuzzy logic approach for the detection of broken rotor bars in squirrel cage induction motors,” *In proceeding IEEE International conference on Fuzzy Systems* , pp. 1987–1991 ,Sept.2008.
- [34]] H. Nejjari, M.E.H. Benbouzid, “Application of fuzzy logic to induction motors condition monitoring,” *IEEE Transaction on Power Engineering* ,vol. 19,no.6 ,pp. 52–54 ,Aug.2002.
- [35] A. Widodo ,B.Yang , D. Gu and B. Choi ,” Intelligent fault diagnosis system of induction motor based on transient current signal ,”*Mechatronics* ,vol.19,no.5, pp.680-689, Aug.2009.
- [36] A. Widodo, B. Yang, “Wavelet support vector machine for machine faults classification,” *Lecture Series on Computer and Computational Sciences*, vol.8, pp. 1-5.2007
- [37] A. Widodo and B.Yang, “Review support vector machine in machine condition monitoring and diagnostics,” *Mechanical System and Signal Processing*, vol.21, pp.2560-2574, 2007.

- [38] S.Poyhonen,P. Jover and H.Hyotyniemi “ Signal Processing of Vibrations for Condition Monitoring of an Induction Motor ”,*In Processing of the 1st IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing*, Hammamet, Tunisia, ,pp. 499-502, March2004.
- [39] S.Poyhonen,P. Jover and H.Hyotyniemi ,”Coupling Pairwise Support Vector Machines for Fault Classification,” *In Proceeding of the 5th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*, Safeprocess2003, Washington, D.C., USA, pp. 705-710, June.2003.
- [40] V.T.Tran, B.S .Yang, M.Suck Oh and A. C. Chiow Tan ,”Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference,” *Expert Systems with Applications*, vol. 36, no.2, pp. 1840-1849, March 2009.
- [41] S. Altug, M.Y. Chow, H.J. Trussell, “Fuzzy inference systems implemented on neural architectures for motor fault detection and diagnosis,” *IEEE Transaction on Industrial Electronics*, vol.46, no.6, pp.1069–1079, Dec.1999.
- [42] W.X.Hong and H.Y.Gang,” Fuzzy Neural Network based On-line Stator Winding Turn Fault Detection for Induction Motors,” *2nd IEEE conference on industrial electronics and applications (ICIEA)*, Harbin, china, pp. 2461-2464, May 2007.
- [43] X.Z. Gao, S.J. Ovaska and Y. Dote, “Motor fault detection using Elman neural network with genetic algorithm-aided training ,” *In Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, Nashville, TN,vol.4, pp. 2386–2392, Aug 2002.
- [44] J.Arrillaga and N.R.Watson ,*power system harmonics* ,second edition :john Wiley ,UK ,2003
- [45] D.F.Walnut ,*an introduction to wavelet analysis* ,second edition : birkhauser ,Boston, USA,2004
- [46] S.Mallat,” A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.11, no.7, pp. 674-693, 1989

- [47] Y. Gao and Q.Zhang, "A Wavelet Packet and Residual Analysis Based Method for Hydraulic Pump Health Diagnosis," *In Proceedings of the Institution of Mechanical Engineers*, vol.220, no.6, pp.735–745, Nov.2006.
- [48] Z. Ye, B. Wu, and A.R. Sadeghian, "Signature analysis of induction motor mechanical faults by wavelet packet decomposition," *In Proceeding of IEEE Confrence on application of Power Electronics*, pp. 1022–1029, 2001.
- [49] Z. Ye, B. Wu, and A. R. Sadeghian, "Current signature analysis of in- duction motor mechanical faults by wavelet packet decomposition," *IEEE Transaction on Industrial Electronics*, vol. 50, no. 6, pp. 1217–1228, Dec. 2003.
- [50] C.J.C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp.121-167, Kluwer Academic Publishers, 1998.
- [51] A. Widodo, B.S. Yang and T. Han, "Combination of independent component analysis and support vector machines for intelligent faults diagnosis of induction motors," *Expert Systems with Applications*, vol. 32 , no.2 , pp. 299–312,2007
- [52] A. Rakotomamonjy ,” Optimizing area under ROC curves with SVMs ,”*first Workshop on ROC anlysisin AI* , Valencia, Spain, August. 2004.
- [53] S. Abbasion , A.Rafsanjani , A.Farshidianfar , N.Irani ,”Rolling element bearings multi-fault classification based on the wavelet denoising and support vector machine,”*Mechanical system Signal Process* ,vol.21, no.7, pp.2933–2945, Oct.2007.
- [54] S.N.S.Mohamed , M.E.H. Benbouzid and A.Benchaib,” Detection of Broken Bars in Induction Motors Usingan Extended Kalman Filter for Rotor Resistance Sensorless Estimation,” *IEEE on Energy Conversion* ,vol.15,no.1,pp.66-70,Aug.2002.
- [55] O.Chapelle ,V.Vapnik,O. Bousquet and S.Mukherjee ,” Choosing Multiple Parameters for Support Vector Machines,” *Machine Learning*, vol. 46, no.1-3, pp. 131-159, 2 January 2002.