APPLICATION OF MACHINE LEARNING TECHNIQUE IN WIND TURBINE FAULT DIAGNOSIS

Afrooz Purarjomandlangrudi

B.Sc. (Electrical Engineering)

Principal supervisor: Dr Ghavameddin Nourbakhsh

Submitted in fulfilment of the requirements for the degree of
Master of Engineering (Research)

Science and Engineering Faculty
Queensland University of Technology
2014
KEYWORDS

Wind turbine, renewable energy, fault detection, condition monitoring, fault diagnosis, rotating components, gearbox, bearing, machine learning, support vector machine, anomaly detection, acoustic emission technique, and data mining.
ABSTRACT

With the increasing demand for electric power, environmental regulations are putting restrictions on the use of thermal power plants and renewable energy sources; in particular, wind farm energy turbines are becoming very popular around the world. As a result, wind turbine availability and the ability to accurately predict faults in advance have become very critical in this industry. Unpredicted failures of an element in a wind turbine, particularly in low speed rotating components such as gearboxes and bearings, can lead to major financial drawbacks. One of the most efficient approaches to prevent catastrophic failures and unplanned outages is by using Condition Monitoring (CM). Although a variety of CM techniques have been used recently, their applications in the power industry are still relatively new. In addition, most CMs require a large number of fault indicators to accurately diagnose the component faults.

Learning techniques can be employed to overcome such problems in CM, as the definition of machine learning is the ability of a program or system to learn, improve and develop its efficiency over time. Machine learning techniques focus on creating a system that improves its performance based on previous results and historical data instead of understanding the process that generated the data. In fact, the machine learning paradigm provides the ability
of changing execution strategy based on newly acquired information from a system. Learning algorithms can be useful in different applications such as prediction of the future value, clustering and detection of anomaly behaviour in the data.

In this study, two learning algorithms called anomaly detection and Support Vector Machine (SVM) are employed to bearing fault diagnosis and CM. Basically the anomaly detection algorithm is used to recognize the presence of unusual and potentially faulty data in a dataset, which contains two phases: a training phase and a testing phase. In the former, the algorithm is trained with a training dataset and in the latter; the learned algorithm is applied to a set of new data. Two bearing datasets were used to validate the proposed technique, fault-seeded bearing from a test rig located at Case Western Reserve University to validate the accuracy of the anomaly detection method. Detecting faults and defects in their early stages is one of the most important aspects of machine CM. The second dataset was a test to failure data of bearings from the NSF I/UCR Centre for Intelligent Maintenance Systems (IMS) which was used to compare anomaly detection with a previously applied method (SVM) for finding the time incipient faults.
List of Publications

Journal papers:


Conference papers:

STATEMENT OF ORIGINAL AUTHORSHIP

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: QUT Verified Signature

Date: 10/03/2014
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my father and mother, Mehdi Purarjomand and Azam Bodaghi, for their love and unwavering support throughout my education, from the 23rd of September 1992, when my mum walked me to the school where I started my primary education, to the present day in 2014, when I am finishing my master’s degree. I attribute whatever achievement I have achieved or will achieve in my life to them. I am also very thankful to my younger sister, Ema Purarjomand, for her love and kindness, and for being with my parents while I have been away.

I would like to gratefully and sincerely thank my supervisor, Dr. Ghavameddin Nourbakhsh for believing in my work and for providing insightful advice and support during all stages of my master’s journey. No words can do justice to my appreciation of his nurturing support and attention. His time, guidance and encouragement have made all the difference. My sincere thanks also go to Professor Andy Tan, whose expertise, understanding, and patience have added considerably to my graduate experience. I appreciate the vast knowledge and skill in many areas that he has shared with me.
I would like to thank the editors and anonymous reviewers of the various journals in which I have published articles associated with this thesis for their precious time in reviewing my works and for their valuable comments and suggestions. As well, professional editor, Ms Diane Kolomeitz, has provided copyediting and proofreading services, according to the guidelines laid out in the University-endorsed national policy guidelines, ‘The editing of research theses by professional editors’ (available at http://iped-editors.org/About_editing/Editing_theses.aspx)

I would also like to thank my other good colleagues and co-authors for their support and insightful suggestions throughout this journey: Dr Amir Hossein Ghapanchi, Dr Mohhamad Esmalifalak and Dr Houman Ghaemmaghami.
# TABLE OF CONTENTS

**CHAPTER 1: INTRODUCTION** ............................................................................................ 1  
1.1 GENERAL INTRODUCTION...................................................................................... 1  
1.2 WIND TURBINE COMPONENTS AND FAILURES ................................................ 3  
1.3 WIND TURBINE CONDITION MONITORING AND RESEARCH QUESTIONS 5  
1.4 RESEARCH PROBLEM........................................................................................... 7  
1.5 OBJECTIVE OF RESEARCH ................................................................................. 9  
1.6 OVERVIEW OF RESEARCH METHODOLOGY.................................................... 11  
1.7 THESIS PRESENTATION AND STRUCTURE....................................................... 13  

**CHAPTER 2: PAPER 1- FAULT DETECTION IN WIND TURBINE: A SYSTEMATIC LITERATURE REVIEW** ........................................................................... 15  
2.1 INTRODUCTION........................................................................................................ 17  
2.2 LITERATURE REVIEW ............................................................................................ 21  
2.2.1 Gearbox and Bearing ......................................................................................... 23  
2.2.2 Power Electronics and Electrical Control Failures.......................................... 24  
2.2.3 Generators .......................................................................................................... 25  
2.3 RESEARCH METHODOLOGY .......................................................................................... 26  
2.3.1 Resources Searched ............................................................................................ 27  
2.3.2 Search terms ....................................................................................................... 27  
2.3.3 Inclusion/Exclusion Criteria .............................................................................. 28  
2.3.4 Data Analysis ...................................................................................................... 29  
2.4 LITERATURE REVIEW FINDINGS AND RESULTS ............................................................ 30  
2.5 CONCLUSION................................................................................................................. 36  

**CHAPTER 3: PAPER 2- WIND TURBINE CONDITION MONITORING USING MACHINE LEARNING TECHNIQUES** ............................................................................. 38  
3.1 INTRODUCTION........................................................................................................ 41  
3.2 FEATURE EXTRACTION ......................................................................................... 43  
3.2.1 Kurtosis............................................................................................................... 44  
3.2.2 Non-Gaussianity Score (NGS) feature ............................................................... 45  
3.3 MACHINE LEARNING APPROACHES .................................................................. 45  
3.3.1 Support Vector Machine (SVM) ......................................................................... 46  
3.3.2 Anomaly detection .............................................................................................. 47  
3.4 EXPERIMENTAL RESULTS..................................................................................... 50  
3.4.1 Model description ............................................................................................... 51  
3.5 CONCLUSION............................................................................................................ 58  

**CHAPTER 4: PAPER 3- APPLICATION OF ANOMALY TECHNIQUE IN WIND TURBINE BEARING FAULT DETECTION** ..................................................................... 60  
4.1 INTRODUCTION........................................................................................................ 62  
4.2 MACHINE LEARNING APPROACHES ................................................................. 66  
4.2.1 One-class Support Vector Machine ................................................................. 66  
4.2.2 Anomaly Detection (AD) .................................................................................... 67  
4.3 EXPERIMENTAL RESULTS..................................................................................... 69
4.3.1 Model description ........................................................................................................ 70  
4.4 CONCLUSIONS .............................................................................................................. 74  

CHAPTER 5: CONCLUSIONS ......................................................................................... 76  
5.1 OVERVIEW ................................................................................................................. 76  
5.2 SUMMARY OF FINDINGS ........................................................................................... 76  
5.3 ADDRESSING RESEARCH QUESTIONS AND CONCLUSION .............................. 79  
5.4 IMPLICATIONS AND FUTURE WORKS .................................................................... 81  
5.4.1 Implications for Industry Practitioners ................................................................. 82  
5.4.2 Implications for Researchers ............................................................................... 83
LIST OF TABLES

TABLE 2.1 ABBREVIATION........................................................................................................... 20
TABLE 2.2. NUMBER OF PAPER EXCLUDE IN EACH STEP.......................................................... 29
TABLE 5.1. AD AND SVM F1 MEASURE FOR BEARING COMPONENTS................................. 77
LIST OF FIGURES

FIGURE 1.1. WIND POWER CAPACITY INSTALLATION FROM AWEA [2] ................................. 2
FIGURE 1.2. FAILURE FREQUENCY AND DOWNTIMES OF COMPONENTS [4] ......................... 5
FIGURE 2.1 THE MAJOR COMPONENT OF A WIND TURBINE .................................................. 22
FIGURE 2.2 FAILURE RATE OF WIND TURBINE COMPONENTS ............................................... 23
FIGURE 2.3 TOOTH BREAKAGE CAUSED BY FREQUENT STOPPING AND STARTING .................. 24
FIGURE 2.4 CONTAMINATION IN A TYPICAL WIND TURBINE .................................................. 26
FIGURE 2.5 STAGES OF THE RESEARCH METHODOLOGY ....................................................... 28
FIGURE 2.6 FREQUENCY OF PAPERS PER YEAR ................................................................. 31
FIGURE 2.7 FREQUENCY OF PAPERS PER CONTINENT .......................................................... 31
FIGURE 2.8 FREQUENCY OF PAPERS PER COUNTRY ............................................................... 33
FIGURE 2.9 FAULT DETECTION TECHNIQUES CLASSIFICATION .......................................... 34
FIGURE 3.1 PHOTOGRAPHY AND SCHEMATIC DESCRIPTION OF THE TEST RIG ................. 51
FIGURE 3.2 VISUALIZATION OF THE PROPOSED ANOMALY DETECTION METHOD FOR AUTOMATIC BEARING FAULT DETECTION ................................................................. 54
FIGURE 3.3 INNER RACE FAULT F1 SCORE TREND FOR 0.007 INCHES, (A) 0 HP AND (B) 1 HP . 55
FIGURE 3.4 OUTER RACE FAULT F1 SCORE TREND FOR 0.007 INCHES, (A) 0 HP AND (B) 1 HP 56
FIGURE 3.5 BALL FAULT F1 SCORE TREND FOR 0.007 INCHES, (A) 0 HP AND (B) 1 HP ......... 57
FIGURE 4.1 ROLLING ELEMENT BEARING COMPONENTS ....................................................... 64
FIGURE 4.2 BEARING TEST RIG AND SENSOR PLACEMENT [62] .......................................... 71
FIGURE 4.3 SVM OUTPUT [83] ................................................................................................. 73
FIGURE 4.4 ANOMALY DETECTION OUTPUT ........................................................................... 74
Chapter 1

INTRODUCTION

1.1 GENERAL INTRODUCTION

Harnessing wind power to generate electricity through wind turbines has gained popularity in recent years. Wind energy is a well-regarded renewable resource due to its abundant availability and environmentally friendly features. According to the European Wind Energy Association (EWEA), each year millions of tonnes of carbon dioxide contribute to climate change and global warming through the burning of fossil fuels (oil, coal and gas). In 2011 EWEA estimated that wind energy had cut carbon emission by 140 million tonnes in the EU continent, which is equivalent to taking 33% of cars in the EU (71 million vehicles) off the road. This reduction in carbon emission has resulted in cost savings of around €1.4 billion [1].

In terms of economy, it was reported in 2010 that onshore wind turbine electricity cost €64.9 MW/h (less than coal at €67.6). By 2020 the gap is predicted to be even wider, estimated at €80.3 for coal and €57.41 for wind.
Introduction

The cost of wind power production can be predicted with a high degree of accuracy, whereas oil, gas and coal prices are subjected to market environment and are expected to increase. For instance the oil price has increased over the past few years from $20 to over $100 and has added $45 billion to the EU’s annual gas import bill. According to the new American Wind Energy Association (AWEA) industry report, the U.S. wind industry’s 45,125 operational utility-scale turbines represent an installed rated capacity of 60,007 Megawatts. That is equivalent to 60 nuclear power plants [2]. Figure 1.1 depicts the wind power capacity installation by quarter in the U.S. from 2008 to 2012. The bar chart illustrates a boost in the 4th quarter in 2012 by 8,385 MW from 4,106 in 2008.

Figure 1.1. Wind power capacity installation from AWEA [2]
As an electricity generator, there are different factors which can influence the wind turbine’s output, such as turbine size and wind speed. An average onshore wind turbine with a capacity of 2.5–3 MW can produce more than 6000 MWh in a year. An average offshore wind turbine of 3.6 MW can power more than 3,312 average households [1]. Wind turbines operate under different wind speed, ranging from 4 to 5 m/s to a maximum of around 15 m/s.

A modern wind turbine has variable outputs depending on the location and wind speed, but generally it generates electricity at 70-85% of the time. It will typically produce about 24% of its rated power (41% offshore) over a year. Since wind turbines generally work in harsh environments with highly variable wind speed, they normally experience several downtimes in a year for maintenance or breakdowns. The downtimes account for the capacity factor of power plants to be in the range of 50%-80%.

1.2 WIND TURBINE COMPONENTS AND FAILURES

Wind turbines consist of various components and the four main parts are: the base, tower and foundation, nacelle, and rotor and rotor blades. The base is made of concrete reinforced with steel bars and there are two types of design for them, shallow flat disk and deeper cylinder. Based on the consistency of the underlying ground, a pile or flat foundation is applied for stability and rigidity of a wind turbine. Typically, towers are designed as a white steel
Introduction
cylinder, about 150 to 200 feet tall and 10 feet in diameter [3]. The tower
construction not only carries the weight of the nacelle, rotor and blades; it also
absorbs static loads created by wind power variation.

The blades capture the wind's energy, spinning a generator in the nacelle.
Their principle is the same as lift, that is, the passing air causes more pressure
on the lower side of the wings and the upper side creates a pull. With the help
of the rotor, the energy in the wind is converted to rotary mechanical
movement.

The nacelle holds all the turbine machinery and contains different components
such as the main axle, gearbox, generator, transformer and control system. The
nacelle is connected to the tower through bearings in order to rotate and follow
the wind direction. Generators convert mechanical energy to electrical energy.
They have to work with a power source (the wind turbine rotor) which
supplies highly fluctuating mechanical power (torque). There are two types of
generators, fixed speed generators and variable speed generators that generate
electricity at a varying frequency to take advantage of different wind speed.

The normal lifetime of wind turbines is 20 years but there is no final statement
regarding actual life expectancy of modern wind turbines [4]. Some features
such as failure rate and downtime can be used to estimate lifetime. The failure
downtimes have different duration and depend on the required repair work,
Introduction

which may last for several weeks. Figure 1.2 illustrates different parts of the wind turbine that contribute to its downtimes and the frequency of them. Different types of failures and their causes are discussed extensively in Chapter 2.

![Failure Frequency and downtimes of components](image)

Figure 1.2. Failure Frequency and downtimes of components [4].

1.3 WIND TURBINE CONDITION MONITORING AND RESEARCH QUESTIONS

Condition Monitoring (CM) and fault diagnosis are critical aspects of wind turbine safety and reliability, which aim to decrease the failure rate and downtimes described in the previous section. Gearbox and bearing faults are
one of the foremost causes of failures in rotating mechanical systems (40–50% in wind turbines [5]), for they include some or numerous bearings to provide smooth rotation with minimal losses, and their faults can be directly contributed to consecutive problems in other major components.

Since the time to principal failures varies for inner race, outer race, ball, and rolling element, the accuracy and sensitivity of the maintenance techniques are essential in detecting incipient faults in bearings. The majority of existing works have focused on classified fault types on the basis of availability of fault samples; in practice collecting all types of faulty data from bearing defects is very difficult if not impossible. This is due to the fact that some components occur very occasionally and also each type of machine has specific failure vibration patterns [6-8].

Some previous studies have overcome the problem by applying data-mining algorithms and machine learning classification technologies, which use a historical database of the system to predict failures. Among the various methods that have been used in machine learning, artificial neural networks (ANN) have experienced the fastest development over the past few years [9]. Nevertheless, there are some drawbacks with neural networks, such as structure identification difficulties, local convergence, and poor generalization
abilities, since they originally applied for Experienced Risk Minimization (ERM).

Support Vector Machines (SVM), were found to offer a better solution to overcome the disadvantages mentioned in [10, 11] and rapidly became the centre of attention in recent research activities. Basically, the SVM algorithm deals with binary classification of problems. However, various kinds of SVM fault classifications suffer from huge amounts of computation, which causes some restrictions. Anomaly detection, however, can detect faults with fewer amounts of data and also is able to detect new defects, which may not exist in historical data sets of the system.

Due to the fact that in many practical systems data collection is limited, access to this information is not always possible. With this in mind, this research work explores the development of design techniques which require limited data and accurately predict incipient defects.

1.4 RESEARCH PROBLEM

The efficiency, maintenance and downtime costs of the wind turbine could be improved by implementing condition monitoring based on accurate and prompt detection of incipient faults. Research in fault diagnosis and condition monitoring is highly important in wind turbines. Therefore, condition
monitoring and fault diagnostics systems (CMFDS) for wind turbines are critical in establishing condition-based maintenance and repair.

Various methods have been applied for fault detection of wind turbines, such as vibration analysis [12-16], oil analysis [17-19], noise analysis [20], [8], data analysis [20-24] and acoustic emission (AE) analysis [25, 26]. To keep the wind turbine in operation, performance of the condition monitoring system (CMS) and fault detection system (FDS) is paramount and for this reason extensive knowledge of these two types of systems is mandatory. The condition monitoring system (CMS) plays a vital role in establishing condition-based maintenance and repair (M&R), which can be more effective than corrective and preventive maintenance. For this purpose, it needs to develop effective fault prediction algorithms and these algorithms would be the basis of CMS. Autonomous online CMSs with integrated fault detection algorithms could detect any mechanical and electrical defects in very early stages to prevent major component failures [27].

For many engineering and science problems, there is no direct mathematical solution. Learning techniques have been used extensively to overcome this problem. Researchers in different fields try to develop algorithms that learn the behaviour of the given problem using historical data [28], [29], [30]. Learning algorithms can be used for different applications such as prediction of future
value and detection of anomaly behaviour in the data. In this research, data analysing would be carried out based on machine-learning methods and using supervised machine-learning algorithms.

1.5 OBJECTIVE OF RESEARCH

For many engineering and science problems there are no direct mathematical solutions. Learning techniques have been used extensively to overcome this problem. Researchers in different fields try to develop algorithms that learn the behaviour of a given problem using historical data [28, 30]. Learning algorithms can be used in different applications such as the prediction of a future value, and clustering and detection of anomaly behaviour in a data [31]. Machine learning provides the ability to learn without being explicitly programmed for systems. This technique is based on computer programs that are able to establish learning formation with training-based algorithms to find patterns in data where programs can detect discrepancies and act according to set of perceived criteria.

In machine learning, anomaly detection, also called outlier detection, is the recognition of observations which do not conform to an expected pattern in a dataset[32]. Anomalies are mainly referred to as outliers, novelties, noise, deviations and exceptions [33]. There are three main categories of anomaly detection technique, namely, unsupervised, supervised and semi-supervised.
Unsupervised anomaly detection techniques find anomalies in an unlabelled dataset. These techniques consider the majority of data as normal data and look for samples that seem to fit least to them. Supervised anomaly detection techniques require a labelled dataset as “normal” and “abnormal” and need to train a classifier. Finally, semi-supervised anomaly detection techniques are designed to look for deviations from a labelled sample of normal data.

Given the failure developments in wind turbine bearings, this research study proposes a fault diagnosis method, based on supervised anomaly detection techniques to create models of normal data, and then attempts to detect abnormalities from the normal model in the observed data. Hence, the anomaly detection algorithm is able to recognize the majority of new types of intrusion [34, 35]. However, this method needs a purely normal data set to train the algorithm. The algorithm may not recognize future failures and will assume they are normal if the training data set includes the effects of the intrusions. The aforementioned feature contributes to diagnosing faults and fatigues in their early stages, and because of the high sensitivity of its nature this method is extremely rigorous in comparison with previous techniques.

The main objectives of this research are to:
• Thoroughly investigate various types of fault in the different rotating components of a wind turbine and become familiar with bearing condition monitoring techniques.

• Employ two vibrational data sets; one is the seeded fault of a different size and the second one is a test to failure experiment [36, 37].

• Analyze these data and implementing machine learning algorithms for detecting faults and anomalies.

• Interpret the results and compare the output of each algorithm to find out the most effective way to detect incipient faults and defects in their early stages.

1.6 OVERVIEW OF RESEARCH METHODOLOGY

According to the recent investigations [38] the most faults found in wind turbines are in the rotating components, especially the bearings, which are of great importance within the others components. Wind turbine downtimes and failure are fully-described in Chapter 1, which shows that 25% of the total of wind turbine downtimes are due to gearbox and bearing failures [18]. Main shaft/bearing and rotor are also important factors in wind turbine failure with
the percentages of 17% and 15% respectively. Therefore, bearings’ CM should be taken into account in wind turbine fault diagnosis and condition monitoring, using the most efficient method to detect incipient faults and failures to enhance system operation.

In machine learning, unsupervised learning refers to the types of algorithms that try to find correlations without any external inputs other than the raw data, trying to find hidden structure in unlabeled data. Supervised learning is when the algorithm input data is "labeled" to help the logic in the code to make suitable decisions. Based on the wind turbine bearing characteristic discussed earlier, in this research work, the supervised learning technique was found best fit to detect faults and defects of rotating components of a wind turbine such as bearings, according to the following steps.

**Step 1:** Conduct a comprehensive literature review of wind turbine components and their failures. Investigate various fault detection techniques and acquire the knowledge of how the techniques work to bearing condition monitoring of wind turbine. Also investigate sensors and their characteristics in using them for this application. Learn data collection and signal analysis, in preparation to use these for vibration analysis.

**Step 2:** Employing two sets of data; the first is a bearing data set with seeded fault in different size and load from Case Western Reserve University test rig
and the second is a test to failure bearing data set from IMS, University of Cincinnati, NASA Ames Prognostics Data Repository, Rexnord Technical Services.

**Step 3**: Analyzing the vibrational signals to extract the relevant features associated with the defects and prepare the features for use in learning techniques.

**Step 4**: Applying supervised machine learning methods, Support Vector Machine (SVM) and Anomaly Detection (AD) algorithms to compare the methods and find the pros and cons of the different techniques.

**Step 5**: Writing up the findings of the research in the form of journal and conference papers, and finally write up the Master thesis.

### 1.7 THESIS PRESENTATION AND STRUCTURE

The organization of this thesis follows QUT rules (which can be found at [www.rsc.qut.edu.au](http://www.rsc.qut.edu.au)) for Master by Research by Publication, which authorises examiners to examine the thesis based on the presentation of relevant published or submitted manuscripts for the body of the work, with introduction and conclusion chapters. The chapters of this thesis are arranged as follows:
Chapter 2, paper 1, provides a systematic literature review in wind turbine fault detection. Different wind turbine components and their failure are discussed in this chapter. All the relevant search terms and data bases applied in this literature review are in this chapter.

Chapter 3, paper 2, provides the implementation of machine learning and anomaly detection techniques in different parts of the bearing data set with various fault sizes to examine the anomaly detection technique operation in fault detection in terms of accuracy and precision.

In Chapter 4, paper 3, a test to failure of a real bearing data set is utilized to validate the anomaly detection technique used in this research to find an incipient fault and compare it with the state-of-the-art SVM technique to validate the anomaly detection technique ability in terms of rapidity of detecting incipient fault.

Chapter 5 draws conclusions on results found, by analysing the proposed method in resolving the research problems. This chapter also contains discussions related to the application of the proposed method, for practitioners and researchers.
Chapter 2

PAPER 1- FAULT DETECTION IN WIND TURBINE: A SYSTEMATIC LITERATURE REVIEW

Afrooz Purarjomandlangrudi, Ghavameddin Nourbakhsh,
Mohammad Esmalifalak, Andy Tan

School of Power Engineering, Science and Engineering Faculty,
Queensland University of Technology
School of Electrical and Computer Engineering, University of Houston, Houston
School of Mechanical Engineering, Science and Engineering Faculty,
Queensland University of Technology

Journal of Wind Engineering volume 37, NO. 5, 2013 PP 535-546
Statement of Contribution of Co-Authors for Thesis by Published Paper

The following is the format for the required declaration provided at the start of any thesis chapter which includes a co-authored publication.

The authors listed below have certified* that:
1. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
2. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
3. there are no other authors of the publication according to these criteria;
4. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
5. they agree to the use of the publication in the student’s thesis and its publication on the QUT ePrints database consistent with any limitations set by publisher requirements.

In the case of this chapter:

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Statement of contribution*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afrooz Purarjomandlangrudi</td>
<td>Wrote the manuscript, experimental design, conducted experiments, and data analysis</td>
</tr>
<tr>
<td>Ghavameddin Nourbakhsh</td>
<td>Assisted with experimental design, manuscript writing and editing</td>
</tr>
<tr>
<td>Mohammad Esmalifalak</td>
<td>Aided experimental design, data analysis</td>
</tr>
<tr>
<td>Andy Tan</td>
<td>Assisted with experimental design, manuscript editing</td>
</tr>
</tbody>
</table>

Principal Supervisor Confirmation
I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Name                                          Signature                                          Date
Wind power has become one of the popular renewable resources all over the world and is anticipated to occupy 12% of the total global electricity generation capacity by 2020. For the harsh environment that the wind turbine operates, fault diagnostic and condition monitoring are important for wind turbine safety and reliability. This paper employs a systematic literature review to report the most recent promotions in the wind turbine fault diagnostic, from 2005 to 2012. The frequent faults and failures in wind turbines are considered and different techniques which have been used by researchers are introduced, classified and discussed.

2.1 INTRODUCTION

Due to the lack of fossil energy and the issue of global warming in the recent years, wind energy regarded as a major source of renewable energy in the world and plays a pivotal role in the future renewable energy sources. Wind power as a kind of “green energy” has experienced a noticeable elaboration in the last decade. Wind farms are coming into account as a significant amount of the electrical generating capacity. In 2011, the world added about 40 GW of wind generation, a 24% increase, to total more than 238 GW (GWEC 2012).
This is enough capacity to cover about 3% of the world’s electricity demand (WWEC 2012) [39].

As the size and number of wind farms increase, the operation and maintenance (O& M) of wind turbines have become a critical issue for power system managers. Only the maintenance cost may constitute 10% of the total generation cost [40]. Generally, the beneficial life of each wind turbine is 52237 hours per year. The factors contributing to these downtimes are installation errors, aging, harsh environment, and variable loading condition [41].

The efficiency of the wind turbine performance could be improved due to the implementing of different maintenance practices and they can reduce the maintenance cost if they are continues and automated. Therefore research in fault diagnosis and condition monitoring is in high importance. Various methods have been applied in fault detection of wind turbines such as vibration analysis [12], [13], [14], [15], [16], oil analysis [17], [18], [19], noise analysis [20], [42], data analysis [21], [22], [20], [23], [24] and acoustic emission (AE) analysis [25], [26].

For many engineering and science problems, there is no direct mathematical solution. Learning techniques have been used extensively to overcome this problem. Researchers in different fields try to develop algorithms that learn the
behaviour of the given problem using historical data [28], [29], [30]. Learning algorithms can be used in different application such as prediction of the future value, clustering and detection of anomaly behaviour in the data.

Performance monitoring is an example of learning methods which is similar to the condition monitoring but it utilizes the historical data of the wind turbine to predict the performance of the different parameters such as gearbox oil temperature and tower acceleration. Performance monitoring is very cost-effective approach for analysing of the wind turbine performance and detecting different faults and fatigues using the data collected by the Supervisory Control and Data Acquisition (SCADA) systems [40]. There are several methods and algorithms that researchers have applied for analysing of the recorded data. For example in [43] the authors use a Support Vector Machine (SVM) paradigm for alarm detection and diagnosis of failures in the mechanical components of power wind mills. Other examples are [44], [45] that use neural network algorithm.

However there is a lack of research that provides a big picture of different methods and techniques used by various researchers in the field of data mining specially using clustering techniques. Thus, this paper aims to present a taxonomy that demonstrates what already has been done in the literature. To do so, the paper employs a systematic review of the current state-of-the-art
research into fault detection methods and techniques. The taxonomy proposed in this paper provides a good starting point for a researcher interested in following up on one or more of the methods discussed in this study.

The remainder of this paper is organized as follows: The literature review is provided in Section 2.2. Research methodology is given in Section 2.3, and the results of the systematic literature review are given in 2.4. Finally the conclusion closes the paper in Section 2.5. For the sake of clarity, we show the abbreviations in Table 2.1.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>O&amp;M</td>
<td>Operation and maintenance</td>
</tr>
<tr>
<td>SCADA</td>
<td>Supervisory Control and Data Acquisition</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>PM</td>
<td>Preventive Maintenance</td>
</tr>
<tr>
<td>PdM</td>
<td>Predictive Maintenance</td>
</tr>
<tr>
<td>CBM</td>
<td>Condition-Based Maintenance</td>
</tr>
<tr>
<td>DFIG</td>
<td>Doubly Fed Induction Generator</td>
</tr>
<tr>
<td>ERA</td>
<td>Excellence in Research for Australia</td>
</tr>
<tr>
<td>SLR</td>
<td>Systematic Literature Review</td>
</tr>
<tr>
<td>BT</td>
<td>Boosting Tree regression</td>
</tr>
<tr>
<td>RF</td>
<td>Random Forest regression</td>
</tr>
<tr>
<td>CART</td>
<td>Classification and Regression Tree</td>
</tr>
</tbody>
</table>
2.2 LITERATURE REVIEW

Wind turbines are mostly located in remote areas and unlike traditional power plants are not very protected facing with highly variable and harsh weather conditions, severe winds, tropical condition, lightning stroke, icing, and etc. These reasons reveal the importance of fault detection techniques in the maintenance of wind turbines. It is obvious that the majority of electrical and mechanical faults in such systems that have high correlation between their components may cause different failures and fatigues.

Figure 2.1 shows the major components of a typical wind turbine that are faced all the above concerns. Further studies showed that the most conventional failures has root causes in subsystems which include gearbox, main shaft and bearings, blades, electrical control, yaw system, generator and rotor brake. Figure 2.2 has depicted the failure rate of wind turbine components [18].
The wind turbines operate until the failure made it to stop working [19] and then based on the nature or severity of damage, it should be maintained or replaced (reactive maintenance). Through the development of wind turbines and increasing their capacity, preventive maintenance (PM) became more approved. This method requires periodic inspections for condition assessment based on empirical measures which are generally very expensive and not very comprehensive.

With the improving technology and implementing condition monitoring and fault detection techniques, predictive maintenance (PdM) and condition-based maintenance (CBM) have gained highly attention from wind farm holders and academia [19]. In the following part we first elaborate the major parts of the
wind turbine that frequent faults would happen and then we describe different methods of monitoring and fault diagnostic in wind turbines.

![Pie chart showing failure rate of wind turbine components.](image)

Figure 2.2 Failure rate of wind turbine components.

2.2.1 Gearbox and Bearing

The majority of literatures regarding fault and fatigues in wind turbine have focused on gearboxes for the most costly repairs are allocated to their damages [27]. The failures are normally gear tooth damage (Fig. 2.3), backlash and bearing faults and they mainly occur for that of high pressure, structure and work environment. Failures are reported as the consequents of frequent stoppage, high loaded and particle contaminations (Fig 2.4) [46].
2.2.2 Power Electronics and Electrical Control Failures

Power electronics are contributed in a very noticeable portion of failures (13%) while it accounts only 1% of the whole cost of maintenance of a wind turbine. The general failures are short-circuit and over voltage of the subsystems, damages in generator winding and transformer wirings. The root causes are generally lightning, poor electrical installation, and technical defects [47]. Semiconductor devices in the power electronic circuits are the major cause of power electronic failures. Particularly [48] reports IGBT malfunctioning as the main reason for open-circuit, short circuit and gate drive in three phase power converters.
2.2.3 Generators

The main objective of the generators in wind turbine is converting rotational energy into electrical energy. There are different kinds of generators used by wind turbines but induction or double fed induction machines are more common [12]. Bearing faults, rotor and stator breakdown are allocated the biggest proportion of failures in this component.

There are many techniques and tools available for fault diagnostic in wind turbine sub-systems. The steady-state spectral components are applied in induction machine of the stator quantities which include voltage, current, and power. They can detect faults in rotor bars, bearings, air gap eccentricities [42].

G. Vachtsevanos, R. Chen, Y. Ho, et al. [17] employed a particle filter (PF) for fault diagnostic and prognostics in gearbox and bearings.

![Contamination in a typical wind turbine.](image)

**Figure 2.4.** Contamination in a typical wind turbine.

### 2.3 Research Methodology

This study has been undertaken as a systematic literature review (SLR) based on the original instruction as proposed by Kitchenham [52]. In this section we are going to expound the steps of the methodology implemented in a systematic review study. There are three main phases that should be taken into consideration: Planning the review, Conducting the review, and Reporting on the review [52]. According to these guidelines, a systematic literature review
process is consist of the following successive stages (1) identify resources; (2) data extraction; (3) data analysis; and (4) writing-up study as a report [53].

2.3.1 Resources Searched

We have used Kitchenham’s (2004) guidelines to show different stages of the research methodology in Figure 2.5. As the fellow chart illustrates, the first step is finding the best resources for starting of the review. In this case Excellence in Research for Australia (ERA) is one of the best references available to estimate the quality of sources by implementing indicators and research experts to rank research works [20]. In this study the following data bases were used to search key words which mentioned in the ‘Search terms’ section: Google scholar; ISI Web of Science; Science Direct; IEEE Explore.

2.3.2 Search terms

The next step is exploring in each data base according to the search engine framework, the titles, abstracts, and keywords of the journal and conference articles using the following keywords:

- “Wind energy” or “Wind turbine”

- “Fault diagnosis in wind turbine”

- “Data analysis techniques in fault detection”
2.3.3 Inclusion/Exclusion Criteria

Selecting criteria plays an important role in conducting research. In this systematic review the following including and excluding criteria are
considered, 1. Published between 2005 up to 2012. 2. Published in power engineering related journals. 3. Focusing on different methods of advanced fault diagnosis in wind turbine. 4. Being in the form of journals, books, editorial notes, article summaries, reviews, and discussions.

### 2.3.4 Data Analysis

Table 2.2 indicates the number of papers published in each databases and the extraction after processing through steps 3 to 5. In the first step, after searching all mentioned databases, we found 1330 papers related to the topic. By reviewing the titles the majority of them were eliminated and the outcome was about 125 papers.

<table>
<thead>
<tr>
<th>Initial number of papers</th>
<th>Stage of exclusion</th>
<th>number of papers excluded</th>
<th>Number of papers remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>1330 (from databases)</td>
<td>Title</td>
<td>1330</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>83</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Full text</td>
<td>13</td>
<td>52</td>
</tr>
</tbody>
</table>

**Table 2.2.** Number of paper exclude in each step
The remaining papers were including the related keywords and should be considered in more details. Hence, all the abstracts regarding those papers were studied and after excluding 83 papers, 65 papers were filtered. In these cases the full text of papers were considered in more details and finally 52 papers, which should be contributed in the literature review section were came out.

2.4 Literature Review Findings and Results

This paper has verified fault diagnosis in wind turbines through a systematic review approach. The results of our systematic review are displayed in figures bellow. Figure 2.6 illustrates that this topic was not very well-researched within 2005 and 2009, but it experienced a dramatic upward trend up to 15 in 2010. However, we can see a moderate reduction into 9 at 2012. So, it can be implied that this is a very open area of research and noteworthy topic in the recent years.
Figure 2.6. Frequency of papers per year.

Figure 2.7. Frequency of papers per continent.
Figure 2.7 depicts the frequency of the studies authored in different continents. Asia has constituted the most percentage of the research area by 42%, then Europe and North America, 31% and 25% respectively. There is a very negligible researches has been done in Africa, 2%, which belongs to Tunisia as figure 2.8 shows the frequency of papers in various countries. It can be seen than China is the most dominant country with 15 papers that of the authors affiliated to Chinese universities. US is the second country with 11 papers which indicates that this topic was taken into consideration there. The UK, Taiwan and France were roughly at the same stage and in the other countries like Canada, Finland, and Korea researchers paid less attention to this area.
The purpose of this study was to classify different methods of fault diagnostics in wind turbines. We identified four main categories: vibration analysis, data analysis, noise analysis, and oil analysis. Some have sub categories that are depicted in the figure 2.9 and some of them are described below.
Vibration analysis is one of the most known and popular approach in condition monitoring of wind turbines which is surveyed comprehensively in [54], [55]. There are two major groups of vibration analysis: 1) broadband analysis and 2) analysis based on the selected spectral lines. Basic broadband analysis parameters are: root mean square, peak values, crest factor and kurtosis. Analysis techniques based on selected spectral lines reflect specific frequencies generated by certain components and some of them are gear mesh, low shaft harmonics and characteristic bearing harmonic [56].
For the gearbox and bearing faults that are the most frequent faults in wind turbine wavelet and the Fourier transformation are the two widely used techniques. Chu et al present a novel morphological undecimated wavelet (MUDW) based on morphological coupled wavelet theories. In [13], [14] the authors used the rotor modulating signals spectra for the stator and rotor fault diagnosing. H. Douglas et al. [15] utilized wavelet analysis for detection of the stator faults and [57], [58] have used wavelet analysis for distinguishing of fault signals and the background noise in wind turbine. Mohanty and Kar [16] represent fault detection of a multistage gearbox by applying discrete wavelet transformation to demodulate the current signal. In [26] the authors describe acoustic emission (AE) techniques based on Hilbert-Huang transform (HHT) that were applied to defined the AE signals released from the wind turbine bearing.

In [23] authors applied the data mining and statistical methods to analyze the data collected from monitoring of the gearbox in a wind turbine. They classified the failed stages of the gearbox into time-domain analysis and frequency domain analysis. Seven data-mining algorithms—neural network (NN) [45], artificial neural network (ANN) [44, 59], boosting tree regression (BT), random forest regression (RF), classification and regression tree (CART) [22], support vector machine (SVM) [43], and k nearest neighbour (kNN) — are selected to construct the data-driven models [23]. Z. Chen, X.

K. Kim, G. Parthasarathy, O. Uluyol, W. Foslien, S. Sheng, and P. Fleming, in National Renewable Energy Laboratory (NREL) [61] applied the SCADA data system for development of fault detection techniques and used anomaly detection algorithm and investigated classification techniques using clustering algorithm and principal component analysis (PCA) for diagnosing faults and failures.

### 2.5 Conclusion

Fault detection and wind turbine maintenance has become one of the critical topics in the recent years for wind farms. We can imply that the fault detection of wind turbines has become more interesting topic recently. The result of the review shows that there are several techniques and methods for diagnosing fault that we classified them into four major categories: vibration analysis, data analysis, noise analysis, and oil analysis. Different methods and algorithms that have been used by researchers are depicted. It can be concluded that some researches has been done related to the data analysis but most of them used
Neural Network algorithms and there is a lack of research in data mining algorithms which could be the area of future researches.
Chapter 3

PAPER 2- WIND TURBINE CONDITION MONITORING USING MACHINE LEARNING TECHNIQUES

Afrooz Purarjomandlangrudi, Ghavameddin Nourbakhsh, Andy Tan, Houman Ghaemmaghami

School of Power Engineering, Science and Engineering Faculty, Queensland University of Technology
School of Mechanical Engineering, Science and Engineering Faculty, Queensland University of Technology

Journal of Expert systems and applications
Statement of Contribution of Co-Authors for Thesis by Published Paper

The following is the format for the required declaration provided at the start of any thesis chapter which includes a co-authored publication.

The authors listed below have certified* that:
6. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
7. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
8. there are no other authors of the publication according to these criteria;
9. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
10. they agree to the use of the publication in the student’s thesis and its publication on the QUT ePrints database consistent with any limitations set by publisher requirements.

In the case of this chapter:
Publication title and date of publication or status:

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Statement of contribution*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afrooz Purarjomandlangrudi</td>
<td>Wrote the manuscript, experimental design, conducted experiments, and data analysis</td>
</tr>
<tr>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>Ghavameddin Nourbakhsh</td>
<td>Assisted with experimental design, manuscript writing and editing</td>
</tr>
<tr>
<td>Houman Ghaemmaghami</td>
<td>Aided experimental design, data analysis</td>
</tr>
<tr>
<td>Andy Tan</td>
<td>Assisted with experimental design, manuscript editing</td>
</tr>
</tbody>
</table>

Principal Supervisor Confirmation
I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Name ____________________ Signature ____________________ Date ______________
ABSTRACT

Wind power has become a popular source of energy due to its renewable nature and the minimal environmental impact of the technology required to harness this energy. In addition, recent advances in technology have provided various means for efficiently harnessing this energy to generate electricity. This has brought about an increase in research efforts directed towards the study of wind turbines as energy converting devices. In order to ensure the safety and reliability of wind turbine devices, condition monitoring (CM) and fault diagnosis is necessary. In particular, CM of the wind turbine bearing (a rotating component of the device) is vital to turbine safety, as their failure is the most common drawback in rotating machinery. In this study we propose the use of machine learning techniques to CM of the wind turbine bearing and evaluate our techniques using real data captured from a fault-seeded bearing test. We extract two features from the raw data, the kurtosis of the data and a measure of non-Gaussianity known as the non-Gaussianity score (NGS). We first apply a state-of-the-art support vector machine (SVM) classifier to detect faulty data. We then employ anomaly detection to carry out this classification and compare the robustness of this approach to the baseline SVM technique. Through evaluation we demonstrate that anomaly detection outperforms SVM with respect to accuracy and sensitivity.
3.1 INTRODUCTION

The need for electricity production has increased all over the world in recent decades. This, together with issues such as the lack of fuel resources and global warming, has brought great attention to sources of renewable energy. Today, wind farms are responsible for a huge proportion of our electricity generation capacity. In 2011 a 24% increase, amounting to about 40 GW of additional electricity generation, was observed over the previous year (GWEC 2012) [39, 62]. One of the most important components of wind turbines is the bearings, which require condition monitoring and maintenance to optimize their reliability and economic efficiency. Hence, bearing condition monitoring (CM) has attracted a lot of attention recently, with researchers attempting to enhance system performance by detecting bearing faults in advance.

There are various types of techniques for conducting CM that have been used in industry such as vibration analysis [12-16], noise analysis [63], oil analysis [17-19], acoustic emission (AE) analysis [25, 26] and data analysis approaches [20-24]. Vibration analysis is a commonly employed technology in data acquisition and fault diagnosis, especially in rotation elements [27, 31]. Data-mining techniques, particularly machine learning technologies that are able to predict defects and abnormal data, can be used to overcome the aforementioned problems. One of the more popular machine learning
techniques that is widely used across various fields and has significantly improved over the past decade, is artificial neural network (ANN) [9]. However, it is important to note that ANN-based approaches are sometimes adversely affected by issues such as structure identification difficulty, identification ability difference and poor generalization [32].

Support vector machine (SVM) is another popular machine learning method [10, 11]. In this technique, the classification algorithm deals with a binary classification problem. Therefore, the problem with this method is computation density and massiveness, which can cause some limitations and computation difficulties. This study proposes a fault diagnosis method based on a supervised learning technique, referred to as anomaly detection. The training dataset used for this method mostly includes normal data collected from various systems. We employ the anomaly detection algorithm and demonstrate that this technique is able to capture new types of intrusions [34, 35].

An important point to note is that when conducting the training phase of the anomaly detection algorithm, the training data should only contain normal data. This is because any abnormal training data would mean that the algorithm then may not recognize future abnormalities at the classification stage. We show that training and applying this method provides the CM
system with the ability to diagnose faults at early stages of occurrence with a higher accuracy. The high sensitivity of this technique, in comparison to previously employed CM methods, allows for this approach to outperform state-of-the-art systems with respect to precision.

This paper is organized as follows. Section 3.2 is discusses the features and feature extraction methods used in this work. Section 3.3 presents the applied machine learning approaches. In Section 3.4, the evaluation results, using real data from the Case Western Reserve University test rig [64], are presented. Finally, the results and outcomes of this study are discussed in Section 3.5.

### 3.2 FEATURE EXTRACTION

Fault diagnosis and CM using vibration analysis are commonly applied in a variety of industrial tasks. In any rotating component, vibration signals exist that can be observed via special sensors. These vibration signals can then provide an input source of information for carrying out feature extraction and signal processing. Various analysis techniques are then conducted on this raw data collected from the sensors. These include time-domain processing, frequency-domain processing and time-frequency techniques. These are the three main classes among numerous techniques that have been developed for waveform and data analysis, and interpretation. Frequency-domain techniques have previously been employed to show that a localized defect can generate a
periodic signal with a singular characteristic frequency \[65\]. This approach identifies and isolates the added frequency component to diagnose faults. However, this approach is suitable for defect information where the collected data contains a waveform signal with certain harmonic attributes. Time-domain techniques are primarily descended from statistical behaviours of signal waveforms. There are several characteristic features like peak-to-peak amplitude, Root-Mean Square (RMS) energy, standard deviation, skewness and kurtosis. These statistical features and their probability density functions would thus experience modifications when abnormalities are observed in the input data.

### 3.2.1 Kurtosis

We employ the kurtosis measure as a time-domain feature for conducting CM using anomaly detection. The kurtosis of the time-domain data is the fourth statistical moment of the data, normalized by the standard deviation to the fourth power. The measure of kurtosis indicates the flattening value of the density probability function close to the average value. It is a measure of how outlier-prone a distribution is. The kurtosis of a normal distribution is three. Distributions that are more outlier-prone than a normal distribution have a kurtosis value greater than three. Subsequently, those that are less outlier-prone will have a kurtosis value of less than three\[31\]. The kurtosis is calculated using:
\[ K = \frac{1}{N\sigma^4} \sum_{i=1}^{N} (x_i - \mu)^4 \quad (1) \]

Where \( x[20] \) represents a sample of our time-domain data, \( \mu \) is mean of \( x[20] \) and \( \sigma \) is the standard deviation of \( x[20] \).

### 3.2.2 Non-Gaussianity Score (NGS) feature

The Non-Gaussianity Score (NGS) is a quantitative measure of the deviation of data from a normal distribution. The NGS feature (\( \psi \)), thus measures the non-Gaussianity of a given segment of data. To do this, the normal probability plot \( \gamma \), of the dataset (\( x \)) is employed to calculate the deviation from the set’s reference Gaussian probability plot \( g \) in (2). In (2), \( g \) and \( \gamma \) represent the probabilities and \( \bar{g} \) and \( \bar{\gamma} \) are the mean of the reference normal data and the analysed data, respectively, with \( i \) ranging from the 1 to \( N \). For detailed information regarding the computation of the NGS feature refer to the study in[66].

\[ \psi = 1 - \left( \frac{\sum_{i=1}^{N} (g[i] - \bar{g})^2}{\sum_{i=1}^{N} (\gamma[i] - \bar{\gamma})^2} \right) \quad (2) \]

### 3.3 MACHINE LEARNING APPROACHES

In many cases there is no absolute mathematical solution to engineering and scientific problems. For this reason, machine learning techniques can provide a way of overcoming this issue. Scientists and researchers in different fields
attempt to employ historical data to develop algorithms that can learn the behaviour of systems [28-30]. Learning algorithms may be utilized in different applications such as data interpolation, clustering and detection of data anomalies [31]. Machine learning techniques provide the ability to learn without being explicitly programmed for systems [67, 68]. This technique develops algorithms are able to find different patterns in data and adjust program actions according to the training dataset.

3.3.1 Support Vector Machine (SVM)

Among various learning approaches SVM is commonly used for different classification scenarios and has thus emerged as a popular machine learning method. SVMs are employed in exceedingly numerous diagnostic applications like bearing faults [69], induction motor [70], machine tools [71, 72], rotating machines [63], etc. In order to train an SVM classifier, training data samples need to be injected to the SVM model to train the algorithm. Then, the model has the ability to classify new and unseen data, assigning them to their specific category, in our case the “normal” or “fault” categories. In fault diagnostics, SVMs are typically used in combination with kernel functions[73]. In this paper, our proposed anomaly detection technique and the state-of-the-art SVM technique are applied to several bearing data sets to evaluate and compute these techniques in terms of accuracy and sensitivity, allowing for the
investigation of the application of learning techniques to bearing fault diagnosis.

### 3.3.2 Anomaly detection

Anomaly detection, also referred to as novelty detection, outlier detection [74] or one-class learning [75], is a classification-based machine learning approach. This method provides the ability to classify data where generally we only have access to a single class of data, or a second class of data is under-represented. It consists of two phases, the first is the training phase and the second is the testing, or evaluation, phase. In the former, the algorithm is trained using labelled data, consisting of mostly normal examples. In the latter case, the algorithm and the trained classifier are applied to new and unseen data to conduct a diagnosis. Anomaly detection thus obtains a functional mapping $f : \mathbb{R}^N \to \{C_0, C_1\}$ of the training data to an unknown probability distribution $p(x, y)$. Where the normal samples are $(x, y) \in C_0$ and the abnormal samples are $(x, y) \in C_1$ and are summarised in (3) as follow:

$$(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^N \times Y, \quad Y = \{C_0, C_1\}$$

(3)

Training of the anomaly detection algorithm is based on the assumption that anomalous data are not generated by the source of normal data and the training set contains a huge percentage of normal data. As such, the algorithm would
detect any other example that is anomalous and has intrusive information, which provides a higher sensitivity to incipient faults [76]. Different methods such as statistical anomaly detection, data-mining based methods and machine learning based techniques are used for anomaly detection, in which the latter is used for machine fault detection.

The procedure of the algorithm and its mathematical concepts can be presented in three phases: 1. Gaussian distribution: In this paper we employed a supervised learning approach and carried out training using a labelled reference dataset. The majority of this data was made up of non-anomalous data samples, with some faulty data in the dataset. In order to apply anomaly detection, we first need to select a training set \( \{x^{(1)}, ..., x^{(m)}\} \) (where \( x^{(i)} \in \mathbb{R}^n \)) and then fit a model to this data distribution. Hence, for each feature \( x_i \) (i=1,..,n) the algorithm estimates the Gaussian parameters \( \mu_i \) and \( \sigma_i^2 \) that fit the data, in the i-th dimension \( \{x_i^{(1)}, ..., x_i^{(m)}\} \). The Gaussian distribution is obtained by (4) where \( \mu \) is the mean, \( \sigma^2 \) is the variance and \( p \) is the probability density function. 2. Estimating parameters: the Gaussian parameters \( (\mu_i, \sigma_i^2) \) of the i-th feature are estimated using (5) and (6).

\[
p(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4)
\]
\[
\mu_i = \frac{1}{m} \sum_{j=1}^{m} x_i^{(j)} \tag{5}
\]

\[
\sigma_i^2 = \frac{1}{m} \sum_{j=1}^{m} \left( x_i^{(j)} - \mu_i \right)^2 \tag{6}
\]

3. Selecting the threshold \( \epsilon \): for every dataset that the Gaussian parameters are estimated, the algorithm provides the probability that the data in the dataset is emitted by the estimated normal distribution. A low probability value associated with features would thus indicate that those features are more likely to be anomalies in the analysed data. We employ a threshold value to then separate anomalies from the normal data and compute the F1 performance measure, as depicted in (7), to evaluate our proposed approach. It must be noted that we carry out evaluations through a cross-validation approach. The precision (\( \text{prec} \)) and recall (\( \text{rec} \)) are used to compute the F1 measure:

\[
F_1 = \frac{2 \cdot \text{prec} \cdot \text{rec}}{\text{prec} + \text{rec}} \tag{7}
\]

The \( \text{prec} \) and \( \text{rec} \) can be obtained from:

\[
\text{prec} = \frac{tp}{tp + fp} \tag{8}
\]

\[
\text{rec} = \frac{tp}{tp + fn} \tag{9}
\]
Where \( tp \) is the number of true positives, referring to the number of ground-truth anomaly labels that are correctly classified by our algorithm, \( fp \) is the number of false positives. False positives occur when our system detects anomalous data where the ground-truth labels do not indicate any anomalies. And \( fn \) is the number of false negatives and represents the number of cases where the ground-truth labels indicate anomalous data but our system falsely classified it as not being anomalous. The algorithm will try several threshold values \((\varepsilon)\) to optimize the \( F_1 \) performance measure over the data. Once the optimum \( F_1 \) is found that \( \varepsilon \) is the selected threshold and then the algorithm finds the cases that lie beyond the threshold boundary, thus identifying anomalies in the data set.

### 3.4 EXPERIMENTAL RESULTS

The vibration data employed in this paper is from the data set of the rolling element bearings [42]. The test rig, shown in Figure 3.1, contains a 2 hp motor (left), three-phase induction, a torque transducer/encoder with self-aligning coupling (centre), a dynamometer (right), and control electronics (not shown). Single point faults were introduced to the test bearings using electro-discharge machining with fault diameters of 7 mils, 14 mils, 21 mils, 28 mils, and 40 mils (1 mil=0.001 inches) and the fault is 7 mil diameter and 11 mil depth. For acquiring the vibration signals, an accelerometer with a bandwidth of up to
5000 Hz and a 1 V/g output is installed on the motor housing with magnetic bases. Vibration signals were collected using a 16 channel DAT recorder with sample rate 12 K/s for drive end bearing faults. The test was conducted using different loads (0, 1, 2, 3 hp) and the types of fault are ball fault, inner race fault and outer race fault and various fault rang (7, 14, 21, 28 mil). The data recorder is equipped with low-pass filters at the input stage for anti-aliasing. In this paper we used the data associated with the 7 mil fault under 0 and 1 hp load vibration signals.

Figure 3.1. Photography and schematic description of the test rig.

3.4.1 Model description

In order to test the diagnostic method proposed in this paper, using both SVM and anomaly detection, seven sets of experimental data were used: normal
baseline data, including a fault on the outer race (with 0 – 1 hp loads), including a fault on the inner race (with 0 – 1 hp loads) and including a ball fault. The experimental motor speed was approximately 1740 rpm. The two indicators presented earlier in this paper, kurtosis and NGS, were extracted over short segments of the raw data. These features are independent from the energy of the signal. The features were extracted using a moving rectangular window over 50 ms segments of the raw data. This was done using a 25 ms window shift to achieve a smoother estimate of the features over short segments of the signal.

The obtained data from feature extraction was then applied as input to the classification algorithm proposed in this paper. The data were assessed using the previously discussed three phase process of Gaussian distribution, estimating parameters for a Gaussian and selecting the threshold. When the optimum threshold $\varepsilon$ is obtained for the inner race, outer race and ball, the anomalies can be determined for each fault type. Figure 3.2 provides a visualization of anomaly detection used for bearing fault diagnosis. In Figure 3.2, the detected anomalies have been circled to distinguish them from normal data.

The F1 measure was used to reflect and compare the accuracy and efficiency of each method. The first measurements were taken for the first 10 datasets,
consecutively 10 more datasets were added and new measurements were taken. As the process proceeded, the F1 scores were observed. As the figures demonstrate, for all displayed fault types; inner race (Fig. 3.3), outer race (Fig. 3.4), ball (Fig. 3.5) and in both no load and 1 hp conditions, F1 for the anomaly detection method is higher than SVM after training using the same data set. The F1 measure, which represents level of accuracy, is consistently more than 90% using anomaly detection while it is only around 80% when using SVM according to the following figures. Moreover, all of them the F1 measure anomaly detection tends to stabilize earlier, indicating that earlier fault detection is possible, using fewer data samples, through employing this technique.
Figure 3.2. Visualization of the proposed Anomaly detection method for automatic bearing fault detection.
Figure 3.3. Inner race fault F1 score trend for 0.007 inches, (a) 0 HP and (b) 1 HP.
Figure 3.4. Outer race fault F1 score trend for 0.007 inches, (a) 0 HP and (b) 1 HP.
Figure 3.5. Ball fault F1 score trend for 0.007 inches, (a) 0 HP and (b) 1 HP.
3.5 CONCLUSION

Wind turbine CM, maintenance, and fault diagnosis have recently become an important matter for wind farm holders and researchers all over the world. There have been several techniques and methods for fault detection that have been applied by researchers and engineers to provide the ability to achieve fast and effective defect detection. Machine learning techniques have improved significantly in the last decade and have thus been employed for automatic classification, especially in fault diagnosis and their ability of learning and training, provide a range of possible applications. Two machine learning methods, anomaly detection and SVM, were employed in this research using real data from a fault seeded bearing test.

Two features, kurtosis and NGS, were extracted, as part of the developed anomaly detection algorithm. The data which is used in this paper is 0.007 inch fault in Inner race, outer race and ball under 0 and 1 HP in a bearing test. The results indicate that learning techniques equipped with a high rate of accuracy, in this case anomaly detection, can achieve a higher percentage of accuracy for bearing fault diagnosis compared to previously applied methods such as the SVM classifier approach. In addition, it was shown that anomaly detection can carry out this accurate fault detection process using fewer data samples compared to the SVM classifier. This indicates that through the
proposed anomaly detection scheme the fault detection process can be achieved more accurately and more efficiently.
Chapter 4

PAPER 3- APPLICATION OF ANOMALY TECHNIQUE IN WIND TURBINE BEARING FAULT DETECTION

Afrooz Purarjomandlangrudi, Ghavameddin Nourbakhsh, Yateendra Mishra, Houman Ghaemmaghami, Andy Tan

School of Power Engineering, Science and Engineering Faculty, Queensland University of Technology
School of Mechanical Engineering, Science and Engineering Faculty, Queensland University of Technology

IEEE Power & Energy Society (PES General Meeting) 2014, Submitted
Statement of Contribution of Co-Authors for Thesis by Published Paper

The following is the format for the required declaration provided at the start of any thesis chapter which includes a co-authored publication.

The authors listed below have certified* that:
11. they meet the criteria for authorship in that they have participated in the conception, execution, or interpretation, of at least that part of the publication in their field of expertise;
12. they take public responsibility for their part of the publication, except for the responsible author who accepts overall responsibility for the publication;
13. there are no other authors of the publication according to these criteria;
14. potential conflicts of interest have been disclosed to (a) granting bodies, (b) the editor or publisher of journals or other publications, and (c) the head of the responsible academic unit, and
15. They agree to the use of the publication in the student’s thesis and its publication on the QUT ePrints database consistent with any limitations set by publisher requirements.

In the case of this chapter:
Publication title and date of publication or status:

<table>
<thead>
<tr>
<th>Contributor</th>
<th>Statement of contribution*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afrooz Purarjomandlangrudi</td>
<td>Wrote the manuscript, experimental design, conducted experiments, and data analysis</td>
</tr>
<tr>
<td>Date</td>
<td></td>
</tr>
<tr>
<td>Ghavameddin Nourbakhsh</td>
<td>Assisted with experimental design, manuscript writing and editing</td>
</tr>
<tr>
<td>Houman Ghaemmaghami</td>
<td>Aided experimental design, data analysis</td>
</tr>
<tr>
<td>Andy Tan</td>
<td>Assisted with experimental design, manuscript editing</td>
</tr>
<tr>
<td>Yateendra Mishra</td>
<td>Aided with editing and formatting</td>
</tr>
</tbody>
</table>

Principal Supervisor Confirmation
I have sighted email or other correspondence from all Co-authors confirming their certifying authorship.

Name ___________________________ Signature ___________________________ Date ___________________________
ABSTRACT

Bearing faults are the most common cause of wind turbine failures. Availability and maintenance cost of wind turbines are becoming critically important in electric networks. Early fault detection can reduce outage time and costs. This paper proposes Anomaly Detection (AD) machine learning algorithms for fault diagnosis of wind turbine bearings. The application of this method on a real data set was conducted and is presented in this paper. For validation and comparison purposes, a set of baseline results are produced using the state-of-the-art one-class m-SVM methods to examine the ability of the proposed technique in detecting incipient faults.

4.1 INTRODUCTION

Wind turbines, as part of renewable energy generation technology, are increasingly deployed throughout electricity networks around the world. Low speed rotating components, such as gearbox and bearing, play an important role in determining wind turbines’ efficiency. According to the Department of Trade and Industry (DTI) in UK, Condition Based Maintenance (CBM) in wind turbine rotating elements has contributed to a saving of up to £1.3 billion per year [77]. Hence, Condition Monitoring (CM) of bearings has become a popular approach to increase performance and reduce costs [27].
As wind turbines are prone to failure due to high stress on the gearbox and the bearing, application of appropriate fault diagnosis techniques (especially for bearings) in wind turbine maintenance has been the topic of many studies in recent literature [27]. Figure 4.1 displays the different components of a rolling element bearing, including the inner race, outer race, rolling element (ball) and cage, which despite their simplicity have a complex internal operation [78]. In recent literature, several studies have been conducted on bearing fault analysis using vibration and Acoustic Emission (AE) techniques [79].

Tandon et al.[80] presented a comprehensive review of the vibration and AE techniques including vibration measurements in both time-and frequency-domains, the shock pulse methodology, sound measurements and AE for bearings CM. Kim et al.[81] presented techniques for vibration and wear debris analysis by studying railway freight cars that include vibration, spike energy, spectrographic oil analysis, shock pulse and chip detection. Watson et al.[50] used wavelet techniques for doubly fed induction generator (DFIG) bearing faults detection based on power output data. R. Sehgal et al.[82] presented different factors that brought about faults and fatigues in bearings. These factors included excessive preloading during installation, overloading and stray electric currents.
The majority of existing methods for bearing fault detection are focused on detecting the type of fault that has already occurred and for which data samples are available. This is while collecting all possible types of faulty data is almost impossible. The existing methods are thus prone to erroneous behavior when presented with previously unseen data. On the contrary, data-mining methods, and machine learning techniques in particular, are able to predict defects and the existence of abnormal data without such conditions. One of the more popular machine learning techniques, which has been widely applied across multiple disciplines and has improved over the past decade, is Artificial Neural Network (ANN) [9]. This technique, however, has some drawbacks such as structure identification difficulty, identification ability
difference, local convergence, owes to learning process and poor
generalization ability, as it originally applies Experience Risk Minimization
(ERM) [32, 68]. Another popular classification technique is the use of Support
Vector Machine (SVM), which has recently become a very popular learning
method[10]. The use of SVM classifiers simplifies a classification task by
reducing the problem to a binary classification. The drawback of using SVM
however, lies in issues with computational density and massiveness that bring
about limitations and difficulties.

In this paper a fault diagnosis method based on supervised learning technique
named anomaly detection (AD) is proposed. The training data set used for this
method mostly includes normal data, which is easier to collect in practice. The
AD algorithm is then able to capture any new types of intrusions or
inconsistencies from normal data[34]. The point that needs to be taken into
account is that if the training data does not contain normal variations of the
data, the algorithm may not recognize future abnormalities and will assume
they are normal. The training and application of this method provides the CM
system with the ability to diagnose bearing fault in its early stages with a
higher accuracy. The high sensitivity of this method means that it can provide
a higher precision in fault detection, when compared to the previous state-of-
the-art bearing fault detection algorithms.
The contents of this paper are organized in the following way. Section I is about previous research and explains some examples in literature. It includes descriptions about the importance of bearing in wind turbine industry, existing analysis techniques and fault diagnosis methods. Section 4.2 presents the details on applied machine learning approaches, and in Section 4.3, the evaluation of the proposed approach and the baseline method over real data, obtained from a test rig of Case Western Reserve University, is carried out to confirm the robustness and efficiency of the proposed technique and conclusions are provided in Section 4.4.

4.2 MACHINE LEARNING APPROACHES

4.2.1 One-class Support Vector Machine

Among various learning approaches SVM is extensively used for different classification purposes and has been established as a popular machine learning method. SVMs are employed in numerous diagnosis applications like bearing faults, induction motor, rotating machines, etc. [63]. Traditionally, many classification techniques try to categorize two or multi-class situations. The aim of applying machine learning techniques is to identify test data between a number of classes, using training data. In the case where only one class of data is available, new data needs to be tested to determine whether or not it
contains faulty data. To overcome this problem, researchers have proposed the use of the One-Class Support Vector Machine.

Training data samples need to be injected to the one-class SVM model to train the classifier. After that, the model has the ability to classify new data and distinguish whether it inconsistent with the trained model or not. Typically in fault diagnostics the combination of one-class SVMs with other techniques, like kernel functions, are used [73]. In this paper, AD and SVM are applied on a real bearing data set for comparing the techniques and to examine the application of learning techniques in fault diagnosis.

4.2.2 Anomaly Detection (AD)

AD is a machine learning approach based on classification techniques, which provides the user with the ability to classify data, where generally only a single class of data is available or a second class of data is under-represented. This method typically consists of two phases, the first is the training phase and the second is the testing phase. In the former phase, the algorithm is trained using a labelled dataset, which consists of mostly normal data. In the latter, the learned algorithm is applied to new and unseen data or a cross validation dataset. To put it more formally, the classifier learns a functional mapping $f : \mathbb{R}^N \rightarrow \{C_0, C_1\}$ of the training dataset to an unknown probability distribution $p(x,y)$, where the normal samples are $(x,y) \in C_0$ and the
abnormal samples are \((x, y) \in C_1\). This can be represented in the following manner:

\[
(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^N \times Y, \quad Y = \{C_0, C_1\}
\]

The AD algorithm is trained based on the presumption that anomalous data are not generated by the source of normal data and that the training set contains a huge percentage of normal data, the algorithm then detects any other data which is not normal and displays intrusive behaviour [76]. The main procedure of algorithm is based on the use of Gaussian distributions and it contains three phases: 1. Gaussian distribution, 2. Estimating parameters, and 3. Selecting threshold, \(\varepsilon\). In phase 1 and 2, all data are modelled and the Gaussian parameters are defined. In phase 3, the algorithm selects a threshold to recognize anomaly cases by using the \(F_1\) score. This can be expressed using the following equations:

\[
F_1 = \frac{2 \cdot \text{prec} \cdot \text{rec}}{\text{prec} + \text{rec}}
\]

While, the \(\text{prec}\) and \(\text{rec}\) are obtained from:

\[
\text{prec} = \frac{tp}{tp + fp}
\]
where \( tp \) is the number of true positives, and refers to the ground truth dataset labels that reflect anomaly and are also correctly classified by our algorithm. The complementary measure \( fp \), is the number of false positives. This is when anomaly does not exist in the ground truth labels, but the data is incorrectly classified as containing anomaly. Finally, \( fn \) is the number of false negatives and reflects the cases of the data that are labelled as anomalies in the ground truth labels but are incorrectly classified as not being anomalous by our algorithm. The algorithm will try several values of \( \varepsilon \) to find the best value based on the \( F_1 \) score, using only the training data. Once the best \( \varepsilon \) is selected, the algorithm then applies this threshold to the evaluation data in order to find the data that fall beyond the threshold boundaries and classifies them as anomalies of the data.

### 4.3 EXPERIMENTAL RESULTS

The vibration data employed in this paper was collected from the dataset of rolling element bearings from the NSF I/UCR Centre for Intelligent Maintenance Systems (IMS – www.imscenter.net) with support from Rexnord Corp. in Milwaukee, WI [62]. As Figure 4.2 illustrates, four bearings were installed on a shaft. An AC motor is coupled to the shaft via rub belts to keep
the rotation speed at a constant 2000 RPM and a radial load of 6000 lbs is added to the shaft and the bearings by a spring mechanism. All bearings are force lubricated. Vibration data was collected every 10 min for 164h with a sampling rate of 20 kHz. An outer race defect occurred at the end of this experiment on bearing 1.

The data of the horizontal accelerometer of bearing 1 have been applied, in which the bearing consists of 16 rollers in each row, with a pitch diameter of 2.815 in., a roller diameter of 0.331 in., and a tapered contact angle of 15.17. The majority of past studies on bearing fault diagnostics have been conducted using simulated or seeded damage. The experiments employing these types of faults are not appropriate for detecting natural early defects. As the main objective of this paper is to present a method with the ability to detect bearing faults at early stages, this dataset is completely appropriate for our purpose.

4.3.1 Model description

In order to test this technique and compare it to the previous applied method for the same bearing, as used in [83], We carried out extensive evaluations. In this experiment 100 normal operating conditions were captured and used as the training dataset (16hr), and then time-domain parameters were extracted for each sub-band of the raw data. These parameters are independent of the energy of the signal and have a proper distribution for machine learning data
analysis. The sub-bands are selected with 50 ms length and 25 ms shift. The shift is employed to provide overlapping of the sub-bands of the dataset.

![Figure 4.2. Bearing test rig and sensor placement [62].](image)

The obtained data, gathered using feature extraction, is then applied as an input to the algorithm. These features are processed by the three AD steps of
Gaussian distribution, estimating parameters for a Gaussian, and selecting the threshold. In step 1 a Gaussian model is applied to the distribution of the data. The Gaussian distribution is obtained by equation (4) where $\mu$ is the mean, $\sigma^2$ is variance, and $p$ is the probability density function:

$$p(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$  \hspace{1cm} (4)

In the next step, Gaussian parameters are calculated by utilizing the training dataset. The algorithm then examines several values of $\varepsilon$ to find the best value based on the $F_1$ score. Once the best $\varepsilon$ is selected, the algorithm finds the data that fall beyond the threshold boundaries, which are then marked as anomalies in the dataset. The $F_1$ score represents the accuracy of the algorithm, in which a score of 1 means 100% accuracy. It can be seen that the algorithm can identify all anomalies in the dataset.

Incipient fault detection is one of the most important aspects of this method. Figure 4.3 and Figure 4.4 display the exact time that each technique has detected the first anomaly point. Figure 4.3 has been adopted from the results of past research of Kaewkongka et al. that has been conducted on the same dataset [83]. In this figure the red dots show the points where the methods detect a change in behaviour for the first time, for the one-class SVM. This is 75 h before the crack makes the machine stop for SVM in Figure 4.3. This is
while in Figure 4, which is for that of AD, the first anomalous data was detected in 100 h before the major defect occurs that causes the breakdown of the bearing.

Figure 4.3. SVM output [83].
4.4 CONCLUSIONS

Bearing condition monitoring and incipient fault detections are very important for wind turbine generating units to cut maintenance costs and increase availability. There are a variety of CM and fault diagnosis techniques that are used to detect incipient faults and failures in wind turbine bearings. However, the majority of these techniques need huge samples of faulty data for the development/training phase of the algorithm. In practice, these kinds of datasets are not easily available. In order to overcome this problem, this paper
proposed anomaly detection as part of a machine learning approach for fault diagnosing of bearings in wind turbines. The results obtained from this method were compared with that of a SVM based state-of-the-art approach. The results indicated that the proposed AD learning technique provides a higher accuracy, as well as the ability to conduct early detection of bearing faults.
Chapter 5

CONCLUSIONS

5.1 OVERVIEW

Chapter 5 contains a summary of this research work, which includes the research problems, research methodology, research validation through experimental results and case study, and outcome analysis. Conclusion and possible applications for the proposed method in this thesis are discussed.

5.2 SUMMARY OF FINDINGS

Application of the proposed method on the first data set from the Case Western Reserve University test rig has come up with interesting results in favour of an anomaly detection technique. Both the anomaly detection technique and the SVM technique (which is a previously applied method [84]) were applied to different parts of bearing, inner race, outer race, and ball under 0 HP (no loud) and 1HP.
Table 5.1. AD and SVM $F_1$ measure for bearing components

<table>
<thead>
<tr>
<th></th>
<th>0 HP</th>
<th>1HP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AD $F_1$ (%)</td>
<td>SVM $F_1$ (%)</td>
</tr>
<tr>
<td>Inner race</td>
<td>95</td>
<td>83</td>
</tr>
<tr>
<td>Outer race</td>
<td>99</td>
<td>84</td>
</tr>
<tr>
<td>Ball</td>
<td>100</td>
<td>80</td>
</tr>
</tbody>
</table>

The results show that with an inner race defect at 0 HP the $F_1$ measure was at 95% for anomaly detection and 83% for that of SVM. With 1 HP they are 92% and 83%, respectively. With an outer race defect on the bearing, the percentage of the $F_1$ measured for no load condition was 99% for the anomaly detection technique and 84% for the SVM technique. With 1 HP it was 96% for the former and 79% for the later. Finally for a ball defect, they are 100% and 80% under 0 HP and under 1 HP they are 97% and 81%.

It can be concluded from the outcome statistics that accuracy of the anomaly detection technique is higher than the SVM technique in regard with bearing
condition monitoring. Particularly with a ball defect, the $F_1$ measures were close to 100%, which means that almost all of the anomalous and defected data can be captured through this technique. This quality is of high importance for the rotating component CM in the wind turbine. Using a technique with higher accuracy and sensitivity would yield less failures and breakdown times. If the fault detection technique is able to detect defects at their incipient stage, they can be fixed before causing major failures.

In Chapter 4, the proposed method was utilized to examine the failure data set of real bearings (IMS Bearing Data [62]). In this test, all the data from the time the bearing was working normally, until failure, were captured via vibrational sensors. An anomaly detection technique was implemented on the data set and the result was compared with the SVM technique result [84]. The comparison shows that the anomaly detection technique for an incipient fault was detected at around 100 hours before the major breakdown, while for the SVM technique was at 70 hours. This experiment indicates that the anomaly detection technique can recognize faults and defects sooner than the state-of-the-art SVM technique. This means the AD method can detect the fault earlier, giving more time for preventive measures.
5.3 ADDRESSING RESEARCH QUESTIONS AND CONCLUSION

Learning techniques have recently become very popular and are widespread in use, especially in fault diagnostics, and the ability of self-teaching and training provides an opportunity for a variety of applications. In this research, a machine-learning technique called anomaly detection is applied to diagnose faults and fatigue in rolling element bearings. This method employs a supervised learning algorithm to categorize the anomalies from normal data via supervised learning techniques.

Two features, kurtosis and Non-Gaussianity Score (NGS), were extracted to develop the anomaly detection algorithms. In order to test the efficiency of the technique in bearing fault detection, a real data set of bearing failure from a Case Western Reserve University test rig [7] and a bearing test to failure data set from the NSF I/UCR Centre [62] were used.

The results indicate that this method is able to detect faults and anomalies in bearing components with higher accuracy, even with small data sets. From Figures 3.3, 3.4 and 3.5 in Chapter 3 it can be concluded that the anomaly detection technique is able to reach a stable output in $F_1$ measure trend, which means this method would be more reliable when we do not have access to the huge amount of data sets in the system. Moreover, this method provides the
ability to detect defects at earlier stages than previously applied methods, which is significant and essential in wind turbine bearing condition monitoring.

The accuracy and sensitivity of the anomaly detection technique was measured in Chapter 3, where Chapter 4 was mainly devoted on formulating the ability of the technique for incipient fault detection. This is due to the fact that, one of the major concerns regarding wind turbines is detecting incipient faults in their early stages, before turning into a major defect causing major maintenance and downtime costs. In fact, both accuracy and celerity abilities are investigated for the proposed technique in Chapters 3 and 4.

Results from paper 2 (Chapter 3) indicate that anomaly detection technique is able to identify most of the faulty data, where the indicator $F_1$ measures are very high above 90% which are much more than the SVM. Furthermore, data analysis of paper 3 (Chapter 4) shows that anomaly detection technique not only is very sensitive, it is also able to recognize the anomalous data and incipient faults sooner that that of SVM.
5.4 IMPLICATIONS AND FUTURE WORKS

Although several techniques have been reported in the literature for bearing fault detection and diagnosis, it is still challenging to implement a reliable condition monitoring system for real-world industrial applications because of complex bearing structures and noisy operating conditions. The theme of this thesis is to develop a novel intelligent system to tackle these related challenges. The strategy is to develop more robust techniques at each processing stage to improve the condition monitoring reliability.

This research investigated fault diagnostics and incipient failure and defect detection in rolling element bearings. The proposed method utilized machine learning techniques to detect abnormalities in system operation. Two real bearings vibration data sets from the Case Western Reserve University and NSF I/UCR Center for Intelligent Maintenance Systems were utilized to test the performance of the anomaly detection algorithm. The former was applied to validate the accuracy and the later was used to measure how fast this method can detect incipient faults. The same data were applied to state-of-the-art SVM algorithm for comparison. The implications of the results of this work for industrial users and researchers are discussed in the following.
5.4.1 Implications for Industry Practitioners

Bearings as a principal element are used in many applications that contain rotating components, to provide a relative motion for other moving machine elements. Therefore, they are widely used in automotive, industrial, marine, and aerospace applications as an effective means to support transmitted power to the load. In industrial applications, rotating machinery components such as gearbox, shafts and bearings degrade slowly with operating time; this can be detected if a robust predictive condition monitoring technology is used. The proposed method can be easily applied for detecting any type of faults and fatigues on gear teeth, shafts, bearings and many other elements. In order to do this, data is collected from the desired part so that one can train the algorithm by a historical data of normal data and then apply it on the online CMS to detect any abnormalities in advance. This method does not require a huge database to train and also is fast in computation. Thus, it is a very efficient and cost effective method for a variety of users.

The automotive industries can be a good example of who may be interested in early fault detection of many moving parts, for instance, pumps bearing, steering systems, air-conditioning compressors, engine rocker arms, throttle butterfly valves, gearboxes, and transmissions. Another example can be the diesel engine bearings in a marine environment used to help reduce friction by converting sliding friction into rolling friction. These bearings are subjected to
different types of forces due to gas pressure, and different reciprocating and rotating motions of engine parts. Early defects and fault diagnostics can prevent major failure, and save time and costs associated with these cases. However, there are many similar examples relating to industries where this technique can contribute to their effective operation and maintenance regimes.

5.4.2 Implications for Researchers

For many science and engineering problems, such as fault detection in moving components, learning techniques have been widely used when there are no direct mathematical solutions. There are several machine-learning methods that have been used for the purpose of fault detection in low speed rotating components such as bearings and gearboxes. Among these methods much research has been done using ANN, which has some disadvantages such as structure identification difficulties, local convergence, and poor generalization abilities. SVM is another technique that has recently become very popular, but often its huge amount of computation makes it impractical for many applications. To overcome these problems, the anomaly detection approach, which is based on classification techniques, was proposed and adapted in this thesis to detect incipient defects. This method contains less computation and training data, while the results indicate a higher precision and efficiency. It also is able to detect the first abnormal behaviour earlier, in comparison with the other techniques applied.
Regarding future work for researchers, anomaly detection for fault diagnosis in rolling element bearings provides a platform for the further use of this method in the condition monitoring of other low speed rotating components, such as gearboxes and shafts. Moreover, there are some other machine-learning approaches that can be investigated for this purpose, such as inductive logic programming, clustering, reinforcement learning, and similarity and metric learning. Also a combination of these methods with principal components analysis, classification techniques or game theory would be some areas of research that may enhance the fault detection and condition monitoring of rotating components.
The following slides which were adapted from an online Machine Learning course [6] would help to give a better understanding of the basic theory of anomaly detection and SVM techniques. There are some explanations and examples which describe the mathematical theory behind these techniques.
**Anomaly detection algorithm**

1. Choose features $x_i$ that you think might be indicative of anomalous examples.
2. Fit parameters $\mu_1, \ldots, \mu_n, \sigma_1^2, \ldots, \sigma_n^2$
   
   $$\mu_j = \frac{1}{m} \sum_{i=1}^{m} x_j^{(i)}$$
   
   $$\sigma_j^2 = \frac{1}{m} \sum_{i=1}^{m} (x_j^{(i)} - \mu_j)^2$$
3. Given new example $x$, compute $p(x)$:
   
   $$p(x) = \prod_{j=1}^{n} p(x_j; \mu_j, \sigma_j^2) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi\sigma_j}} \exp\left(-\frac{(x_j - \mu_j)^2}{2\sigma_j^2}\right)$$

   Anomaly if $p(x) < \varepsilon$

**Anomaly detection example**

- $\mu_1 = 5, \sigma_1 = 2$
- $\mu_2 = 3, \sigma_2 = 1$

$\varepsilon = 0.02$

$$p(x_{test}^{(1)}) = 0.0426$$

$$p(x_{test}^{(2)}) = 0.0021$$
Anomaly detection

Developing and evaluating an anomaly detection system

The importance of real-number evaluation
When developing a learning algorithm (choosing features, etc.), making decisions is much easier if we have a way of evaluating our learning algorithm.

Assume we have some labeled data, of anomalous and non-anomalous examples. \( y = 0 \) if normal, \( y = 1 \) if anomalous.

Training set: \( x^{(1)}, x^{(2)}, \ldots, x^{(m)} \) (assume normal examples/not anomalous)

Cross validation set: \( (x_{cv}^{(1)}, y_{cv}^{(1)}), \ldots, (x_{cv}^{(m_{cv})}, y_{cv}^{(m_{cv})}) \)

Test set: \( (x_{test}^{(1)}, y_{test}^{(1)}), \ldots, (x_{test}^{(m_{test})}, y_{test}^{(m_{test})}) \)

Algorithm evaluation
Fit model \( p(x) \) on training set \( \{x^{(1)}, \ldots, x^{(m)}\} \)

On a cross validation/test example \( x \), predict

\[
y = \begin{cases} 
1 & \text{if } p(x) < \varepsilon \text{ (anomaly)} \\
0 & \text{if } p(x) \geq \varepsilon \text{ (normal)}
\end{cases}
\]

Possible evaluation metrics:
- True positive, false positive, false negative, true negative
- Precision/Recall
- \( F_1 \)-score

Can also use cross validation set to choose parameter \( \varepsilon \)
Anomaly detection
Choosing what features to use

Machine Learning

Non-Gaussian features

Error analysis for anomaly detection
Want $p(x)$ large for normal examples $x$.
$p(x)$ small for anomalous examples $x$.

Most common problem:
$p(x)$ is comparable (say, both large) for normal and anomalous examples
Anomaly detection

Anomaly detection using the multivariate Gaussian distribution

Machine Learning

Multivariate Gaussian (Normal) distribution

Parameters $\mu, \Sigma$

$$p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} \exp \left( -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)$$

Parameter fitting:

Given training set $\{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$

$$\mu = \frac{1}{m} \sum_{i=1}^{m} x^{(i)} \quad \Sigma = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)} - \mu)(x^{(i)} - \mu)^T$$
Anomaly detection with the multivariate Gaussian

1. Fit model $p(x)$ by setting

$$
\mu = \frac{1}{m} \sum_{i=1}^{m} x^{(i)}
$$

$$
\Sigma = \frac{1}{m} \sum_{i=1}^{m} (x^{(i)} - \mu)(x^{(i)} - \mu)^T
$$

2. Given a new example $x$, compute

$$
p(x) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)
$$

Flag an anomaly if $p(x) < \varepsilon$

Relationship to original model

Original model: $p(x) = p(x_1; \mu_1, \sigma_1^2) \times p(x_2; \mu_2, \sigma_2^2) \times \cdots \times p(x_n; \mu_n, \sigma_n^2)$

Corresponds to multivariate Gaussian

$$
p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} \exp \left( -\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right)
$$

where
Support Vector Machines

Optimization objective

Machine Learning

**Alternative view of logistic regression**

\[ h_\theta(x) = \frac{1}{1 + e^{-\theta^T x}} \]

If \( y = 1 \), we want \( h_\theta(x) \approx 1 \), \( \theta^T x \gg 0 \)
If \( y = 0 \), we want \( h_\theta(x) \approx 0 \), \( \theta^T x \ll 0 \)

**Alternative view of logistic regression**

Cost of example:

\[ -y \log h_\theta(x) + (1 - y) \log(1 - h_\theta(x)) \]

\[ = -y \log \frac{1}{1 + e^{-\theta^T x}} - (1 - y) \log \left(1 - \frac{1}{1 + e^{-\theta^T x}}\right) \]

If \( y = 1 \) (want \( \theta^T x \gg 0 \)):

\[ -\log \frac{1}{1 + e^{-z}} \]

If \( y = 0 \) (want \( \theta^T x \ll 0 \)):

\[ -\log \left(1 - \frac{1}{1 + e^{-z}}\right) \]
Support vector machine

**Constrained regression:**

\[
\min_{\theta} \frac{1}{m} \sum_{i=1}^{m} y^{(i)} \left( -\log h_\theta(x^{(i)}) \right) + (1 - y^{(i)}) \left( -\log(1 - h_\theta(x^{(i)})) \right) + \frac{\lambda}{2m} \sum_{j=1}^{n} \theta_j^2
\]

Support vector machine:

\[
\min_{\theta} C \sum_{i=1}^{m} \left[ y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{j=1}^{n} \theta_j^2
\]

**Support Vector Machine**

\[
\min_{\theta} C \sum_{i=1}^{m} \left[ y^{(i)} cost_1(\theta^T x^{(i)}) + (1 - y^{(i)}) cost_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{j=1}^{n} \theta_j^2
\]

If \( y = 1 \), we want \( \theta^T x \geq 1 \) (not just \( \geq 0 \))

If \( y = 0 \), we want \( \theta^T x \leq -1 \) (not just \( < 0 \))
CONCLUSIONS

**SVM Decision Boundary**

\[ \min_{\theta} C \sum_{i=1}^{m} \left[ y^{(i)} \text{cost}_1(\theta^T x^{(i)}) + (1 - y^{(i)}) \text{cost}_0(\theta^T x^{(i)}) \right] + \frac{1}{2} \sum_{i=1}^{n} \theta_i^2 \]

Whenever \( y^{(i)} = 1 \):

\[ \theta^T x^{(i)} \geq 1 \]

Whenever \( y^{(i)} = 0 \):

\[ \theta^T x^{(i)} \leq -1 \]

**SVM Decision Boundary: Linearly separable case**

Large margin classifier

**Large margin classifier in presence of outliers**

\[ \text{\( C \) very large} \]

\[ \text{\( C \) not too large} \]
Support Vector Machines
The mathematics behind large margin classification (optional)

Machine Learning

\[ u = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \]

\[ u^T v = \mathbb{R} \]

\[ ||u|| = \text{length of vector } u \]
\[ = \sqrt{u_1^2 + u_2^2} \]

\[ p = \text{length of projection of } v \text{ onto } u \]
\[ u^T v = p \cdot ||u|| \]
\[ = u_1 v_1 + u_2 v_2 \]

\[ u^T v = p \cdot ||u|| \]
\[ p < 0 \]
CONCLUSIONS

**SVM Decision Boundary**

\[
\min_{\theta} \frac{1}{2} \sum_{j=1}^{n} \theta_j^2 = \frac{1}{2} (\theta_1^2 + \theta_2^2) = \frac{1}{2} \left( \theta_1^2 + \theta_2^2 \right) \leq \frac{1}{2} \|\theta\|^2
\]

s.t. \( \theta^T x^{(i)} \geq 1 \) if \( y^{(i)} = 1 \)

\( \theta^T x^{(i)} \leq -1 \) if \( y^{(i)} = 0 \)

Simplification: \( \theta_0 = 0 \quad n=2 \)

\[
\theta^T x^{(i)} = p^{(i)} \cdot \|\theta\| = \theta_1 x_1 + \theta_2 x_2
\]

**SVM Decision Boundary**

\[
\min_{\theta} \frac{1}{2} \sum_{j=1}^{n} \theta_j^2
\]

s.t. \( p^{(i)} \cdot \|\theta\| \geq 1 \) if \( y^{(i)} = 1 \)

\( p^{(i)} \cdot \|\theta\| \leq -1 \) if \( y^{(i)} = 1 \)

where \( p^{(i)} \) is the projection of \( x^{(i)} \) onto the vector \( \theta \).

Simplification: \( \theta_0 = 0 \)
REFERENCES


CONCLUSIONS


N. Pican and Y. Lallement, "The OWE Artificial Neural Network and its Parallel Implementation."


Sun, M. Rahman, Y. Wong, and G. Hong, "Multiclassification of tool wear with support vector machine by manufacturing loss consideration,"


