ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

BUILDING SENSOR-BASED REAL-TIME PREDICTIVE MAINTENANCE SYSTEM BY UTILIZING ARTIFICIAL INTELLIGENT TECHNIQUES

Ph.D. THESIS

Raghad Mohammed KHORSHEED

Department of Industrial Engineering

Industrial Engineering Programme

JUNE 2021

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

BUILDING SENSOR-BASED REAL-TIME PREDICTIVE MAINTENANCE SYSTEM BY UTILIZING ARTIFICIAL INTELLIGENT TECHNIQUES

Ph.D. THESIS

Raghad Mohammed KHORSHEED
(507162111)

Department of Industrial Engineering

Industrial Engineering Programme

Thesis Advisor: Assoc. Prof. Dr. Ömer Faruk BEYCA

JUNE 2021

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

YAPAY AKILLI TEKNİKLERİ KULLANARAK SENSÖR TABANLI GERÇEK ZAMAN TAHMİNLİ BAKIM SİSTEMİ KURULMASI

DOKTORA TEZİ

Raghad Mohammed KHORSHEED
(507162111)

Endüstri Mühendisliği Anabilim Dalı

Endüstri Mühendisliği Programı

Tez Danışmanı: Dr. Ömer Faruk BEYCA

HAZİRAN 2021

Raghad Mohammed KHORSHEED, a Ph.D. student of ITU Graduate School student, ID 507162111, successfully defended the dissertation entitled "BUILDING SENSOR-BASED REAL-TIME PREDICTIVE MAINTENANCE SYSTEM BY UTILIZING ARTIFICIAL INTELLIGENT TECHNIQUES", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor : **Assoc. Prof. Dr. Ömer Faruk BEYCA** .....
Istanbul Technical University

Jury Members: **Prof. Dr. Şule ITIR SATOĞLU** .....
Istanbul Technical University

Prof. Dr. Alp ÜSTÜNDAĞ .....
Istanbul Technical University

Prof. Dr. Selim ZAİM .....
Sabahattin Zaim University

Prof. Dr. İhsan KAYA .....
Yıldız Technical University

Date of Submission : 06 April 2021
Date of Defense : 22 June 2021

To my *parents*

Whose affection, love, encouragement, and prayers to achieve much success and honor

Also to my beloved *husband*, sweet *daughter Ay*şe, loving *sister* and *brother*

To *all* whom I love and respect

I dedicate this humble effort

FOREWORD

‘Praise to Allah for giving me the strength to complete this thesis’

I would first and foremost like to express my sincere gratitude and thanks to my supervisor, Dr. Ömer Faruk BEYCA, for his faithful help, wide opinions, invaluable advice, guidance, and suggestion during the preparation of this dissertation.

I would like to extend my thank Prof. Dr. Selim ZAİM for his assistance and offers a chance to work for a postgraduate degree in Industrial Engineering.

A word of thanks should go to Sewerage Treatment Company/QATAR, for the provision of various data and information on the machines and equipment. I am also very grateful to my colleague Dr. Mohanad AL-ANİ for his assist during the data collecting stage.

Special thanks go to both Prof. Dr. Şule ITIR and Prof. Dr. İhsan KAYA, for their support and guidance throughout all the phases of this work.

I also would like to thank my friend Ph.D. candidate Israa ISMAİL for her support and help during the Ph.D. study.

Special thanks and appreciation to my husband Şahin, who stood, and still, beside me all the way, who without his help and encouragement this work wouldn’t come to light.

Finally, last but by no means least, I want to express my deepest gratitude to my family for their unlimited love, unceasing encouragement, support, and attention giving me the possibility to complete my studies. I owe them everything.

April 2021

Raghad Mohammed KHORSHEED
(Industrial Engineer)

TABLE OF CONTENTS

Page

ABBREVIATIONS

ABCPD : Adaptive Bayesian Change Point Detection
AE : Acoustic Emission
AEFD : Aircraft Engine Fault Diagnosis
AI : Artificial Intelligence
ANN : Artificial Neural Network
AUC : Area under Curve
BPNN : Back Propagation Neural Network
CBM : Condition-Based Maintenance
CM : Condition Monitoring
CNN : Convolutional Neural Networks
CVRMSE : Coefficient of Variation Root Mean Square Error
DBNs : Deep Belief Networks
DE : Driving End
DFN : Deep Feed Forward Networks
DL : Deep Learning
DNN : Deep Neural Network
DSAE : Deep Stacked Auto Encoder
DT : Decision Trees
DTCWT : Dual Tree Complex Wavelet Transform
EMD : Empirical Mode Decomposition
FFT : Fast Fourier Transform
FN : False Negative Rate
FP : False Positive Rate
GA : Genetic Algorithm
GB : Gradient Boosted
GRU : Gated Recurrent Unit
HDN : Hierarchical Diagnosis Network
HELM : Hysteretic Extreme Learning Machine
HI : Health Indicator
ICA : Independent Component Analysis
I-CBM : Intelligent Condition-Based Maintenance
ICS : Industrial Control System
IoT : Internet of Things
I-PdM : Intelligent Predictive Maintenance
IWPT : Improved Wavelet Package Transform
KNN : k-Nearest Neighbors
LMU : Local Monitoring Unit
LSTM : Long Short-Term Memory
MAE : Mean Absolute Error
MHMS : Machine Health Monitoring Systems
MIT : Massachusetts Institute of Technology
ML : Machine Learning
MLP : Multilayer Perceptron

NDE : Non Driving End
PCA : Principal Component Analysis
PCC : Pearson's Correlation Coefficient
PdM : Predictive Maintenance
PHM : Prognostics and Health Management
PLC : Programmable Logic Controller
PLS : Partial Least Square
PM : Preventive Maintenance
PNN : Probabilistic Neural Network
PT&I : Predictive Testing and Inspection
PV : Photovoltaic
R2F : Run-to-Failure
RBF : Radial Basis Function
RBM : Restricted Boltzmann Machine
RCM : Reliability Centered Maintenance
REP : Rotating Equipment Performance
RF : Random Forest
RFID : Radio Frequency Identification
RMS : Root Mean Square
RMSE : Root Mean Square Error
RNN : Recurrent Neural Network
RTD : Resistance Temperature Detectors
RUL : Remaining Useful Life
SD : Standard Deviation
SDAE : Stacked Denoising Auto-encoder
STFT : Short-Time Fourier Transform
SVC : Support Vector Classifier
SVM : Support Vector Machines
TCM : Tool Condition Monitoring
TL : Transfer Learning
TN : True Negative Rate
TP : True Positive Rate
TPM : Total Productive Maintenance
TQM : Total Quality Management
TSP : Time to Start Prediction
WPT : Wavelet Packet Transforms
WT : Wavelet Transform
WVD : Wigner-Ville Distribution

LIST OF TABLES

Page

LIST OF FIGURE

Page

BUILDING SENSOR-BASED REAL-TIME PREDICTIVE MAINTENANCE SYSTEM BY UTILIZING ARTIFICIAL INTELLIGENT TECHNIQUES

SUMMARY

Rolling elements are commonly used in heavy-duty machinery, oil firms, water treatment processing, transportation, and aeronautical equipment among other applications. Bearing failures can lead to a complete breakdown of machinery, resulting in a disastrous accident and financial losses for the owner. As a consequence, correctly detecting the presence of these vulnerabilities as early as possible is crucial. Predictive maintenance is therefore of a great importance for lowering the cost of repairing or replacing capital-intensive infrastructure. The definition of predictive maintenance and condition monitoring techniques is extensively discussed in this dissertation.

Over the past few decades, a considerable attention has been paid to predictive maintenance policies as the basis for production management in many leading companies. Under the concept of zero failure manufacturing, predictive maintenance seeks to reduce downtime and maintenance costs by using real-time data to detect potential faults. The predictive maintenance principle states that maintenance is only done when it is required, which means that it is only done after analytical models have detected impending failures or degradations. In other words, the defects do not appear suddenly, they advance in time and increase the critical state of equipment aging. Considering this incremental and progressive nature, aging starts in the machine components when the operating condition changes from normal to critical.

This dissertation spotlights on the concept of predictive maintenance and condition monitoring techniques. Specifically, it presents a review of the most popular condition monitoring methods applied on rotary machines like pumps, motors, gearboxes, turbines, etc. The pump is chosen as a case study as it plays a vital role in our everyday lives. Besides, pump machines are considered the most used mechanical equipment after motors, and an immense amount of money is spent annually on their maintenance activities. The current study provides various intelligent monitoring techniques for detecting bearings defects at early stages so that predictive maintenance actions can be taken timely to prevent major pumping systems' failures.

Numerous publications discussing the topic of condition monitoring have been reviewed modules during this study. They give an overview of emerging capabilities in predictive maintenance by applying the Internet of Things technology. Achieving efficient predictive maintenance requires access to the machining process data (historical data and real-time), industrial network, and communication layers. These activities are accomplished by an intelligent condition monitoring system and industrial communication protocols.

Fault prediction models and predictive maintenance suggested in this research have been applied on forwarding pumping stations run by the SEWERAGE TREATMENT COMPANY (STC), one of the largest sewage treatment firms in Qatar. The stoppage

in any of those pumping systems leads to significant financial loss consequences such as repair cost, replacement costs, consequential damage, etc. In addition, risk consequences such as the potential for safety or environmental incident, breach of statutory or license requirement can rise.

The most challenging task this research attempts to attain is to keep pump machines in a functional state by estimating their operating condition in order to perform the necessary maintenance interventions. This would minimize machines downtime and achieve the maximum availability and reliability of pumping stations. The author proposes networked monitoring systems by IoT technology, specifically by SKF@ptitude observer monitoring, an expert diagnostics software commonly used for pump monitoring systems. Temperature measurement and vibration signal analysis are used to track rolling bearing conditions and provide more accurate detection results.

Temperature measurements help in identifying potential temperature-related equipment faults, such as excessive mechanical friction (faulty bearings, inadequate lubrication, fouling in a heat exchanger, and shoddy electrical connections). Variable vibration signals may indicate wear, imbalance, misalignment, or damage. These measurements assist in identifying the causes of bearing failure, which is caused mainly by temperature and/or vibration. The required maintenance action can be carried out based on the findings of these observations, the professional experience of maintainers, as well as the rotary machinery maintenance manual.

With artificial intelligence's massive regeneration, predictive maintenance has become the most effective process to deal with the vast amounts of data collected from smart manufacturing and complex engineering processes, particularly for implementing fault prediction systems based on data-driven approaches. This thesis presents two different case studies that utilize condition monitoring data and artificial intelligence techniques (namely machine learning and deep learning) as effective procedures for intelligent fault detection.

In the first case, supervised machine learning is combined with decision-making techniques to anticipate potential bearing failures and improve overall manufacturing operations by performing necessary maintenance actions at the right time. The integrated model has been applied in this research where the data fed (mainly temperature and vibration) belong to the labeled type. In this regard, a comparison of four different types of classifiers is conducted. These classifiers are: decision trees, random forests, gradient boosted, and support vector machines. The comparison is achieved using python programming package to investigate which type provides the highest detection accuracy. The predictive maintenance module's accuracy is tested using real-world industrial development datasets. Since the binary classification output of the applied machine learning algorithms would generate the pseudo probability of an observation belonging to a class, we decided to use the utility theory to leverage the likelihood of failures and thus help to perform correct maintenance behavior.

The second case study introduces four different deep architectures, which are mostly used in predictive maintenance field, namely, the Deep Feedforward Networks (DFN), a standard Long Short-Term Memory (LSTM), gradient boosted, and an LSTM with Convolutional Neural Networks (CNN). These models are implemented using a vibration signal dataset for roller bearings to assess their superiorities in fault identification and prediction. The vibration signals are first processed and extracted using the statistical time-domain method. The extracted statistical parameters are then fed to the suggested DL approaches to classify the bearings operating conditions.

As a result, we tried to take advantage of CNN and LSTM complementarity by merging them into a single unified architecture to train the model jointly. The experimental results are then compared with the other suggested deep learning models to confirm the best performing model in terms of fault detection and operation assessment. Five performance indicators are evaluated to measure the performance of the tested ML and DL algorithms: accuracy, F-score, precision, recall, and area under the curve (AUC).

Our research's novelty lays in the new perspective on predictions and the suggestion and comparison of several classifier models. This comparison is conducted on two real-world datasets from the pumping systems. Experimental results revealed that our proposed classifier models produced promising results.

YAPAY AKILLI TEKNİKLERİ KULLANARAK SENSÖR TABANLI GERÇEK ZAMAN TAHMİNLİ BAKIM SİSTEMİ KURULMASI

ÖZET

Yuvarlanma elemanları, diğer uygulamaların yanı sıra ağır iş makinelerinde, petrol firmalarında, su arıtma işlemlerinde, nakliye ve havacılık ekipmanlarında yaygın olarak kullanılmaktadır. Yatak arızaları, makinenin tamamen bozulmasına yol açarak, feci bir kazaya ve mal sahipleri için mali kayıplara neden olabilir. Sonuç olarak, bu güvenlik açıklarının varlığını olabildiğince erken doğru bir şekilde tespit etmek çok önemlidir. Bu yüzden kestirimci bakım; tamir, tadilat ve altyapı yenileme maliyetlerini düşürdüğünden büyük önem arz etmektedir. Kestirimci bakım ve süreç takip tekniklerinin tanımı bu tez içinde geniş bir biçimde tartışılacaktır.

Geçtiğimiz birkaç onyılda sektöre yön veren büyük şirketler, üretimin yönetilmesi açısından dikkatlerini kesitirimci bakım politikalarına vermişlerdir. Sıfır hata üretim politikaları altında, kestirimci bakım; arıza süreleri ve bakım maliyetlerini kısmak için, gerçek zamanlı verileri kullanarak potansiyel arızaları tespit etmek için çalışır. Kestirimci bakım prensibi şunu beyan eder; genel olarak bakım sadece gerekli olduğu zaman gerçekleştirilir, bu da analitik modellerin yaklaşmakta olan bir arıza veya bozulmayı tespit etmesinden sonra yapılmasıdır. Diğer bir deyişle arızalar aniden ortaya çıkmaz, zaman içerisinde gelişirler ekipmanların kritik yıpranma sürecini artırırlar. Bu sürecin artan ve ilerleyen doğasını göz önüne aldığımızda, makine parçalarındaki yıpranma makinelerin işleyiş koşullarının normalden seviyeden kritik seviyeye gelmesiyle başlar.

Bu tezde kestirimci bakım ve durum izleme konseptlerine dikkat çekilecektir. Özellikle dönen parçalardan ve aksamlardan oluşan pompalar, motorlar, vites kutuları, ve türbinlerdeki en bilindik süreç izleme methodları sunulacaktır. Çalışma konusu olarak pompa seçilmiştir. Pompaların günlük hayatımızda önemli rol oynadığı malumdur. Bunun dışında pompalar motorlardan sonra en sık kullanılan mekanik ekipmanlardır ve yıllık bakım faaliyetlerine ciddi miktarda para ödenmektedir. Yakın zamanda yapılan çalışmalar akıllı takip tekniklerinin erken safhalarda rulman kusurlarını tespit edebilmesini sağlayarak pompa sistemlerinin büyük bir arızaya girmeden kestirimci bakım aksiyonlarının alınmasını sağlamaktadır.

Bu çalışmada durum takibi başlığı altında yayımlanmış bir çok yayın gözden geçirilerek değerlendirilmiştir. Bu çalışmalar; Nesnelerin İnterneti teknolojisini uygulayarak artan kapasiteleri ile kestirimci bakım hakkında genel bakış sağlamaktadır. Etkili bir kestirimci bakıma ulaşmak için makine işlem verileri (tarihsel ve anlık), endüstriyel ağ ve iletişim katmanlarına erişim gereklidir. Bu aktiviteler akıllı durum takibi sistemi ve endüstriyel iletişim protokolleri ile tamamlanmış olur.

Bu araştırmada sunulan arıza tahmini modelleri ve kestirimsel bakım çalışmaları Katar'ın en büyük pis su artıma tesislerinden biri olan SEWERAGE TREATMENT COMPANY (STC) 'nin pompa nakil istasyonlarından birinde uygulanmıştır. Bu sistemlerde meydana gelebilecek en ufak bir aksama veya durma hem mali anlamda

ciddi bir kayba (tamir masrafı, değiştirme masrafları ve takibinde meydana gelebilecek olasu hasarlar) bunun yanında çevresel, güvenlik ve görev ihlali veya lisans gereksinimleri risklerini artırabilir.

Araştırmada en zorlayıcı kısım; pompa makinelerini işler düzeyde tutacak şartların tahmini ve bakım müdahalelerinin uygun anda yapılmasının sağlanmasıdır. Bu şekilde zorunlu duruş süreleri minimize edilir ve pompa istasyonlarının güvenilir ve maksimum kapasitede çalışması sağlanır. Yazar nesnelerin internet teknolojili ağ izleme sistemleri; özellikle pompa izleme sistemlerinde sıklıkla kullanılan SKF@ptitude gözleyici izleme, uzman tanı yazılımı önermektedir. Sıcaklık ölçümleri ve titreşim sinyal analizleri döner rulman sistemlerinin durumunun takibinde daha kesin sonuçların elde edilmesi için kullanılır.

Sıcaklık ölçümleri yüksek mekanik sürtünmeden dolayı aşırı ısınma sebepli ekipman arızalarının tespitine yardım eder. (Arızalı rulmanlar, yetersiz yağlama, eşanjörde oluşan tortular ve gelişigüzel yapılmış elektrik bağlantıları) Değişken titreşim sinyalleri; yorulma, dengesizlik, hatalı hizalanma veya hasarları göstermekte faydalı olurlar. Bu tip önlemler başlıca rulman arıza sebebi olan ısınma ve titreşimin tespit edilmesine yardımcı olurlar. Gerekli bakım onarım faaliyetleri bu gözlemlerden sonra profösyonel bakımcılar tarafından veya bakım kılavuzlarına göre gerçekleştirilebilir.

Yapay zekanın muazzam gelişimi, kestirimci bakım faaliyetinin; akıllı üretim ve karmaşık mühendislik işlemlerinden elde edilen (genellikle veriye dayalı arıza tespit system yaklaşımlarında) devasa boyuttaki verilerin en etkili şekilde işlenmesini mümkün kılar. Bu tezde akıllı arıza tespitinde etkili prosedür olarak süreç takip verileri ve yapay zeka teknikleri (makine öğrenmesi ve derin öğrenme) olmak üzere iki farklı vaka çalışması sunulacaktır.

İlk durumda, kontrollü makine öğrenmesi, karar verme teknikleri ile birlikte kullanılarak potansiyel rulman arızalarının tespitini ve genel üretim operasyonunu geliştirmek için doğru zamanda gerekli bakım faaliyeti aksiyonunu almaktadır. Bu çalışmada işaretli tipte veri beslemeye (genel olarak sıcaklık ve titreşim) ait olan tümleşik model uygulanmıştır. Bu bağlamda 4 farklı sınıflandırıcı mukayesesi yapılmıştır. Bu sınıflandırıcılar: karar şemaları (DT), rastgele orman (RF), gradian artırma (GB) ve destek vektör makineleridir (SVM). Python bilgisayar programı kullanarak hangi tipin en yüksek kesinliği sağladığı tespit edilmiştir. Kestirimci bakım modülünün kesinliği endüstiryel gerçek dünya endüstriyel gelişim veri setleri kullanılarak test edilmiştir. Ikili sınıflandırma çıktıları uygulanmış makine öğrenme algoritmaları ilgili sınıfa ait yalancı gözlem ihtimalleri oluşturacağından, arıza olasılığından yararlanmak ve böylece doğru bakım davranışını gerçekleştirmeye yardımcı olmak için fayda teorisini kullanmaya karar verdik.

İkinci vaka çalışması, çoğunlukla kestirimci bakım alanında kullanılan dört farklı derin mimariyi tanıtmaktadır: Derin İleri Beslemeli Ağlar (DFN), standanrt Uzun Kısa Süreli Bellek (LSTM), gradyan artırılmış (GB) ve Evrişimli Sinir Ağlarına sahip (CNN) ve LSTM. Bu modeller, arıza tespiti ve tahminindeki üstünlüklerini değerlendirmek için makaralı rulmanlar için bir titreşim sinyali veri seti kullanılarak uygulanır. Titreşim sinyalleri ilk olarak istatistiksel zaman alanı yöntemi kullanılarak işlenir ve çıkarılır. Çıkarılan istatistiksel parametreler daha sonra rulman çalışma koşullarını sınıflandırmak için önerilen DL yaklaşımlarını beslemek için kullanılır.

Sonuç olarak, modeli birlikte eğitmek için CNN ve LSTM tamamlayıcılığını tek bir birleşik mimaride birleştirerek yararlanmaya çalıştık. Daha sonra deneysel sonuçlar, hata tespiti ve operasyon değerlendirmesi açısından en iyi performans gösteren modeli

doğrulamak için önerilen diğer derin öğrenme modelleriyle karşılaştırılır. Test edilen ML ve DL algoritmalarının performansını ölçmek için beş performans göstergesi değerlendirilir: doğruluk, F-skoru, hassasiyet, geri çağırma ve eğri altındaki alan (AUC).

Araştırmamızın yeniliği, tahminler ve birkaç sınıflandırıcı modelin önerisi ve karşılaştırması üzerine yeni perspektifte yatmaktadır. Bu karşılaştırma, pompalama sistemlerinden iki gerçek dünya veri setleri üzerinde yapılmıştır. Deneysel sonuçlar, önerilen sınıflandırıcı modellerimizin umut verici sonuçlar verdiğini ortaya koydu.

1. INTRODUCTION

This chapter provides a general overview of the maintenances in industries and a review of the maintenance strategies relevant to this study. It introduces the motivation, framework, and objectives of this thesis. Finally, the thesis organization is outlined.

1.1 Background

In modern business, reliability and maintenance have a major effect on the three core elements of competitiveness: quality, expense, and time to market. Machines that are well-maintained retain tolerances better, minimize scrap and rework, increase component accuracy and efficiency, and lower overall production costs. Owing to a lack of recognition of the machine's output actions and the need to increase production quality, many factories still conduct reactive maintenance on machinery today.

Systems that work in industrial environments must be highly reliable and accessible. Complex equipment's operating capacity is essential for facilities such as power plants, assembly lines, and oil and gas firms. The costs of missed output income aren't the only consequences of unscheduled maintenance. Failures affect the bottom line because of the costs of maintenance and clean-up, while critical failures can damage people and the environment. Reliability engineering and production management have long tried to reduce the risk of failure in these settings [1]. According to surveys on equipment reliability issues undertaken over the last 30 years, maintenance is to blame for about 17% of manufacturing interruptions and quality matters. The remaining 83% is completely beyond the reach of traditional maintenance [2].

To understand the maintenance concept, there are several definitions, which are necessary to follow the subsequent chapters:

The term "maintenance" comes from the dictionary. "the work of keeping something in proper condition" In the literature on maintenance management, maintenance is characterized as a series of technical and managerial actions aimed at maintaining or

restoring an asset or system to a state where it can perform the necessary functions[3]. In[4] maintenance is described as having a significant impact on the three critical elements of important factors: cost, quality, and product lead-time. Machines that are regularly observed and maintained retain tolerances better, minimize scrap and rework, and increase component accuracy and efficiency. While Sethiya [5] depicted maintenance as a collection of actions taken to keep a factory or equipment from falling or to restore deteriorated equipment. According to the current maintenance classification, maintenance is split into three important categories: corrective, preventive, and predictive maintenance, also known as condition-based maintenance.

Maintaining a system regularly can help it meet its availability, reliability, product quality, and safety requirements. For meeting production goals, ensuring safety, and lowering costs, equipment reliability is critical. It also helps prevent catastrophic failures which result in extended downtime and labor/spare parts costs. Maintenance costs account for a significant portion of all manufacturing or production facility running costs. Maintenance costs can vary from 15% to 60% of the cost of products produced, depending on the industry. In the food industry, for example, average maintenance costs account for around 15% of the cost of products produced.

Maintenance costs for iron and steel, pulp and paper, and other heavy industries, on the other hand, can account for up to 60% of overall production costs [2]. Practically, all operating systems are subjected to deterioration and random failure in performing their assigned tasks. The maintenance is reflected by the renewal of the equipment deterioration state, and after maintenance, the manufacturing system is restored to an as-good-as-new condition [6]. Martin provides a short description of the evolution of machine tool repair techniques [7].

Run-to-failure maintenance, which happens only when something fails, is one of the earliest maintenance strategies. Time-based preventive maintenance is a later maintenance technique that defines a daily period for conducting preventive maintenance (PM) regardless of a physical item's health status. Products are becoming more complex as new manufacturing technology advances at a rapid rate, necessitating higher quality and reliability. As a result, PM costs continue to increase. PM has gradually become a major burden for many manufacturing firms. Different components and subsystems may work together to accomplish a specific purpose. It is important to understand the consequences and causes of a malfunction as soon as

possible and take effective maintenance steps. Only a small portion of a system's downtime can be spent determining the root cause of the problem. A catastrophic failure may result in a serious accident as well as significant financial consequences for the firm.

Consequently, early prediction capability, which can prevent failures from developing and ultimately turn into a serious problem, is meaningful and imperative for industrial situations [8]. Therefore, more effective maintenance approaches are being introduced to manage the situation, such as predictive maintenance and condition-based maintenance (CBM). According to the available information, the concept is to schedule maintenance, indicating the equipment current state or predicting specific degradations. However, it was difficult to achieve the goal because of technological capability and difficulties acquiring and integrating all essential information in the last few decades. With the fourth industrial revolution trend, we also see the potentials and challenges of predictive maintenance in this new era.

1.2 Maintenance Strategies

A maintenance strategy covers the entire maintenance management process, which includes defining maintenance priorities, deciding the maintenance schedule, and establishing the maintenance organization. PM is credited with developing the maintenance management discipline in the 1950s. It has grown into a sophisticated area of study over time. Today, maintenance management covers terms such as reliability-centered maintenance (RCM), predictive maintenance (PdM) or CBM, and total productive maintenance (TPM). The principles of maintenance management are used by the majority of businesses to operate their equipment [9].

According to [10], Maintenance management can be categorized into three groups, which increase complexity and productivity in order.

1) Run-to-failure (R2F), also known as (unplanned maintenance), is a maintenance technique in which maintenance interventions are carried out only after failures have occurred. This is the best strategy for dealing with maintenance (and for this reason, it is frequently adopted). Still, It is the least efficient. Intervention costs and downtime following the failure are normally

significantly higher than those associated with expected corrective actions conducted ahead of time.

2) Maintenance steps are taken out according to a planned schedule depending on time or procedure iterations in preventive maintenance (PM), also known as (planned maintenance). Failures are typically avoided in this method, often known as scheduled maintenance, however excessive corrective steps are often undertaken, resulting in ineffective resource usage and higher operating costs. By reducing the number of failures and avoiding unplanned corrective maintenance, this approach aims to increase equipment reliability and availability. Since the correction is made before the malfunction, the equipment can be shielded from faults in PM. Even so, due to preventive maintenance, equipment elements can need to be replaced earlier than expected [11].

3) Predictive maintenance (PdM) is when equipment is monitored continuously and maintenance is done only when it is required [12]. The PdM framework enables advanced identification of suspension failures and allows timely pre-failure interventions using prediction tools based on historical data (e.g., ML and DL strategies), integrity factors (e.g., visual features, wear, coloration different from the original, among others), statistical inference methods, and engineering processes. Another technique worth mentioning is RCM, which was developed in the 1960s but was initially geared toward aircraft maintenance and was used by aircraft manufacturers, airlines, and the government. It wasn't until two decades later that it started to spread to other industries. RCM is an analysis method for guiding maintenance activities on machines and components where reliability (i.e., safety) is critical [9]. RCM incorporates PM, Predictive Testing and Inspection (PT&I), Repair (also known as reactive maintenance), and Proactive Maintenance to improve the likelihood that equipment will perform as expected during its design life cycle with minimal maintenance and downtime [13].

When a small machine stops and requires minor maintenance, such as lubrication or screw tightening, the operator must wait for the technician to complete the task, resulting in lost production time. As a result, it's logical to assume that the operator can be taught to make such minor adjustments. This way of thinking gave birth to the

TPM philosophy, which has since become an integral part of the Total Quality Management (TQM) philosophy [14]. TPM aims to reduce losses associated with equipment failure, startup and adjustment time, operations at reduced speed, and low quality to improve equipment efficiency [15].

Predictive maintenance, as the most common and current maintenance technique, tests parameters in the state of equipment to perform the necessary tasks, extending the life of equipment and processes while reducing the probability of failure risk. The technique is based on the objective phenomenon that when machinery begins to malfunction, it is possible to detect various forms of symptoms, such as temperature fluctuations, vibration, or noise, if sharp eyes, ears, and noses are used to detect failure precursors [16]. In this era of rapid technological growth, sensors have progressed to the point that they can replace sharp eyes, ears, and noses. However, it does not function as a trigger for constructive maintenance. Like the human brain, we must also derive information from these signals, uncover the intelligence behind information, and collect data about possible failures. As a result, condition-based predictive maintenance's crucial challenges in the last few decades have been accessed to the information necessary for condition monitoring and the accuracy of fault prediction. A machine's life is represented in (Figure 1.1) as a classic bathtub curve. Since a system's malfunction is typically followed by a rise in vibration and/or noise, the vibration level follows the same bathtub curve shape. During the initial run-in period, the vibration level decreases, then gradually increases due to normal wear during the normal operating condition period. Finally, it develops rapidly until failure or breakdown during the wear-out phase due to prolonged wear [17,18].

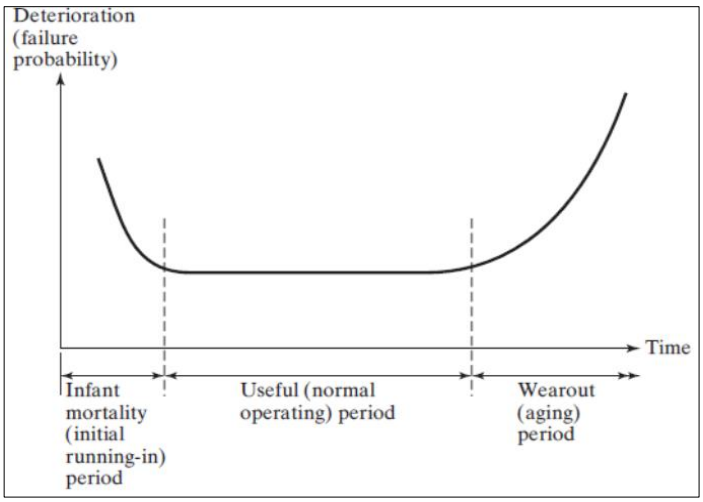


Figure 1.1 : The bathtub curve for the life of the equipment [19].

It is worth noting that some of the advantages of PdM and active approaches to system monitoring include the following:

- Maintenance affects production by increasing production capacity and reliability of machines and the quantity of output.
- Reduced quality goods and an inability to meet high-quality standards result from improved quality machine deterioration.
- Machines and processes have been given a longer life span.
- Unproductive downtime is reduced by significantly reducing the total time spent maintaining machines and providing adequate lead-time for proper maintenance and repair schedule.
- Improve customer relations by improving machine performance consistency and staying on schedule.
- Reduce and predict machine failures in advance to improve operator safety.
- During the process, there should be as little equipment interference as possible.
- Reduce production costs by eliminating unnecessary maintenance. Machine failures that are costly and catastrophic are avoided. It is possible to reduce or eliminate safety stock.
- The use of CBM modelling in maintenance decision-making can help to reduce incorrect maintenance activities by utilizing measured CM information [20].

The new maintenance patterns, as represented by machinery status tracking and CBM techniques, lean toward a predictive approach. As a consequence, intelligent predictive maintenance (I-PdM) or Intelligent condition-based maintenance (I-CBM) is a cutting-edge maintenance technique which can anticipate future failures and take prompt and effective maintenance steps [21]. It has gradually replaced traditional maintenance policies such as breakdown maintenance, RCM, PM, and others, which cannot eliminate faults and may no longer meet the demands of the modern industrial world. Wang et al. [22] proposed a new maintenance strategy classification, as illustrated in (Figure 1.2). They have separated the PdM from PM. This time corrective maintenance, PM, and PdM are three parallel kinds.

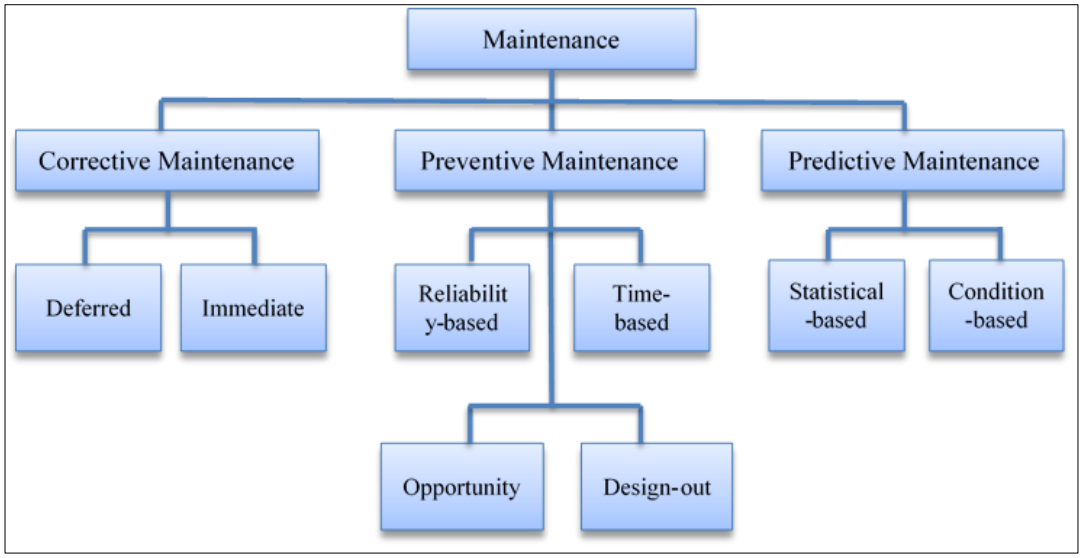


Figure 1.2 : Maintenance Strategy.

Corrective maintenance is defined as a kind of maintenance achieved to recognize and rectify the cause for failures in a flawed system. It focuses on identifying failures that could involve symptom failures from the failure phenomenon [23]. Under this strategy, failure is allowed to occur as maintenance is carried out, which means that it is only acceptable if the effect of failure is minimal. Some issues do not matter, such as whether the equipment falls or how long the repair can take. This type of maintenance is classified into two categories: immediate and deferred.

Interestingly, the availability of vibration, temperature, load, and other forms of condition monitoring data for electrical and mechanical devices has increased dramatically over the last few years. As a result, these data have been functionalized according to particular operating conditions. More explicitly, when the operating conditions reach a certain critical level, alerting signals are displayed by developed models to apply PdM [24]. Industrial equipment is connected as a community and independently exchanges information, particularly in the definition of Industry 4.0, which means that abundant industrial data can be obtained conveniently for condition-based maintenance. So, Artificial Intelligence (AI) techniques, in particular, Machine Learning (ML), Deep Learning (DL), and Transfer Learning (TL), have been extensively included in the latest PdM program in the face of industrial big data [21].

The use of machine learning in condition monitoring is becoming more prevalent than traditional approaches due to the increased availability of computing resources and vast algorithmic advances. In the field of AI, ML has emerged as an effective tool for

constructing intelligent predictive models in a variety of applications [25]. ML methods are computer programs employed to solve a given problem using data or experience. These programs can learn from data, where learning is the process of obtaining new knowledge. Hence, for PdM applications, machine learning offers strong predictive approaches. As for DL, it has grown over time as computing capacity and big data have increased. Data mining attempts to derive valuable knowledge from a large volume of data by modelling high-level abstraction. In PdM, several DL methods have been employed.

The final phase in a PdM schedule is maintenance decision-making. Maintenance employees decisions on maintenance actions will benefit from adequate and reliable decision support. Diagnostics and prognostics are the two primary types of maintenance decision support strategies in a PdM programme. When faults occur, fault diagnostics focus on finding, isolating, and recognizing them. However, prognostics seek to foresee defects or failures before they happen. Since prognostics can prevent errors, they are superior to diagnostics. If this isn't necessary, be prepared for problems (with replacement parts on hand and human resources set aside) to prevent further unplanned maintenance costs [26].

1.3 Research Motivation

Predictive maintenance gathers equipment parameters, detects changes in the physical state of equipment, and discovers fault details, such as when, where, and what sort of fault can occur, as an ideal maintenance policy. PdM may arrange appropriate maintenance activity through fault information to optimize the service life of equipment without raising the risk of failure. Predicting a possible future fault gives adequate time before the damage occurs for maintenance preparation (tools, spare parts, and technicians). Ideally, how to reduce downtime and maintenance costs and achieve zero equipment breakdowns is always a critical issue for a company to be efficient and sustainable.

Sensor data is becoming more widely accessible, technology is becoming more complex, and many data-driven approaches are evolving, allowing for new PdM innovations. Through the use of artificial intelligence in smart factories, this research establishes a framework for identifying PdM's benefits and condition monitoring in rotating equipment, especially bearing components. Rotary machines are commonly

used and are important to the majority of engineering processes. Stator, rotor, and bearing failures are all common machine failures. Approximately half of all machine failures are caused by bearing failures [27]. As a result, the SEWERAGE TREATMENT COMPANY (STC) forwarding pumping stations were chosen as a case study of such a fault prediction application in the current study. Pumps, as we all know, play an major portion in our daily lives. Pumps are kept going by continually tracking their condition to avoid downtime. It also assists production managers in preparing maintenance tasks. As a consequence, it is important to keep track of bearing conditions in rotary machines. To control the bearing state, different techniques should be used depending on the application.

Pumping equipment and the pumping station of a sewerage pumping station are essential elements. Pumping equipment is subject to wear, tear, degradation, and corrosion as a result of its use, making it prone to failure. More faults or interruptions in Sewerage Pumping Stations are caused by pumping machinery than any other component. Proper operation and timely maintenance and upkeep of pumping stations and pumping equipment are necessary to ensure an uninterrupted sewerage pumping station. Inspections carried out on time, steps taken in response to inspection results, and planned routine maintenance can also help to avoid unexpected failures. Holding a stock of fast-moving spare parts on hand will help you cut down on downtime. Due to normal wear and tear, the efficiency of pumping equipment degrades over time. If timely action is taken to restore efficiency, energy bills can be kept within reasonable limits. It's also vital to maintain correct records.

All of these variables must take into account for pumping machines to operate efficiently and consistently. This study looks at how to apply fault prediction models, as well as the issues that occur when working with pumping machinery and its electrical and mechanical components.

Finally, rather than relying solely on off-line sensor data, this study's main motivation is to combine on-line sensor measurement and off-line model training in condition monitoring of the bearing component. Off-line model training is done with historical data, while on-line data is gathered from real-time sensor measurements to predict the process state.

1.4 Research Framework

The predictive maintenance program framework is presented in the thesis to achieve accurate fault detection, efficient maintenance planning, an extension of data sources, and equipment condition monitoring. As a consequence, the efficiency of maintenance execution could be improved. The framework also offers facility managers and researchers an overall understanding and helpful guidance to implement predictive maintenance in the manufacturing industry. Figure 1.3 shows the study framework and summarizes the main procedures and steps.

In this thesis, the supervised machine learning method was used, with the data fed (mostly temperature and vibration) being of the labelled type. Decision Trees, Random Forest, Gradient Boosted trees, and Support Vector Machine are the four classifiers compared in this study. Besides, four DL architectures, Deep Feedforward Networks, standard Long-Short Term Memory, Gradient Boosted, integrated Convolutional Neural Networks with Long-Short Term Memory are established using a new sample of vibration signal belong to the pumping system to interpret their superiorities in fault recognition and prediction in the predictive maintenance. With the experimental results, a comparison of ML and DL methods is presented. This comparison was made with the help of the Python programming language, which was used to see which type has the best detection accuracy.

Since the applied ML algorithms' binary classification performance can generate the pseudo probability of an observation belonging to a class, decision making using utility theory is used to exploit the probability of failures and thereby assist in the implementation of successful maintenance interventions. This offers a logical framework for determining the correct action with the greatest expected benefit to decision-makers.

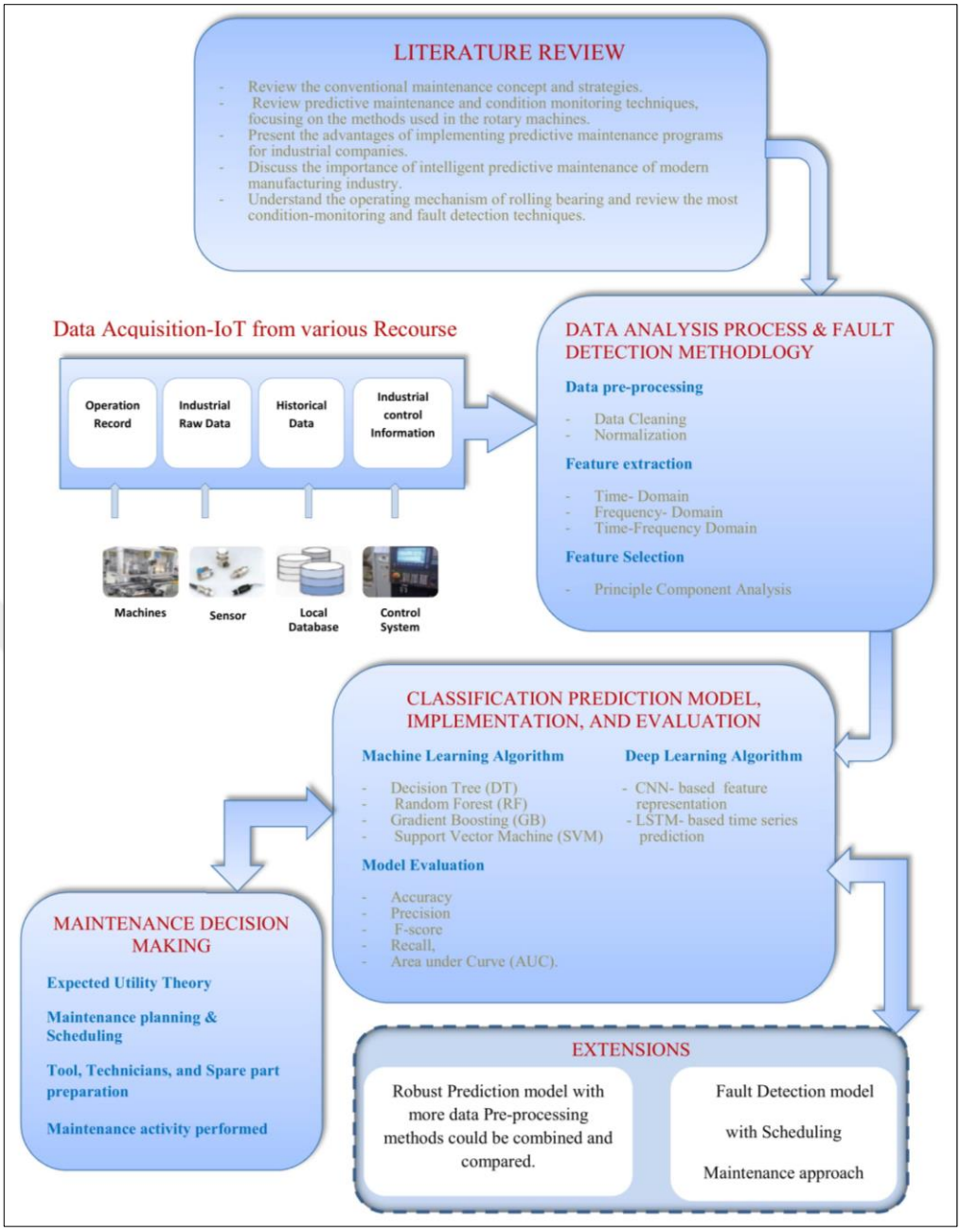


Figure 1.3 : Research framework.

1.5 Research Aim and Objectives

This study aims to understand whether ML and DL approaches can be used in the field of predictive maintenance of rotating systems to detect common mechanical faults in bearing components. The following are the main objectives for achieving this aim:

Objective 1: To address condition monitoring methods and their applications to rotary machines, as well as to study predictive maintenance strategies.

Objective 2: To concentrate on using an ideal data-driven model to detect impending faults and maintenance requirements in rotary machines, especially in a pumping system.

Objective 3: To identify rolling bearings types in general, focus on the types used in the research as a case study and locate their operational characteristics, degradation mechanisms, and fault reasons.

Objective 4: To present a framework for predictive maintenance concerning the artificial intelligent approaches to achieve accurate fault detection, efficient maintenance planning, an extension of data sources, and condition monitoring of equipment.

Objective 5: To establish a predictive model using Internet of things technology to collect the process data integrated with machine learning techniques.

Objective 6: To review various machine learning techniques and their important parameters on detecting bearing faults.

Objective 7: To realize a prediction model to detect the critical operational conditions commonly present in bearing components of the single-stage forwarding pump before they lead to actual malfunction and/or stop the manufacturing process.

Objective 8: To conduct a comparison of different machine learning methods to illustrate the most useful predictive methods for the application in the industry. Moreover, various techniques are applied to show additional accuracy and precision according to the case study.

Objective 9: To execute feature selection and extraction to identify necessary signals and characteristics before training the data-driven model. This task is accomplished by a statistics method and the principal component analysis solution to perform dimension reduction.

Objective 10: To provide a general framework for integrating the concept utility theory with machine learning to improve PdM's decision-making.

Objective 11: To overcome the dependence of classical machine learning algorithms on feature extraction methods utilizing deep learning methods.

Objective 12: To present various deep learning algorithms regarding vibration analysis in the roller bearing elements to predict the malfunctions in rotating machinery (using different sample data for the same case study).

Objective 13: To introduce an experiment of fault classification for pumping systems to provide a comprehensive comparison of various deep learning methods. A novel CNN-LSTM approach is proposed for bearing fault prediction to train a data-driven model in different cases.

Objective 14: To propose two case studies for two sample data as an experiment of the fault classification model for bearing components to achieve PdM in practical application.

Objective 15: To construct a predictive maintenance platform for rotating machinery in the modern industry depended on experimental results from fault prediction models.

Objective 16: To assess the performance of the proposed prediction models several evaluation measures are intended.

1.6 Thesis Organization

The remainder of this thesis report is organized as follows. Chapter 2 provides a literature review related to predictive maintenance policy development and highlights online process data monitoring. It provides a comprehensive description of the data collection technique used to collect the data required to implement the prediction model. The development and application of the internet of things technology in smart manufacturing are also discussed in this chapter. Chapter 3 provides helpful guidelines about rolling bearing and describes the fundamental operating mechanisms for thrust ball bearings. It gives a general overview of different condition monitoring and fault detection techniques for roller bearings, such as vibration analysis and temperature monitoring techniques. Chapter 4 presents a literature review relevant to the development of prediction models using machine learning algorithms in rotary machines. It introduced a case study implementing a bearing fault detection model for pumping systems and gives the experimental results for detection accuracy. Chapter 5 demonstrates an experiment of fault classification and fault detection for rotating machinery through deep learning approaches. Finally, chapter 6 concludes the thesis and highlights future research directions.

2. PREDICTIVE MAINTENANCE

The primary maintenance principles in Chapter 1 include general guidance for handling maintenance operations as well as a study of maintenance management and strategies. Maintenance managers must master tactical skills to organize and run the logistics activities of the maintenance role to efficiently address maintenance operations and apply ideas. As a consequence, this chapter aims to go through the pumping system's predictive maintenance and express condition monitoring techniques.

Moreover, it describes the data sample that was used to implement the prediction model. Finally, it presents a review of online process data monitoring and the internet of things, the most relevant strategy in this study.

2.1 Introduction

Condition-based maintenance, or PdM, is a technique for detecting impending faults before they become critical, allowing for more precise PM preparation [5]. PdM is commonly recognized as the most newest policy in the evolution of maintenance management, and its worth is better known when compared to conventional management policies. PdM maintains track of the mechanical condition, operating performance, and other measures of a process's health to provide the details needed to ensure the longest possible time between repairs. Furthermore, to reduce the number of unscheduled maintenance activities triggered by system failures and to make reasonable maintenance decisions to prolong the equipment's service life [2,28]. As a result, PdM can lead to substantial improvements in equipment and facility availability, protection, efficiency, and productivity [29]. Since PdM guarantees adequate time for maintenance planning (operators, equipment, spare parts, and so on), industrial organizations can schedule maintenance activities more effectively and flexibly.

Conventional maintenance management approaches are not replaced by predictive maintenance. It is, however, an essential component of a systematic, long-term

planning and maintenance programme. Conventional maintenance management systems depend on routine maintenance of all equipment and a swift responding to unplanned damages. A programme for PdM schedules complex maintenance activities as required by plant machinery. It will not be able to remove the need for conventional maintenance programmes (i.e., R2F and PM). PdM, on the other hand, will assist in the avoidance of unanticipated errors as well as provide a more effective scheduling system for repetitive PM activities.

Before a system breaks down, there are many PdM techniques for assigned faults, and new techniques are created every year. Chemical/particle analysis, vibration analysis to detect incipient problems such as bearings and gearboxes, temperature control, infrared image monitoring for electrical switchgear, engines, and electrical equipment, ultrasonic inspection, acoustic emission, lubricating oil analysis, and advanced visualization techniques are some of the most common PdM techniques [9]. As a result of these measures or assessments of machine state, maintenance activities are planned, typically trending parameters and forecasting lead time to failure. The premise is that most mechanical components show warning signs that they are about to malfunction. Use these warning signals to determine the machine's condition and boost preparation to fix equipment issues. Schedule system maintenance and upgrades during scheduled outages to prevent unplanned worker overtime triggered by reactive equipment management [30,31].

Condition monitoring (CM) is a simple PdM implementation that integrates equipment condition measurements into maintenance planning. PdM characteristics are often reported by inspecting the device regularly or by tracking it with various sensors regularly [32]. Industry machines have sensors mounted that output real-time data, which is then sent over the network to a monitoring device. The developed data is also preserved indefinitely so that historical views can be produced. The PdM strategy is based on historical and real-time data. As a result, the majority of PdM studies concentrate on identifying early warning signs of degradation and taking adequate precautions. It's good to realize that these indicators appear before 99% of system failures [4]. The CM procedure is divided into two sections. It begins by gathering the equipment's condition data (information). Second, it enhances awareness of failure causes and consequences, as well as equipment degradation patterns, which can be used to identify and evaluate equipment condition during service. The three kinds of

instrumentation systems which can be used for CM of equipment are the basic system, portable system, and computer-based system [17].

Data collection, data pre-processing, and maintenance decision-making are the three key steps in PdM's related operations, as shown in (Figure 2.1), [33,34]. Data is registered and collected from the sensors in the first phase, data acquisition. PM schedule decisions, on the other hand, are based on the knowledge and instincts of the individuals involved, as well as alerting systems, spreadsheets, operator logs, and change transition discussions. The second step in a PdM's work is data pre-processing, which changes and interprets the data obtained in the first step with noise reduction. The final step includes making a conclusion based on the sequence of data analysis, which is primarily dictated by the type of data being analyzed. Scientists have created a variety of models and sophisticated algorithms that can be used to better understand and interpret different datasets. These decisions can alter operating procedures or maintenance plans, necessitating the collection and analysis of additional data. Reports can be created and archived for future reference after a decision has been taken, and assessments can be conducted if necessary [35].

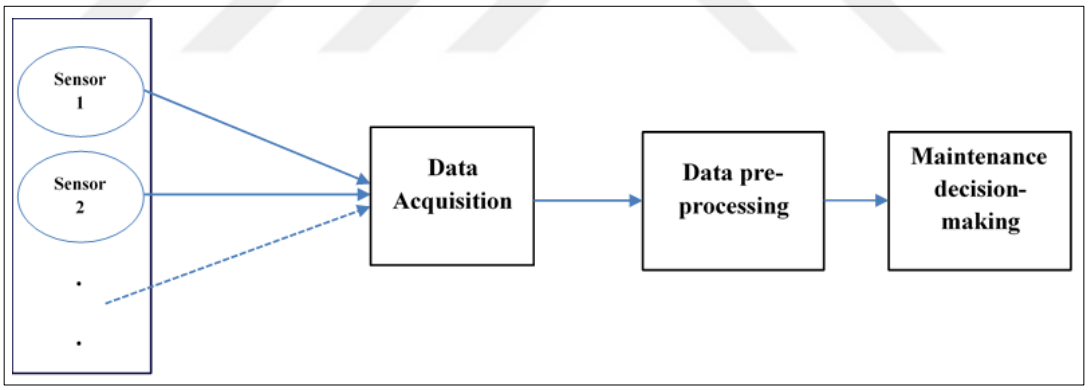


Figure 2.1 : Various steps in the PdM program.

In general, PdM's main goal is to anticipate a device malfunction by predicting early warning signs of deterioration and make maintenance more constructive PdM aims to forecast a system's failure cycle using experience, physical law, and machine learning methods to patch faulty components until they fail, minimizing system latency, maintenance costs, and product consistency [36]. As a result, using efficient PdM will save 8% on maintenance costs while also increasing productivity by 8% [37].

2.2 The Internet of Things

The Internet of Things (IoT) is a network of physical objects and computers that assists in data collection and sharing. These devices have designed a portal to connect to machines and their subcomponents to capture process data and parameters, as well as physical health aspects of the machine such as vibration, temperature, temperature, viscosity, flow rate, acoustics, and displacement signals, and provide the most up-to-date information for the network system of processing, transmission, analysis, and feedback [38]. Early fault identification and identification, machine health assessment, and future state prediction are all common uses for this information. Machine learning algorithms, which are applicable across a variety of learning realms, make this possible [39]. IoT has been versioned for use in maintenance, particularly PdM, because of the two most significant reasons for PdM's growth. The first is that modern equipment often has embedded computer chips for reading and controlling, allowing for data collection. Second, the cost of incorporating embedded sensors and other new information technology has decreased and is continuing to do so. [40]. To put it another way, the IoT refers to the growing trend of equipping physical objects and devices with sensing (all forms of sensors and wireless sensor networks), computing, and networking capabilities, then linking them to form a network and exploiting the collective effect of networked devices.

The Internet of Things idea was born out of a network of radio frequency identification (RFID) systems established by the Massachusetts Institute of Technology's (MIT) automatic identification center in 1999. All of the items in this system, such as sensing devices, can be linked to the Internet through radio frequency identification information. It's all about achieving intelligent identification and management. The primary functions are knowledge collection, transmission, retrieval, and application of information [41].

With the emergence of the IoT, PdM can now be extended to all forms of computers in all industries, resulting in a paradigm shift that generates major new business opportunities. Furthermore, integrating IoT into PdM task will provide new possibilities for E-maintenance, Remote maintenance, and Tele-maintenance applications [32].

2.3 Online Process Data Monitoring (Data Acquisition System)

An inevitable necessity for the PdM application is data, from which to analyze the past asset usage and extrapolate to the future. These data were collected from different sources such as Programmable Logic Controller (PLC) units and Data Acquisition (SCADA) systems that can be subdivided under the term Industrial Control System (ICS) [42].

The CM process can be performed in two ways: online or offline. On-line processing occurs when the equipment is in use, whereas offline processing occurs when the equipment is not in service [43]. The voltages of the sensor's data, like temperature, vibration, and pressure, are captured by a data acquisition device for online and offline monitoring. In industrial machinery, process data acquisition is moving toward automated systems (continuous or on-line systems), which in some cases provide more substantial benefits than data acquisition with handheld data collectors for the following reasons:

- Since data is collected constantly, a significant reduction in data collection intervals allows for the identification of any system status changes.
- Automatic data acquisition reduces the labor expense of collecting computer data, resulting in lower operating costs.
- Since the data is measured simultaneously with the same sensor, the data measurement precision is higher, and the data collection can be adjusted to particular machine operating conditions (speed and load).

The current research presents a fault detection model in the roller bearings component using data collected by IoT technology, expressly, by SKF@ptitude observer monitoring, an expert diagnostics software commonly used for pump monitoring system illustrated in (Figure 2.2). It maximizes the rotating equipment performance (REP) via allowing more agile business, delivers greater output, reliability, and optimizes safety. Various sensors are installed along with the pumping system components used to measure the data needed as input to the prediction model. Various sensors are installed along with the pumping system components used to measure the data needed as input to the prediction model. Furthermore, other pertinent information is presented in a user-friendly manner. Live data, which is modified every second, as well as long-term history, can be viewed in several formats. Live data and warning

indicators for pumps are shown in descriptive pictures in the process overview pane. The SKF@ptitude observer monitors bearing temperature and vibration directly. Furthermore, the SKF@ptitude observer detects bearing noise, detecting defects that could contribute to bearing overheating. These measurements must be registered to protect the system, as the observer will sound an alarm if the recorded data reaches a pre-determined threshold. The critical limit is normally set per the pump's specifications and manufacturing standards. When the bearing temperature reaches 120°, for example, a warning is activated, and the pump must be stopped immediately. When unfavorable events occur, the programme sends a warning message to the maintenance management department, which investigates the causes of the critical machine state. Due to their simplicity, they can only reliably detect imminent overheating and bearing failure. This does not allow for sufficient preparation and resource optimization for PdM. This illustrates the importance of the current analysis, which integrates the benefits of SKF@ptitude and the predictive ability of ML to predict the occurrence of abnormal conditions in advance.

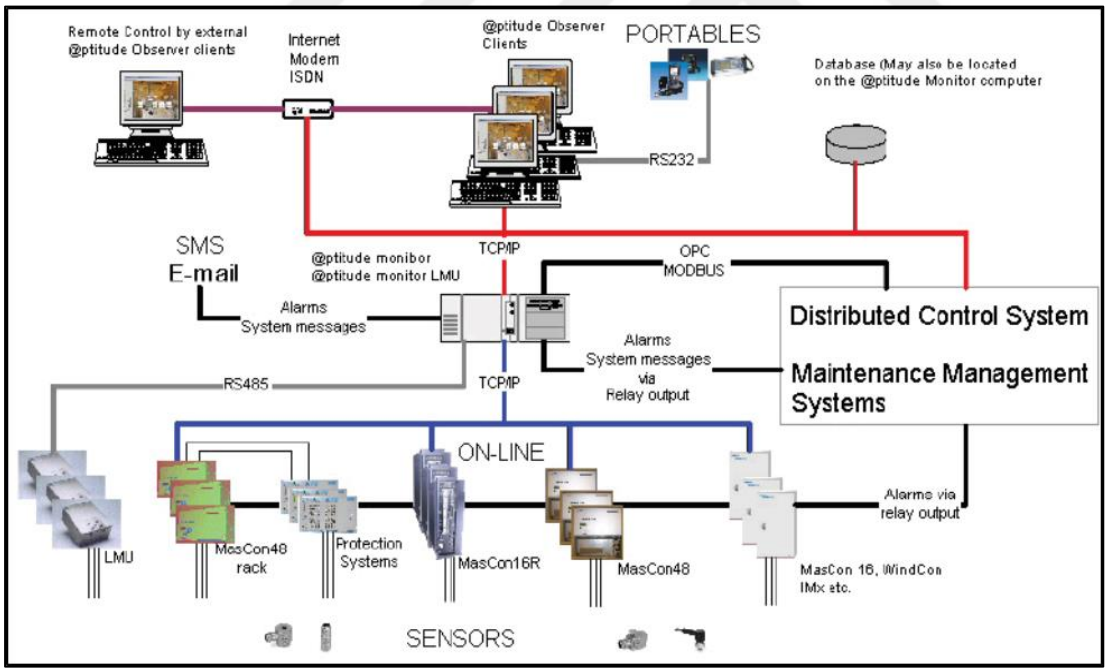


Figure 2.2 : SKF@ptitude observer monitoring system [44].

Using the SKF@ptitude observer monitoring system, the pump system is continuously monitored and measurements from specific sensors mounted along with the pumping system are processed. User-friendly displays reveal measurement data as well as other related information. The bulk of device data is extracted from specific data that is accessed either online or offline. In general, the two other kinds of data captured during

data collecting are failure data and process data. The failure data consists of a system component's direct-address failure modes, such as vibration signals and lubrication oil ingredients. Process data related to functional elements such as pressure, flow, and temperature, on the other hand, can predict failure mode.

2.4 Data Description

Initially, a sample of real data obtained for model creation is called row-data, which must be arranged and prepared. Reserved-dataset, on the other hand, is not specifically appropriate for developing a predicting model since it includes a lot of noise and missing function values. As a consequence, data preprocessing is the first step in data cleaning to prevent errors caused by data anomalies and preparing raw data for further analysis before feeding it into prediction modeling.

It is worth noting that the data gathered is split into two categories: condition monitoring data and event data. The condition monitoring data is used to assess the current state of the physical asset's health. Event data provides historical details such as installation, breakdown, overhaul, and so on. The error in event data may come from various sources, including human factors, while unreliable or inaccurate sensors cause condition monitoring data errors. Both event data and condition tracking data are equally relevant in the PdM application. Nonetheless, in practice, more focus is generally put on collecting condition monitoring data, with event data collection being wholly ignored in some cases [26].

Data analysis (data pre-processing) is the next step in the process, which varies depending on the type of data being analyzed. Researchers have developed various signal processing and sophisticated algorithms to help in the understanding and analysis of various datasets. Additionally, different feature extraction techniques are used to quantify multiple vibration characteristics such as amplitude, frequency, displacement, velocity, acceleration, phase, and period is common in the literature.

In general, two approaches can be used to make predictions: model-based and data-driven methods. These are competing approaches, but they can be combined in practice, resulting in hybrid prognostics [45]. Due to the rapidly growing condition monitoring data, data-driven fault detection has become the most popular approach and a hot research subject in the PdM framework. It has piqued the interest of both

academic and industry. To explain the operating state of industrial equipment, model-based prediction methods depend on analytical models. Real-world ageing mechanisms are typically non-linear, randomized, and complex, making it difficult to achieve precise results using an empirical model. Data-driven methods, on the other hand, can be extended to the situation that is reluctant to construct an analytical model. The data-driven method aims to convert the machine's operational data into degradation information, exposing the system's functional status as well as the degradation mechanism model [46].

As an example, the real-world experimental dataset used in this research is the pumping system collected from several sensors that monitor the bearings' operation conditions in the sewerage treatment company. The pump station is monitored through its operating state, measured based on vibration and temperature monitoring parameters. The retrieved information belongs to a three-month operation of the pumping system, including 130,956 data recorded by specific sensors for every one minute. Four accelerometers are installed vertically and horizontally to pick up the vibration (acceleration) signals generated at the Driving End (DE) and Non-Driving End (NDE) bearings in terms of vibration. Similarly, there are two types of temperature sensors: Resistance Temperature Detectors and Infrared Temperature Detectors (RTD). These sensors are mounted in the same directional mode as the DE & NDE Bearing temperature sensors.

Vibration condition monitoring is the most commonly used PdM technology in rotating equipment. It is used to detect a bearing fault and a deterioration process that has progressed to a certain damage point. In particular, increased vibration is often a sign of irregular conditions. Vibration measurements are typically the first indication of emerging failures on rotating machinery. This general method, known as the P-F curve, is depicted in (Figure 2.3). It depicts how a failure starts (potential failure) and progresses to the point where it happens (actual failure) (functional failure). The points in between illustrate how a condition deteriorates to the point that it can be detected (Point P) and then deteriorates further to reach (Point F) if it is not detected and corrected. A functional failure occurs when a system cannot perform a particular function to the desired quality level.

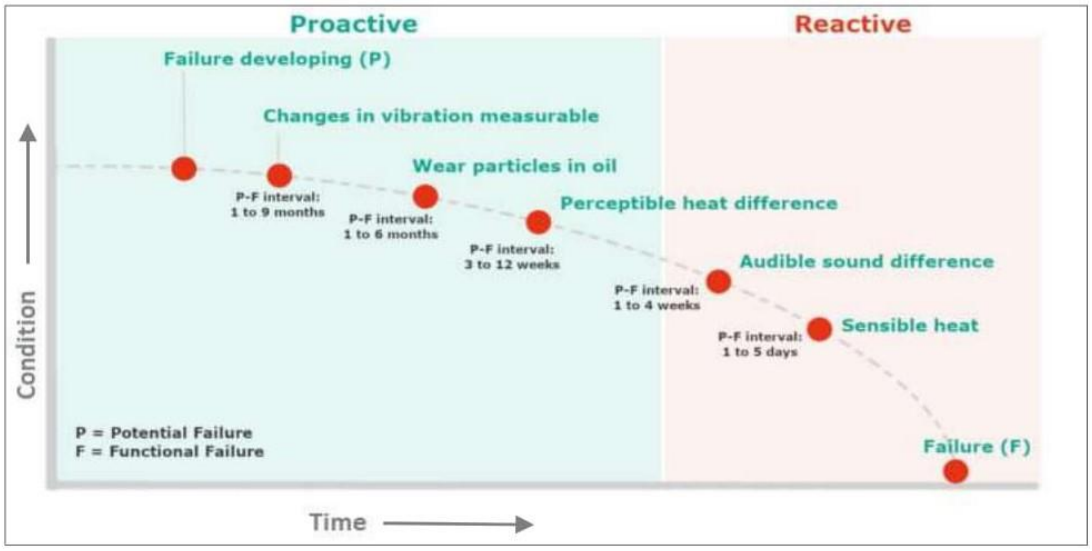


Figure 2.3 : P-F curve represents an asset's behavior (pump, motor), or asset element bearing until functional failure has occurred [47].

Another commonly used technique in PdM for rotating equipment is temperature measurement. It assists in the detection of possible equipment failures caused by a temperature change. Excessive mechanical friction can be demonstrated by temperature changes (e.g., faulty bearings, insufficient lubrication). It's worth noting that vibration monitoring and thermography have also been shown to effectively predict failures and provide appropriate alert time for impending maintenance. When this strategy is in place, equipment repair is only undertaken when it is necessary. Since this study focuses on analyzing maintenance actions in pumping systems, the following section spotlights on the PdM tasks related to the pumping component.

2.5 Predictive Maintenance for the Pumping System

This study addresses the PdM and applies the case study to one of the largest water treatment firms in Doha/ Qatar. While equipment failure can disrupt workflow in any sector, it is especially critical when it comes to the Sewerage Treatment Plant, which treats 245,000 m^3 of wastewater each day for irrigation and other non-potable uses. The sludge from the treatment facility, on the other hand, is used as a soil conditioner in nearby farm fields and as a source of renewable energy. Vertical and single-stage forwarding pumps (TORISHIMA, Korea) are used for this. Driver output: 840 kW, flow capacity: 4738 m^3/h, total head: 47 m, speed: 730 rev/min, and frequency: 50 Hz are the technical requirements for these pumps. The vibration and temperature data from one of these several pumps are used to implement ML algorithms. The

forwarding pump adopted in our research is illustrated in the following (Figure 2.4). Moreover, a photo of the pumping station and sensor types is shown in (Figure 2.5).

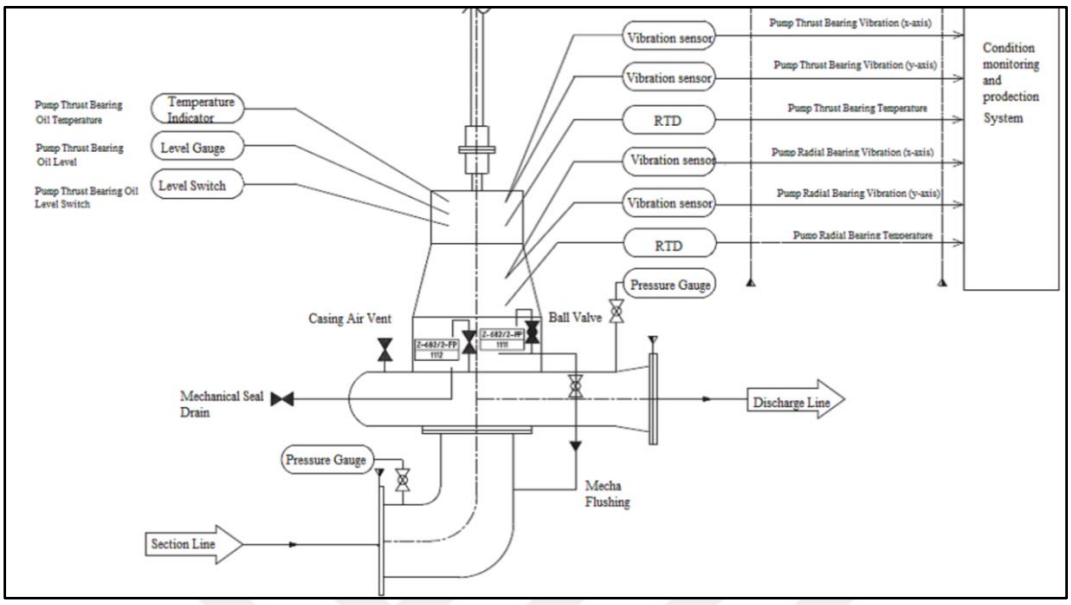


Figure 2.4 : A schematic drawing of the forwarding pump and specific location of the sensors[48].

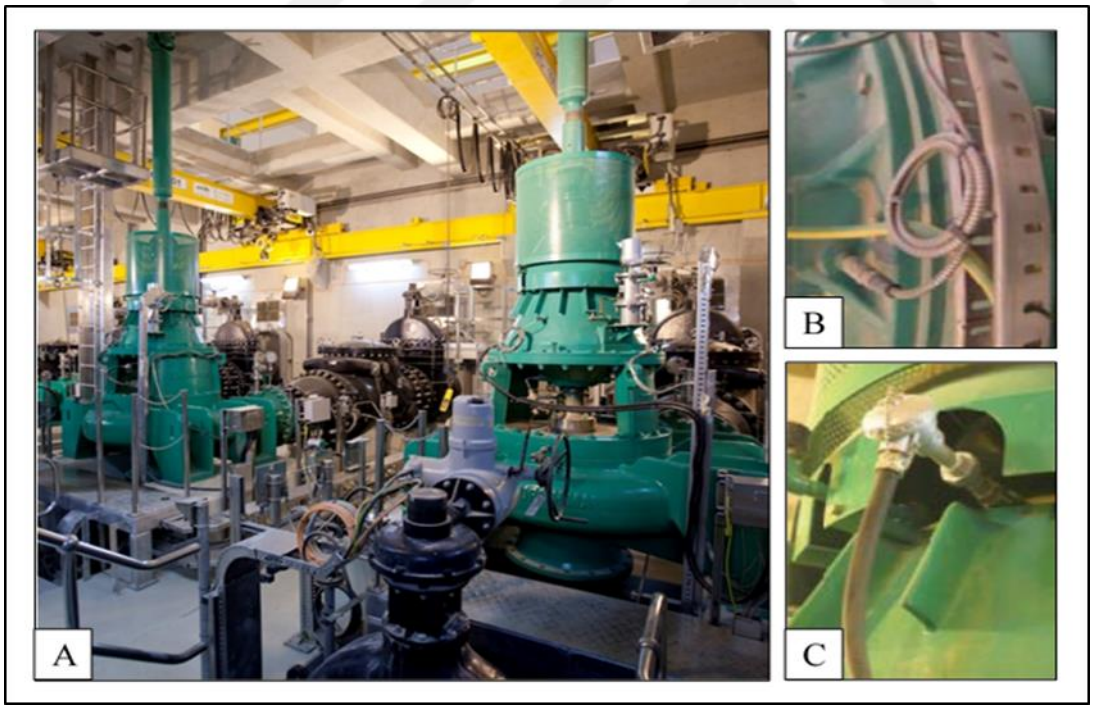


Figure 2.5 : Photographs of the (A) pumping station. (B) Vibration sensor. (C) Temperature sensor [48].

To keep the pumping system going for as long as possible and improve pump operation performance, a prediction model was developed using ML algorithms to predict a system's failure. A company's ability to mitigate failure risks, reduce maintenance

costs, and optimize asset availability is aided by accurate information on pump conditions. CBM, a predictive tool in which data assessments and data trending are used to assess future maintenance needs at an early stage, is made possible by vibration and temperature measurements.

In concrete terms, the aim is to enhance data-driven prediction methods' accuracy and precision in complex pumping system maintenance problems. Two case studies are presented to discuss this viewpoint, and they are used to examine the various AI methodologies suggested in this thesis. In the following chapters, we will go over how to apply the prediction model and how it performs.

3. CONDITION MONITORING AND FAULT DETECTION OF ROLLER ELEMENT BEARING

The classifications of rolling bearings and their basic operating mechanisms are explained in this chapter. The radial and thrust ball bearings are discussed in greater depth, and the standard failure modes in roller bearings. Furthermore, it provides a description of the specific rolling element bearing condition monitoring and fault detecting strategies, such as vibration analysis and temperature monitoring that were used in the current research, as well as the advantages and disadvantages of each process. Finally, this chapter discusses time domain, frequency domain, and time-frequency domain vibration signal processing methods that are widely used to track roller bearings. Various measuring instruments are also depicted.

3.1 Introduction

Most rotary machines require rolling element bearings, which serve as the interface between the machine's stationary and rotating parts. The bearing's working condition has a direct impact on the machine's function and operation. As a result of the importance of bearings, detecting their running condition is extremely important. Many monitoring techniques and fault diagnosis procedures have been developed to minimize maintenance costs, increase productivity, and prevent catastrophic component failure during operation, resulting in machine downtime.

Bearings come in a range of shapes and sizes. Fundamentally, bearings are classified into: (1) journal bearings, also known as sliding or plain surface bearings, and (2) rolling part bearings are often referred to as ball bearings. Sliding bearings are those that produce only sliding friction. The shaft is usually protected by the sliding board, with a tinny layer of lubricant in between to make the shaft slip smoothly. Journal bearings are small and compact, with long service life and no vibration or noise [48]. Moreover, unlike a rolling bearing, a sliding bearing is a bearing that does not include any rotating element. See (Figure 3.1).

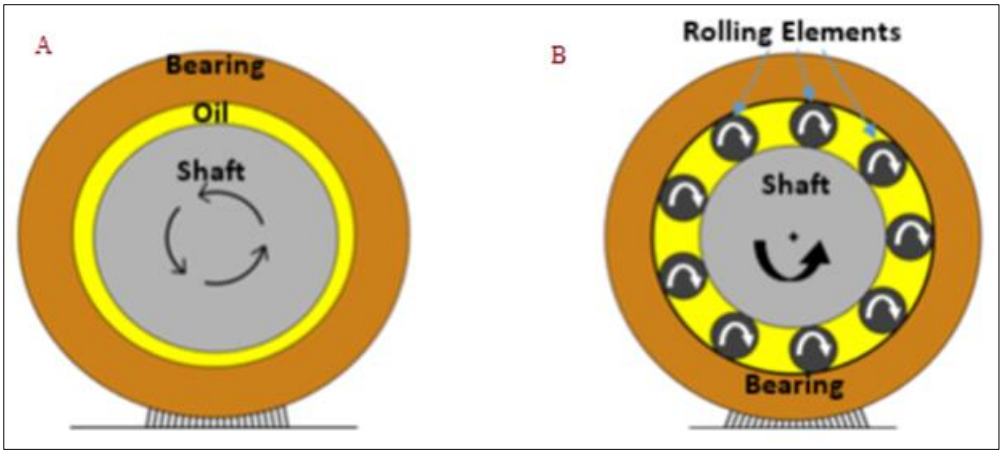


Figure 3.1 : Bearing type (A) Sliding Bearing. (B) Roller Bearing [48].

Rolling bearing is considered an irreplaceable component in automation systems. Failure of these components may cause catastrophic breakdown and costly downtime for the machine's maintenance action. Hence, an appropriate fault detection system must be established to prevent bearing damage and breakages during service, leading to catastrophic failure. A rolling-element bearing involves inner and outer races, a cage, and a roller, unlike sliding bearings which have just a sliding motion. The rolling element is described in (Figure 3.2). The roller between the inner and outer tracks and all wear will occur on the bearing's internal parts under normal running conditions.

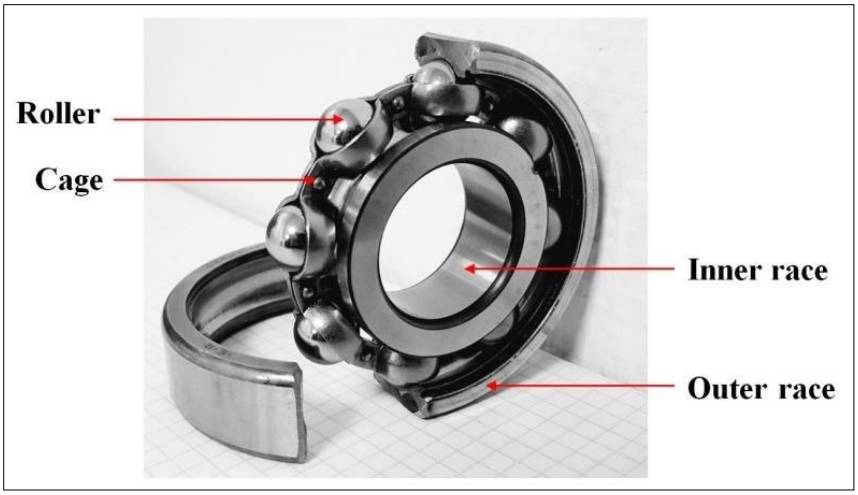


Figure 3.2 : Element of the roller bearing [49].

Rolling bearings are generally classified by the type of the rolling element, i.e., balls, cylindrical rollers, spherical rollers, tapered rollers, and needle rollers [49]. (Figure 3.3) illustrates elements of common types of rolling bearings.

Figure 3.3 : Types of rolling bearings [50].

While all roller bearings reduce rotational friction to enable a component to move at excessively high speeds and carry large loads efficiently and effectively, some high precision bearings are specifically designed to bear support bearing loads in different directions, such as the deep-groove single-row type shown in (Figure 3.4). These bearings are inexpensive and can handle radial and axial (thrust) loads. In most new machine designs, they are normally given top priority. This thesis focuses on a boll bearing named a thrust bearing and radial bearing which will be used as a case study in our research.

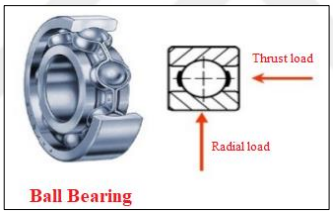


Figure 3.4 : Radial and Thrust loads in the deep-groove ball bearing [51].

3.2 Radial &Thrust Ball Bearings

These bearings minimize friction between the shaft and bearing surfaces and accommodate radial and axial loads in rotating machinery. A thrust bearing is used for axial loading paths, which is ideal for heavy applications and high-speed machines that require precise control, low noise, and long tool life [51]. They are divided into two types: single-direction thrust bearings and bidirectional thrust bearings (double direction). A radial load cannot be supported by either unidirectional or bidirectional thrust ball bearings. A shaft washer, a housing washer, and a ball and cage assembly set up a single direction thrust ball bearing. Since they can only bear axial load in one direction, they are typically used in matched sets. One shaft washer, two housing washers, and two ball and cage assemblies form a double-direction thrust ball bearing. They can withstand the axial loads in both directions. Tool spindles, rock grinding,

rolling mill stands, trucks, gearboxes, pumps, and farm machinery all use thrust ball bearings.

Radial bearings are rolling bearings that are mostly used for radial load-bearing. In other words, the load acts perpendicular to the moving element's direction of motion, while the load acts along the thrust bearings' axis of rotation, as shown in (Figure 3.5). These bearings are most commonly used in high-speed spinning machines like compressors, engines, pumps, and gas turbines.

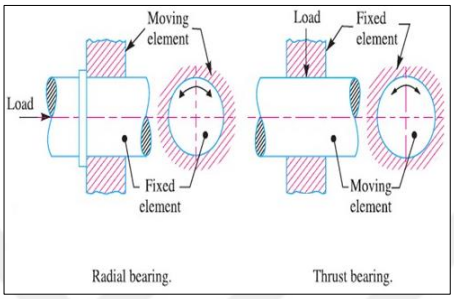


Figure 3.5 : Radial and Thrust bearings load direction according to the moving element [52].

Some ball bearings can support radial and axial loads on the shaft; Axial angular contact is used to achieve these axial/radial load bearings. The axial and radial loads are distributed more uniformly around the axial angular contact ball bearing because of the axial radial bearings' angle. The load line in angular contact ball bearings, for example, forms an angle with the bearing axis at the contacts between balls and raceways, as illustrated by (Figure 3.6). The inner and outer rings act as stabilizers for each other, and the bearings are primarily designed to handle radial and axial loads [53]. Heavy-duty machinery, electric motors, aircraft gas turbines, rotary tables, dental drills, and pump processing all use angular contact bearings.

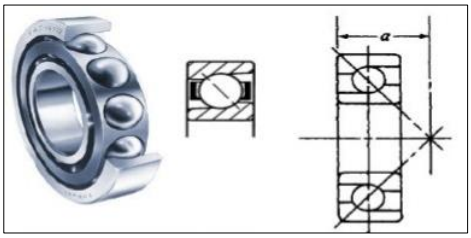


Figure 3.6 : Angular ball bearing [53].

Low-friction bearings are often crucial for efficiency, to reduce wear, and to simplify high speeds. A bearing can reduce frictional resistance by the effect of its design, material, or by introducing and containing a layer of fluid (known as a lubricant)

between its surfaces. The lubricant used to separate the contact surfaces, either rolling elements or any guiding surfaces on outer or inner bearing rings, is usually a mineral oil refined from petroleum, synthetic oil, vegetable oil, silicon oils, greases, etc. Inadequate lubrication often leads to high friction, wear, high temperature, sliding and fretting at rolling element contacts, and early cage breaking with a jammed bearing.

It's essential to focus on a machine's bearings when running to avoid redundant bearing failure. Since each ball is subjected to cyclic loading as it makes repetitive trips through the highly loaded position, the rolling element bearings wear out over time. Fatigue wear or single point defects such as chips or dents are caused by cyclic loading. Vibration levels in rolling element bearings steadily increase as wear progresses until they fail. Bearings can be replaced well before they fail to maintain optimal performance. Wear failure of ball bearing elements (i.e., outer race, inner race, and rolling parts) causes most defects, which can be sensed mainly by mechanical noise, temperature, and vibration [30]. As a such, the most commonly encountered bearing issues can be summarized as follows [54]:

(1) Lubricant is either insufficient or unnecessary,

(2) The bearings were installed incorrectly,

(3) Bearing clearance is restricted, or a heavy load is applied,

(4) There is a lot of tension between the lip and the seal groove,

(5) Using the wrong lubricant, and

(6) Between the fitting surfaces, there is a creep.

However, bearing failure can be expensive due to various reasons, including lost productivity, maintenance costs, and serious harm to other rotating machinery components [55]. As a result, continuous condition monitoring has emerged as a promising way of avoiding catastrophic component breakdown in machines by detecting, calculating, and recording physical variables obtained from sensor-mediated components [56]. As a result, the data is functionalized for specific operational conditions. In more detail, built models show alerting signals when operational conditions hit a critical threshold, allowing PdM to be applied [24]. This explanation is demonstrated in (Figure3.7), which illustrates a typical trend curve and alerts an

incipient failure. It also gives a lead time in which to schedule and implement maintenance.

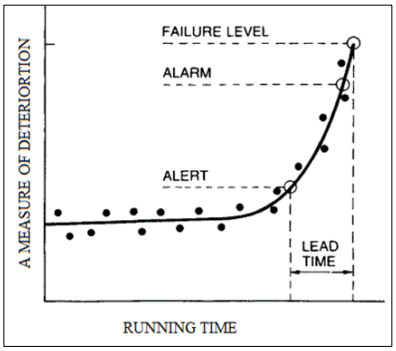


Figure 3.7 : The theory of condition monitoring indicators, which shows the degradation of the bearing [57].

Various tools, such as monitoring and diagnostic methods, can be used as part of a robust programme for PdM. Vibration analysis is one of these strategies [58], acoustic emission[59], thermographic inspection [60], oil analysis, radiographic inspection [61], shock pulse [62], ultrasonic leak detectors [31], performance testing, wear and dimensional measurements [63], signature analysis [64], time and frequency domain [65, 66]. Vibration analysis has gotten a lot of coverage as a time domain and time-frequency domain method for measuring machinery working conditions, which essentially diagnoses faults and improves the machinery's life.

3.3 Condition Monitoring Techniques in Bearings

Prediction of bearing faults has received considerable attention and is remains one of the state-of-the-art topics in recent years. Scientists have focused on developing various techniques and methodologies to improve bearing fault detection, typically modeling the bearing signal using sensor data. Local failures in roller bearings can be identified in the time domain using statistical features (RMS, mean, variance, skewness, kurtosis, crest factor, impulse factor, shape factor, spectrum, median, and range) [67].

The following (Table 3.1) shows the various condition monitoring methods used on bearing defects during the last few years. A depth study established four different methods for detecting and diagnosing bearing defects, categorized as vibration monitoring, acoustic measurements, temperature monitoring, and wear debris analysis. Vibration measurements are the most commonly used of these.

Table 3.1 : Monitoring methods which are suitable for the bearing components.

Technique	Principle	Application examples
Vibration analysis [68]	Vibration analysis is a method of detecting irregular vibration events caused by changes in dynamic forces and evaluating the overall values and individual frequencies associated with machinery anomalies by monitoring the levels and patterns of vibration signals within moving elements.	• Bearings, gears, shafts, freewheel. • Gearboxes, engines, fans, drive-trains, high-speed rotors in turbines, and pumps are examples of rotating machines.
Temperature monitoring [69]	Temperature allows for the gathering of information on bearing state and working conditions to comprehend all bearing anomalies. Consequently, a temperature increase above the normal value serves as a useful indicator of equipment failure. The greater the bearing system's safety margin if the running temperature value is below the appropriate limit [57]. Several critical parameters are affected by bearing temperature, including lubricant viscosity, load-carrying capability, power loss, and load distribution. The temperature of a bearing determines whether it is working with low friction; as an effect, as the temperature rises, lubrication deteriorates, and internal clearance decreases.	The temperature control technique is used to forecast how transient temperatures in bearing elements will evolve. In turn, the control system keeps track of the temperature and ensures that nothing goes wrong. A bearing's temperature shows whether or not it is working with low friction.
Wear debris analysis [57]	The accumulation and analysis of wear debris forms as metallic particles on component surfaces and is carried away in the lubricating oil. The amount of material, shape, and size of the debris particles may indicate the source and failure mechanism, allowing for early detection and control of the wear process.	Wear debris analysis is used to diagnose the health of rotary machines. Online debris tracking is currently being used to improve system reliability and lower maintenance costs in commercial engines, fighter engines, helicopter gearboxes, and wind turbines.
Acoustic measurements [70][71]	In structural diagnosis, the acoustic emission (AE) approach is a widely used and standardized method. The stress waves caused by a sudden redistribution of internal stresses within a material are known as AE. When a material is exposed to friction, wear, corrosion, phase transition, cavitation, crack, fracture, and other stresses, the lattice structure changes.	The AE measurement is used to detect defects in rolling element bearings early. Furthermore, the AE technique can detect changes in machinery elements at the lattice level that would otherwise go undetected by a low-frequency traditional vibration technique. The frequency range is usually limited to 20 kHz [72].
Oil analysis [73]	Oil particle counting and moisture measurements, as well as viscosity, acidity, and temperature, can all be used in this analysis. Many types of machine failures can be caused by lubricant degradation. In operation, a lubricant is exposed to a variety of conditions that can cause its base oil and additive system to fail. Heat, entrained air, incompatible pollutants, internal or external emissions, the presence of soil or water, and inadvertent mixing of a different fluid are examples of such causes.	Oil analysis is most often used to assess the mechanical wear state of insulating machines such as a gearbox, transformer, and other electrical distribution equipment by testing the oil condition for bearing lubrication.

Condition monitoring and predictive maintenance technique for the rolling element bearings is considered an essential activity. When the component fails, they cause the whole machinery system's failure, possibly with some consequential damage and financial losses. Among the techniques mentioned in the previous table with the adoption of the published research in this area, the operating temperature and vibration signal are considered important aspects of bearing condition monitoring. They directly affect the performance and service life of the machine. Thus when these readings are higher than a pre-set critical level, the machine monitored is announced faulty, and a maintenance intervention is triggered. (Figure 3.8) depicts the distribution of pump failure using pie chart to show how vital roller bearings are for keeping pumps up and running. When a rotating machine fails, it's often due to a shaft seal failure or bearings. Roller bearings failures cause about 15% of all unscheduled critical pump shutdowns.

According to a report prepared by Swagelok Company under the title of (Rely on Swagelok Seal Support Systems to Reduce Pump Failures) Luke Wurban in Jun 2020, displayed a pie chart of the pump failures distribution.

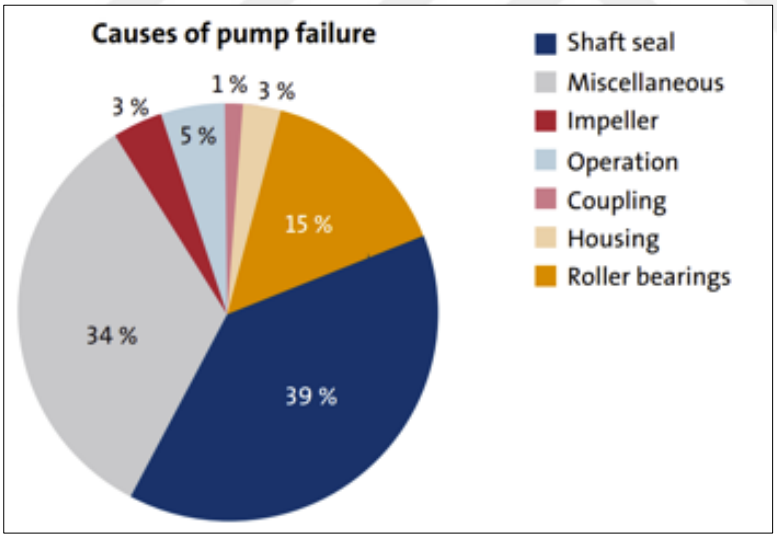


Figure 3.8 : Analysis of pump failure [74].

As a response, the most serious pump failures are the primary focus of further analysis and review. Pump seal failures account for the most failures, followed by downtime due to overhaul, bearing failures, impeller, shaft, and coupling malfunctions, and so on. Seals and bearings are non-repairable components that must be replaced when they alfunction. Seals, bearings, and overhaul are considered to be leading causes of pump failure [75].

3.3.1 Vibration analysis

Vibration analysis is a useful tool for measuring the state of roller bearings in machine components. By calculating the entire vibration range and conducting frequency analysis, vibration-based signal analysis is used to detect bearing faults. The ability to predict when a component will fail based on data analyzing increases uptime, production, and maintenance efficiency. The fit and runout (radial and axial) of the bearing and mounting surfaces, as well as the elastic behaviour of bearing elements and related system parts, all of which are affected by rotor unbalance [51].

Local and distributed defects are the two types of rolling bearing defects. Pits, spalls, flakes, and cracks on rolling components caused by fatigue on the roller surface are examples of local defects. Surface roughness, surface distress, waviness, smearing, waviness, misaligned races, and uneven diameter of rolling components are all examples of distributed defects. Operating conditions, installation position, improper bearing design, improper manufacturing, improper lubrication, overloading, uneven wear, and other factors may cause these defects. Bearing faults, whether local or distributed, will normally produce consecutive and periodic impulse terms in machine vibration as the ball bearing passes through the defect points. Knowing the rotating rate, location of faults, operating area, types of measurements, and bearing dimensions can be used to determine these terms [76]. Vibration analysis makes use of vibration signals produced while a machine is in operation. These signals are collected using different types of sensors at an accessible location on a computer, and then signal analysis is performed to predict the state of the bearings inside the machine at an early stage.

The vibration analysis technique can detect, diagnose, and predict rolling element bearing faults. Depending on which transducers are used, vibration measurements are conveyed in displacement, envelope signal, velocity, or acceleration. A transducer is a system that converts physical variables into electrical quantities (such as current or voltage) [17]. A signal transformation instrument is used to amplify the signal to the appropriate value because the output signal (voltage or current) of the transducer is too small to be captured directly. The basic features of a vibration measurement pattern are shown in Figure . Due to a peculiar vibration frequencies that are excited in most measurements, it is possible to distinguish the defective bearing elements. The same cannot be said, however, for the defects that occur in bearing cages [49].

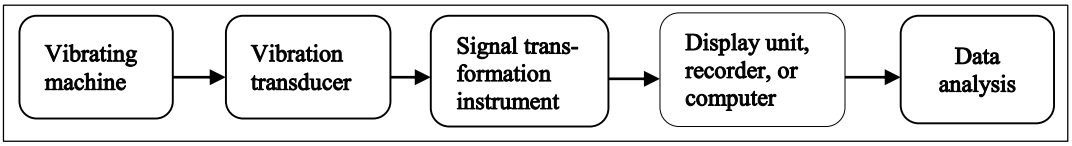


Figure 3.9 : Basic principle of vibration measurement.

The data from a vibration sensor's raw measurement is shown as a function of time and amplitude. The amplitude is a measurement of a bearing's vibration intensity. As a response, a certain amplitude value is used as a safety limit to denote the risk if the vibration exceeds this limit [77]. To put it another way, each measure has its own set of warning limits. When the measurement's value exceeds the set alarm limits, the predictive maintenance software or data collection sends an alert to the analyst. The vibration pattern of a broken bearing is made up of low-frequency collision components as well as high-frequency collision components. The structural details of the bearing part or machine are preserved [69].

Most fault identification methodologies in rotary equipment, like pumps, motors, and gearboxes, use vibration data [78]. The mathematical time domain, frequency domain, and time-frequency domain features are the most frequently used research techniques. In this regard, vibration signals are firstly gathered and processed using vibration analyzers equipped with sensors in the time domain. These signals are then converted into the frequency domain using Fast Fourier Transform (FFT) to extract the frequency signature. The information obtained from a vibration signals has significant advantages in predicting catastrophic failures [79].

As previously mentioned, vibration data from the DE and NDE thrust ball bearings of a pumping system were collected using an accelerometer and velocity sensor in our study. Sensor vibration signals were measured and analyzed. The statistical parameters are then measured and compared to detect bearing faults in the roller component, outer and inner races. The most commonly used vibration analysis (waveform data analysis) and fault prediction techniques for rolling element bearings are briefly discussed and outlined in the following sections.

3.3.1.1 Time domain analysis

The time domain analysis evaluates the performance product of vibration analysis from finite component software by looking at the signal's time history. The root mean square (RMS) is a condition monitoring parameter that measures the overall level of a

discrete signal. The time signal descriptors are peak to peak amplitudes, which are determined from the top of the positive peak to the bottom of the negative peak. Some statistical parameters, such as the crest factor, kurtosis factors, spectrum, kurtosis, and so on, are critical in determining the bearing's wellbeing [80,81].

3.3.1.2 Frequency domain analysis

The most common method for detecting bearing defects is frequency domain or spectral analysis. Using a FFT, frequency-domain techniques translate time-domain vibration signals into discrete frequency components. Simply put, FFT transforms time domain vibration signals into a sequence of discrete frequency components using mathematics [79].

The FFT spectrum can be used to distinguish various frequency signals or components. The frequency is plotted on the X-axis, while the signal components' amplitude, velocity, or acceleration is plotted on the Y-axis. The key benefit of frequency domain analysis over time domain analysis is that specific frequency components of interest may be easily detected. Furthermore, the FFT spectrum is an extremely useful method for analyzing machinery vibration and diagnosing most bearing problems by providing data that can be used to pinpoint the source and cause of the problem.

3.3.1.3 Time-Frequency domain analysis

Both stationary and non-stationary vibration signals can be handled using time frequency domain techniques. This is the key benefit of time domain techniques over frequency domain techniques. The signal frequency components can be seen using time-frequency analysis, as well as their time-variant features. The Short-Time Fourier Transform (STFT) [82], Wigner-Ville Distribution (WVD), and Wavelet Transform (WT) [83] are some of the time-frequency analysis methods that have been adopted.

3.3.2 Temperature monitoring

The operating temperature of a bearing element system is critical to its overall performance. The bearing temperature affects several critical parameters, including lubricant viscosity, load-carrying capability, load distribution, material thermal expansion, and power loss. As a result, temperature data is regarded as one of the most important calculated parameters for monitoring the bearing situation and operating status. Temperatures fluctuate rapidly during system startup and shutdown, so

temperature monitoring is advantageous for obtaining information on machine performance [84]. Most bearings in the industry are designed to work well at specific temperatures, but if the measured temperature exceeds this limit, the alarm message emerges.

Temperature rises in bearing structures play a key role in rotary machine degradation (pumps, motors, gearboxes, and so on), resulting in catastrophic failures of the machinery. As a result, it's important to keep an eye on rotor components to make sure they stay within certain temperature ranges. As a result, successful thermal control prevents rotary machines from overheating and improves the overall drive process performance. To build a fault prediction model, we selected a sample of temperature data obtained from embedded temperature sensors mounted along with the vertical and horizontal directions of the bearings in this study.

3.4 Measurement Instrumentations

3.4.1 Accelerometers

High-quality, rugged, and cost-effective accelerometers are used with the SKF Local Monitoring Unit (LMU) on-line systems, protection systems, and the ever-versatile portable data collection instruments for vibration measurements shown in the (Figure 3.9). The thrust ball bearing has an optional banded frequencies measure from (3.0 Hz to 5,000 Hz or 1.0 Hz to 9,000 Hz) accelerometer (model CMSS 2100) set up on the bearing housing case. A data acquisition system (SKF@ptitude observer monitoring), was used to record the signal sample in the computer software system for further analysis. To maximize use in a variety of applications, the nominal resonance frequency is 30 kHz, with a high sensitivity of the 100 mV/g and a sensitivity precision of 5% at +25 °C. These accelerometers, in particular, are designed at extremely low noise levels for low frequencies at high temperatures, are corrosion resistant, and are hermetically sealed for use in high humidity environments.

Via a threaded bronze stud base that is glued to the bearing house, accelerometers were installed vertically and horizontally to the NE and NDE bearing housings. In centrifugal pumps case, the accelerometers are mounted horizontally on the pump bearing casing, perpendicular to the shaft movement, as near to the load regions as possible to provide high-quality signals. Similarly, the vertical accelerometer should

be mounted near the bearing housing, but two accelerometers should be mounted 90° apart. Due to the effect of structural resonances, an axial measurement in the vertical direction close to the pump casing can also produce good signals to the lubrication mechanism.

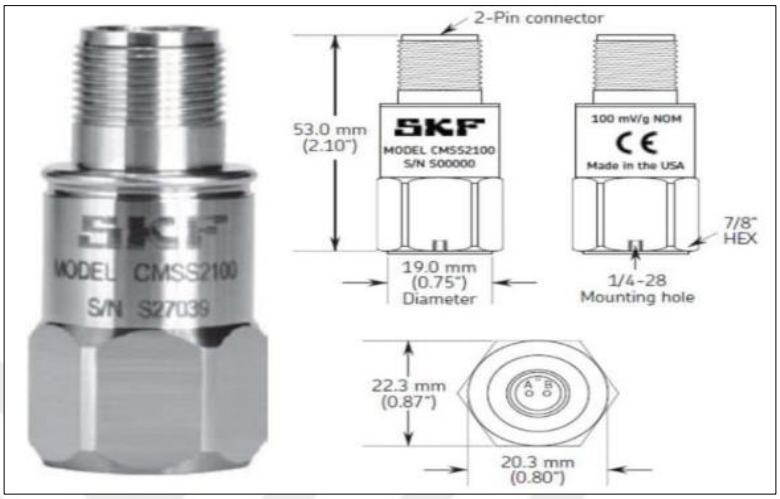


Figure 3.10 : SKF vibration sensor (CMSS 2100) [44].

3.4.2 Temperature sensors

Resistance Temperature Detectors (RTD) is among the most precise temperature sensors on the market. In our case study, this sensor is used to calculate temperature data applied by the detection bearings fault model. Two RTD-PT100 temperature sensors are mounted in the vertical and horizontal directions of the pump bearing housing to collect temperature data with a measuring range of -50 to +180°C for DE and NDE Bearings. The resistance of these sensors varies with temperature, which can then be correlated to provide a temperature reading. Furthermore, the most common type of sensor (RTD Pt100) has a resistance of 100 ohms at 0°C. Instead of an alert on the measurement point, a device alarm will be triggered if the calculated value is beyond the measurement range. For instance, suppose the range is set to 0-300 °C, and the temperature sensor output is above 300 °C. In that case, this value will be treated as an unrealistic value, and the MasCon/IMx[1] system will generate a system alarm in the system alarm list. The cause of this alarm could be a sinful earth connection or surrounding interference that disturbs the output signal from the sensor.

[1]An on-line condition monitoring system like MasCon/IMx and @ptitude observer can be successfully operated only on an installed and tested network infrastructure. Even though the MasCon/IMx devices and @ptitude observer monitor are equipped with several fault-tolerant routines and procedures, they can ultimately only be as reliable and effective as the network to which they are connected.

4. MACHINE LEARNING IN PREDICTIVE MAINTENANCE

This chapter presents an introduction to machine learning (ML) in general and gives a brief explanation of the ML types. It then focuses explicitly on the ML approaches in PdM and provides the recent related work in this field. Section (4.3) describes intelligent predictive maintenance architecture illustrated with a flow chart. Moreover, it introduces feature selection and extraction methodology focus on the application methods used in the prediction model.

Also, this chapter covers the research methodology followed to construct a data-driven approach based on an ML techniques. Moreover, it presents an overview of the main steps in developing the fault prediction model; various ML algorithms are established using sample bearings data from the pumping system. Concretely, four ML algorithms are tested, which include: Decision Trees, Random Forest, and Gradient Boosted trees, as well as Support Vector Machine. Finally, it introduces the evaluation measure of the tested ML approaches is based on five performance indicators: accuracy, precision, F-score, recall, and an area under curve (AUC).

4.1 Introduction

At its most basic level, machine learning (ML) is the process of using algorithms to evaluate data, learn from it, and then determine or predict. Reverse the process of manually coding programme routines with a given set of instructions to complete a task. Huge quantities of data and algorithms are used to "train" the computer, allowing it to learn the association between input and desired output, as well as how to execute the task [40].

Innovations in specific domains such as mathematics and computer science, such as mathematical learning and the availability of easy-to-use, often freely accessible (software) tools, have the ability to turn the industrial domain and sustainably grasp increased manufacturing data depots. In the field of ML, one of the most exciting advances is taking place. However, several different and diverse algorithms, theories, and methods are adopted for applying ML structures [85].

Unlike conventional software programming, these algorithms are not built from a set of predefined rules. These algorithms, on the other hand, are self-learning. They achieve at rules by running a series of tests on training data and building an application field model. Each new set of data is then used to fine-tune the model and enhance its predictive power [86].

4.2 Machine Learning Types

The ML model is a mathematical model that finds trends in data to produce predictions. Many big data issues, speech recognition, vision, and robotics can all benefit from machine learning [87,88]. As briefly described in the following sections, there are three distinct forms of learning: supervised, unsupervised, and reinforcement.

4.2.1 Supervised learning

Supervised Learning is a ML pattern for learning a system's input-output relationship information from a collection of paired input-output training data. An input-output training set is also known as labelled training data or supervised data because the output is treated as a label of the input data [89].

"Regression" and "classification" problems are two types of supervised learning problems. A regression problem aims to predict outcomes from a continuous output or to map input variables to a continuous function. Instead, a classification problem aims to predict outcomes to a discrete output or to map input variables into discrete categories.

To evaluate the predictor characteristics, supervised ML aims to construct a predictive model of class label classification. The resulting classifier is then used to assign the testing dataset as class labels, which has known predictor feature values but unknown class label values [10]. As far as this work is concerned, supervised ML has chosen to analyze the collected data as a prediction model.

4.2.2 Unsupervised learning

When a model or system is not given clear input, it is referred to as unsupervised learning. On its own, the machine must learn patterns or structures in the data [39].

There are no branded examples of unsupervised learning processes, and the output is unknown during the learning process. In other words, no external input is needed during the training phase [4].

4.2.3 Reinforcement learning

Reinforcement learners engage with their surroundings and use what they have learned to select or prevent specific actions depending on the effects. If the same problem arises, situation tends to replicate actions that resulted in high rewards in the past. Choices that result in smaller rewards, on the other hand, are more likely to be avoided.

Semi-supervised learning is another form of learning that is commonly discussed. It's a mix of supervised and unsupervised learning with some labelled data. Getting labelled data for supervised learning can be expensive, but getting large quantities of unlabeled data is not. Semi-supervised learning takes advantage of both at the same time and is useful when only a limited of labelled data is obtainable. Furthermore, where only a portion of the historical data is available, the semi-supervised learning approach will assess the situation and detect failure, which is advantageous [90].

4.3 Intelligent Predictive Maintenance Architecture

Figure 4.1 depicts the I-PdM platform's architecture. This platform includes modules that transform rotary equipment data into helpful information for maintainers so they can take corrective steps, inspect the states, and repair the defective component to avoid equipment breakdowns. Many appropriate algorithms could be correctly chosen at each stage to achieve the best result and performance [90]. Artificial neural network (ANN) is the most usual method for intelligent diagnostics and prognostics of machines. Inputs to an ANN can be data gathered from vibration sensors, pressure, flow, power, and other performance features. ANNs are simulators inspired by a human nervous system that can handle various tasks like pattern recognition and classification [91]

Operating observers can manually input working conditions or collect data from sensors installed on the devices for condition monitoring. Following that, suitable data processing algorithms, feature extraction, and feature selection is used to transform these data into features. A proper algorithm is employed in the feature space to identify the types of faults, predict degradation, and forecast the remaining lifetime of

machinery [55]. The deterioration data can be useful in making maintenance decisions and optimizing maintenance schedules. We may examine and discover patterns, rules, and information from data captured from various sources using decision making. As a consequence, depend on the analysis results and real-time data, we have the ability to make the right decision at the right time and in the right part [92].

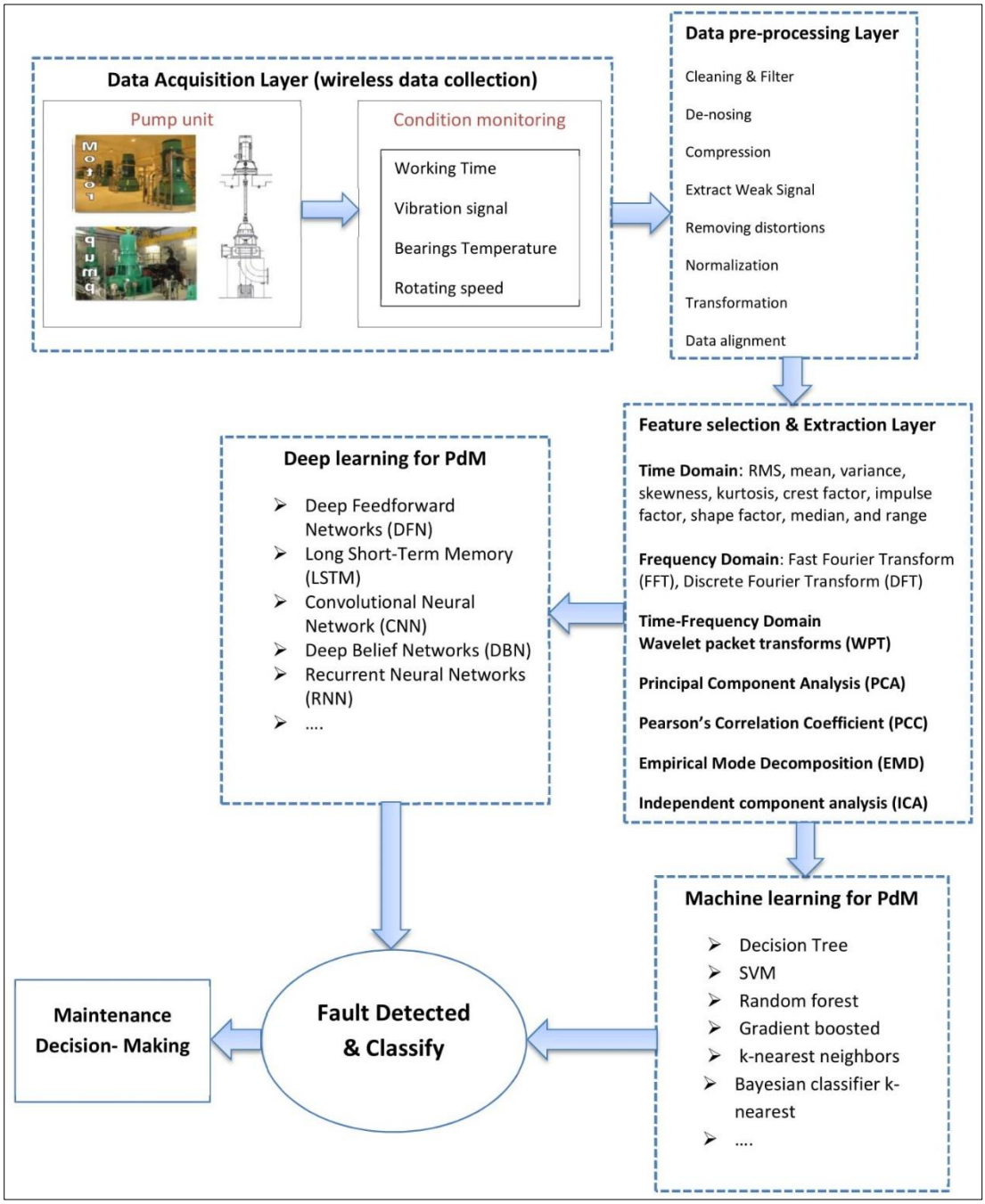


Figure 4.1 : The architecture of I-PdM platform.

Some primary concepts of intelligent predictive decision support systems have been presented in [93]. Different techniques can be applied in various PdM implementation

phases, such as data processing, diagnostics, practical real-time data monitoring, and prognostics. Therefore, evaluating the condition of components, the degradation status, and the processing system's quality-reliability chain may have a decision-making policy to determine the production cost and ways to reduce cost according to maintenance efficiency.

There are two ways to formulate the PdM dilemma. The first step is to formulate the PdM as a classification problem to identify the part that is failing before the machine fails. The second option is to frame it as a regression problem in which the goal is to estimate remains of time before the next failure. Both of these cases are examples of supervised learning techniques. Generalized linear models, tree-based ensemble methods (random forest, gradient boosted trees), and deep learning techniques would be the most appropriate PdM problems [94]. AI methods can learn from the dataset delivered to them and eventually predict or classify unknown situations. Diagnosis using condition monitoring methods can be solved as a pattern recognition problem, comprising of three stages; Feature extraction, Feature selection, and classification. We discuss these methods briefly in the next sections.

4.4 Machine Learning Approaches in PdM

The accurate prediction of a machinery failure is one of the most exciting and challenging tasks for production planning managers. As a result, ML approaches are becoming increasingly common among industrial researchers. These methods can accurately predict the possible outcomes of faults by discovering and identifying patterns and interactions between them from complex datasets.

Different ML methods are used in the thesis to predict the equipment operating condition. Classification algorithms inclusive Decision Tree, Ensemble Classification, Random Forest, Gradient Boosting, Logistic Regression, k-Nearest Neighbors, and Support Vector Machine algorithms are used to find the best classifier for the data in the comparative studies.

Machine learning algorithms are commonly used in the field of PdM. Nam et al. [95] have utilized a data-driven approach to develop a health monitoring and diagnosis framework for a fused deposition modeling process based on an ML method. For a data-driven approach, three accelerometers, an acoustic emission sensor, three

thermocouples are installed, and associated data are collected from those sensors and processed to obtain RMS values. The frame's acceleration data were most effective for diagnosing the fused deposition modeling process's health states with the non-linear support vector machine-based model among various RMS values. Xayyasith et al. [96] presented the ML application for PdM of a cooling water system using the trained model's classification learner application. Twenty two classifier types were organized in six main comparable classification techniques, involve Decision Trees (DT), Discriminant Analysis, Support Vector Machines (SVM), Logistic Regression, k-Nearest Neighbors (KNN), and Ensemble Classification. It was shown that the SVM and DT are better at predicting failures than the other methods used in the study.

Qiao et al.[97] Presented a fault diagnosis model based on an improved wavelet package transform (IWPT), a distance evaluation technique, and the SVMs ensemble. The proposed model consisted of three phases. Firstly, the feature of impact fault in vibration signals was investigated; secondly, the optimal features were selected from the statistical characteristics of raw signals with the distance evaluation technique. Finally, to distinguish the various abnormal cases, the ideal characteristics were entered into the SVMs ensemble with the AdaBoost algorithm. Yan et al. [98] applied random forest (RF) as a classifier on Aircraft Engine Fault Diagnosis (AEFD). The methodology proved to be a reliable classification tool for various machine faults. They have also made some efforts to improve RF performance specifically for the AEFD problem. Yiakopoulos et al. [99] submitted a method for diagnosing faulty rolling element bearings using K-means clustering. The method was implemented as a two-stage procedure. In the first step, the method was deciding whether a bearing fault exists or not. In the second step, the method was identifying the type of defect (e.g., inner or outer race). They showed the advantages of a presented model by the ease of programming, simplicity, and robustness.

Recently, Z. Allah Bukhsh et al. [100] employed the ML techniques to develop PdM models based on DT, RF, and gradient boosted (GB) using existing data from a railway agency. For the prediction of maintenance need, the GB model performed most optimally compared to other methods with 86% accuracy. For maintenance activity type and trigger status prediction, the RF model attains 70% and 79% accuracy on the held-out test set, respectively. Gutschi et al. [101] introduced a data-driven approach to estimate the machine breakdown probability during a specified time interval in the

future. The authors described applied data-mining, feature-extraction, and ML methods and concluded that machine failures could be reliably predicted up to 168 hours in advance. Soualhi et al. [102] utilized artificial ant clustering to predict broken rotor bars and bearing fault at diverse load levels in the induction motors. The experimental results indicate the effectiveness of the presented method compared with supervised classification methods.

Kroll et al. [103] presented a system for PdM of manufacturing stations by using timed-hybrid automata of the machine's normal operation.They have demonstrated that this method has an advantage over traditional, static limit testing. This advantage is a model of continuous dynamics, which reduces it to separately modeled state vectors. This, in turn, allowed powerful anomaly detection by using a hybrid data acquisition and anomaly detection strategy, as well as they presented an outlook for other applications, such as PdM scheduling. While, Wei et al.[104] introduced a new CBM strategy to determine the optimal action (e.g., no action, imperfect repair, and corrective replacement) based on the system status by reducing the average long run cost average.

Cline et al. [1] revealed the potential of ML strategies for improving the activities of a service department for oil and gas machinery. Analyzing significant data sets of individual machine performance resulted in substantial improvements in the customer’s ability to identify risky assets up to one year in advance. Paolanti [105] depicted the ML architecture for PdM on the base of the RF algorithm. The model was tested on a real manufacturing example by implementing the data collection and data analysis system, applying the ML algorithm, and comparing it to the simulation tool analysis. Data collection has been done using different sensors, PLC, and communication protocols before being available to data analysis tools on the Azure Cloud architecture. Preliminary results show that the technique is capable of effectively detecting various machine conditions.

The current research, which is concerned with the fault detection in the bearings components, helps the production managers plan the maintenance activities, i.e., technicians and spare part availability. The supervised ML method has been applied in this study where the data fed (mainly temperature and vibration) belong to the labeled type. In this respect, a comparison of four different types of classifiers; DT, RF, GB, and SVM. This comparison was achieved by utilizing the python

programming language to investigate which type provides the highest detection accuracy. Since the binary classification output of the applied ML algorithms can generate the pseudo probability of an observation belongs to a class, the authors choose to use the utility theory to utilize the likelihood of failures and perform correct maintenance actions.

4.5 Feature Selection & Extraction Methodology

Data pre-processing, feature extraction and selection, and fault detection are the three key steps in a learning protocol for intelligence prediction models. According to a specific feature selection criterion that selects the dataset's relevant features, locating a subset from an initial data set is referred to as feature selection. It contributes to reducing the size of data processing by removing redundant and irrelevant characteristics. Feature selection strategies can increase learning performance, minimize learning time, and simplify learning outcomes by pre-processing learning algorithms and selecting useful features [106].

The feature selection methods can be based on statistics, data theory, manifold, and rough set and can be categorized according to various criteria [107].

- According to the used training data (labeled, unlabeled, or partially labeled), feature selection methods can be divided into supervised, unsupervised, and semi-supervised models.
- According to their relationship with learning methods, feature selection methods can be classified into a wrapper, filter, and embedded models [108].
- According to the evaluation criterion, feature selection methods can be derived from correlation, Euclidean distance [109], consistency, dependence, and measure information.
- According to the search approaches, feature selection methods can be divided into forward increase, backward deletion, random, and hybrid models [45].
- According to the nature of the output, feature selection methods can be divided into subset selection models and feature rank (weighting) [110].

The objective of using feature selection methods is to choose features that allow for an accurate explanation of the equipment's state and, as a result, effective defect

classification and prediction. It is critical to devise a comprehensive scheme capable of selecting the most proper features to optimize the classification model's performance for fault detection evaluation. Researchers have previously investigated the principal component analysis (PCA) technique for signal processing. The PCA-based feature selection scheme for machine condition monitoring is built on the assumption that the amplitude of vibration signals faulty machine components rises in proportion to the defect's magnitude. This study explores PCA suitability to identify the most representative function as inputs to train ML models due to its ability to distinguish directions with the most significant variance in the original dataset [111].

The feature extraction process is considered a critical stage that directly affects fault detection effectiveness and precision. As mentioned in the previous chapter, several sensors are installed at various machine locations connecting in a condition monitoring device to collect numerous potential fault data. The data collected by these sensors is disordered and associated with a variety of sources. As a result, these signals' characteristics are inconsistent, making artificial feature selection complex and increasing ambiguity in the prediction model. An efficient and sensitive method for extracting statistical characteristics from sensor signals to determine equipment status is needed as an outcome [112].

Furthermore, feature extraction often necessitates raw data transformation into features with high pattern recognition capacity, while raw data is regarded as features with low recognition ability [110]. In recent years, a plethora of existing approaches have been created and classified into various categories. Execution of feature extraction by time domain, frequency domain, time-frequency representation, and phase-space dissimilarity are only a few examples. In the roller element bearings case, the time domain features extraction method is the most commonly utilized. For non-stationary and non-linear signals, the third and fourth methods are appropriate [113,114].

4.6 Machine Learning Algorithms

Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), and Support Vector Machine (SVM) algorithms are used to find the best classifier for the data under study. A brief introduction about ML algorithms is given in sections 4.6.1 through 4.6.4.

4.6.1 Decision trees (DT)

In respect of this, Y. Sheng and S. M. Rovnyak [115] and Kotsiantis et al.[116] have created a summary of decision trees in which the benefits of the DT in ML are discussed. The decision tree, which is a well-known methodology for creating logic-based rules as well as classification rules by tracing down the tree's nodes and branches, was selected as a consideration by the authors of the current thesis. The decision tree model often yields strong results, satisfies the consistency criteria, and creates simple logic rules that operators can understand.

For learning patterns, decision trees usually require a large number of data samples. This denotes that a time series should not be too short, and multiple time series should be obtainable. Both related conditions and events should be included in the sample time sequence. The method can only be trained using historical data. As a consequence, it is unable to detect events that did not occur in the dataset.

In the proposed ML classification model, the binary classification (0,1) dependent-decision trees display a stronger propensity in terms of results, and thus the decision/classification can be easily calculated [117].

4.6.2 Random Forest (RF)

Random Forest, also known as ensemble decision trees, was used as a classifier algorithm for several reasons: (1) it provides better predictive results. (2), it allows for constructing multiple decision trees. (3) it will enable fault detection with more excellent reliability and accuracy than DT, particularly when the data is initially extended [118, 119]. RF is often used to minimize differences in actual and predicted values, such as variance, bias, and noise, which are not functionally included in RF.

For nodes splitting, each tree is built under the consideration of a random subset of the features. Because of the usage of only features subset, the decision tree forest can handle a larger number of features. Within these features, the random subset selection makes the decision tree such a random subspace method, which in turn led to prevent overfitting. Furthermore, sampling on the dataset, trees are randomized using bagging and boosting techniques to generate splits [101].

4.6.3 Gradient Boosted (GB)

Gradient Boosted is an ensemble learning strategy that generates weak tree classifiers in a stage-wise fashion, close to how other boosting algorithms do for a different base model. To apply a GB algorithm to a particular problem, we must first specify the optimal tree size, and the number of iterations (trees) needed to achieve the best prediction accuracy. Each iteration attempts to reduce the loss function, such as cross-entropy or overall squared errors, ensuring the number of iterations should be sufficient to minimize the error function [120]. RF, on the other hand, creates each tree by random sampling and replacement. In the GB model, the learning protocol often sequentially suits new models to provide a more reliable estimate of the response variable. With various model designs, boosting algorithms are relatively easy to implement [121].

A gradient boosted model is a generalization of tree boosting that aims to create a reliable and efficient off-the-shelf data mining procedure compared to a single powerful predictive approach, such as neural networks.

4.6.4 Support Vector Machines (SVM)

Support Vector Machines (SVM) are currently the most popular tool for the classification and regression of large sample sizes, owing to their high classification precision, also for nonlinear problems, and the availability of optimized algorithms for their computation [10, 122–124]. It solves for separating hyperplanes when used in conjunction with learning algorithms. It provides the best separating hyperplane for maximizing the margin between two classes on either side. The hyperplane margin theorem separates the two data classes to minimize an upper bound on the predicted generalization error. In recent years, SVMs have received a lot of attention in various science, particularly in machine health diagnosis and monitoring [125].

4.7 Research Methodology

Figure 4.2 summarizes the research methodology, which is based on integrating an online fault detection algorithm with the decision theory for PdM. As the figure indicates, historical data are used for the offline model training, whereas online data are observed from instantaneous sensor measurement for predicting the process state. The utility theory is finally used for planning PdM.

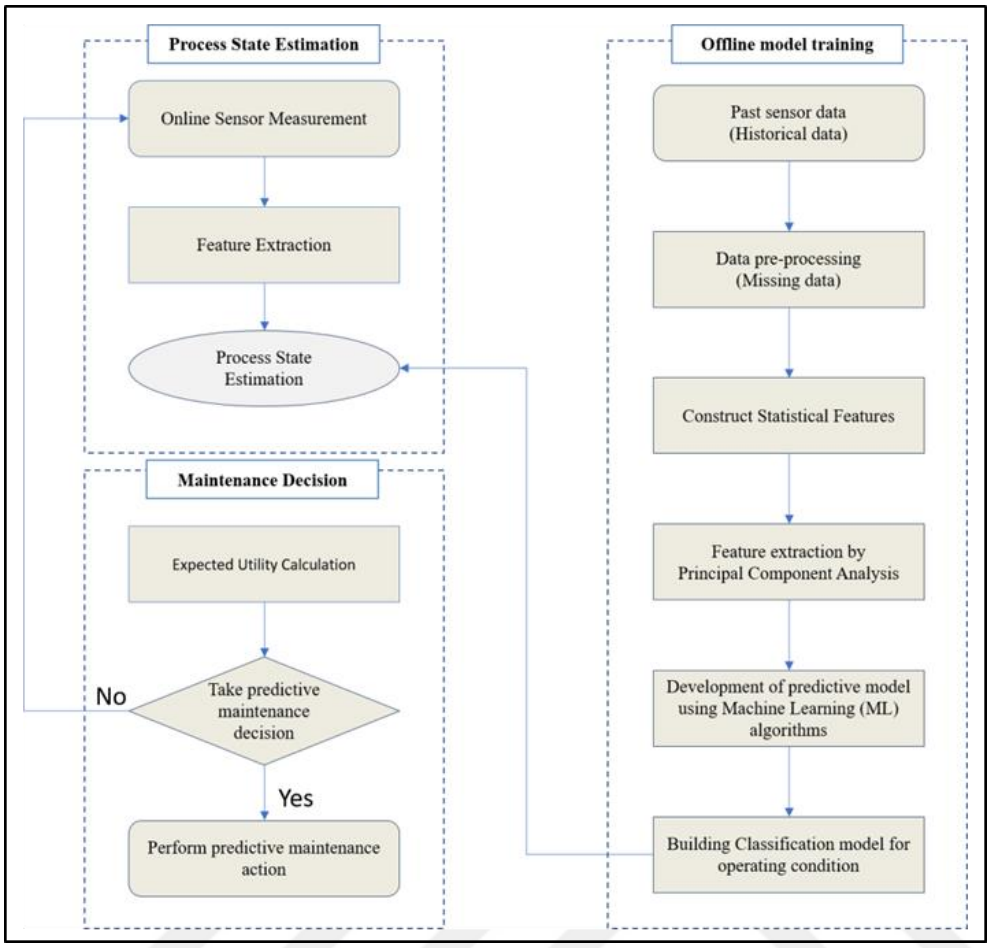


Figure 4.2 : The flow chart of the proposed method for roller bearing fault prediction model.

The first stage of building the prediction model depends on data acquisition, collecting and storing useful data from the target system to monitor the condition and diagnose the faults. The input for the data acquisition process is vibration signals and temperature readings. These signals are extracted to reduce feature space's dimension where the reduced features are fed to several ML algorithms to classify the operating conditions. Performance comparison among the tested ML algorithms is performed to select the one with the most accurate bearings faults prediction. The chosen model is then used to process state estimation for online sensor data measurements after feature extraction. Besides, we utilize utility theory coupled with the probability scores resulted from ML to guide decision-makers on when to implement the maintenance activities efficiently and cost-effectively. Therefore, our decision model provides a well-defined framework for selecting the correct maintenance action. The following sections explain more details of the study's main stages.

4.8 Offline ML Models Training

Several statistical features were extracted to train the ML models that, in turn, generate the final fault predictions. Seven descriptive statistical features for each sensor signal were constructed from the selected dataset of the bearing component; they are mean, skewness, kurtosis, maximum and minimum values representing the upper and lower ends of our data. The standard deviation (SD) and RMS was also included [55]. These statistical features are calculated for each selected attribute (i.e., temperature and vibration) gained from six different sensors. Among those features, the RMS values are considered the most effective to distinguish between healthy and faulty states [95]. The attributes used in the model listed in (Table 4.1).

Table 4.1 : Selected attributes for ML prediction model.

Type of sensor	Attribute
Accelerometer (vibration signal reading)	NDE Bearing (x-axis)
	NDE Bearing (y-axis)
	DE Bearing (x-axis)
	DE Bearing (y-axis)
RTD (temperature reading)	NDE Bearing
	DE Bearing

The binary classification is viable for PdM, being able to estimate whether the machine will fail over a future period of time. To use a binary classification, it is necessary to identify two classes, represented by zero and one. Each class is a record of a unit of time for an asset that conceptually defines the operating situations, considering the pump design's technical data and specifications.

In the PdM binary classification context, the class “1” denotes the faults while the “0” class stands on the normal operation condition. This classification aims to find a model that identifies which bearing may fail or typically work in the future. In the present work, two different operating conditions have been considered. The first condition was labeled as normal, where no faults were present in the bearings. Whereas the second one is known as fault indication condition (announced when the operating conditions of the bearings, i.e., temperature or vibration, reach to or go over the critical limit value). The ranges of the standard and critical limits are listed in (Table 4.2).

Table 4.2 : Binary classification of the bearing operating condition for ML model.

Description	Range	Critical Limit	Model Classification
Pump DE &NDE Temperature	0-200 deg	80 deg	0: Normal operating condition <80 deg 1: Fault indication when ≥80 deg
Pump DE & NDE Vibration	0-10 mm/sec	6 mm/sec	0: Normal operating condition <6 mm/sec 1: Fault indication when ≥ 6 mm/sec

Thus, The application of machine learning to temperature and vibration data has primarily focused on defining associations between normal and critical operating conditions to extract the most possible root causes for fault classification [113].

Temperature measurements help in potential failure estimating related to the temperature change in the equipment, such as excessive mechanical friction (faulty bearings, inadequate lubrication, fouling in a heat exchanger, and shoddy electrical connections). Those vibrations can indicate wear, imbalance, misalignment, and damage [78]. These measurements contribute to determining the causes of the faults that occur in the bearings, mainly due to either temperature and/or vibration. Following the results that came from these observations, expert knowledge of maintainers as well as the maintenance manual of pumping machinery, the right maintenance action can be executed.

4.9 Data Collection and Preprocess

Data collecting is the most important step in applying ML algorithms. As mentioned previously, this work is based on a real data-set collected from several types of sensors that monitor the pumping processes in the sewerage treatment company. The sensor data stream-in at an interval of one minute, which is equivalent to 1440 rows of data per day, describes the original data sets as a time series of accelerations and temperature shown in (Figure 4.3). However, a reserved-dataset is not directly suitable for creating a predicting model because it mostly contains noise and missing feature values. Therefore, the second step of data preparation and data preprocessing is applied before feeding it to the ML algorithm in order to convert the raw data into a clean data set and make them more suitable for further analysis.

In this respect, feature extraction is used for data preprocessing that focuses on modifying the data for better fitting in a specific ML method. It also involves

condensing the data by producing a smaller set of predictors that aim to capture most of the initial variables' information [10]. In this way, the actual data are replaced by fewer variables providing a reasonable fidelity.

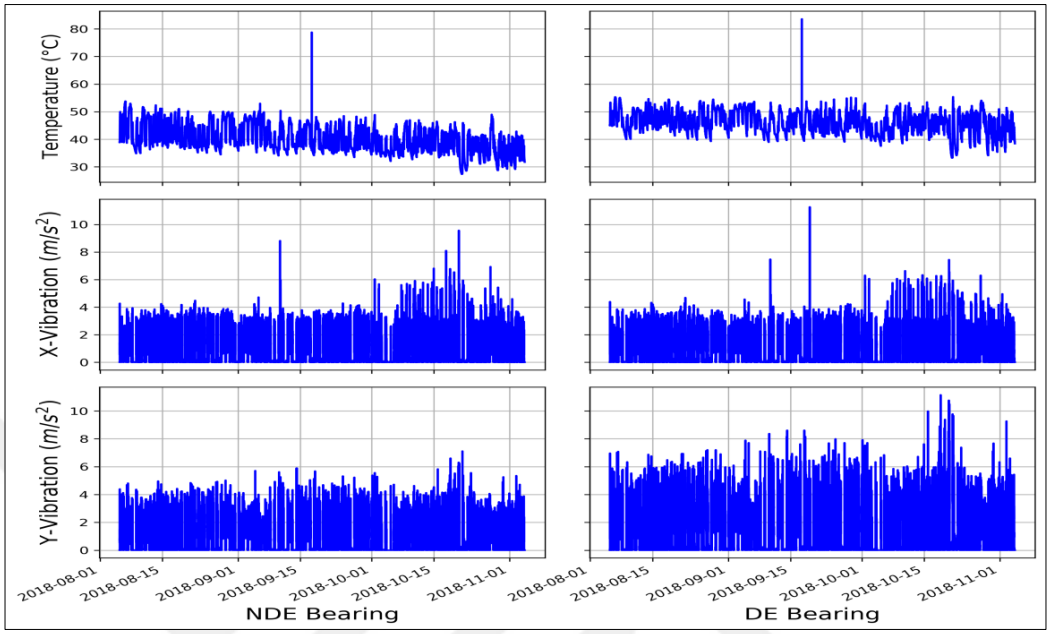


Figure 4.3 : Features of the original data sets.

Basing on this unique characteristic, PCA is finally used for the classification of variables and hence early identification of abnormalities in the data structure. In respect of this explanation and according to the available data, a set of 42 statistical features for six attributes listed in (Table 4.1) was reduced to a smaller set of seven uncorrelated final features corresponding to a 95% variance of the original data set.

Figure 4.4 shows the change of the explained variance ratio of the 42 variables selected for PCA vs. the principal component. It is can be seen that seven of these 42 variables have explained 95% of the variance. This means that the seven PCs subspace contains enough information about the variation of the original features which is sufficient to construct the model that can detect the faults in the bearing component.

Later, the analyzed frame of the target timestamp has been adequately sized by a limited-analysis approach. In current work, the time series is split into sub smaller periods in which the above-described features are extracted from sliding windows with a size of ten hours and a sliding length of one hour. These strategies could be performed using weekly or monthly time periods depending on the PdM requirements [126].

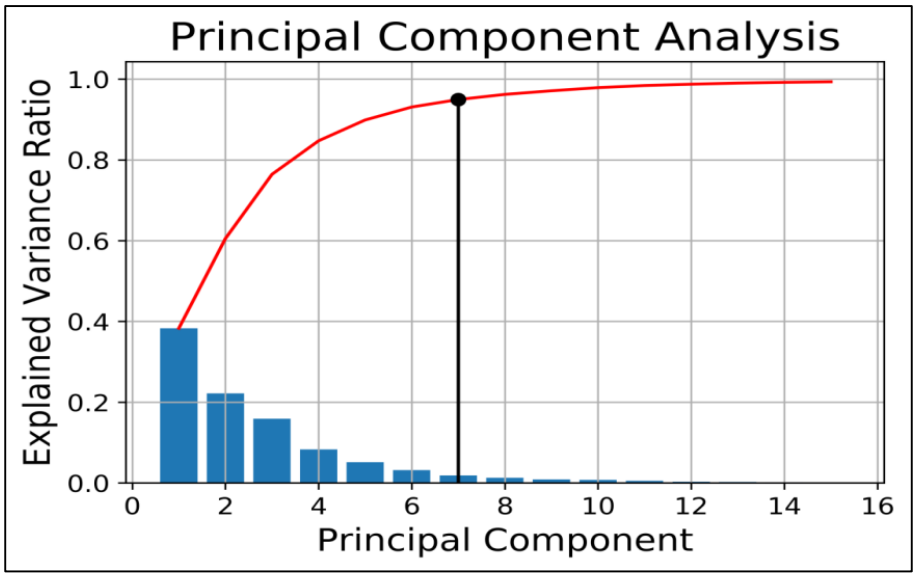


Figure 4.4 : Principal Component selection.

The use of previous time steps to predict the next time step is called the sliding window method. In some literature, it's referred to as the window process. In statistics and time series analysis, this is called a lag or lag method.

A classification model is then generated from the training set while its accuracy is estimated on the test set. Among the most commonly used methods for evaluating a classifier's performance by splitting the original data set into subsets is k-fold cross-validation. A subset of the training set is used to construct the classifier, called a validation set, used as a test set with which the original training set is learned to tune the model or obtain the model parameters [87].

In our model, we performed five-fold cross-validation using the raw data set. The training set is divided into five equal parts; one is used as the validation set, whereas the remained ones formed the training set. We have repeated this process five times, considering a different part as a validation set at each time and compute the validation data accuracy. The final accuracy results are the average of all different validation cycles.

4.10 Building Classification Algorithm

While the Process state is used as input for ML algorithms, y_t is the output as presented in the following equation (4.1):

$$y_t = f(X_{t-q}) \tag{4.1}$$

Where y_t is the machine condition which is defined as a normal condition and critical condition, X_t is the process state represented by extracted time-series features at time (t) and time lag (q). In this formulation, we would like to predict the process condition at q periods ahead.

DT, RF, GB, and SVM are used to determine the best ML technique that will predict process conditions. In DT, the Gini index has a dual function: it is utilized to find the feature splits the training set that would be a root node of the tree; moreover, it can be used in evaluating the quality of a particular split [127].

The Gini index is determined by:

$$G = \sum_{k=1}^{K} \hat{p}_{mk}(1 - \widehat{P}_{\mathrm{mk}}) \tag{4.2}$$

where $\widehat{P}_{\mathrm{mk}}$ denotes the proportion of training observations in the mth region in the kth class.

The maximum depth of a tree is set to five to prevent overfitting where max depth gives the maximum depth up to which a tree can grow [87]. To achieve the best results in the test data set, we tried values from two to ten for the maximum depth parameter so as to cover a wide range of possibilities.

In the RF algorithm, the number of trees (number of iteration) is set to 100, which used the same parameters of the splitting decision and the maximum depth of the DT. Using more than 100 models in RF algorithm did not improve the results.

The learning rate is taken as 0.12 and 100 models are built in a GB Algorithm. As the case of RF algorithm; 100 models did not improve the results of GB algorithm. Among the tested learning rates (0.01 to 0.5), a learning rate of 0.12 gave the best accuracy results for GB.

For the SVM algorithm, the radial basis kernel function outperforms the kernel functions of linear, polynomial of order two and three, and sigmoid function. Another important support vector classifier (SVC) parameter is regularization parameter C changing the regularization parameter affects the shape of the function. While High values of C results in more smooth functions, low values result in more complex functions leading to overfitting problems. In our experiments, we found that the best

C value is 1.0. Results of the algorithms are summarized in the figures and tables below.

Next, we will fit a classification method to predict pump condition using delay (lag) functions that need to be created from data sources including timestamps. Lag features are the classical way that time series forecasting problems are transformed into supervised learning problems. The most straightforward approach is to predict the value at the next time (t+1) given the value at the previous time (t).

The discussion begins with analyzing the numerical and graphical summaries that resulted from applying ML algorithms for the bearings data. For each recorded data, we have predicted the fault occurrence recognized by pump operating conditions for the nine previous hours, Lag 1 through Lag 9. Now we compare the algorithms' performance across five random train-test splits of the data using classification accuracy. Figure 4.5 presents the output of accuracy for every nine lags expressed as the probability of correct classification. As the figure indicates, the GB and RF achieved slightly more than 88% mean accuracy in Lag 1, associated with the correct detection of critical bearing conditions before 1 hour. On the other hand, SVM and DT respectively resulted in 82.2% and 81.9% mean accuracy giving an initial indication that DT gives the worse accuracy compared with the other three algorithms. A more extensive analysis of the algorithm’s performance is presented later in section (4.11).

In general, we can see that the prediction accuracy for all models is decreasing meaningfully with increasing a lag number from one to nine, reaching minimum prediction accuracy of less than 53% at lag9. This is a logical consequence since many unexpected circumstances might appear when the prediction took place earlier. (Table 4.3) summarizes the mean and SD for all ML models for nine lags.

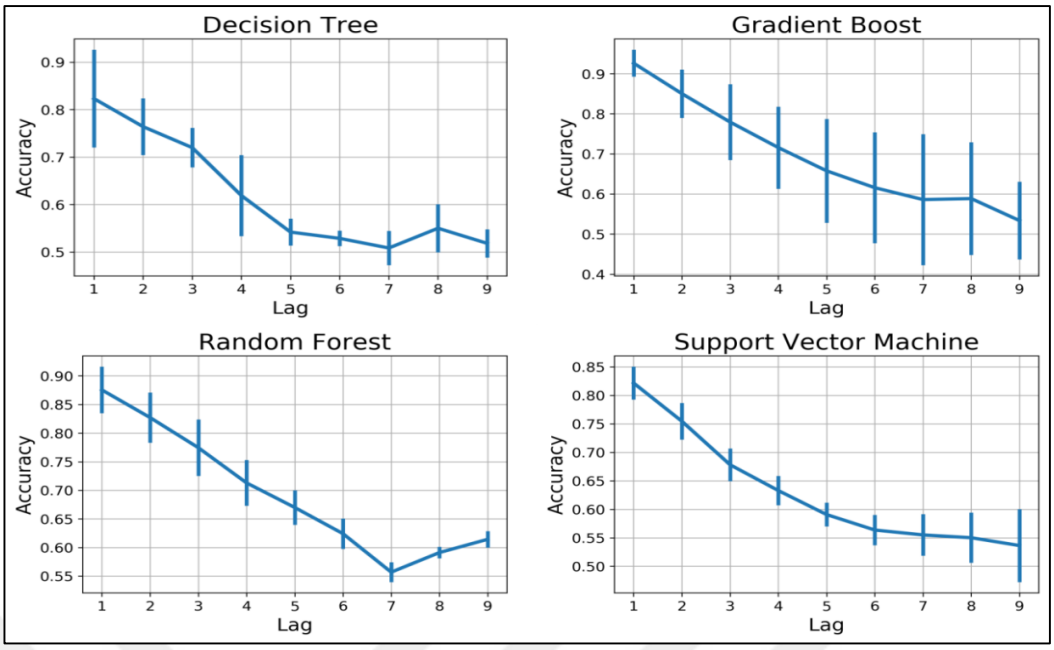


Figure 4.5 : Comparison of ML algorithms performance with respect to their Accuracy-Lag.

Table 4.3 : The mean and standard deviation for all ML models for nine lags.

DT algorithm	mean	SD	GB algorithm	mean	SD
Lag1	0.819	0.102	Lag1	0.926	0.034
Lag2	0.759	0.058	Lag2	0.85	0.06
Lag3	0.707	0.031	Lag3	0.779	0.095
Lag4	0.616	0.075	Lag4	0.716	0.102
Lag5	0.543	0.032	Lag5	0.658	0.129
Lag6	0.518	0.022	Lag6	0.616	0.139
Lag7	0.511	0.043	Lag7	0.586	0.163
Lag8	0.544	0.051	Lag8	0.589	0.141
Lag9	0.524	0.017	Lag9	0.534	0.097
RF algorithm	**mean**	**SD**	**SVM algorithm**	**mean**	**SD**
Lag1	0.876	0.031	Lag1	0.822	0.029
Lag2	0.826	0.045	Lag2	0.755	0.032
Lag3	0.767	0.041	Lag3	0.678	0.029
Lag4	0.729	0.027	Lag4	0.633	0.026
Lag5	0.672	0.025	Lag5	0.591	0.021
Lag6	0.611	0.042	Lag6	0.564	0.027
Lag7	0.569	0.023	Lag7	0.555	0.037
Lag8	0.595	0.014	Lag8	0.55	0.044
Lag9	0.61	0.015	Lag9	0.536	0.064

For the purpose of comparing algorithms performance, it is important to consider both mean and SD values. The higher the SD, the less precise is the prediction estimate. For example, although the mean value for GB is better than RF, the SD in RF is less indicating a more precise estimation.

4.11 Performance Measures for ML Models

Several performance measures are used to compare and evaluate the intensity of model prediction. As mentioned earlier, four distinct models are developed to predict if the operating condition is critical or normal where the maintenance is performed or delay accordingly. For the testing accuracy of the classifiers method, the training outcomes of an application are compared in terms of predictive efficiency. In the analysis, accuracy (%) is considered as a performance index and is calculated as:

$$Accuracy = (TP + TN)/(TP + FP + TN + FN) \tag{4.3}$$

TP, TN, FP, and FN are true positive, true negative, false positive, and false negative rates. Thus, the accuracy is an integration of precision (or positive predictive value) and recall (sensitivity) measures [128]. The precision determines the exactness of the model. It is a ratio of correctly predicted positive instances (TP) to the total positively predicted instances (TP+FP). Precision is denoted as:

$$Precision = TP/(TP + FP) \tag{4.4}$$

Recall, on the other hand, is an indicator of the model's completeness. It is a ratio of a correctly predicted positive instance to the total instance of the positive class (TP+FN) in the test set. A recall is determined as:

$$Recall = TP/(TP + FN) \tag{4.5}$$

Precision represents the model's performance concerning false positives, whereas recall represents the performance with respect to false negatives. The F_1-score conveys the balance between precision and recall by taking their weighted sum. F_1-score is calculated as follows:

$$\mathrm{F}_1 = (2 * Precision * Recall)/(Precision + Recall) \tag{4.6}$$

Similar to the accuracy, F_1-score performs well with the reasonably balanced dataset. Given the performance evaluation measures, the idea is to maximize the TP and TN and minimize the FN and FP. Generally, a reasonable tradeoff between the FP and TN risks is needed for better predictability. However, in the case of maintenance, the false negative (i.e., when the model predicts no need for maintenance, where it is needed)

is more critical. Finally, in our experiments, we also compute the receiver operating characteristic curve ROC (a.k.a Area under the curve AUC), which is a measure of the model's performance based on the tradeoffs between TP and TN rates over all possible risk thresholds between 0% and 100%. It is worth noting that a ROC over 0.70 is considered good, and a ROC over 0.80 is very good in the ML community [1].

Table 4.4 shows the model evaluation results tested on a cross-validation dataset. All models offer a negligible difference in performance. The GB model performs best in terms of all performance indicators where it reaches an accuracy of 92%; precision of 92.6% has an F-score and Recall of 91% and 89.5%, respectively. This supports our previous conclusion that the GB approach outperforms the other tested ML models. The SVM model, on the other hand, shows the lowest accuracy rate nearly to 82% as such as to the other evaluation criteria. Therefore, GB and RF showed the best performance indicators, with almost identical measures, and are found to outperform the other two models. It is also noteworthy that even with the worst ML model, DT, the AUC measure is considered acceptable (>0.7). ROC curves plots are used in order to evaluate the distinctive ability of the detection, the GB model exhibits an even higher AUC (92.8%) while the DT model gives a lower AUC (74.5%) as shown in (Figure 4.6).

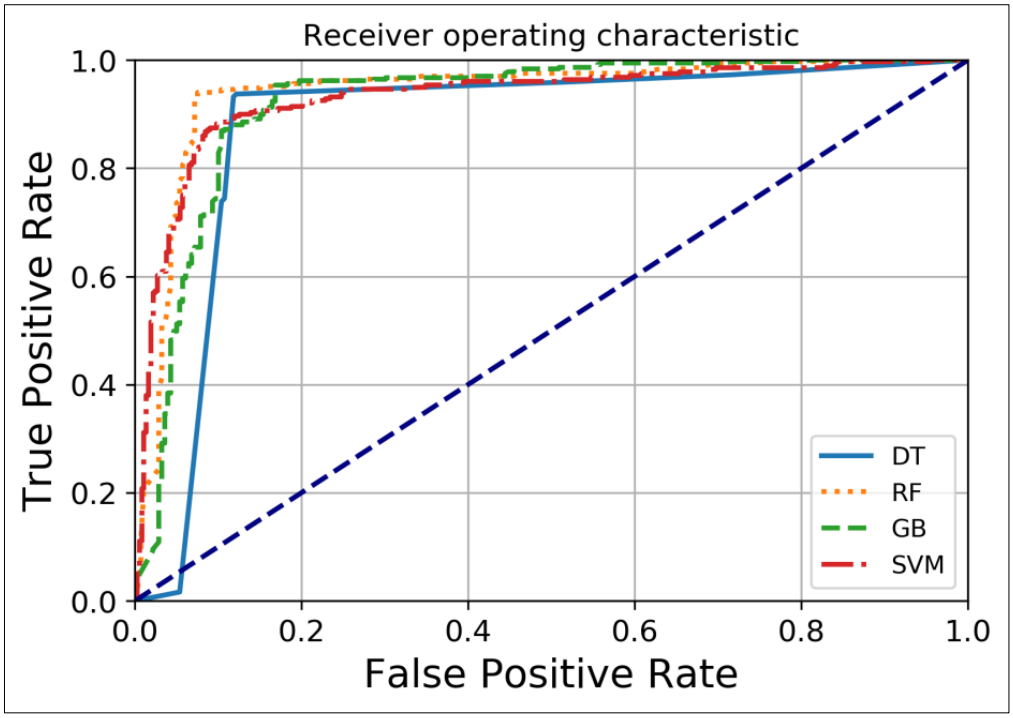


Figure 4.6 : ROC Curve of ML Algorithms.

Table 4.4 : Result of prediction models.

Test Type	DT	GB	RF	SVM
Accuracy	0.829	0.923	0.876	0.822
F_1-score	0.775	0.909	0.847	0.785
Precision	0.847	0.926	0.906	0.824
Recall	0.721	0.895	0.799	0.758
ROC_AUC	0.745	0.928	0.912	0.885

4.12 Decision-Making Theory

Aiming to determine the optimal strategy alternatives, decision-making and utility theory have been comprehensively used to plan manufacturing and production activities [129]. In our case study, there is a list of $d_1, d_2 \dots d_{\mathrm{m}}$ of decisions (such as taking, or not, the maintenance action) and $\emptyset_1, \emptyset_2 \dots \emptyset_{\mathrm{n}}$ of events (such as normal or critical conditions) with the uncertainty of probability $p(\emptyset_{\mathrm{j}})$ of event $\emptyset_{\mathrm{j}}$ ($j = 1, 2 \dots n$).

Among the possible decisions ($d_1, d_2 \dots d_{\mathrm{m}}$), the optimal one is chosen to avoid the extra costs of incorrect maintenance that arise from unreal predictions which can be achieved by maximizing the expected utility function [130]. The utility of the consequence (u_{ij}), in correspondence to a decision (i) on an event (j), is determined by a utility function. As for fault prediction of bearings in the current work, there are two decisions, namely, d_1: no maintenance action (continue working) and d_2: perform maintenances action along with two events $\emptyset_1$ and $\emptyset_2$ which present normal and critical conditions, respectively. The two correct decisions are: (a) to do maintenance if it is a critical condition and (b) to continue operating pumps if the normal condition has a significant utility.

To compute the expected utilities for diverse decisions, formula (4.7) is considered for which the probabilities for each event $p(\emptyset_j)$ shall be calculated [131]. They can be initially estimated based on historical data. After that, once the in-situ sensor data y is available, $p(\emptyset_j)$ will be updated as $p(\emptyset_j|y)$ using ML algorithms. The optimal decision (d_i) will thus be chosen based on maximal expected utility. This aims to obtain the optimum maintenance action which combines the reliability and availability for each possible action [132].

$$\max_{i} \sum_{j=1}^{N} u_{ij}\, p(Øj\,|y) \tag{4.7}$$

4.13 Decision Making

Once the ML algorithms are tested and the best approach is selected, the utility theory as described in the previous section - is integrated into our model to plan the maintenance action based on the probability of fault occurrence. (Table 4.5) summarizes the utility matrix uij expressing the corresponding consequence for taking decision *i* given event j. Since it is less desirable to continue the machine work when the process is under a critical condition compared with taking a maintenance action when the process is normal, the utility (cost) for consequence u_{12} is chosen to be less (higher) than that of u_{21}.

Table 4.5 : Decision Table for critical condition prediction.

	Normal Condition ($Ø_1$)	Critical Condition($Ø_2$)
Probability	$p(Ø_1)$	$p(Ø_2)$
Continue Working (d_1)	$u_{11} = 1$	$u_{12} = -1$
Maintenance Action (d_2)	$u_{21} = -0.8$	$u_{22} = 1$

The probability $p(Ø_j)$ represents how likely event *i* is to happen given the status of the active features extracted in the offline training phase. There are several approaches to detect these probabilities such as Bayesian networks and neural network algorithms. In this paper, however, we utilize the score resulted from the ML, which is a reflection of the status of all extracted features, and use it as input to the utility theory-based decision making. It is worth mentioning that the final output of ML is binary 0, 1 classification depending on whether the resulted decimal score is less or greater than 0.5, respectively. However, the decimal score (before binary classification) can be utilized in the application of utility theory as probability of normal and critical conditions.

The choice of the utility value u_{ij} is based on the decision maker's knowledge of the system under investigation. Obviously, continue working under normal condition and take a maintenance action under critical condition will receive the maximum utility 1. The worst consequence, on the other hand, happens when work continues on a machine

under a critical condition which not only produces nonconforming items but also may cause extra damage in the machinery and production system. Therefore, a utility of (-1) is selected for such a consequence. The choice of utility value for taking a maintenance action under normal conditions depends on the cost of unnecessary maintenance and its impact on the production flow but in most cases, it is less serious than working under critical conditions. In this study, a value of (-0.8) has been arbitrarily chosen for computational illustration.

It is also worthy to mention that the attitude of decision makers toward risk can significantly influence the way utility scores are defined. For example, in situations where the production management seeks to maximize profit through increasing the production volume within short period (risk-seeking approach), an unnecessary maintenance act will be avoided and critical condition signs with relatively small probabilities will be disregarded. On the other hand, risk-averse decision-makers who follow a more conservative approach will be more protective against any risk of machine failure even with low probability and thus will assign higher utility to unnecessary maintenance action.

Therefore, in the current thesis we provides a general framework for integrating the concept utility theory with ML to improve decision making regarding PdM. Considering maintenance costs including inspection, repair, failure, and replacement costs; as well as decision-makers' attitude to risks will help to provide an accurate estimate of the expected utility for each decision alternative. We highlight this as a gap for further extension and investigation.

Figure 4.7 shows the two expected utility curves for d_1 and d_2. Based on utility maximization rule we should take maintenance action when the expected utility for d_2 (maintenance action) becomes more significant than d_1 (continue working). Thus, the maintenance action time can be recognized by the intersection of two expected utility functions.

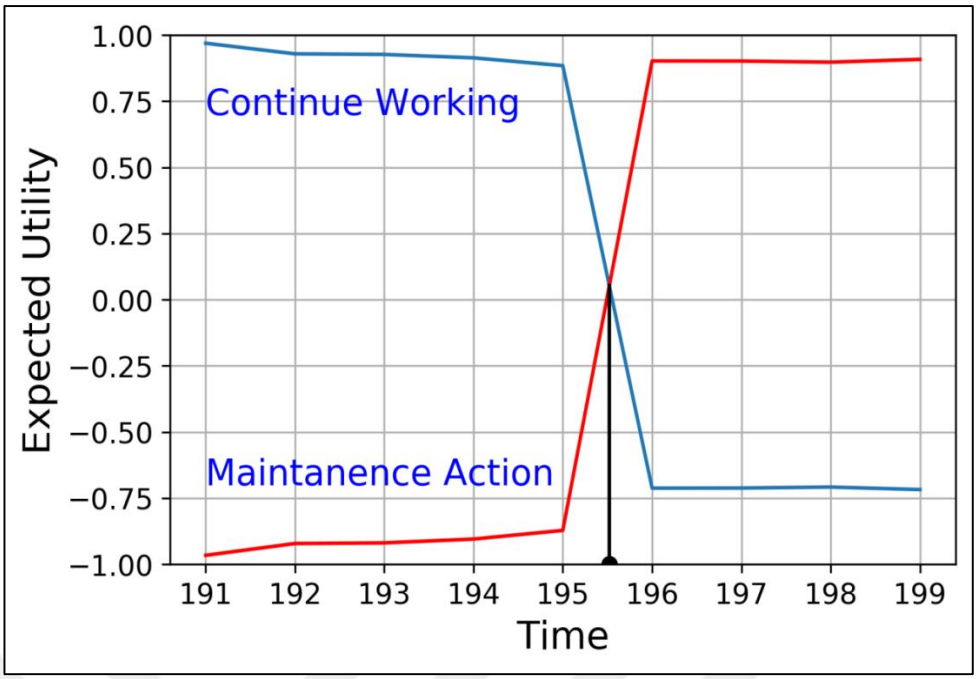


Figure 4.7 : Expected utility for "continue working" and "maintenance action".

5. DEEP LEARNING IN PREDICTIVE MAINTENANCE

This chapter introduces the impact of AI on the PdM, which is a significant piece of future progressed creation frameworks. First, it presents a general background of DL algorithms and reviews the most of methods implemented in the PdM strategy and provides a description of the data set that was used to build a prediction model. Besides, it talks about the most procedure for data pre-processing, extraction, and selection features which is performed on the vibration signals for building classification algorithm. It then provides a comprehensive review for recent research into the field of machine health assessment and fault prediction using LSTM and CNN techniques, which that adopted to build the prediction model of this research. Finally, this chapter provides a comparing result of predictive accuracy between ML and DL algorithms. Finally, it presents the experimental results of both models that have been evaluated by five performance indicators: accuracy, precision, recall, F-score, recall, and an area under curve (AUC).

5.1 Introduction

Deep learning is a sort of Artificial Neural Network (ANN), or, to put it another way, it is the application of ANNs to learning tasks with multiple hidden layers [28]. Due to its possible benefits in data classification and feature extraction issues, DL has received a lot of attention. System health management, computer vision, natural language processing, voice recognition, power installations, and aerospace specialties are all examples of emerging research areas. Similarly, it provides solutions for detecting faults in electromechanical equipment, classifying deterioration, pattern recognition, and predicting part Remaining Useful Life (RUL) [90]. DL algorithms' primary advantage is that the highlights are not built by human expertise however gained from information itself through a generalized self-learning process. The aim is to model high-level abstractions in data to evaluate a high-level context, which can be done using supervised, partially supervised, and unsupervised learning techniques.

Various DL algorithms, such as Deep Neural Network (DNN) [133], Convolutional Neural Network (CNN) [134], Deep Feedforward Networks (DFN) [135], Long Short-Term Memory (LSTM) [136], Back Propagation Neural Network (BPNN) [137], and Deep Belief Networks (DBNs) [138], have been successfully introduced in a fault detection and PdM research fields. Numerous researches on intelligent fault identification of rotating machinery, as well as rolling part bearings methods for reliable prognostics, have been performed. The methods based on ANNs are frequently used in these studies, which used signal processing methods to extract features and then fed the features into ANNs to classifying failures [139]. Uncertainties, of course, play a significant role in the overall process, influencing the predictive performance. As a response, it's normal to set certain minimum requirements for the health management system in use. Information about the operating environment, sensor resistance, trust levels, and so on can be included. Additionally, a variety of techniques can be used to process all of the data, including optimization algorithms like genetic algorithms (GAs), artificial immune systems (AIS), and Monte Carlo methods, learning algorithms like SVMs and ANNs, and reasoning algorithms like fuzzy logic systems, clustering algorithms, particle filtering algorithms, wavelet analysis algorithms, and the PCA algorithm. Data pre-processing, feature extraction, and feature selection can all be done with these approaches [140].

Samanta et al.[141] presented a performance comparison of the bearing fault detection; they used time domain features of vibration signals and used three sorts of ANNs: multilayer perceptron (MLP), radial basis function (RBF) network, and probabilistic neural network (PNN). The characteristics were derived from finite segments of two signals, one with regular gears and the other with defective gears. In [142] by merging WPT and DBN, a Hierarchical Diagnosis Network (HDN) was created to address the challenge of consecutive bearing damage position and severity recognition. The HDN is made up of two layers of DBNs that aid a device in determining the data's basic structure. With a mixture of fault intensities, the first layer was qualified to distinguish bearing fault positions. The second layer, on the other hand, obtained the first layer's result to further separate the internal fault severities. Khaled et al. [143] combined a PCA with ANNs to detect faults in the manufacturing processes. They have been presented a model entails of three-parts. To begin with, data analysis helps to distinguish between normal and abnormal data clusters (data with and without defects).

Second, using Partial Least Squares (PLS), fault visualization in the principal component space 2-D was performed. Finally, by measuring the contribution, faults were located.

Many works have developed to prove the superiority of DL algorithms for system health management uses in the last few years due to the impact of increasing overall system versatility or potential cost benefits for maintenance, repair, and replacement. As a result, the authors took a realistic approach and focused their efforts on the area of system health management. For instance, Khan and Yairi [140] given a well-organized and comprehensive review of DL research on system health management, that covers a wide range of technology fields. The process of diagnosing and preventing system failures while forecasting the reliability and RUL of its components is referred to as health management. RUL is the amount of time between now and the end of useful life. Precise RUL estimation plays a crucial role in Prognostics and Health Management (PHM). For example, in [144] for predicting the RUL structures, a novel Restricted Boltzmann Machine (RBM) was proposed. A new regularization term was introduced to model the hidden nodes' trend ability. This was combined with an unsupervised self-organizing map algorithm, which was used to convert the representation into a health attribute that could be used by a similarity-based life prediction algorithm to predict RUL. Ahmad et al.[145] suggested a scheme for detecting the health of rolling element bearings. A bearing's health was determined using a dimensionless health indicator (HI), and the bearing's RUL was calculated with dynamic regression models. The RUL of a bearing component was determined after calculating the time to start prediction (TSP) using an alarm bound method. They used a gradient-based methodology to determine the fault threshold, on the other hand.

Yao et al. [46] Via time Empirical Mode Decomposition (EMD) and CNN, a novel DL method for bearing the RUL estimation approach was presented. The EMD method was used to decompose time sequence data without any character limitations, and it has a major advantage in dealing with non-stationary and nonlinear data. The featured information was then used as an input to the convolution layer of the suggested models. Ensemble models with various weighting methods were proposed for accurate prediction. The experimental results showed that prediction accuracy had improved. Zhang et al.[35] developed the new data-driven structure for estimating RUL in Tool Condition Monitoring (TCM). The structure included many modular components. The

proposed Adaptive Bayesian Change Point Detection (ABCPD) was used to perform data preprocessing for automatic data alignment and normalization. Then, the time window method was used with feature extraction from preprocessed signals by time-frequency domains. Finally, there are two approaches proposed for selecting features. The efficacy of each function extraction process on the chosen dataset was tested using Pearson's Correlation Coefficient (PCC) and PCA.

In [146] the RUL prediction model was developed using a DNN and statistical features in the time frequency domain. They did not define offline or online scenarios when dividing the deterioration data of multiple bearings into training and test sets. Shao et al.[147] to detect the roller bearing fault, researchers used a continuous deep belief network (CDBN) with locally linear embedding and a deep stacked auto encoder (DSAE). The findings revealed that the proposed approach was more successful than other approaches currently in use. Shao et al. [148] introduced a novel deep auto encoder feature learning method for gearbox and electrical locomotive roller bearings fault identification. The findings revealed that the proposed approach has been both reliable and efficient. Zhao et al. [149] employed the machine health monitoring systems (MHMS) based on four DL architecture categories such as Auto-encoder, RBM models, CNN, and Recurrent Neural Network (RNN). Although the applications of several DL models have also been reviewed and summarized the recent achievements in MHMS.

To deal with the previously presented researches, DL techniques have been used to produce increasingly effective solutions to detect the failure and predictions of the future operating conditions of components accurately, which is considered the most challenging problem in the PdM's programs. In this chapter, we attempted to propose a new data-driven model reached from DL algorithms to detect a roller bearing's pending failures and prognosticate its future operating condition to achieve correct maintenance actions.

5.2 Overview of the Long Short -Term Memory (LSTM)

LSTM is a special form of RNN for sequence learning tasks and has achieved great success in prediction and forecasting [150]. LSTM networks firstly were introduced by Hochreiter & Schmidhuber [136] as a new recurrent network architecture combined with a suitable gradient-based learning algorithm A variable estimator memory is

present in RNN, a data-driven process. These networks function by maintaining a persistent, continuous state. This latent state allows taking advantage of the information given by the previous processing steps. Forms a short- and long-term memory that aids in understanding a single prediction in a series of predictions. To put it another way, the RNN's memory serves as a context for the calculations. In others word, the key distinction between the LSTM and standard RNN is that the standard RNN's concealed units arrangement has been replaced by LSTM cells, which solves the problem of gradient disappearance and gradient eruption.

The promising results obtained from using RNN to make predictions demonstrated the models' ability to capture the important temporal information contained in sensor data. The LSTM is one of the recurring approaches discussed here [151], Gated Recurrent Unit (GRU) [152], and the simple RNN [46]. Long-range dependencies and non-linear dynamics in time series data can be captured using LSTMs. Speech recognition, handwriting recognition, machine translation, image captioning, genomic analysis, and natural language processing are just a few of the applications where LSTMs have been popular. A unique LSTM has recently become common for PdM, prognostics, and machine health monitoring.

Long Short-Term Memory networks can deal with variable-length data sequences as well as learn long-term dependencies. LSTM is a form of neural network that combines representation learning and model training, requiring no additional domain knowledge. Furthermore, this construction can allow us to uncover some hidden structures, allowing us to improve model generalization. Raw sensor data normally contains noise, except when temporal information is needed [153]. As a result, we utilize CNN in conjunction with LSTM to extract local features. The first purpose of LSTM is to keep back propagated errors from disappearing or exploding. In LSTMs, forget gates are used to prevent the issue of long-term dependence. Because of their ability to catch long-term dependencies, LSTMs should be superior to conventional RNNs in terms of knowledge utilization in cell states and learning useful representations of system conditions [154]. Recently, many researchers have combined CNN and LSTM models to extract temporal and unique features.

Kim and Cho [155] introduced a hybrid CNN-LSTM model for electric energy consumption achieving superior results than other conventional forecasting methods for the dataset. They found that extracting first the local features and then temporal

ones worked better than an LSTM-CNN model, performing with a mean square error (MSE). They also found that time series decomposition with DL models provides useful visualizations to predict better and analyze energy consumption. In [156] for photovoltaic (PV) power prediction, a one-dimensional hybrid DL model (LSTM-Convolutional Network) was proposed. The LSTM network was used to derive the temporal features from the data in the proposed hybrid prediction algorithm. The CNN model was used in the second phase to extract the data's spatial attributes. In contrast, Tovar et al. [157] presented a CNN-LSTM hybrid model with a stronger multi-layer architecture, this included a 5D-CNN model with max-pooling and a 5D-LSTM model. The results indicate that the presented five-dimensional CNN-LSTM model can consume more computational resources for training than a uni-dimensional model, but high accuracy has been achieved.

The LSTM neural networks have been commonly applied for intelligent fault diagnosis and machinery condition monitoring in recent years. Therefore, it’s adopted in several predictive kinds of research. For instance, a novel nonlinear hybrid model was established using the hysteretic extreme learning machine (HELM), LSTM network, differential evolution (DE) algorithm, and nonlinear combined mechanism to investigate and manipulate the implicit information of wind speed time series for wind speed forecasting. The aim is to improve wind speed detection accuracy and address the drawbacks of a linear combined process [158]. Zheng et al. [151] proposed an LSTM approach to estimating RUL that used several layers of LSTM cells along with regular feedforward layers to uncover concealed patterns from the sensor and operational data under a variety of operating conditions, faults, and degradation models.

The LSTM neural networks commonly applied for intelligent failure diagnosis and machinery condition monitoring in recent years. Therefore, it’s adopted in several predictive kinds of research. For instance, a novel nonlinear hybrid model was established using the hysteretic extreme learning machine (HELM), LSTM network, differential evolution (DE) algorithm, and nonlinear joined mechanism to investigate and manipulate implicit information of wind speed time sequences for wind speed forecasting. The aim is to improve wind speed detection accuracy and address the drawbacks of a linear combined process [158]. Zheng et al. [151] proposed an LSTM approach to estimating RUL that used several layers of LSTM cells along with regular

feedforward layers to uncover concealed patterns from the sensor under a variety of operating conditions, faults, and degradation models.

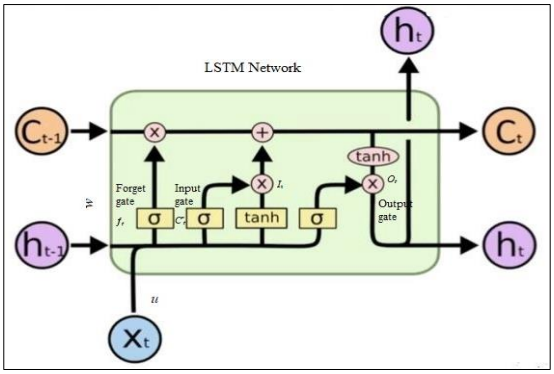
Wu et al. [159] Vanilla LSTM neural networks, which are typically used for supervised learning on language modelling and related state-of-the-art feature extraction technologies, were used to increase accuracy in RUL prediction problems involving complex industrial items. The vanilla LSTM was put to the test on NASA datasets for health testing of aircraft turbofan engines, which had four problems. Performance comparison of the used model with regular RNN and GRU-LSTM was also conducted, with vanilla LSTMs demonstrating excellent performance in the RUL estimation sector. Wu et al. [160] proposed a new data-driven paradigm for PdM that uses LSTM and RNN approaches to anticipate predicted defects and identify future health conditions in manufacturing systems. They used a motor bearing failure method to test the proposed model's accuracy and performance. Moreover, for combined demand-side prediction models over short and medium-term monthly horizons, an LSTM-RNN-based model was proposed by Bouktif et al. [161]. They executed ML techniques to compare with their proposed model. Using function selection and the GA, they introduced critical predictor variables, optimal latency, and layer selection. Several measurement metrics were used to assess the efficiency of the presented model, including the Coefficient of Variation RMSE (CVRMSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

Most of the research above have exposed that the LSTM outperforms DL models in terms of prediction accuracy. As a result, the current research suggests the LSTM approach for bearing fault detection, which uncovers hidden trends in sensor data under a diversity of operating conditions (normal and critical). More detail about the LSTM method is provided in the following section.

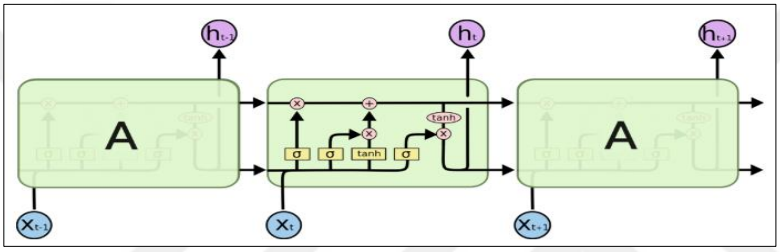
5.2.1 Architecture of LSTM

As previously mentioned, the LSTM methodology usually has dominant performance in dealing with time series due to its superior ability to address long-term dependence issues by managing information flow using three gate structures, i.e. the input gate, forget gate, and output gate. These gates control information passage along with the sequences, which can acquire long-range dependencies with more precision. Furthermore, LSTMs has a capability to capture nonlinear dynamics in time sequence

data and avoid the fundamental problem of gradient vanishing. (Figure 5.1a) depicts a standard LSTM cell architecture, while (Figure 5.1b) depicts LSTM cells at various time steps [153].



(a)



(b)

Figure 5.1 : LSTM architecture (a) for single cell (b) sequence of LSTM cell at time series [162].

The forget gate f_t, input gate I_t and output gate O_t are single-layered neural networks that have sigmoid activation function σ (yields output between 0 and 1), 0 indicates that no data is passed while 1 denotes that all data is passed. Simultaneously, a candidate layer uses the tan*h* activation function (yields output between -1 and 1). These gates take the input vectors (u) and previous output vectors (w), concatenate them, and finally apply the sigmoid activation function[162,163].

The current time input vector x_t is the input data to the LSTM model at time t; (h_{t-1}) is the last time output vector; and c_{t-1} represents the previous time cell state. The forget gate and the input gate, which are used to control the model's cell state. Basically, forget gates and input gates are designed to restrict the information flow. The forget gate controls the last cell state information c_{t-1} transmitted to the current cell state c_t. This process is defined in the following equation.

$$f_t = g(wf \cdot [h_{t-1}, x_t] + bf) \tag{5.1}$$

Where g(·) represent the activate function that achieves the sigmoid nonlinear function, w_F is the forget gate weight matrix, b_F represents the bias vector of the forget gate, and $[h_{t-1}, x_t]$ consists of the combination vector of the last time output vector h_{t-1} and the current time input vector x_t.

The input gate regulates the current input information x_t transmitted to the current cell state c_t, depicted in the equation (5.2).

$$I_t = g(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{5.2}$$

Where w_i represents the input gate's weight matrix and the input gate's bias vector is b_i. Then $\acute{c}_t$ is determined to get the current input state as seen in the equation below:

$$\acute{c}_t = \tanh(w_{\acute{c}} \cdot [h_{t-1}, x_t] + b_{\acute{c}}) \tag{5.3}$$

Where $w_{\acute{c}}$ is the weight matrix, and $b_{\acute{c}}$ is the bias vector, tan*h* represents a hyperbolic tangent function. It is then possible to achieve the current cell state c_t from equation (5.4), considering input gate and forget gate.

$$c_t = f_t * c_{t-1} + I_t * \acute{c}_t \tag{5.4}$$

The data flowing from the current cell state c_t is regulated, represented by the output gate O_t, to the current output.

$$O_t = g(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{5.5}$$

Where w_o represents the weight matrix, b_o is the bias vector.

Finally, the output gate O_t and the current cell state c_t calculate the LSTM model output displayed in the following equation.

$$h_t = O_t * \tanh(c_t) \tag{5.6}$$

By stacked memory cells, can hold information of previous input x in the output to some degree, carried by cell state, making LSTM an excellent tool to imitate time series. This is a reason why we will implement this approach for fault prediction.

5.3 Overview of the Convolutional Neural Network model

The fundamental theory of CNN is that convolutional kernels and the pooling operation will extract abstract features. The convolutional kernels in CNN create invariant local features by convolving multiple local filters with raw sequential data, and the pooling layers extract the most relevant features inside fixed length sliding windows [153]. The CNN technique has been actively explored for this purpose in a variety of applications, comprising image processing, voice recognition, and natural language processing. For the learning picture, CNN learns the weights of each layer's feature maps, extracts abstract visual characteristics including (input data points, lines), and faces, and maintains relationships between pixels for the learning image [164,159]. In addition to all these mentioned applications, CNN is widely used in many PdM fields for fault detection and diagnosis [160].

Ren et al. [165] used a CNN to develop a new approach for predicting bearing RUL. The spectrum-principal-energy-vector was introduced as a modern feature extraction tool for obtaining the eigenvector. The results demonstrate that the proposed model would significantly improve the accuracy of bearing RUL estimation. Chen et al.[166] introduced the DL technique based on the CNN method to detect and distinguish failures in the gearbox using vibration data determined with an accelerometer. With a vector generated by RMS values, SD, skewness, kurtosis, and rotation frequency, the feature representation was chosen as CNN's input parameters. An uncertainty matrix was used to evaluate the success of the presented model. Finally, the authors contrasted the CNN approach and the SVM algorithm to determine which method produced better experimental results for gearbox fault identification. Janssens et al. [167] suggested the CNN-based function learning model for condition tracking that uses vibration data to learn autonomously useful features for bearing failure prediction in a rotary machines.

The feature learning model is depend on CNN, which has been proven to be effective in many fields [168,169]. When compared to other feature-learning approaches, CNNs have numerous advantages:

1. Through their layered structure, CNNs learn multiple levels of data representations on their own.

2. Since CNN is considered as an end-to-end learning system, only one system has to be improved.
3. CNN is utilized to manipulate a data's spatial structure. For example, in a vibration signal's frequency spectrum, the spatial structure is described as a frequency sequence.

The majority of the DL architecture used to predict faults is based on CNN so it can accomplish the same precision or functionality with fewer parameters. As a result, CNN is thought to be a good choice for high-dimensional data and its ability to eliminate noise from vibration signals [170]. Furthermore, a large number of academic achievements emerged in the field of bearing fault detection using the CNN model. In [170] for carrying fault size prediction and severity determination, a hierarchical adaptive deep CNN was suggested. They prove that the proposed approach achieves adequate precision in both fault pattern recognition and fault size assessment. Sun at al. [171] utilized a dual-tree complex wavelet transform (DTCWT) and CNN to automatically classify a gear damage feature from multiscale signal features. The experiments' results showed that the proposed method was feasible and reliable, especially in the gear's poor fault features. Pan et al. [172] as a coherent frame, a one-dimensional CNN with LSTM was used. To define the bearing fault types, they used the CNN output as input to the LSTM. The proposed model was compared to various ML algorithms using the same extracted elements, with prediction accuracy rates ranging from 75% to 90%. In contrast, the findings show that the presented model's average accuracy score in the research dataset is over 99%. Another methodology used by Guo et al. [173] for a rolling bearing fault trained in a greedy layer-wise manner, an optimized deep fault classifier approach based on the stacked de-noising auto encoder (SDAE) was used to de-noise random noises in the initial signals and reflect fault features in defect pattern diagnosis. The experimental findings indicate that the presented model outperforms DBN in terms of diagnostic precision, with more than 99% accuracy.

Magar et al. [174] introduced a CNN FaultNet that can efficiently evaluate the bearing fault with a high degree of precision. Different signal processing was combined with the ML algorithm to identify different types of bearing faults by analyzing bearings' vibration signal data. They also demonstrated that the distinctive aspect was the concept of channels to extract additional information from the signal. Li et al.[175]

unified the CNN and Demspter-Shafer (D-S) proof theory based on data fusion model for bearing fault detection. The results indicate that the suggested model cans adaptability to different loads and reported an accuracy rate of nearly 99%. Hoang et al. [176] submitted a systematic survey on DL-based bearing fault diagnosis. They provided a comprehensive description of three types of standard DL algorithms: Auto encoder, RBM, and CNN, for the bearing fault detection. Finally, Jiao et al. [177] presented a work to survey fault diagnosis methods based on CNN more broadly. In the mean time they have been provided useful guidance for individuals who want to comprehend the development of CNN technologies for equipment fault prediction to implement predictive maintenance program.

To drive further along this line and rely on established research about CNN model in fault detection and the advantages presented in the previous literature reviews by following intelligent prediction flow, explicitly, from data to model and evaluation. Likewise, it is noteworthy that the shift of learning technologies using CNN as a backbone has begun to gain growing interest as more realistic diagnostic problems can be addressed. With these points in mind, this thesis intends to pull out a sequence of local features from the original data based on CNN to build a fault detection model more effectively.

5.4 The Accelerated Data Set Description

In the industry, the art of predicting faults in rotary machines by vibration monitoring is frequently used. Vibration monitoring detects approximately 80% of common rotating equipment problems related to misalignment and imbalance [37]. If a rotary component has a defect, the vibration signal levels can change, and these measurements could indicate the severity of the damages. At specific vibration frequencies, vibrations develop by the defects take place. In other words, the components' characteristics are changing by their operation, assembly, and wear. This may probably be why the most common technique for fault detection and classification in rotary machinery is vibration condition monitoring. Therefore, it is often a significant issue to implement DL approaches for maintenance decision-making with greater precision in vibration signals. Actually, vibration patterns are the most often used to infer rotary machine working conditions. They have a wealth of information and are simple to measure with low-cost, off-the-shelf sensors.

During the DL models experiment, the same pump type used for applying ML algorithms (vertical and single-stage forwarding TORISHIMA pump) is utilized to predict the real working condition of roller bearings. Four model (CMSS 2100) accelerometers are installed in a bearing house. The technical specifications of these accelerometers are: sensitivity 100 mV/g, sensitivity precision; +/- 5% at 25C° and acceleration range; 0-80 peak, were mounted in two directions (x & y), to measure the vibration signal from Driving End (DE) & Non-Driving End (NDE) Bearings. The vibration (acceleration) signals stream-in at an interval of one second, giving a total of 7,776,000 data points, which corresponds to three months. A description of the original data sets as a time series of accelerations shown in (Figure 5.2).

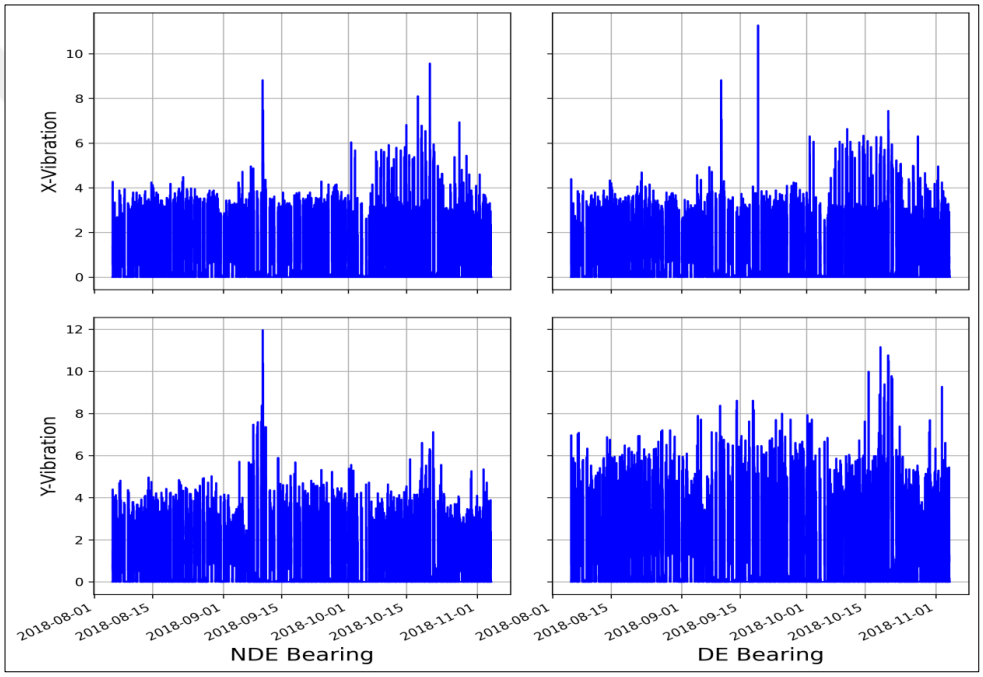


Figure 5.2 : Features of the original data sets.

After the data acquisition, all the collected data will be stored in the data warehouse for diagnosis and prediction. The data pre-processing method is applied to the acquired raw data before the subsequent step to improve the feature information to obtain better fault detection results. Generally, data preprocessing major functions include data cleaning, data de-nosing, data normalization, data integration, data reduction, and data transformation [89]. In our model, the data pre-processing step comprises data transformation (normalization), data cleaning (missing data) to be processed efficiently by the DL-based prediction model. The pre-processing data approach manages and analyses the collected data for a better understanding and interpretation.

Therefore, the data pre-processing approach is essential for extracting features and achieving high fault prediction model accuracy.

5.5 Programming Language

Artificial intelligence approaches vary from conventional software approaches. The variance lies in the technology infrastructure, the skills needed for applying AI models, and the importance of in-depth analysis. Coelho and Richert [178] presented ML with Python as a perfect team. They introduced how an ML algorithm works to learn real data classification, how to apply python programs to train ML models proposed different application examples. They have illustrated the ML algorithm as an iterative process, making Python the right language for ML.

Python enables designers to be more efficient and secure in the applications they develop. Benefits and capabilities that make the Python program the perfect fit for proposed prediction models in this thesis based on ML and DL approaches include:

- Simplicity, stability, and coherence
- Access to outstanding AI libraries and frameworks
- Versatility and flexibility
- Independency of platform
- Tools availability
- A diverse culture

These contribute to the language's overall success, so we used Python in our AI prediction models.

Besides, Python is considered an open-source language commonly used in the industry or for academic purposes. Python has many useful modules for simpler operations, such as NumPy, Pandas, Sklearn, and SciPy. It also has several deep learning frameworks that run on Python, Tensorflow, Keras, PaddlePaddle.

5.6 Methodology

The proposed data-driven models based on integrating offline (history data) and online data are used for prediction model training. It makes maintenance decisions based on

the online measurements collected from accelerometers fixed on the machines to estimate the process state. When the operating condition of machines is changed, in real-time the model parameters can be changed also. The input data is first processed by time-domain statistical analysis, which diminishes the input data before feeding to the prediction models to decrease the model's required training time and prediction time. Four DL-based models are implemented for the classification of the operating conditions. Then, a Performance comparison is performed for the tested DL models to select the one with the most accurate bearings faults prediction. Finally, the DL classification results for prediction models are utilized to guide decision-makers when planning the PdM activities. The structure of the presented method is summarized as a flowchart shown in the (Figure 5.3).

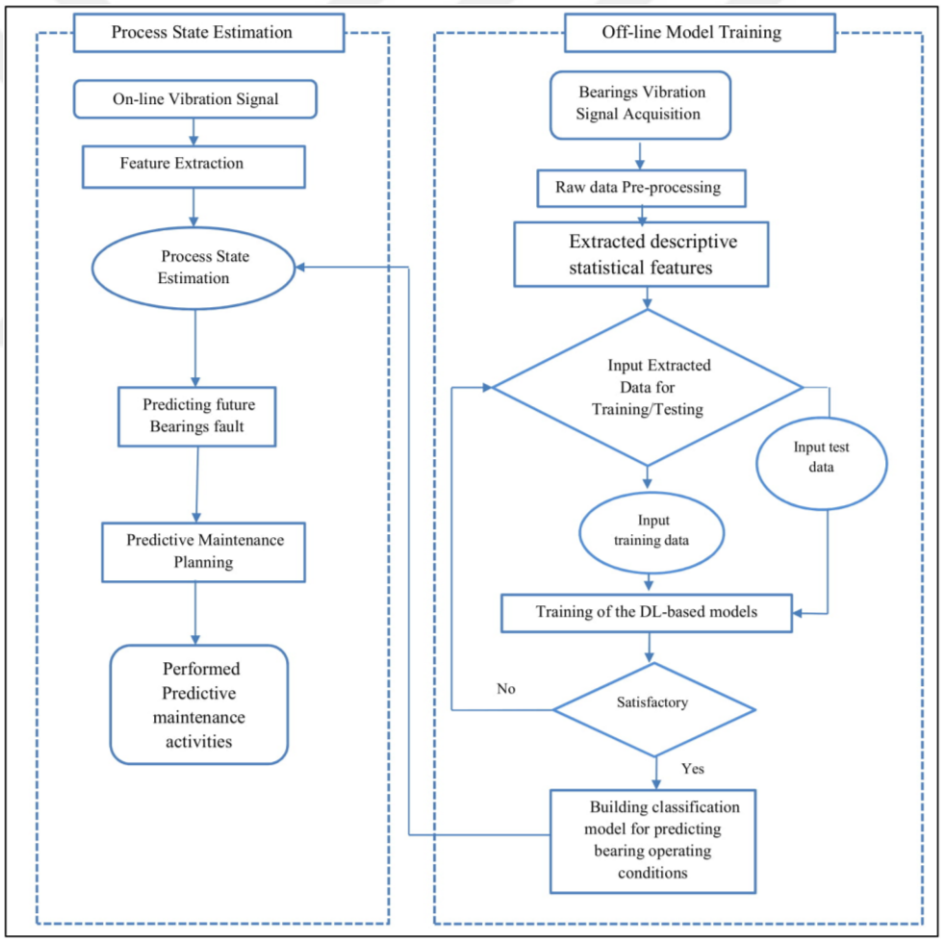


Figure 5.3 : The flow chart of the proposed DL-based method for roller bearing fault prediction model.

5.6.1 Offline DL models Training

As discussed previously, many analysis methods such as FFT, STFT, EMD, etc., are processed on the raw vibration data. These techniques have been commonly applied

to extract patterns from vibration signals in either the time domain or frequency domain, which can subsequently be effective for fault prediction [172]. Generally, statistical parameters provide good indications for extracting the condition information and pattern recognition. Thus, several statistical characteristics are extracted from rolling bearing signals in this study that are used to train the DL classification model for fault identification. Eleven statistical features are constructed for each sensor signal from the initial dataset of the bearing component and used to train the proposed DL models. These features are: mean, rang, skewness, kurtosis, maximum and minimum values, SD, quartiles (95%, 90%, 80%), and RMS. Concerning this explanation, and based on the available data, 44 statistical features from the four attributes are constructed to feed DL algorithms to classify the operation bearings condition. Selected attributes are listed in the (Table 5.1). In other words, these statistical features obtained from the vibration signals serve as input parameters of the DL-based models for bearings fault detection. Seventy percent of the samples set are used to train the prediction model, and thirty percent are used for testing. Before that, sliding windows are used to split the time series into a small time window to train the prediction models at each window and classify each signal point within a given class. In the DL approach, a four-hour window is selected, and the sliding length is chosen as one hour.

Table 5.1 : Selected attributes for DL prediction model.

Type of sensor	Attribute
Accelerometer (vibration signal reading)	NDE Bearing (x-axis) NDE Bearing (y-axis) DE Bearing (x-axis) DE Bearing (y-axis)

The binary classification is being built with the DL-based prediction model, where the class labels can only have two possible values: 0 or 1. Each class represents a specific operating condition, considering the pumping system design's technical information and specifications. As mentioned in the previous chapter, binary classification is commonly used for PdM, distinguishing between only two states. Furthermore, the binary classification can tune hyper-parameters, and their aim is not just discriminating between two classes— the classification of the operating conditions is specified in the table (Table 5.2).

Table 5.2 : Binary classification of the bearing operating condition for the DL model.

Description	Range	Critical Limit	Model Classification
Roller bearing DE & NDE Vibration	0-10 mm/sec	6 mm/sec	0: Normal operating condition <6 mm/sec 1: Critical operating condition ≥ 6 mm/sec

Thus, the DL algorithms applied to vibration signals have mainly focused on finding the bearing's normal and critical operation conditions to identify the most likely root causes for the pump's failure.

To achieve the above-mentioned structure, this thesis presents four-fault classification DL methods. First, a Deep Feedforward Networks (DFN) is built and the network is trained using the statistical parameters to classify the bearings operating condition. The second model, a standard LSTM, has been applied for developing a detection model using the same parameters. It aims to improve the recognition of the operating conditions and increase the prediction accuracy. The LSTM is adopted as a prediction model due to its ability to remember the observation results of long-term series intervals and many advantages. Thus, It has been usually used in the field of PdM [179] as discussed in section (5.2).

In the third proposed model, the advantage of both CNN and LSTM is taken by combining them into one structure for enhancing the fault prediction performance. The CNN is introduced to extract the most significant features from the input statistical measurements bypassing through a few convolutional layers and pooling layers. In addition, CNN has an outstanding ability to reduce frequency variation as well as its ability to compact the length of the series. This leads to improving the prediction model capability to capture temporal information that is highly beneficial for bearings condition identification. Afterwards, LSTM follows to attain a good feature representation of the input signal and model training. More importantly, the LSTM is considered as an effective technology to solve the problem of long sequence dependency and to improve of the fault identification accuracy. Consequently, this integration can reveal several hidden structures to improve the model generalization capabilities.

Finally, the gradient Boosted (GB) algorithm is also applied to the new data sample to demonstrate whether it outperform other DL models in terms of accuracy. It was considered the most effective prediction model in the previous case study. It is also

noteworthy that these models are different from the ML-based models presented in chapter 4, in terms of there is no dimension reduction methods are used, and the DL-based models are applied to the vibration signal statistical measurements so that network can learn features on itself. The proposed fault prediction models based on DL architecture are shown in (Figure 5.4).

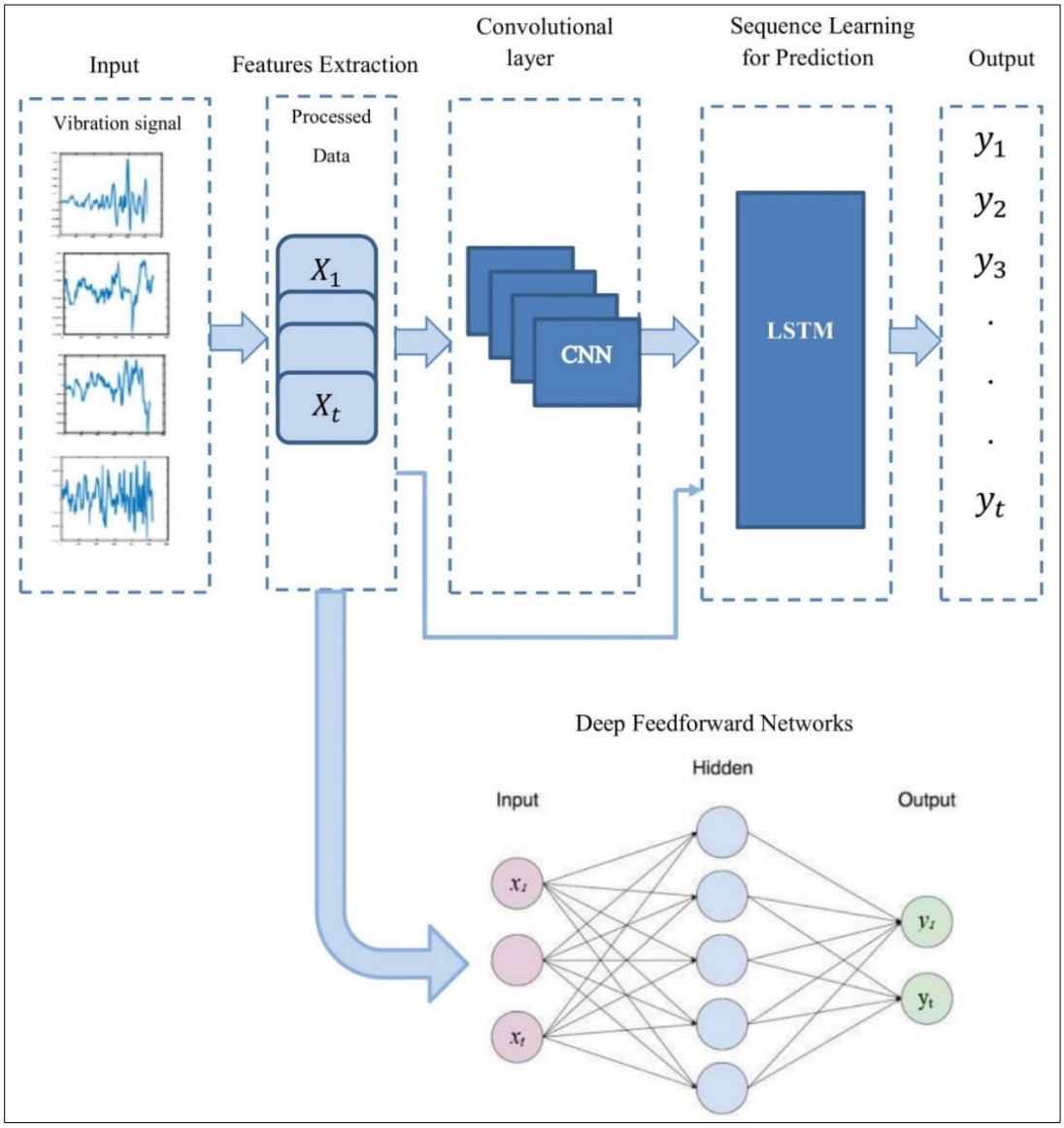


Figure 5.4 : The architecture of the DL-based Models.

5.6.2 Building classification algorithm

As mentioned previously, the DFN, GB, LSTM, CNN-LSTM models are used to determine the best DL technique to detect bearings operating conditions. This section introduces the key parameters that are used to implement the above-mentioned models. It is noteworthy to mention that a comprehensive comparison between DL models have

been performed utilizing the same sample dataset to identify the distinctions between the presented models in the state of prediction performance.

There are three layers in the DFN: input, hidden, and output. The normalized features extracted from the recorded vibration signals are represented by nodes in the input layer. The hidden layers consist of three layers with 64, 32, and 16 nodes. The target values of two output nodes can have only binary levels representing "normal" (0) and "critical" (1) bearings, which is the number of classes. The sigmoidal activation function is used in the output layers to preserve the outputs close to 0 and 1; while "relu" activation is used in the hidden layers to allow signals to travel one way only, from input to output. The model is trained iteratively using 100 epochs for the training data set to maximize the mean's accuracy function to the corresponding target values. Interestingly, literature show that each condition's output fluctuations are minimized with the increasing number of hidden layer nodes [180], but they reported that no mathematical procedure was proven to find the best number of the hidden layer nodes. Nevertheless, in the current case study changing the number of hidden layer nodes did not affect the output's accuracy. Therefore, there is not much need to consider the number of hidden layer nodes.

The learning rate is taken as 0.01 and 100 models are built in a GB Algorithm. 100 models did not improve the results of GB algorithm. Among the tested learning rates (0.01 to 0.5), a learning rate of 0.01 gave the best accuracy results for GB.

LSTM is built by supplanting each hidden neural with a memory cell and adjusting the relevant parameters through repeated tests to prevent the vanishing gradient problem. The network consists of three LSTM layers, a dense layer, input layer, and output layer. The learning rate is started to 0.001, with a dropout of 0.2 to avoid overfitting. We are setting the maximum number of training epochs to 100. The ultimate model structure and parameters are shown in (Table 5.3).

Table 5.3 : Key structural parameters of LSTM model.

Layer Type	Parameter for Layer	Other parameters
Input Layer	Feature Size= 4x11	Epoch = 100 Classifier =’Sgimoid’ Activation =’relu’ Optimizer = ’AdamOptimizer’
LSTM layer	(32 nodes) with Lag = 10	
Dropout Layer	Rate = 0.2	
Dense Layer	(100 nodes)	
Output Layer	Output channel=1	

In the structure of CNN-LSTM, the learning performance is modified by adjusting the parameters of the layers composing the network. The proposed model entails of multiple layers such as a convolutional layer, pooling layer, LSTM layer, and dense layer. Each layer can adjust the number of filters, the kernel size, and the number of strides. Designing of the CNN-LSTM model parameters are presented in (Table 5.4), and the Adam optimizer is chosen to minimize the loss function[181].

Table 5.4 : Key Parameters for the CNN-LSTM model.

Layer Type	Parameter for Layer	Other parameters
Input Layer	Feature Size= 4*11	
Convolutional Layer	(64 filters) with filter size= 3	
Pooling Layer	Pooling length = 2 Stride= 1	Epoch = 50 Dropout =0.2 Activation='relu' Classifier ='Sgimoid' Optimizer ='AdamOptimizer'
LSTM Layer	(32 nodes) with lag = 10	
Dense Layer	(100 nodes)	
Output Layer	Output channel =1	

To train the DL classification model, samples of input and output pairs must be ordered according to features that define the machine condition. A standard PdM model typically utilizes every feature produced from one data sample as an input and the relevant components operating condition as an output. To solve this issue and enhance the information contained in an input data sample, delay functions (lag) is used to make a prediction for the next time step (t+1) given the value at the previous time (t) as previously assumed in the ML model.

For each type of roller bearings used in the DL models case study, eleven features are removed from the original vibration signals, the details of the features mentioned in the previous section (5.6.1). The conditions are labeled into two states: normal and critical, the time series window size is chosen to be four. The sample data are randomly separated into two sets: training set (70%) and testing set (30%); this structure used to train the DL-based models, and then is used to verify the prediction precision. The results of training and testing of the presented DL models are shown in (Figure 5.5). The model also predicts the fault event identified by bearing operating conditions for the nine previous hours, Lag 1 to Lag 9.

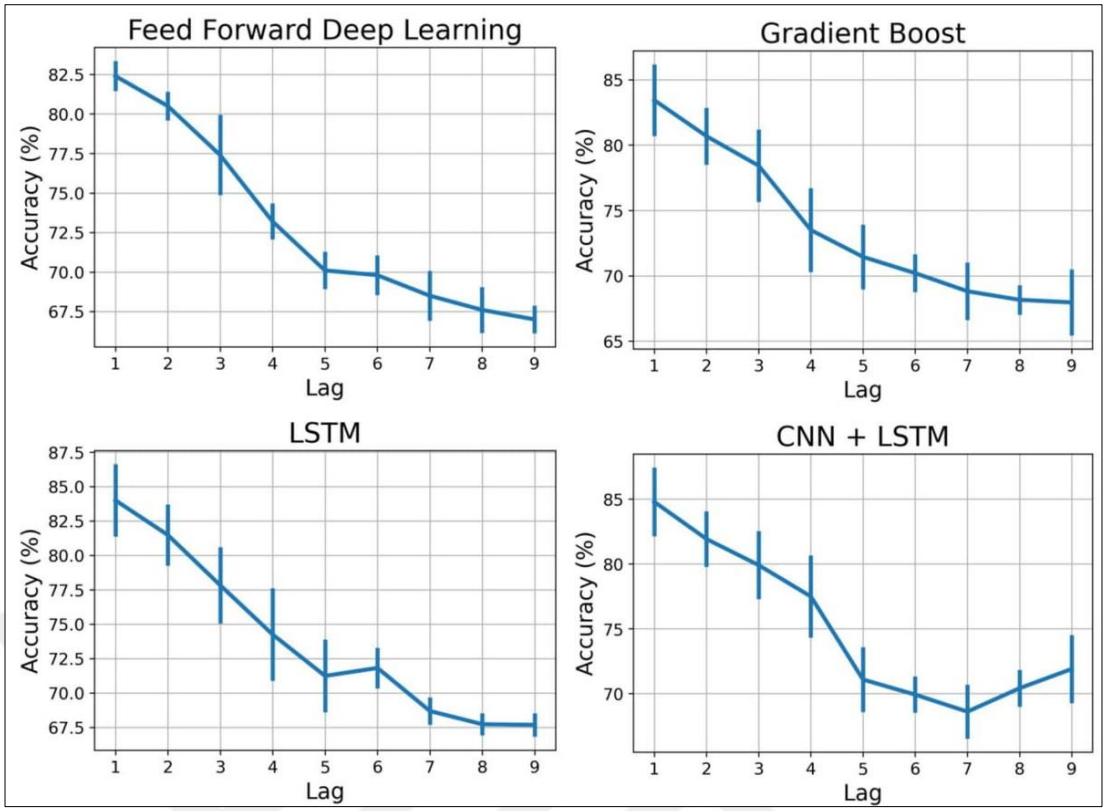


Figure 5.5 : Comparison of DL algorithms performance with respect to their Accuracy-Lag.

As the previous figure shows, the CNN-LSTM achieves higher accuracy than other proposed models which reach about nearly 85% mean accuracy in Lag 1, associated with the correct detection of critical bearing conditions before 1 hour. While, LSTM, GB, and DFN respectively resulted in 84%, 83.4%, and 82.4% mean accuracy giving an initial impression that DFN provides the worst accuracy compared with the others model. A more detailed analysis to the model's performance is presented in the next section.

In general, we can see that the prediction accuracy for all models is decreasing significantly with increasing the lag number from one to nine reaching minimum prediction accuracy. This is a logical consequence since many unexpected circumstances might appear when the prediction takes place earlier. (Table 5.5) summarizes the mean and SD for all DL models for nine lags.

Table 5.5 : The mean and standard deviation for all DL models for nine lags.

DFN algorithm	mean	SD	GB algorithm	mean	SD
Lag1	82.4	0.950	Lag1	83.4	2.734
Lag2	80.5	0.913	Lag2	80.7	2.174
Lag3	77.4	2.543	Lag3	78.4	2.760
Lag4	73.2	1.148	Lag4	73.5	3.218
Lag5	70.1	1.182	Lag5	71.4	2.478
Lag6	69.8	1.248	Lag6	70.2	1.440
Lag7	68.5	1.579	Lag7	68.8	2.201
Lag8	67.6	1.450	Lag8	68.2	1.131
Lag9	67	0.879	Lag9	68.0	2.527
LSTM algorithm	**mean**	**SD**	**CNN-LSTM algorithm**	**mean**	**SD**
Lag1	84.0	2.627	Lag1	84.8	2.644
Lag2	81.5	2.218	Lag2	81.9	2.142
Lag3	77.8	2.786	Lag3	79.9	2.628
Lag4	74.2	3.374	Lag4	77.5	3.176
Lag5	71.2	2.645	Lag5	71.1	2.494
Lag6	71.8	1.481	Lag6	69.9	1.394
Lag7	68.7	1.008	Lag7	68.6	2.077
Lag8	67.7	0.803	Lag8	70.4	1.418
Lag9	67.7	0.855	Lag9	71.9	2.616

5.7 Performance Measures for DL Models

For the performance evaluation in the DL-based models, the same performance measurements (accuracy, precision, F-score, recall, and AUC) that are tested in ML-based algorithms are utilized to evaluate the power of models prediction and compare it with the previously reported performance results of the proposed ML models.

When compared to the previous ML models, the experimental results showed that the tested DL models are ineffective. (Table 5.6) provides a list of the model evaluation results tested on the selected dataset.

Table 5.6 : Result of DL prediction models.

Test Type	DFN	GB	LSTM	CNN-LSTM
Accuracy	82.4	83.4	84	84.8
F_1-score	78.44	79.17	81.28	82.44
Precision	80.6	80.8	84.5	86.2
Recall	76.4	77.6	78.3	79

The evaluation of the effects of the integrated CNN-LSTM architecture is implemented to show that if the CNN can efficiently modify the model's detection capability. Unfortunately, this integration shows a negligible advance in the prediction performance compared with other presented DL models. It performs slightly better than three other compared models widely used as predicting methods, where it reaches an accuracy of 84.8%; precision of 86.2 % has an F-score and Recall of 82.4% and 79% respectively. In contrast, the DFN model, shows the lowest accuracy and precision rate which are 82.4%, 80.6% respectively, as well as the lowest performance of 78.4% F-score and 76.4% Recall.

Finally, the area under the curve (AUC) is also could be computed from model, which is a measure of a classifier model's overall performance based on the compromises between TP and TN rates over all possible risk thresholds between 0% and 100%. A perfect ROC curve would reach the top left corner, so the greater area under the AUC the better the classifier. As mentioned previously, a ROC over 0.70 is considered good, and a ROC over 0.80 is very good. Besides, ROC curves are useful for comparing different classifiers, since they take into account all possible thresholds.

It should be noted that the proposed DL-based models achieved success rates of between 82 and 85 percent in the classification of vibration data, according to the obtained results, which are considered unsatisfactory results for the performed correct preventive maintenance tasks.

6. CONCLUSION

The work presented in this thesis is generally overviewed in this chapter and some suggestions for future research are presented.

6.1 Conclusions

Increased complexity in industrial equipment and production processes has resulted from smart manufacturing evolution, making it difficult or almost impossible to recognize and detect critical conditions in a proper time using traditional methods. Therefore, it is essential to establish a framework for predictive maintenance and equipment-condition monitoring employing artificial intelligence concepts to improve their processes and increase machinery performance. It has been observed that prediction models aid the production managements in providing automated tools for scheduling PdM, which is flexible and easy to use. Today's intelligent predictive maintenance in the modern industry has the potential to be totally capable of applying maintenance with less effort and with reduced downtimes. Therefore, it is possible to have a successful PdM program from a maintenance time reduction viewpoint as it is essential when the proper periodic inspection and corrective actions are taken, thus the overall equipment performance can be maximized, as presented in this dissertation.

Chapter two provides helpful guidelines for steps to construct intelligent predictive maintenance on the rotary machines. The dissertation contains the entire condition monitoring process and its data analysis applications, including sensor and data preprocessing, features extracting and selecting, fault classification, fault identification and detection, and performance indicator evaluation. It also introduces detailed explanations for the pumping system that was used as a case study to develop the proposed prediction model.

The basic fundamental operating mechanisms for roller bearings are presented in chapter 3. The state of rolling element bearings can be easily recognized using vibration and temperature monitoring. Vibration signals reveal essential information about fault progress, while temperature data provides information on bearing working

conditions to understand all phenomena affecting bearings. Therefore, the implementation of our proposed classification algorithm focused on vibration signals and temperature readings on the state of the data acquisition process. The dissertation also presents many techniques for vibration analysis used to identify rolling bearing faults.

A systematic investigation about machine learning and deep learning approaches applied for fault identification and prediction are presented. Our models are based on estimations taken one to nine hours (Lag) in advance, giving operator's sufficient time to prepare for inspections. This helps in taking the correct maintenance action on bearing components (e.g., checking the lubricant, cleaning the bearing housing, or preventing overheating), which will, in turn increase bearings durability.

The prediction model has been implemented on a real industrial company which provided us with the necessary data collected using an online sensor measurement system and internet of things technology. The proposed model has been achieved by training several artificial intelligent algorithms on a python coded program.

Chapter 4 demonstrates a case study of applying machine learning approaches to predict the roller bearings' operating conditions. The computational analysis showed that the four machine learning approaches resulted in acceptable fault detection power. However, GB and RF gave the best performance in terms of accuracy among the tested algorithms: 92% and 87.5%, respectively. Using machine learning with the recorded maintenance data demonstrated that predictive maintenance could be done and provides good and reliable criteria for the maintenance planned interventions. The model aids operators to quickly visualize and monitor the pumping system. The proposed prediction models of this dissertation can be applied easily and flexibly in all industrial processes.

A systematic investigation about machine learning and deep learning approaches applied for fault classification and prediction are presented. Our models are based on estimations taken one to nine hours (Lag) in advance, giving operator's sufficient time to prepare for inspections. This helps in taking the correct maintenance action on bearing components (e.g., checking the lubricant, cleaning the bearing housing, or preventing overheating), which will, in turn increase bearings durability.

The prediction model has been implemented on a real industrial company which provided us with the necessary data collected using an online sensor measurement system and internet of things technology. The proposed model has been achieved by training several artificial intelligent algorithms on a python coded program.

Chapter 4 demonstrates a case study of applying machine learning approaches to predict the roller bearings' operating conditions. The computational analysis showed that the four machine learning approaches resulted in acceptable fault detection power. However, GB and RF gave the best performance in terms of accuracy among the tested algorithms: 92% and 87.5%, respectively. Using machine learning with documented maintenance data revealed that predictive maintenance could be done and offers good and reliable guidelines for the maintenance planned interventions. The model aids operators to quickly visualize and monitor the pumping system. The proposed prediction models of this dissertation can be applied easily and flexibly in all industrial processes.

Chapter 5 introduces an experiment of deep learning algorithms for fault classification in the roller bearings depending on vibration signals collected from the sensor-mediated components. The experimental results exposed that the novel model is, to some extent, more efficient for fault prediction than traditional deep learning algorithms and makes mechanical fault prediction move toward real artificial intelligence. An integrated CNN-LSTM model is also proposed and compared with a standard LSTM, GB, and DFN models to evaluate the change in fault detection performance on the bearing component. Additionally, CNN-LSTM integrated model gave the best accuracy among the tested algorithms nearly to 85 %. Thus, this research is intended on the model evaluation to assess the prediction models' performance, which is the key to achieving the correct predictive maintenance planning.

Furthermore, while most of the related literature depends on their maintenance action only based on machine learning results (0, 1 binary classification), the current model is distinguished by its ability to show the probability of critical conditions through the use of utility theory. This helps to avoid false-positive alarms and thus reduces unnecessary maintenance costs. Therefore, this research significantly contributes to achieving a more trustworthy maintenance management system for different Industrial applications such as power plants, oil companies, water treatment companies, the aerospace industry, manufacturing facilities, and the like.

Finally, the key results show that machine learning algorithms provide more efficient solutions for most fault detection problems than deep learning algorithms. The experimental findings suggest that the proposed machine learning models are a competitive alternative to the conventional deep learning models.

Via experiments that were performed using various methods of artificial intelligence, and by their application on the two data samples, it is concluded that to obtain high accuracy prediction model need: (1) large data sets to increase the size of training data and contain enough label information, i.e., normal and critical data for the presented case study. (2) better features tuning when applying AI methods. (3) finding the optimum value for each parameter used for training model, and (4) proper choice of AI algorithm that best fit with the data set.

6.2 Future Work

1. Alarms system based on technical information rules could be created to detect a particular machine fault, specifically, the critical failure that could cause the malfunction of the pump. An alarming system can be integrated with the proposed prediction model and set for individual machines or a group of machines. These alarms can be displayed in many different formats like descriptive pictures, visual Alarming, specific sounds, or text messages appearing on the user window. In the bearing critical condition case, the system will automatically alarm triggers to alert the operator and maintainers to act the proper inspection and maintenance jobs.
2. We may depend on multiple fault classification instead of binary classification output of the implemented fault detection models of rotating machinery; for example, there are various fault classes based on fault types and severity.
3. We could also obtain high prediction accuracy by training several classification models with various operational conditions (normal, critical, trip or shutdown) to provide additional performance tradeoffs in terms of frequency of unexpected breaks and unexploited lifetime, and then use this information in an operating cost-based maintenance decision system to reduce expected costs.
4. The ability to improve the prediction model in the future can be investigated by adding new label data such as pressure, flow rate, load, etc., and extend the

period of the collected sample data to make the model more effective and comprehensive.

5. RNN and LSTM may be considered as an option for the fault prediction DL model. More data pre-processing methods could be integrated and compared.
6. The future models shall not only be implemented on bearings but also on other pumping elements such as gearbox, seals, motors, etc.

REFERENCES

[1] **Cline B., Niculescu RS., Huffman D., Deckel B.** (2017) Predictive maintenance applications for machine learning. *Annual Reliability and Maintainability Symposium (RAMS). IEEE*, pp 1–7.

[2] **Butler J., Smalley C.,** (2002) *An Introduction to Predictive Maintenance*.(Butterworth-Heinemann), Elsevier.

[3] **British Standards Institution.** (2010). *BSI Standards Publication Maintenance — Maintenance terminology. 36.* https://doi.org/10.1007/s00168-003-0173-6

[4] **Wang JL& Ben.** (1999). *Computer-aided Maintenance Methodologies and Practices*. (Springer Science+Business Media, B.V.), UK.

[5] **Sethiya S.** (2006). Condition-based maintenance (CBM). Secy.to CME/WCR/JBP.

[6] **Kang K., Subramaniam V.** (2018). Integrated control policy of production and preventive maintenance for a deteriorating manufacturing system. *Comput Ind Eng 118*:266–277.

[7] **Martin KF.** (1994). A review by discussion of condition monitoring and fault diagnosis in machine tools. Int J Mach Tools Manuf *34*:527–551.

[8] **Henriquez P., Alonso JB., Ferrer MA., Travieso CM.** (2014) Review of automatic fault diagnosis systems using audio and vibration signals. *IEEE Trans Syst Man, Cybern Syst 44*:642–652.

[9] **Peng K.** (2012). Equipment management in the post-maintenance era : a new alternative to total productive maintenance (TPM). pp 132–136

[10] **Susto GA., Schirru A., Pampuri S., et al.** (2015). Machine Learning for Predictive Maintenance: A Multiple Classifier Approach. *IEEE Trans Ind Informatics 11*:812–820.

[11] **De Faria H., Costa JGS., Olivas JLM.** (2015). A review of monitoring methods for predictive maintenance of electric power transformers based on dissolved gas analysis. *Renew Sustain Energy Rev 46*:201–209.

[12] **Carvalho TP., Soares FAAMN., Vita R., et al.** (2019). A systematic literature review of machine learning methods applied to predictive maintenance. *Comput Ind Eng 137*:106024.

[13] **NASA.** (2002). Reliability Centered Maintenance. Engineering Maintenance. *CRC Press*, p 472.

[14] **Agustiady TK., Cudney EA.** (2018) Total productive maintenance. *Total Qual Manag Bus Excell*, pp 1–8.

[15] **Lesage A., Dehombreux P.** (2012). Maintenance & quality control: A first methodological approach for maintenance policy optimization. *Proceedings of the 14th IFAC Symposium on Information Control Problems in Manufacturing Bucharest, 14*:1041-1046.

[16] **Hashemian HM., Bean WC.** (2011). State-of-the-art predictive maintenance techniques. *IEEE Trans Instrum Meas, 60*:3480–3492.

[17] **Rao singiresu S.** (1986)**.** Mechanical Vibration.*Vibration Measurement and Applications* (3rd ed., pp 870–938).

[18] **Motaghare O., Pillai AS., Ramachandran KI.** (2018). Predictive Maintenance Architecture. *The International Conference on Computational Intelligence and Computing Research (ICCIC). IEEE*, pp 1–4.

[19] **ABS (American Bureau of Shipping).** (2016). *Guidance Notes on Equipment Condition Monitoring Techniques,* Houston, USA.

[20] **Al-Najjar B.** (2007). The lack of maintenance and not maintenance which costs: A model to describe and quantify the impact of vibration-based maintenance on company's business. *Int J Prod Econ, 107*:260–273.

[21] **Xu G., Liu M., Wang J., et al.** (2019). Data-driven fault diagnostics and prognostics for predictive maintenance: A brief overview. *IEEE Int Conf Autom Sci Eng*., pp103–108.

[22] **Wang KS., Li Z., Braaten J., Yu Q.** (2015). Interpretation and compensation of backlash error data in machine centers for intelligent predictive maintenance using ANNs. *Adv Manuf, 3*:97–104.

[23] **Wang Y., Deng C., Wu J., et al.** (2014). A corrective maintenance scheme for engineering equipment. *Eng Fail Anal, 36*:269–283.

[24] **Zhao Z., Wang F., Jia M., Wang S.** (2010). Predictive maintenance policy based on process data. *Chemom Intell Lab Syst*em, *103*:137–143.

[25] **Wuest T., Weimer D., Irgens C., Thoben KD.** (2016) Machine learning in manufacturing: Advantages, challenges, and applications. *Prod Manuf Res, 4*:23–45.

[26] **Jardine AKS., Lin D., Banjevic D.** (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Process, 20*:1483–1510.

[27] **Zhou W., Habetler TG., Harley RG.** (2007). Bearing condition monitoring methods for electric machines: A general review. *IEEE Int Symp Diagnostics Electr Mach Power Electron Drives.*,pp 3–6.

[28] **Wang K., Wang Y.** (2018). How AI Affects the Future Predictive Maintenance: A Primer of Deep Learning. In: *Lecture Notes in Electrical Engineering*. pp 1–9

[29] **Bana E Costa CA., Carnero MC., Oliveira MD.** (2012). A multi-criteria model for auditing a Predictive Maintenance Programme. *Eur J Oper Res. 217*:381–393.

[30] **Hashemian HM.** (2011). Wireless sensors for predictive maintenance of rotating equipment in research reactors. *Ann Nucl Energy, 38*:665–680.

[31] **García Márquez FP., Tobias AM., Pinar Pérez JM., Papaelias M.** (2012). Condition monitoring of wind turbines: *Techniques and methods. Renew Energy, 46*:169–178.

[32] **Selcuk S.** (2017). Predictive maintenance, its implementation and latest trends. *Proc Inst Mech Eng Part B J Eng Manuf, 231*:1670–1679.

[33] **Salehi M.** (2018). *Bayesian-Based Predictive Analytics for Manufacturing Performance Metrics in the Era of Industry 4.0*. (Master Thesis). Karlsruhe Institute of Technology (KIT).

[34] **Lessmeier C., Enge-Rosenblatt O., Bayer C., Zimmer D.** (2014). Data Acquisition and Signal Analysis from Measured Motor Currents for Defect Detection in Electromechanical Drive Systems. *Eur Conf Progn Heal Manag Soc*, PP 1–10.

[35] **Zhang C., Hong GS,. Xu H., et al.** (2017). A data-driven prognostics framework for tool remaining useful life estimation in tool condition monitoring. *IEEE Int Conf Emerg Technol Fact Autom ETFA*, PP 1–8.

[36] **Voronov S.** (2020). *Machine Learning Models for Predictive Maintenance.* (Doctoral thesis)**.** Linköping University, Faculty of Science & Engineering.

[37] **Ede Kn., Ogbonnaya Ea., lilly Mt., et al.** (2010). Vibration Monitoring of Rotating Systems. *Engineering, 2*:46–54.

[38] **Xu Y., Chen M.** (2018). An Internet of Things based framework to enhance just-in-time manufacturing. *Proc Inst Mech Eng Part B J Eng Manuf., 232*:2353–2363.

[39] **Amruthnath N., Gupta T.** (2018). A research study on unsupervised machine learning algorithms for early fault detection in predictive maintenance. *5th International Conference on Industrial Engineering and Applications (ICIEA). IEEE*, pp 355–361

[40] **Lærum KH.** (2018). *A study of Machine Learning for Predictive Maintenance*. (Master Thesis). Norwegian University of Science and Technology. NTNU Faculty of Engineering.

[41] **Dong L., Mingyue R., Guoying M.** (2017). Application of Internet of Things Technology on Predictive Maintenance System of Coal Equipment. *Procedia Eng, 174*:885–889.

[42] **Meyer VFA.** (2019). *Challenges and Reliability of Predictive Maintenance*. (Master Thesis). Faculty of Communication and Environment.

[43] **Hameed Z., Hong YS., Cho YM., et al.** (2009). Condition monitoring and fault detection of wind turbines and related algorithms: A review. *Renew Sustain Energy Rev 13*:1–39.

[44] Url-1 <http://www.skf.com>, date retrieved 2/12/2020

[45] **Liao L., Köttig F.** (2014). Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction. *IEEE Trans Reliab., 63*:191–207.

[46] **Yao Q., Yang T., Liu Z., Zheng Z.** (2019). Remaining useful life estimation by empirical mode decomposition and ensemble deep convolution neural networks. *The IEEE International Conference on Prognostics and Health Management ICPHM*. DOI: 10.1109/ICPHM.2019.8819373.

[47] **Calixto E.** (2016). Reliability and Maintenance. In: Gas and Oil Reliability Engineering. *Elsevier*, pp 159–267.

[48] **Childs PRN.** (2019). *Mechanical Design Engineering Handbook.* (chapter 5-Journal bearings). Elsevier, pp 167–230.

[49] **Alsalaet J.** (2020). Detecting Rolling Elements Bearings Faults. College of Engineering- University of Basrah, p57. https://www.researchgate.net/publication/339123818_Detecting_Rolling_Elements_Bearings_Faults.

[50] **SKF Group** (2018). Rolling bearings. 1–1722. https://www. skf.com/go/17000.

[51] **Khonsari MM., Booser ER.** (2017). *Applied Tribology: Bearing Design and Lubrication.* (John Wiley & Sons, Ltd, Chichester),Curtin University. UK.

[52] **Gupta R. & Khurmi R.S.** (2005). *Machine Design*. A Textbook for the Students of B.E. / B.Tech., (Eurasia Publishing House LTD.), pp 1–1251.

[53] **Simmons CH., Phelps N., Maguire DE.** (2012). Manual of Engineering Drawing. Fourth Edition, *Bearings and Applied Technology*(Ch.35). *Elsevier*, pp 315–330. https://doi.org/10.1016/C2009-0-64285-9.

[54] **Cong F., Chen J., Dong G., Pecht M.** (2013). Vibration model of rolling element bearings in a rotor-bearing system for fault diagnosis. *J Sound Vibration, 332*:2081–2097.

[55] **Tran VT., Yang B-S.** (2012). An intelligent condition-based maintenance platform for rotating machinery. *Expert System Appl., 39*:2977–2988.

[56] **Carnero MC.** (2005). Selection of diagnostic techniques and instrumentation in a predictive maintenance program. A case study. *Decis Support Syst 38*:539–555.

[57] **Neale M.** (1996). *The Tribology handbook*. Second Edition, Elsevier, Butterworth-Heinemann.

[58] **Hassin OAA.** (2017). *Condition Monitoring of Journal Bearings For Predictive Maintenance Management Based on High Frequency Vibration Analysis*. (Doctoral thesis), University of Huddersfield.

[59] **Caesarendra W., Kosasih B., Tieu AK., et al.** (2016). Acoustic emission-based condition monitoring methods: Review and application for low speed slew bearing. *Mech Syst Signal Process, 72*:134–159.

[60] **Giannenas I., Triantafillou E., Stavrakakis S., et al.** (2012). A Basic Guide to Thermography. *Land Instruments International*, pp 26–32.

[61] **Fantidis JG., Potolias C., Bandekas D V.** (2011). Wind Turbine Blade Nondestructive Testing with a Transportable Radiography System. *Sci Technol Nucl Install*, pp1–6.

[62] **Zhen L., Zhengjia H., Yanyang Z., Xuefeng C.** (2008). Bearing condition monitoring based on shock pulse method and improved redundant lifting scheme. *Mathematics and Computers in Simulation*. *79*:318–338.

[63] **Vianna WOL., Yoneyama T.** (2018). Predictive Maintenance Optimization for Aircraft Redundant Systems Subjected to Multiple Wear Profiles. *IEEE Systems Journal, 12*:1170–1181.

[64] **Salomon CP., Ferreira C., Sant'Ana WC., et al.** (2019). A study of fault diagnosis based on electrical signature analysis for synchronous generators predictive maintenance in bulk electric systems. *Energies, 12*: 1506.

[65] **Tse PW., Peng YH., Yam R.** (2001). Wavelet Analysis and Envelope Detection For Rolling Element Bearing Fault Diagnosis—Their Effectiveness and Flexibilities. *J Vib Acoust 123*:303–310.

[66] **Pappachan B., Caesarendra W., Tjahjowidodo T., Wijaya T.** (2017) Frequency Domain Analysis of Sensor Data for Event Classification in Real-Time Robot Assisted Deburring. *Sensors 17*:1247.

[67] **Kuncan M., Kaplan K., Minaz MR., et al.** (2020) A novel feature extraction method for bearing fault classification with one dimensional ternary patterns. *ISA Trans 100*:346–357.

[68] **Saruhan H., Sandemir S., Çiçek A., Uygur I.** (2014). Vibration Analysis of Rolling Element Bearings Defects. *J Appl Res Technol, 12*:384–395.

[69] **Nabhan A., Ghazaly NM., Samy A., M.O M.** (2015). Bearing fault detection techniques - a review. *Turkish J Eng Sci Technol, vol 3*.

[70] **Ferrando Chacon JL., Kappatos V., Balachandran W., Gan T-H.** (2015). A novel approach for incipient defect detection in rolling bearings using acoustic emission technique. *Appl. Acoust., 89*:88–100.

[71] **Tandon N., Choudhury A.** (1999). Review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribol Int., 32*:469–480.

[72] **Poddar S., Tandon N.** (2019). Detection of particle contamination in journal bearing using acoustic emission and vibration monitoring techniques. *Tribol Int., 134*:154–164.

[73] **Raposo H., Farinha JT., Fonseca I., Galar D.** (2019) Predicting condition based on oil analysis – A case study. *Tribol Int., 135*:65–74.

[74] **Wurban L.** (2020). Rely on Swagelok Seal Support Systems to Reduce Pump Failures. http://machining.grundfos.com. Accessed 21 Mar 2021

[75] **Khalil Sheikh A., Younas M., Matar Al-Anazi D.** (2002). Weibull Analysis of Time Between Failures of Pumps Used in an Oil Refinery. *The 6th Saudi Eng Conf (KFUPM). 4*:475–491

[76] **Zarei J., Tajeddini MA., Karimi HR.** (2014). Vibration analysis for bearing fault detection and classification using an intelligent filter. *Mechatronics, 24*:151–157.

[77] **Orhan S., Aktürk N., Çelik V.** (2006). Vibration monitoring for defect diagnosis of rolling element bearings as a predictive maintenance tool: Comprehensive case studies. *NDT & E International, 39*:293–298.

[78] **Finley WR., Hodowanec MM., Holter WG.** (2000). An analytical approach to solving motor vibration problems. *IEEE Trans Ind Appl., 36*:1467–1480.

[79] **Patidar S., Soni P.** (2013). An Overview on Vibration Analysis Techniques for the Diagnosis of Rolling Element Bearing Faults. *Int J Eng Trends Technol., 4*:1804–1809

[80] **Cai J., Xiao Y.** (2017). Time-frequency analysis method of bearing fault diagnosis based on the generalized S transformation. *J Vibro engineering, 19*:4221–4230.

[81] **Tahir MM., Khan AQ., Iqbal N., et al.** (2017). Enhancing Fault Classification Accuracy of Ball Bearing Using Central Tendency Based Time Domain Features. *IEEE Access, 5*:72–83.

[82] **Kumar A., Ramkumar J., Verma NK., Dixit S.** (2015). Detection and classification for faults in drilling process using vibration analysis. *2014 Int Conf Progn Heal Manag PHM,* pp 3–8.

[83] **Chebil J., Hrairi M., Abushikhah N.** (2011). Signal Analysis of Vibration Measurements for Condition Monitoring of Bearings. *Aust J basic Appl Sci., 5*:70–78

[84] **Li Z., Zhao C., Lu Z., Liu F.** (2019). Thermal Performances Prediction Analysis of High Speed Feed Shaft Bearings Under Actual Working Condition. *IEEE Access, 7*:168011–168019.

[85] **Liulys K.** (2019) Machine Learning Application in Predictive Maintenance. *The Conference of Electrical, Electronic and Information Sciences, eStream - Proceedings. IEEE,* pp 1–4.

[86] **Haarman M., Mulders M., Vassiliadis C.** (2017). *Predictive Maintenance 4.0: Predict the unpredictable*. (Master Thesis). Norwegian University of Science and Technology.

[87] **Murty MN., Devi VS.** (2015). *Introduction to Pattern Recognition and Machine Learning*. Indian Institute of Science, India.

[88] **Baptista M., Sankararaman S., de Medeiros IP., et al.** (2018) Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Comput Ind Eng 115*:41–53.

[89] Bruce M., Scheuer O. (2012). Educational Data Mining. *Encyclopedia of the Sciences of Learning, Springer,* pp 1075–1079

[90] **Li Z., Wang Y., Wang KS.** (2017). Intelligent predictive maintenance for fault diagnosis and prognosis in machine centers: Industry 4.0 scenario. *Adv Manuf , 5*:377–387.

[91] **Ahadi E., Riahi M.** (2019). Intelligent Prognostics and Diagnostics of pumps using condition monitoring methods to improve reliability. *The 27th Annual International Conference of Iranian Society of Mechanical Engineers.* Tehran, Iran.

[92] **Cachada A., Moreira PM., Romero L., et al.** (2018). Maintenance 4.0: Intelligent and Predictive Maintenance System Architecture. *The 23rd International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE*, pp 139–146.

[93] **Yam RCM., Tse PW., Li L., Tu P.** (2001). Intelligent predictive decision support system for condition-based maintenance. *Int J Adv Manuf Technol., 17*:383–391.

[94] **Butte S., Prashanth AR., Patil S.** (2018). Machine Learning Based Predictive Maintenance Strategy: A Super Learning Approach with Deep Neural Networks. *Workshop on Microelectronics and Electron Devices (WMED). IEEE*, pp 1–5

[95] **Nam J., Jo N., Kim JS., Lee SW.** (2020). Development of a health monitoring and diagnosis framework for fused deposition modeling process based on a machine learning algorithm. *Proc Inst Mech Eng Part B J Eng Manuf, 234*:324–332.

[96] **Xayyasith S., Promwungkwa A., Ngamsanroaj K.** (2018). Application of Machine Learning for Predictive Maintenance Cooling System in Nam Ngum-1 Hydropower Plant. *The 16th International Conference on ICT and Knowledge Engineering (ICT&KE). IEEE*, pp 1–5.

[97] **Hu Q., He Z., Zhang Z., Zi Y.** (2007). Fault diagnosis of rotating machinery based on improved wavelet package transform and SVMs ensemble. *Mech Syst Signal Process 21*:688–705.

[98] **Yan W.** (2007). Application of Random Forest to Aircraft Engine Fault Diagnosis. *IMACS Multiconference on "Computational Engineering in Systems Applications"(CESA)*, China. October, pp 468–475.

[99] **Yiakopoulos CT., Gryllias KC., Antoniadis IA.** (2011) Rolling element bearing fault detection in industrial environments based on a K-means clustering approach. *Expert System Appl., 38*:2888–2911.

[100] **Allah Bukhsh Z., Saeed A., Stipanovic I., Doree AG.** (2019). Predictive maintenance using tree-based classification techniques: A case of railway switches. *Transp Res Part C Emerg Technol 101*:35–54.

[101] **Gutschi C., Furian N., Suschnigg J., et al.** (2019) Log-based predictive maintenance in discrete parts manufacturing. *Procedia CIRP, 79*:528–533.

[102] **Soualhi A., Clerc G., Razik H.** (2013). Detection and Diagnosis of Faults in Induction Motor Using an Improved Artificial Ant Clustering Technique. *IEEE Trans Ind Electron, 60*:4053–4062.

[103] **Kroll B., Schaffranek D., Schriegel S., Niggemann O.** (2014). System modeling based on machine learning for anomaly detection and predictive maintenance in industrial plants. In: *Proceedings of the IEEE Emerging Technology and Factory Automation (ETFA)*, pp 1–7.

[104] **Wei G., Zhao X., He S., He Z.** (2019). Reliability modeling with condition-based maintenance for binary-state deteriorating systems considering zoned shock effects. *Comput Ind Eng., 130*:282–297.

[105] **Paolanti M., Romeo L., Felicetti A., et al.** (2018). Machine Learning approach for Predictive Maintenance in Industry 4.0. *The 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pp 1–6.

[106] **Blum AL., Langley P.** (1997). Artificial Intelligence Selection of relevant features and examples in machine. *Artif Intell*, *97*:245–271

[107] **Lane MC., Xue B., Liu I., Zhang M.** (2014). Gaussian Based Particle Swarm Optimisation and Statistical Clustering for Feature Selection. *European Conference on Evolutionary Computation in Combinatorial Optimization*, pp 133–144.

[108] **Kwon D., Hodkiewicz MR., Fan J., et al.** (2016). IoT-Based Prognostics and Systems Health Management for Industrial Applications. *IEEE Access, 4:*3659–3670.

[109] **Melia T., Cooke A., Grayson S.** (2016). Machine Learning Techniques for Automatic Sensor Fault Detection in Airborne SHM Networks. *The 17th Aust Int Aerosp Congr AIAC*, Melb, pp 5–8.

[110] **Cai J., Luo J., Wang S., Yang S.** (2018). Feature selection in machine learning: A new perspective. *Neurocomputing, 300*:70–79.

[111] **Malhi A., Gao RX.** (2004). PCA-based feature selection scheme for machine defect classification. *IEEE Trans Instrum Meas, 53*:1517–1525.

[112] **Chen Z., Li W.** (2017). Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Trans Instrum Meas, 66*:1693–1702.

[113] **Caesarendra W., Tjahjowidodo T.** (2017). A Review of Feature Extraction Methods in Vibration-Based Condition Monitoring and Its Application for Degradation Trend Estimation of Low-Speed Slew Bearing. *Machines*. *5*:21.

[114] **Ahmed Medjahed S.** (2015). A Comparative Study of Feature Extraction Methods in Images Classification. *Int J Image, Graph Signal Process 7:*16–23.

[115] **Sheng Y., Rovnyak SM.** (2004). Decision tree-based methodology for high impedance fault detection. *IEEE Trans Power Deliv., 19*:533–536.

[116] **Kotsiantis SB., Zaharakis ID., Pintelas PE.** (2007). Supervised Machine Learning : A Review of Classification Techniques General Issues of Supervised Learning Algorithms. *Inform., 31*:249–268.

[117] **Gerdes M.** (2013). Decision trees and genetic algorithms for condition monitoring forecasting of aircraft air conditioning. *Expert Syst Appl., 40*:5021–5026.

[118] **Prytz R., Nowaczyk S., Rögnvaldsson T., Byttner S.** (2015). Predicting the need for vehicle compressor repairs using maintenance records and logged vehicle data. *Eng Appl Artif Intell., 41*:139–150.

[119] **Patel S V., Jokhakar VN.** (2016). A random forest based machine learning approach for mild steel defect diagnosis. *International Conference on Computational Intelligence and Computing Research (ICCIC). IEEE*, pp 1–8.

[120] **Kejela G., Esteves RM., Rong C.** (2014). Predictive Analytics of Sensor Data Using Distributed Machine Learning Techniques. In: *2014 IEEE 6th International Conference on Cloud Computing Technology and Science. IEEE*, pp 626–631.

[121] **Natekin A., Knoll A.** (2013). Gradient boosting machines, a tutorial. *Frontiers in Neurorobotics*. *vol 7*.

[122] **Susto GA., Schirru A., Pampuri S., et al.** (2013). A predictive maintenance system for integral type faults based on support vector machines: An application to ion implantation. *International Conference on Automation Science and Engineering (CASE). IEEE,* pp 195–200.

[123] **Kotsiantis SB., Zaharakis ID., Pintelas PE.** (2006). Machine learning: A review of classification and combining techniques. *Artif Intell Rev., 26*:159–190.

[124] **Hu RL., Granderson J., Auslander DM., Agogino A.** (2019). Design of machine learning models with domain experts for automated sensor selection for energy fault detection. *Appl Energy, 235*:117–128.

[125] **Baccarini LMR., Avelar VS., Silva VVRE., Amaral GF V.** (2013) Intelligent System Design for Stator Windings Faults Diagnosis: Suitable for Maintenance Work. J *Softw Eng Appl., 06*:526–532.

[126] **Manco G., Ritacco E., Rullo P., et al.** (2017) Fault detection and explanation through big data analysis on sensor streams. *Expert Syst Appl., 87*:141–156.

[127] **James G., Witten D., Hastie T., Tibshirani R.** (2013). *An Introduction to Statistical Learning with Applications in R*. Springer, *vol 103*. http://www.springer.com/series/417.

[128] **Sokolova M., Lapalme G.** (2009). A systematic analysis of performance measures for classification tasks. *Inf Process Manag, 45*:427–437.

[129] **Hatamura Y.** (2006). *Decision-making in engineering design: theory and practice*. Springer, p.275.

[130] **Kong Z., Beyca O., Bukkapatnam ST., Komanduri R.** (2011). Nonlinear Sequential Bayesian Analysis-Based Decision Making for End-Point Detection of Chemical Mechanical Planarization (CMP) Processes. *IEEE Trans Semicond Manuf, 24*:523–532.

[131] **D. V. Lindley.** (1985). *Making Decisions*, 2nd edition. UK Wiley, London

[132] **De Almeida AT., Bohoris GA.** (1995). Decision theory in maintenance decision making. *J Qual Maint Eng, 1*:39–45.

[133] **Wang L., Zhang Z., Long H., et al.** (2017). Wind Turbine Gearbox Failure Identification with Deep Neural Networks. *IEEE Trans Ind Informatics, 13*:1360–1368.

[134] **Chen J., Liu Z., Wang H., et al.** (2018). Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network. *IEEE Trans Instrum Meas., 67*:257–269.

[135] **You D., Gao X., Katayama S.** (2015). WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM. *IEEE Trans Ind Electron, 62*:628–636.

[136] **Hochreiter S., Schmidhuber J.** (1997). Long Short-Term Memory. *Neural Comput, 9*:1735–1780.

[137] **Yang Y., Tang W.** (2011). Study of remote bearing fault diagnosis based on BP Neural Network combination. *Proc - 7th Int Conf Nat Comput ICNC, 2:*618–621.

[138] **Qin Y., Wang X., Zou J.** (2019). The Optimized Deep Belief Networks With Improved Logistic Sigmoid Units and Their Application in Fault Diagnosis for Planetary Gearboxes of Wind Turbines. *IEEE Trans Ind Electron, 66*:3814–3824.

[139] **Jia F., Lei Y., Lin J., et al.** (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mech Syst Signal Process,73*:303–315.

[140] **Khan S., Yairi T.** (2018). A review on the application of deep learning in system health management. *Mech Syst Signal Process, 107*:241–265.

[141] **Samanta, B; Khamis R., et al.** (2004). Bearing Fault Detection Using Artificial Neural Networks and Genetic Algorithm. *EURASIP J Appl Signal Processing,* pp 366–377.

[142] **Gan M., Wang C., Zhu C.** (2016). Construction of hierarchical diagnosis network based on deep learning and its application in the fault pattern recognition of rolling element bearings. *Mech Syst Signal Process, 73*:92–104.

[143] **Khaled O., Hedi D., Lotfi N., et al.** (2010). Fault detection and localization with Neural Principal Component Analysis. *18th Mediterr Conf Control Autom MED - Conf Proc,* pp 880–885.

[144] **Liao L., Jin W., Pavel R.** (2016). Enhanced Restricted Boltzmann Machine With Prognosability Regularization for Prognostics and Health Assessment. *IEEE Trans Ind Electron, 63*:7076–7083.

[145] **Ahmad W., Khan SA., Islam MMM., Kim JM.** (2019). A reliable technique for remaining useful life estimation of rolling element bearings using dynamic regression models. *Reliab Eng Syst Saf, 184*:67–76.

[146] **Ren L., Cui J., Sun Y., Cheng X.** (2017). Multi-bearing remaining useful life collaborative prediction: A deep learning approach. *J Manuf Syst 43*:248–256.

[147] **Shao H., Jiang H., Li X., Liang T.** (2018). Rolling bearing fault detection using continuous deep belief network with locally linear embedding. *Comput Ind , 96*:27–39.

[148] **Shao H., Jiang H., Zhao H., Wang F.** (2017). A novel deep autoencoder feature learning method for rotating machinery fault diagnosis. *Mech Systems Signal Process, 95*:187–204.

[149] **Zhao R., Yan R., Chen Z., et al.** (2019). Deep learning and its applications to machine health monitoring. *Mech Syst Signal Process, 115*:213–237.

[150] **Liu H., Mi X., Li Y.** (2018). Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. *Energy Convers Manag, 156*:498–514.

[151] **Zheng S., Ristovski K., Farahat A., Gupta C.** (2017). Long Short-Term Memory Network for Remaining Useful Life estimation. In: *2017 IEEE International Conference on Prognostics and Health Management (ICPHM). IEEE*, pp 88–95

[152] **Xu C., Shen J., Du X., Zhang F.** (2018). An Intrusion Detection System Using a Deep Neural Network with Gated Recurrent Units. *IEEE Access, 6*:48697–48707.

[153] **Zhao R., Yan R., Wang J., Mao K.** (2017). Learning to monitor machine health with convolutional Bi-directional LSTM networks. *Sensors (Switzerland),vol 17*. https://doi.org/10.3390/s17020273

[154] **Zhao R., Wang J., Yan R., Mao K.** (2016). Machine health monitoring with LSTM networks. *Tenth International Conference on Sensing Technology ICST*. https://doi.org/10.1109/ICSensT.2016.7796266

[155] **Kim TY., Cho SB.** (2019). Predicting residential energy consumption using CNN-LSTM neural networks. Energy, *182*:72–81.

[156] **Wang K., Qi X., Liu H.** (2019). Photovoltaic power forecasting based LSTM-Convolutional Network. *Energy, 189*:116225.

[157] **Tovar M., Robles M., Rashid F.** (2020). PV Power Prediction, Using CNN-LSTM Hybrid Neural Network Model . Case of Study : Temixco-Morelos, México. *Energies, 13*:6512.

[158] **Hu Y-L., Chen L.** (2018). A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm. *Energy Convers Manag., 173*:123–142.

[159] **Wu Y., Yuan M., Dong S., et al.** (2018). Remaining useful life estimation of engineered systems using vanilla LSTM neural networks. *Neurocomputing, 275*:167–179.

[160] **Wu H., Huang A., Sutherland JW.** (2020) Avoiding Environmental Consequences of Equipment Failure via an LSTM-Based Model for Predictive Maintenance. *Procedia Manuf., 43*:666–673.

[161] **Bouktif S., Fiaz A., Ouni A., Serhani MA.** (2018). Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies, 11*:1–21.

[162] **Jang-Wook Hur UEA.** (2020). A deep learning approach to prognostics of rolling element bearings. *Int J Integr Eng., 12*:178–186.

[163] **Zou P., Hou B., Lei J., Zhang Z.** (2020). Bearing Fault Diagnosis Method Based on EEMD and LSTM. *Int J Comput Commun Control, 15*:1–14.

[164] **Donahue J., Hendricks LA., Rohrbach M., et al.** (2017). Long-Term Recurrent Convolutional Networks for Visual Recognition and Description. *IEEE Trans Pattern Anal Mach Intell., 39*:677–691.

[165] **Ren L., Sun Y., Wang H., Zhang L.** (2018). Prediction of bearing remaining useful life with deep convolution neural network. *IEEE Access, 6:*13041–13049.

[166] **Chen Z., Li C., Sanchez R-V.** (2015). Gearbox Fault Identification and Classification with Convolutional Neural Networks. *Shock & Vibration*, pp 1–10.

[167] **Janssens O., Slavkovikj V., Vervisch B., et al.** (2016). Convolutional Neural Network Based Fault Detection for Rotating Machinery. *J Sound Vib., 377*:331–345.

[168] **Dieleman S., Willett KW., Dambre J.** (2015). Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Mon Not R Astron Soc., 450*:1441–1459.

[169] **Passricha V., Kumar Aggarwal R.** (2018). Convolutional Neural Networks for Raw Speech Recognition. In: From Natural to Artificial Intelligence - Algorithms and Applications. *IntechOpen*, pp 1–15

[170] **Guo X., Chen L., Shen C.** (2016). Hierarchical adaptive deep convolution neural network and its application to bearing fault diagnosis. *Meas J Int Meas Confed., 93*:490–502.

[171] **Sun W., Yao B., Zeng N., et al.** (2017). An intelligent gear fault diagnosis methodology using a complex wavelet enhanced convolutional neural network. *Materials,vol 10*.

[172] **Pan H., He X., Tang S., Meng F.** (2018). An improved bearing fault diagnosis method using one-dimensional CNN and LSTM. *J Mech Eng, 64*:443–452.

[173] **Guo X., Shen C., Chen L.** (2017). Deep fault recognizer: An integrated model to denoise and extract features for fault diagnosis in rotating machinery. *Applied Sciences,vol 7*.

[174] **Magar R., Ghule L., Li J., et al.** (2020). FaultNet: A deep convolutional neural network for bearing fault classification. *IEEE access, 9*: 25189 - 25199.

[175] **Li S., Liu G., Tang X., et al.** (2017). An ensemble deep convolutional neural network model with improved D-S evidence fusion for bearing fault diagnosis. *Sensors, 17*:1708-1729.

[176] **Hoang DT., Kang HJ.** (2019). A survey on Deep Learning based bearing fault diagnosis. *Neurocomputing, 335*:327–335.

[177] **Jiao J., Zhao M., Lin J., Liang K.** (2020). A comprehensive review on convolutional neural network in machine fault diagnosis. *Neurocomputing, 417*:36–63.

[178] **Richert W., Coelho LP.** (2013). *Building Machine Learning Systems with Python*. Packet publishing, (Birmingham - Mumbai), UK.

[179] **Aydin O., Guldamlasioglu S.** (2017). Using LSTM networks to predict engine condition on large scale data processing framework. *The 4th Int Conf Electr Electron Eng. ICEEE*, pp 281–285.

[180] **Zhang Z., Wang Y., Wang K.** (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *J Intell Manuf*, *24*:1213–1227.

[181] **Kingma DP., Ba J.** (2014) Adam: A Method for Stochastic Optimization. *Conference paper at ICLR*, pp1–15.

CURRICULUM VITAE

Name Surname **:** Raghad Mohammed KHORSHEED

EDUCATION **:**

- **B.Sc.** **:** 1994, University of Technology, Mechanical Engineering Department
- **M.Sc.** **:** 2001, University of Baghdad, İndustrial Engineering Department

PROFESSIONAL EXPERIENCE AND REWARDS:

- 1994-2014 Mechanical Engineer: working of the various mechanical engineering departments like a Gas turbine, main oil line, air conditioning and refrigeration, pump stations, etc., in the North Oil Company / Kirkuk.
- 2001-2004 Lecturer: Lecturing in the college of Technology /Kirkuk University.
- 2010-2014: working as a head of the Enterprise resource planning (ERP) team to represent the Engineering Authority/ north of company after taking special courses in the SAP academy of learning MENA/Dubai.
- In 2014, I got a granted scholarship from the Iraqi Ministry of Higher Education and Scientific Research to accomplish a Ph.D. degree from Istanbul Technical University.

PUBLICATIONS ON THE THESIS:

- **Raghad M Khorsheed,** Omer Faruk Beyca. "An integrated machine learning: Utility theory framework for real-time predictive maintenance in pumping systems," 2020, Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, DOI: 10.1177/0954405420970517.2020.

OTHER PUBLICATIONS AND PRESENTATIONS:

- Research about (Computer-Aided of Preventive Maintenance Scheduling System CAPMSS Design and Development) presented in the Iraqi second National Conference of Mechanical Engineering and Material Engineering /Kufa University 2001.

- Research in the topic (Design of a Computer System for Condition Monitoring and Diagnosis Machine Faults by Vibration Signal Analysis CMDMF) published in Arabic Universities Union of Engineering researches magazine, No.1, Edition 12, 2005.
- Research about (Computer-Aided for Maintenance Management System) Lectures given in fifth scientific conference Baghdad University College of Engineering held in 25-27/3/2003.
- Research about (Design and Implementation of Computer-Aided Preventive Maintenance Scheduling System) published in Engineering magazine, Baghdad University, College of Engineering, No 5, Edition 16, 2003.
- Research about (Computer-Aided Inventory Control and its Influence upon Maintenance Management) presented in Iraqi third National Conference of Mechanical Engineering and Material Engineering /Kufa University 2003.
- Research in the title (Computer-Aided Maintenance Planning System) presented in the 5th scientific engineering conference in Baghdad, college of engineering, Baghdad University, 2003.