DEVELOPMENT OF A PLANNED PREVENTIVE MAINTENANCE (PPM) MODEL USING A MACHINE LEARNING APPROACH

Puspendu Kumar Fouzder
Student ID: 0416082021
Session: April 2016

DEPARTMENT OF INDUSTRIAL AND PRODUCTION ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING & TECHNOLOGY
DHAKA-1000, BANGLADESH

DEVELOPMENT OF A PLANNED PREVENTIVE MAINTENANCE (PPM) MODEL USING A MACHINE LEARNING APPROACH

By
Puspendu Kumar Fouzder

A Thesis Submitted to the
Department of Industrial and Production Engineering
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY
In Partial Fulfillment of the Requirements for the Degree of
MASTER OF SCIENCE IN INDUSTRIAL AND PRODUCTION ENGINEERING

DEPARTMENT OF INDUSTRIAL AND PRODUCTION ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING & TECHNOLOGY
DHAKA-1000, BANGLADESH

October 24, 2020

CERTIFICATE OF APPROVAL

The thesis titled **"Development of a Planned Preventive Maintenance (PPM) Model Using a Machine Learning Approach"** submitted by Puspendu Kumar Fouzder, Student No: 0416082021, Session- April 2016, has been accepted as satisfactory in partial fulfillment of the requirements for the degree of Master of Science in Industrial & Production Engineering on October 24, 2020.

BOARD OF EXAMINERS

Dr. Ferdous Sarwar
Associate Professor
Department of IPE, BUET, Dhaka. — Chairman (Supervisor)

Dr. Nikhil Ranjan Dhar
Professor & Head
Department of IPE, BUET, Dhaka. — Member (Ex-Officio)

Dr. A. K. M. Masud
Professor
Department of IPE, BUET, Dhaka. — Member

Shuva Ghosh

Dr. Shuva Ghosh
Associate Professor
Department of IPE, BUET, Dhaka. — Member

14.02.2021

Dr. A.K.M. Ashikur Rahman
Professor
Department of CSE, BUET, Dhaka. — Member (External)

CANDIDATE'S DECLARATION

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Puspendu

Puspendu Kumar Fouzder

This work is dedicated to my

Loving Parents

ACKNOWLEDGEMENT

All credits go to the Almighty God, for His boundless grace in the successful completion of this thesis.

At the very beginning, the author expresses his deepest gratitude to the supervisor of this thesis- Dr. Ferdous Sarwar, Associate Professor, Department of Industrial & Production Engineering (IPE), Bangladesh University of Engineering and Technology (BUET). Without his affectionate guidance, valuable suggestions and inspirations throughout this work, this research would not have been possible.

Moreover, the author is thankful to the board members- Dr. Nikhil Ranjan Dhar, Professor & Head, Department of IPE, BUET, Dr. A. K. M. Masud, Professor, Department of IPE, BUET, Dr. Shuva Ghosh, Assistant Professor, Department of IPE, BUET, and Dr. A.K.M. Ashikur Rahman, Professor, Department of CSE, BUET, for their constructive remarks and kind evaluations of this research.

Finally, the author would like to thank all of his colleagues, friends, and well-wishers for their co-operation and motivation to complete the thesis work. The author would also like to extend his thanks to his parents whose continuous inspiration, sacrifice and support encouraged him to complete the thesis successfully.

ABSTRACT

Planned preventive maintenance with some expert system is essential for appropriate planning and utilization of maintenance policy effectively and efficiently. A number of preventive maintenance model have been developed that have identified several factors which performed the models by subjective means. However, these models often lack robustness due to bias and variance. Now, the increased availability of data opens the scope of applying machine learning technique to predict the maintenance requirement more accurately and cost effectively. The aim of this research work is to develop a planned preventive maintenance model by using machine learning algorithms (SVM and SVR) that can forecast the maintenance requirements more accurately and cost effectively. To develop the model machine reliability is considered and the reliability depends on various subjective and objective measures which is a data driven approach. The subjective and objective features of Diesel Generator (DG) have been selected from literature and expert opinions and the data are collected from field survey. Two separate feature selection methods have been used to select the best feature set to improve the model accuracy. Wrapper method used correlation-based features selection to rank the features and generate eight different feature sets following backward elimination process. Filtering method eliminates the insignificant features by ANOVA test and selects the significant feature sets. All these feature sets are generated a total 54 number of different models with different accuracy level. Among them the best feature set have been selected with an accuracy of 92.5% from Wrapper method. Finally, a regression model has developed using Support Vector Regression (SVR) to determine the machine reliability value.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

PPM	Planned Preventive Maintenance
ML	Machine Learning
SVM	Support Vector Machine
SVR	Support Vector Regression
MLR	Multivariate Linear Regression
PM	Preventive Maintenance
CM	Corrective Maintenance
TBM	Time-Based Maintenance
PdM	Predictive Maintenance
CBM	Condition-Based Maintenance
CFS	Correlation-based Feature Selection
PCA	Principal Component Analysis
MAPE	Mean Absolute Percentage Error
MAE	Mean Absolute Error
rRMSE	relative Root Mean Squared Error
MSE	Mean Squared Error
DG	Diesel Generator
MTBF	Men Time Between Failure
MTTR	Mean Time to Repair
DT	Downtime
MA	Machine Age
MRE	Machine Room Environment
AOT	Average Operating Time
M	Manufacturer
PMP	Periodic Maintenance Practice
AM	Alternative Machine

CHAPTER 1: INTRODUCTION

1.1 Background of the Study

The failure of machineries, equipment's or system not only reduced the productivity but also affect the quality, availability and customer service. It may even lead to safety and environmental problems. The aim of maintenance management is to keep the assets (i.e. machineries, equipment, tools, building facilities or systems) in its good aesthetic and operating condition to maximize its utility with more effectively and efficiently. Proper planning and implementation of maintenance actions can ensure the machinery availability and reliability to deliver its on-time quality output with a safe working environment.

Since 1960s, many researchers have worked on analysis and modeling of maintenance operations **[Coria et al. 2014]** and developed lots of preventive maintenance policies **[Do et al. 2015; Caballé et al. 2014; Xiao et al. 2015; Sherwin 2000; Garg and Deshmukh 2006; Mabrouk et al. 2016b, El-Ferik et al. 2004]**. In 1960, Barlow and Hunter have proposed a simple periodic replacement model with minimum repair aimed to restore the system to its prior state before failure **[Sherwin 2000]**. After that a lot of scheduled maintenance policies have developed from this basic time-based maintenance model **[Sherwin 2000]**.

Maintenance can be planned and carried out in different ways. The three common planning paradigms are corrective, preventive and predictive maintenance **[Alaswad and Xiang 2017]**. The simplest and most frequently adopted approach to dealing with maintenance is corrective or run to failure or breakdown maintenance where the maintenance action is performed only after the occurrence of failures **[Do et al. 2015]**. This maintenance policy, or actually lack of policy, is common for infrequent failures or where the repair is very expensive or in case the system has redundancy **[Alaswad and Xiang 2017]**, but it is also the least effective one, as the cost of interventions and associated downtime after failure are usually much more substantial than those associated with planned corrective actions taken in advance **[Do et al. 2015]**.

Preventive maintenance is carried out according to a planned schedule based on time or process iterations in order to reduce the corrective maintenance and keep the machines at the desired level of quality and performance. Compared with the failure-based maintenance, scheduled maintenance shows that it is more positive and efficient.

Predicative maintenance which is also called condition-based maintenance is a positive and useful condition-based maintenance methodology that system's condition can be collected through continuous monitoring and maintenance is performed based on an estimate of the health status of a piece of equipment **[Do et al. 2015; Caballé et al. 2014]**. It is a complex initiative that requires the assets to communicate their actual status in real-time with the integrated maintenance management system. It also requires training the plant personnel to operate the system as well as to interpret the analytics or data. Therefore, a high initial investment is associated with this condition-based maintenance. However, this type of maintenance is feasible for important, expensive and time sensitive machineries, equipment or system because of those expensive maintenance tools and system. The purpose of distinguishing the maintenance policy is to reduce unexpected failure and optimize the maintenance cost. The development of different maintenance model is an ongoing process to solve various real-life problems.

In both preventive and predicative maintenance, the most important thing is to predict the maintenance schedule. For predicative maintenance the maintenance requirement is determined by continuous monitoring of the system. But the conventional preventive maintenance policies have the same time interval that may easily neglect system's reliability **[Sherwin 2000]**. As a result, the maintenance actions are performed either too often or not often enough. If it performed too early, although the system remains in good condition, maintenance cost will be much higher or if it is performed too let, although it can reduce maintenance actions so as to reduce maintenance cost, system reliability will be lower and more failure will occur, which may lead to higher breakdown cost **[Sherwin 2000]**.

Thus, neither too long nor too short interval time is suitable for maintenance model **[Graham et al. 1979]**. So, without continuous monitoring the most accurate time can be predicted by considering the machine reliability. The machine reliability highly depends on various external and internal parameters of the system.

Therefore, a model that can consider the above parameters to predict the maintenance requirements is needed. So that, the maintenance actions can be performed more accurately and cost effectively. Machine learning, an application of artificial intelligence (AI), can build the classification model by using supervised learning algorithms **[Lin et al. 2001]**. In this research work machine learning algorithms (SVM & SVR) are used to develop a classification model to forecast the maintenance schedule and machine reliability more effectively and efficiently.

To develop the model, this research has considered various external and internal parameters of Diesel Generator and Boiler. Diesel Generator is widely used as a secondary or backup source of power in different residential, commercial buildings and industries where as Boiler is used for producing stream to perform various operations in industries.

1.2 Rationale of the Study

Perfect maintenance actions restore completely the system to the 'as good as new' state. Optimal maintenance policies aim to provide optimal system reliability and safety performance at lowest possible maintenance costs. In Bangladesh, the overall understanding of maintenance requirement is still very low, however in industrial level it seems to perform some maintenance actions for crucial and time sensitive manufacturing equipment. Most of the cases the common practice is corrective actions except some random inspection and cleaning. So, there is a scope to design an optimal maintenance policy to keep the assets in its best operating condition at lowest possible cost.

For most industrial plants, preventive maintenance (PM) is still a dominant maintenance policy as it is easy to implement and not many systems can be condition- monitored **[De Jonge et al. 2017]**. In the reliability and maintenance literature, PM policies are commonly classified as **[Wang et al. 2018]**: periodic and sequential PM. Periodic PM is executed at fixed time interval whereas sequential PM is implemented at intervals of unequal time lengths. Sequential PM is more suitable when the system requires more frequent maintenance as its ages. Currently the periodic and sequential maintenance requirements are determined by company standard or in-house maintenance experience without any statistical analysis. As a consequence, the realistic operating conditions of the system over time are not be taken into account. As a result, the maintenance actions are performed either too late or too early.

The aim of this work is to develop an optimal preventive maintenance model. To develop the model, system reliability has been considered and it is a data driven approach which has considered the actual condition of the system to determine the maintenance requirements.

1.3 Objectives with Specific Aims

Determination of a maintenance policy is an important issue in optimal maintenance planning **[Lin et al. 2001]**. Researchers developed some maintenance policies and therefor developed different maintenance models. However, some models are effective and some are efficient and most of them are case specific. In case of, traditional PM policies, it is efficient compared to failure-based maintenance, but not effective, because it doesn’t consider the systems (machineries, equipment or tools) reliability, as a result the maintenance action are performed either too early or too late. On the other hand, Condition Based Maintenance (CBM) is more effective but not efficient, because the initial investment is too high which is feasible only for important, expensive and time sensitive machineries. Therefore, this work aims to develop the robust Planned Preventive Maintenance (PPM) model by improving its effectiveness.

The specific objectives of this research work are listed below:

i. Development of a classification model by using Support Vector Machine (SVM) algorithm to predict the preventive maintenance schedule.
ii. Development of a regression model using Support Vector Regression (SVR) algorithm based on prepared dataset to predict the machine reliability.

1.4 Outline of Methodology

The proposed research methodology is outlined below:

i. Machine reliability parameters has been identified for both subjective and objective aspects. The parameters have been identified based on literature and expert opinions.
ii. Two separate datasets have been prepared for maintenance schedule prediction and machine reliability prediction model development. The data has been collected from the field survey. Expert opinions and record keeping registers are the two sources of qualitative and quantitative data for predetermined features sets.
iii. Mathematical models have been formulated for both classification and regression by using machine learning algorithms (SVM and SVR) to predict the maintenance class and machine reliability values.
iv. The models have been implemented for two different real-life problem for Diesel Generator and Boiler.
v. Features (predictor variables) have been selected by using Wrapper method and Filtering method to find the best subsets of features to improve the interpretability and accuracy of the models.

vi. Correlation based feature ranking has been used in Wrapper method that follows backward elimination process for subset generation. Whereas, ANOVA test has been applied for Filtering method to determine the significant features set.

vii. The features sets have been used in these proposed models to determine the model's accuracy. The predictive models have been developed with the help of MATLAB 2018a and the best model has been selected based on the highest accuracy level.

viii. Finally, the regression model for machine reliability have been developed by Support Vector Regression (SVR) which has been compared to the performance of Multivariate Linear Regression (MLR) models.

1.5 Organization of the Thesis

This thesis has been structured into seven chapters along with a list of references & appendices as follows:

Chapter 1 consists the background of the study, research objectives, and research methodologies. Under introduction section, general concepts of maintenance management are discussed. Proper reason for the research work has been demonstrated. Finally, the research objectives are also outlined with some guidelines of research methodologies.

Chapter 2 presents the literature review which includes the related literature on machine learning algorithms and its application in maintenance management.

Chapter 3 represents the theoretical framework of Support Vector Machine (SVM) and its details, different classes and current form. Moreover, the basic theory for the Support Vector Regression (SVR) and Multivariate Linear Regression (MLR) algorithms are included.

Chapter 4 formulates the models through the steps of problem identification, problem formulation and model development both for classification and regression.

Chapter 5 implements the model where the formulated model for classification and regression are implemented in a case study on Diesel Generator. Some Boiler data is also used to build a classification model in a separate case study as a part of model implementation.

Chapter 6 presents a discussion on results and important findings on the developed models. At the bottom, some comparison is placed among different models used in this thesis paper.

Chapter 7 provides major contribution and recommendation for future work associated with this research work.

CHAPTER 2: LITERATURE REVIEW

Support vector machine is one of the most robust techniques in machine learning which is supervised learning model with associated learning algorithms that analyzes data for classification and regression analysis **[Cortes and Vapnik 1995].** Cortes and Vapnik **[1995]** proposed VC theory which forms SVMs as a robust prediction method, being based on statistical learning frameworks. In addition to performing binary classification, SVM can efficiently perform a non-linear classification by using kernel trick, implicitly mapping their inputs into high dimensional feature spaces. The original maximum-margin hyperplane algorithm was proposed by Vapnik in 1963 which constructed a linear classifier. However, in 1992, Bernhard Boser, Isabelle Guyon and Vladimir Vapnik suggested a way to create nonlinear classifiers by applying the kernel trick **[Aiserman et al. 1964]**. Vapnik, Drucker et al. **[1996]** proposed an extension of SVM for regression. This method is called support vector regression (SVR).

2.1 Machine Learning in Maintenance

The increasing availability of data is changing the way decisions are taken in industry in important areas such as scheduling, maintenance management and quality improvement **[Susto et al. 2015]**. Machine learning (ML) approaches have been shown to provide increasingly effective solutions in these areas, facilitated by the growing capabilities of hardware, cloud-based solutions, and newly introduced state-of-the-art algorithms. At the same time the efficient management of maintenance activities is becoming essential to decreasing the costs associated with downtime and defective products, especially in highly competitive advanced manufacturing industries such as semiconductor manufacturing. Among statistical inference-based methods, those based on ML are the most suitable for dealing with modeling of high-dimensional problems, such as those arising in semiconductor manufacturing where hundreds or thousands of physical variables (pressures, voltages, currents, flows, etc.) act on the process **[Lenz and Barak 2013]**.

The discipline that predicts health condition and remaining useful life (RUL) based on previous and current operating conditions is often referred to as prognostics and health management (PHM). Prognostic approaches fall into two categories: model-based and data-driven prognostics **[Gao et al. 2015]**. To complement model-based prognostics, data-driven prognostics refer to approaches that build predictive models using learning algorithms and large volumes of training data.For example, classical data-driven prognostics are based on autoregressive (AR) models, multivariate adaptive regression, fuzzy set theory, ANNs, and SVR. The unique benefit of data-driven methods is that an in-depth understanding of system physical behaviors is not a prerequisite. In addition, data-driven methods do not assume any underlying probability distributions which may not be practical for real-world applications. While ANNs and SVR have been applied in the area of data-driven prognostics, little research has been conducted to evaluate the performance of other machine learning algorithms **[Sick 2002]** Schwabacher and Goebel **[2007]** conducted a review of data driven methods for prognostics. The most popular data-driven approaches to prognostics include ANNs, decision trees, and SVM in the context of systems health management. Cho et al. **[2005]** developed an intelligent tool breakage detection system with the SVM algorithm by monitoring cutting forces and power consumption in end milling processes. Linear and polynomial kernel functions were applied in the SVM algorithm. It has been demonstrated that the predictive model built by SVM can recognize process abnormalities in milling. Benkedjouh et al. **[2013]** presented a method for tool wear assessment and remaining useful life prediction using SVM. The features were extracted from cutting force, vibration, and acoustic emission signals. The experimental results have shown that SVM can be used to estimate the wear progression and predict RUL of cutting tools effectively. Shi and Gindy **[2007]** introduced a predictive modeling method by combining least squares SVM and principal component analysis (PCA). PCA was used to extract statistical features from multiple sensor signals acquired from broaching processes. Experimental results showed that the predictive model trained by SVMs was effective to predict tool wear using the features extracted by PCA **[Wu et al. 2017]**. In the area of condition monitoring and predictive maintenance, some work has been done to provide failure predictions using statistical and machine learning approaches. Liao **[2005]** presented reliability modeling to estimate machine failures. Sharma et al. **[2018]** used neural network classifier for condition monitoring of rotating mechanism systems. In railway applications, Yang et al. **[2017]** adopted a pattern recognition approach to classify the condition of the sleeper into classes (good or bad). Yang et al. **[2010]** proposed an approach to predict train wheel failures but only using one type of detectors, Wheel Impact Load Detector (WILD), without considering the impacts of multiple detectors. Recently, develops a logistic regression model to classify wheel failures based on WILD and Wheel Profile Detector (WPD). They claim that the classification accuracy is 90% with 10% false alarm rate. However, only two detectors are taken into account in that study. The problems that those papers have worked on are not as complicated as what we face and none of them has addressed all the challenges we describe above **[Li and Parikh 2014]**.

In this paper, a preventive maintenance model has been developed using machine-learning approaches to predict the maintenance requirements more accurately and cost effectively. The prediction will drive proactive inspections and repairs, reducing operational equipment failure. Feature selection is one of the core concepts in Machine learning which hugely impacts the performance of the models. Feature selection is the process to automatically or manually select those features which contribute most to prediction variables or output in which the research work is needed. Feature subset selection is of immense importance in the field of data mining. The increased dimensionality of data makes testing and training of general classification method difficult. Mining on the reduced set of attributes reduces computation time and also helps to make the patterns easier to understand **[Karegowda et al. 2010]**. Feature selection, the process of selecting a feature subset from the training examples and ignoring features not in this set during induction and classification, is an effective way to improve the performance and decrease the training time of a supervised learning algorithm **[Das 2001]**. Feature selection is considered a problem of global combinatorial optimization in machine learning, which reduces the number of features, removes irrelevant, noisy and redundant data, and results in acceptable classification accuracy **[Yang et al. 2010]**. Feature selection is one of the important and frequently used techniques in data preprocessing for data mining. Data preprocessing includes data cleaning, data integration, data transformation and data reduction. These data processing techniques, when applied prior to mining, can substantially improve the overall quality of the patterns mined and/or the time required for the actual mining. The goal of data reduction is to find a minimum set of attributes such that the resulting probability distribution of the data classes is as close as possible to the original distribution obtained using all attributes. Mining on the reduced set of attributes has additional benefits. It reduces the number of attributes appearing in the discovered patterns, helping to make the patterns easier to understand. Further it enhances the classification accuracy and learning runtime **[Karegowda et al. 2010].**

Lu et al. **[2018]** proposed a new feature selection model where the effects of different feature selection methods on model performance were compared and discussed. In their research two classification models including logistic regression (LGR) and support vector machine (SVM) were used, and two representative feature selection methods including analysis of variance

(ANOVA) and LGR filter were utilized to improve the performance of estimation models. The results showed that the model performance of LGR and SVM can be improved to a certain degree by all three feature selections methods. Feature selection is sometimes essential to the success of a learning algorithm. For example, the points out that it is not feasible to use a nearest-neighbors algorithm on the Internet Advertisements dataset (described later) because of the overabundance of features. Feature selection can reduce the number of features to the extent that such an algorithm can be applied **[Das 2001].** In real world situations, relevant features are often unknown a priori. Hence feature selection is a must to identify and remove are irrelevant/redundant features. It can be applied in both unsupervised and supervised learning.

The goal of feature selection for unsupervised learning is to find the smallest feature subset that best uncovers clusters form data according to the preferred criterion **[Dy and Brodely 2004]**. Feature selection in unsupervised learning is much harder problem, due to the absence of class labels. Feature election for clustering is the task of selecting important features for the underlying clusters. Feature selection for unsupervised learning can be subdivided in filter methods and wrapper methods. There are strong arguments in favor of both methods. Filter methods are general preprocessing algorithms that do not rely on any knowledge of the algorithm to be used **[Das 2001]**. This method in unsupervised learning is defined as using some intrinsic property of the data to select feature without utilizing the clustering algorithm **[Dy and Brodely 2004]**. Entropy measure has been used as filter method for feature selection for clustering **[Dash et al. 1997]**. Wrapper methods wrap the feature selection around the induction algorithm to be used, using cross-validation to predict the benefits of adding or removing a feature from the feature subset used **[Das 2001]** Wrapper approaches in unsupervised learning apply unsupervised learning algorithm to each candidate feature subset and then evaluate the feature subset by criterion functions that utilize the clustering result **[Dy. and Brodely 2004]**. Volker Roth and Tilman Lange proposes a wrapper method where Gaussian mixture model combines a clustering method with a Bayesian inference mechanism for automatically selecting relevant features **[Roth and Lange 2003]**.

In supervised learning, feature selection aims to maximize classification accuracy **[Kohavi and John 1997]**. It is easier to select features for classification/supervised learning than for clustering, since the classification uses class label information. Though domain experts can eliminate few of the irreverent attributes, selecting the best subset of features usually requires a systematic approach. Feature selection method generally consists of four steps described below **[Dash et al. 1997]**. Liao et al. **[2006]** notes that feature selection algorithms that search through the space of feature subsets must address four main issues: the starting point of the search, the organization of the search, the evaluation of feature subsets and the criterion used to terminate the search. Different algorithms address these issues differently **[Das 2001].** It is intractable to look at all possible feature subsets, even if the size is specified. Feature selection algorithms usually proceed greedily. They can be classified into those that add features to an initially empty set (forward selection) and those that remove features from an initially complete set (backward elimination). Hybrids both add and remove features as the algorithm progresses. A major problem of forward selection methods is that it is difficult for them to select sets of features that are good co-predictors of the class if none of these predictors is a good predictor of the class by itself. On the other hand, forward selection is much faster than backward elimination and therefore scales better to large datasets. The major approaches to the problem of when the greedy search should terminate are specifying the size of the feature set to be selected or evaluating the goodness of each feature set in some manner and stopping when further search results in a decrease in goodness **[Das 2001]**.

Wrapper Model approach uses the method of classification itself to measure the importance of features set; hence the feature selected depends on the classifier model used. Wrapper methods generally result in better performance than filter methods because the feature selection process is optimized for the classification algorithm to be used. However, wrapper methods are too expensive for large dimensional database in terms of computational complexity and time since each feature set considered must be evaluated with the classifier algorithm used **[Karegowda et al. 2010].** Wrapper methods, search through the space of feature subsets using a learning algorithm to inform the search. They calculate the estimated accuracy of the learning algorithm for each feature that can be added to or removed from the feature subset. Accuracy is estimated using cross validation on the training set. In forward selection, a wrapper estimates the accuracy of adding each unselected feature to the feature subset and chooses the best feature to add according to this criterion. These methods typically terminate when the estimated accuracy of adding any feature is less than the estimated accuracy of the feature set already selected **[Das 2001]**.

The Filter Approach actually precedes the actual classification process. The filter approach is independent of the learning induction algorithm computationally simple fast and scalable. Using filter method, feature selection is done once and then can be provided as input to different classifiers. Various feature ranking and feature selection techniques have been proposed such as Correlation-based Feature Selection (CFS), Principal Component Analysis (PCA), Gain Ratio (GR) attribute evaluation, Chi-square Feature Evaluation, Fast Correlation-based Feature selection (FCBF), Information gain, Euclidean distance, Markov blanket filter. Some of these filter methods do not perform feature selection but only feature ranking hence they are combined with search method when one needs to find out the appropriate number of attributes. Such filters are often used with forward selection, backward elimination, bi-directional search, best-first search, genetic search and other methods **[Karegowda et al. 2010].** Filter methods, on the other hand, select a feature set for any learning algorithm to use when learning a concept from that training set. The biases of the feature selection algorithm and the learning algorithm do not interact. The search proceeds until a pre-specified number of features is selected or some thresholding criterion is met **[Das 2001]**. Several methods have been used to perform feature selection, e.g., ANOVA test, genetic algorithms, branch and bound algorithms, sequential search algorithms, mutual information, tabu search, entropy-based methods, regularized least squares, random forests, instance-based methods, and least squares support vector machines **[Yang et al. 2010]**. Shaw and Mitchell-Olds **[1993]**, has recommended that when the response variables have continuous distributions and the conditions are discrete, whether inherently or by design, then it is appropriate to analyze the data using analysis of variance (ANOVA). Estévez-Pérez and Vilar **[2013],** applied ANOVA for discrete data to analysis the air quality data where A nonparametric functional approach is proposed to compare the mean functions of k samples of curves. In practice, curves data are usually collected in a discrete form and hence they must be pre-processed to use purely functional techniques.

2.2 Planned Preventive Maintenance

A Planned Preventive Maintenance (PPM) has an significant impact for operational performance of machineries by extending its service life with safe and efficient operations. Maintenance refers to the set of necessary operations applied to a system so that it can work properly. Nowadays, maintenance plays an important role in most companies, which try to provide high quality products but minimizing the production cost **[Do et al. 2015]**. Maintenance activities performed on an industrial system can increase not only its safety, but also ensure its availability and correct functioning **[Caballé et al. 2014]**. Maintenance optimization and the selection of maintenance strategies play an important role in the effectiveness of any industrial system's operation **[Alaswad and Xiang 2017]**. Perfect maintenance actions restore completely the system to the ’as good as new’ state **[Do et al. 2015]**. Their related costs are however often high. Maintenance involves preventive and corrective actions carried out to retain a system in or restore it to an operating condition **[Do et al. 2015]**.

Traditionally, these maintenance tasks were allocated based on requirements in legislation, company standards or in-house maintenance experience. However, in the early 1960s, authors as Barlow and Hunter started developing mathematical models, which aim to quantify costs and to find the optimum balance between the maintenance cost on one side, and the associated cost (benefit) on the other **[Caballé et al. 2014]**. Maintenance strategies regulate the different maintenance tasks which will be performed on systems. In general, maintenance tasks are classified in corrective maintenance and preventive maintenance. Corrective maintenance is defined as the maintenance which is required when a system has failed. Preventative maintenance is a maintenance planned in order to prevent the occurrence of future failures. It is performed when a system is still working. Preventive maintenance activities in single-unit systems were classified by Roth and Lange **[2003]** in: i) Age-based maintenance, where maintenance tasks are performed when the system exceeds a certain age. ii) Calendar-based maintenance, where maintenance tasks are performed at fixed time instants. iii) Condition-based maintenance, where maintenance tasks are based on one or several variables which measure the state of the system **[Xiao et al. 2016]**. The state variables are monitored continuously or on a regular basis. The maintenance of the system is performed when some state variable exceeds a certain fixed level. Because of its properties, this kind of maintenance is widely used in the current literature **[Caballé et al. 2014]**. Optimal maintenance policies aim to provide optimum system reliability, availability and safety performance at lowest possible maintenance costs **[Do et al. 2015]**. Maintenance actions can be generally classified into two categories: corrective maintenance (CM) and preventive maintenance (PM) **[Alaswad and Xiang 2017]**. CM are actions performed when the system fails. The most common form of CM is “minimal repair”, where the state of the system after repair is nearly the same as that just before failure **[Coria et al. 2015].**

PM is a maintenance policy based on replacing, overhauling or remanufacturing a system at fixed or adaptive time intervals, regardless of its condition at the time. The periodic PM policy can be considered as the most common maintenance policy in which a system is preventively maintained at fixed time intervals, regardless of the failure history of the system **[Coria et al.**

2015]. Traditionally, PM takes the form of system overhaul or unit replacement based one lapsed time, which is often referred to as time-based maintenance (TBM). TBM schedules are typically determined based on a probabilistic model of system failure **[Alaswad and Xiang 2017]**. PM policy has been considered by many researchers as one of the most studied maintenance policies. For most industrial plants, PM is still a dominant maintenance policy as it is easy to implement and not many systems can be condition- monitored **[Caballé et al. 2014]**. A more comprehensive definition is: PM policy is a planned maintenance that reduces or eliminates accumulated system deterioration, and is executed according with planned schedules. In the reliability and maintenance literature, PM policies are commonly classified as **[Sheu and Chang 2009]**: periodic and sequential PM. Periodic PM is executed at integer multiples of some fixed time interval. On the other hand, sequential PM is implemented at intervals of unequal time lengths. Sequential PM is more suitable when the system requires more frequent maintenance as its ages, whereas periodic PM is more convenient to schedule **[Caballé et al. 2014]**. Several researchers also categorized maintenance into three groups: (1) corrective maintenance (CM), (2) preventive maintenance (PM) and (3) predictive maintenance (PdM). In such maintenance models, PM decision is however based on the system age and on the knowledge of the statistical information on the system lifetime **[Do et al. 2015]**. As a consequence, the realistic operating conditions of the system over time are not be taken into account **[Do et al. 2015]**.

PdM is an advanced preventive approach where maintenance is deferred until it is actually needed. The objective of this approach is to monitor the system in order to detect incipient faults before they can cause a part to fail **[Jiang et al. 2013]**. This maintenance strategy has been implemented as condition-based maintenance in systems where certain performance indices are periodically or continuously monitored **[Coria et al. 2015]**. Condition-based maintenance (CBM) is a maintenance strategy that collects and assesses real-time in-formation, and recommends maintenance decisions based on the current condition of the system **[Alaswad and Xiang 2017]**. In recent years, condition-based maintenance (CBM) has received much attention in the maintenance research community. Unlike TBM policies that are developed based on historical failure data, CBM is a maintenance approach that emphasizes on combining data-driven reliability models with sensor data collected from monitored operating systems to develop strategies for condition monitoring and maintenance **[Alaswad and Xiang 2017]**. In recent decades, research on CBM has been rapidly growing due to the rapid development of computer- based monitoring technologies. Research studies have proven that

CBM, if planned properly, can be effective in improving equipment reliability at reduced costs **[Alaswad and Xiang 2017]**. Various CBM policies have been proposed and applied for many industrial systems **[Do et al. 2015]**.

Maintenance actions can also be performed both perfect and imperfect maintenance. In the literature, perfect maintenance actions (or replacement actions) which can restore the system operating condition to as good as new have been considered in various maintenance models **[Barker and Newby 2009]**. The implementation of "perfect" maintenance policies seems quite simple; however, perfect maintenance actions are often expensive **[Do et al. 2015]**. Imperfect maintenance implying that the system condition after maintenance is somewhere between the condition before maintenance and as good as new has grown recently as a popular issue to researchers as well as industrial applications **[Do et al. 2015]**. From a practical point of view, imperfect maintenance can describe a large kinds of realistic maintenance actions imperfect maintenance cost are usually low **[Do et al. 2015]**. A fixed number of allowable imperfect maintenance actions is introduced in maintenance models in and considered as a decision parameter. However, the value of this decision parameter is arbitrary chosen and they do not describe how the imperfect repair actions affect the deterioration evolution of the system **[Do et al. 2015]**.

Among various reliability and maintenance models incorporating internal deterioration and external shocks, Degradation-Threshold-Shock (DTS) model is addressed the most with abundant real-world applications. Sherwin **[2010]** proposed an opportunistic condition-based maintenance policy for offshore wind turbine blades subject to stress corrosion cracking and environmental shocks. Peng et al **[2010]** applied DTS models to micro electro-mechanical systems (MEMS) whose failures are triggered by gradually wear and debris from shock loads. Ye et al. **[2013]** established reliability models under extreme shocks and natural graduation for automobile tires, laser devices and hard disks. Zhou proposed a periodic preventive maintenance method for leased equipment subject to competing failures **[Yang et al. 2017]**.

2.3 Use of Machine Learning in Planned Preventive Maintenance

The aim of periodic PM optimization is to determine the optimal maintenance interval Tn and the optimal number of maintenance actions Nn, such that the total mean cost of repairs, PM, and replacement activities is minimal.

In general, the impact of PM actions can be classified into one of the following situations: perfect, minimal, and imperfect. A perfect PM restores the system to the state “as good as new”. A minimal PM restores the system to the state that it was just before the maintenance action, or “as bad as old”. An imperfect PM takes the system to any state between “as good as new” and “as bad as old”. In practice, PM is usually imperfect. Imperfect PM has grown recently as a popular issue to researchers as well as industrial applications **[Alaswad and Xiang 2017]**. In order to model the impact of imperfect PM, the hazard rate function of the system under maintenance is generally used. Infect, the hazard rate usually is more informative about the underlying mechanism of failure than the other representatives of a life time distribution. For this reason, consideration of the hazard rate may be the dominant method for modeling imperfect PM. Most of the hazard rates used in imperfect PM models are based on univariate analysis, where the single random variable under analysis is the failure time **[Ferreira et al. 2009]**. Recently, several attempts have been made to extend the concept of the univariate hazard rate to the multivariate analysis in order to include variables that influence the failure time of the system under study, for example cumulative load applied, time varying stress, and environmental factors. However, the hazard rate concept is somewhat difficult to extend to the multivariate situation and frequently the observed life time data set is not big enough in relation to the dimension of the hazard rate model in order to find good estimates of the model and its parameters. A number of PM models have been developed in order to describe the impact of imperfect PM on the hazard rate of repairable systems. These PM models can be classified into three groups **[Liao et al. 2006]**: age reduction models, hazard rate models, and hybrids of both. Age reduction models assume that there is an effective age reduction right after a PM action, and that the hazard rate continues to be a function of the effective age **[Dekker 1996]**. The hazard rate models assume that right after a PM action, the hazard rate reduces to zero, and then increases faster than it did in the previous PM interval **[Grall et al. 2002]**.

The above models have made important contributions to this research field, however, in practice the main problem is how to take decisions or make inferences about these unknown age reduction and hazard rate increase factors. Numerous approaches have been proposed based on guessing the values of these factors by subjective means, which is fine, as long as there is enough expert knowledge to perform this task properly. Other approaches are based on estimating these factors from observed data. These statistical inference techniques are very good if there are sufficient data to estimate the factors accurately. However, in practice, few data are available in many areas of maintenance and replacement **[Dekker 1996]**.

The literature on the use of mathematical modeling for the purpose of analyzing, planning, and optimizing TBM is abundant. In contrast, CBM has only received increasing attention recently, and only a few survey papers have considered CBM models extensively. The majority of existing CBM survey papers limits the scope within the existing CBM survey papers limits the scope within the diagnostic and prognostic methods and algorithms. For example, Jardine et al. **[2006]** review recent studies and developments in CBM with emphasis on models, algorithms, and technologies for data acquisition and data processing. Peng et al. **[2010]** divide the prog-nostic models into four categories: physical model, knowledge- based model, data-driven model, and combination model, and review various techniques and algorithms by this category. Ah- mad and Kamaruddin **[2012]** present an overview of time-based and condition-based maintenance in industrial applications, and summarize the most recent condition monitoring techniques. Shin and Jun **[2015]** review CBM approach and address several aspects of CBM, such as definition, advantages and disadvantages elated international standards, procedures and techniques. De- spite the recent rapid development of sensor technology that facilitates CBM, there exists increasing pressure on reducing unnecessary inspection and PM actions and the associated costs incurred from additional data collection, documentation, and analysis through optimal design of CBM policies. Therefore, mathematical modeling and optimization of CBM has become a major concern to operations and maintenance managers, and a review in this particular area is now more relevant. The popularity of CBM in the research community and industrial applications relies heavily on the development of stochastic deterioration models **[Alaswad and Xiang 2017]**.

Another data-driven method for prognostics is based on decision trees. Decision trees are a nonparametric supervised learning method used for classification and regression. The goal of decision tree learning is to create a model that predicts the value of a target variable by learning decision rules inferred from data features. A decision tree is a flowchart-like structure in which each internal node denotes a test on an attribute, each branch represents the outcome of a test, and each leaf node holds a class label. Jiaa and Dornfeld **[1998]** proposed a decision tree-based method for the prediction of tool flank wear in a turning operation using acoustic emission and cutting force signals. The features characterizing the AE root-mean-square and cutting force signals were extracted from both time and frequency domains. The decision tree approach was demonstrated to be able to make reliable inferences and decisions on tool wear classification. Elangovan et al. **[2011]** developed a decision tree-based algorithm for tool wear prediction using vibration signals. Ten-fold cross-validation was used to evaluate the accuracy of the predictive model created by the decision tree algorithm. The maximum classification accuracy was 87.5%. The effects of machining parameters on surface microhardness and microstructure such as grain size and fractions using a random forests-based predictive modeling method along with finite element simulations. Predicted microhardness profiles and grain sizes were used to understand the effects of cutting speed, tool coating, and edge radius on the surface integrity **[Wu et al. 2017]**.

2.4 Summary of Literature Review

Throughout the review, it can be concluded that various researchers have focused on solving maintenance problem with various approaches. The very initial approaches were started with quantifying costs to find the optimum balance with other associated costs. After that several maintenance strategies were identified which regulate different maintenance tasks and sub tasks. Therefore, various maintenance policies were developed to provide optimum system reliability and safety at lowest possible maintenance cost. Among the all, Preventive Maintenance (PM) has been considered as one of the most studied maintenance policies that dominate most of the industrial plant till now. A number of preventive maintenance model have been developed that identified several factors which performed the models by subjective means and somewhat statistical inference techniques. Various techniques, such as probability plotting, moment estimation, modified moment estimation and maximum likelihood estimation (MLE) has been used to estimate and modeling the maintenance model. Condition-based maintenance (CBM) is most advanced preventive approach that received much attention in the maintenance research community in recent years. For several decades, Fuzzy Neural Network (FNN), ANNs and branch-and-bound algorithm are widely used for maintenance planning and joint optimization of PM and production scheduling. Furthermore, the increasing availability of data open the scope of applying Machine Learning (ML) approaches to deal with modeling of high dimensional problems. Some recent work focused on prognostic approaches that build predictive models using learning algorithms and large volumes of training data like classical data-driven prognostics are based on autoregressive (AR) models, multivariate adaptive regression, fuzzy set theory, ANNs, and SVR. All the previous works described in the above section give descriptive knowledge on maintenance planning study and all are relevant to real world problem.

CHAPTER 3: THEORETICAL FRAMEWORK

In machine learning, Support Vector Machines (SVMs) is a relatively new computational supervised learning model for classification and regression analysis by using associated learning algorithm for analyzing input data. It was introduced by Vapnik in 1995 **[Cortes and Vapnik 1995]**.

3.1 Support Vector Machines (SVMs)

SVM were originally designed for binary classification, but now it can efficiently perform nonlinear or multi-class classification by using kernel trick. The basic idea in the SVM is to map the original input data into a high dimensional dot product space called a feature space in order to find a hyper-plane which can separate the two classes. The optimal hyper-plane is found by exploring the optimization theory and respecting insights provided by the structural learning theory.

3.1.1 Binary Classification using SVM

The aim of SVM is to create a line or hyper plane between two sets of data for classification. Figure 3.1 shows how to classify a series of points into two different classes of data, class I (Triangular) and class II (Circular). The attempt of SVM is to create a boundary line denoted by a solid line between the two different classes and organized in such a way that the dotted lines margin is maximized. The SVM tries to orient the boundary such that the distance between the boundary and the nearest data point in each class is maximal. The boundary is then placed in the middle of this margin between the two points. The nearest data points are used to define the margins and are known as Support Vectors (SVs) represented by green circle and square. Once the SVs are selected, the rest of the feature sets can be discarded, since the SVs have all the necessary information for the classifier.

Let's consider a training dataset S is given as (X_i, Y_i) where X_i denotes the input points for i = 1,….., N. and Y_i is the corresponding desired output {-1, 1} for classification model. Each point X_i belongs to one of the classes with label $y_i \epsilon \{1,-1\}$. φ be the kernel function which maps the data from the input space R^n to the feature space F. The goal is to find a hyper-plane which divides S, in a way that the same level points are on the same side of the hyper-plane while maximizing the distance between the two classes. Let's $w\varphi(x)+b$ is the hyper-plane that discriminate data in the feature space and the decision function is

$$f(x) = \text{sign}\ (w\varphi\ (x)) + b \tag{3.1}$$

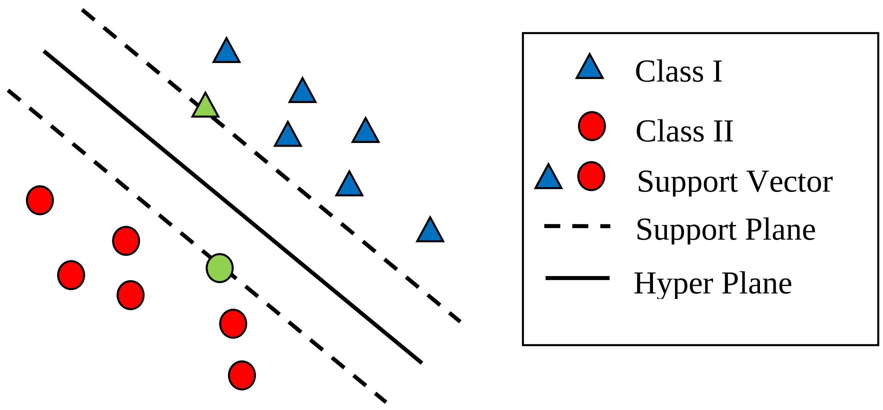

Figure 3.1: Binary classification using SVM

Therefore, the hyper-plane with minimum error and also the maximum margin is found as the optimum hyper-plane by the SVM. Margin is defined as the distance of the closest data point to the hyper-plane. The cost function ψ(w) =0.5(w.w) should be minimized with a view to maximize the margin subject to the constraint in Equation (n). Using Lagrange multiplier method, this constraint optimization problem is solved and the decision function is obtained as in equation 3.2.

$$y = \text{sign}\left[\sum_{\mathrm{i}\in\mathrm{sv}}^{\mathrm{l}} \alpha_i \alpha_{\mathrm{i}}^{0} \left(\varphi(x_i) \times \varphi(x_j)\right) + \mathrm{b}\right] \tag{3.2}$$

where α is the result of the constrained optimization problem and SV denotes the support vectors.

3.1.2 Nonlinear Support Vector Machine

In the last section, all data set considered are linear. However, to deal with more general decision surfaces, Kernel transformation is used to nonlinearly transform the data set $x_1, x_2, \ldots \ldots, x_m \in \mathcal{X}$ into high-dimensional feature space. For linear separation in the feature space map ϕ: $x_i \rightarrow x_i^*$ is used.

Cover’r theorem characterizes the number of possible linear separations of m points in an N-dimensional space. If m < N + 1, then 2^m is possible. According to this Theorem number of separation can be given by $2\sum_{i=1}^{N}\binom{m-1}{i}$ **[Berge 1957].**

With the increase in the number of N, the number of elements is the sum increases. So, it can be said that the number of separations increases with the increase in dimensionality.

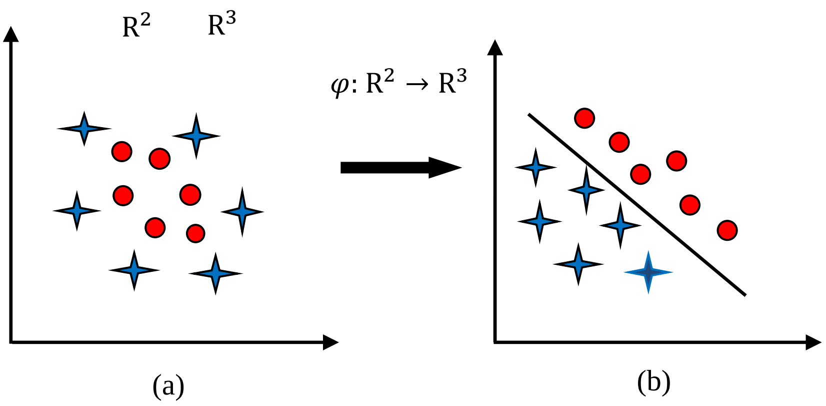

Figure 3.2: Mapping non-linear data into higher dimensional feature space

In Figure 3.2 (a) nonlinear data is shown and in Figure 3.2 (b) those data are transferred to higher dimensional feature space and separating it by a hyperplane.

In order to make equation 3.3 suitable for a general decision surface, $< x, x_i >$ is substituted by $< \phi(x), \phi(x_i) >$ in higher dimensional space. As this substation is computationally expensive, a positive kernel is used to make the calculation easier.

$$maximize\ W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j < \phi(x_i), \phi(x_j) > \tag{3.3}$$

$$\forall\ \alpha_i \in \mathbb{R}^m$$

$$f(x) = sgn(\sum_{i=1}^{m} \alpha_i y_i < \phi(x), \phi(x_i) > +b) \tag{3.4}$$

$$(< \phi(x), \phi(x_i) >) = k(x, x_i) \tag{3.5}$$

Using this transformation of equation 3.5 into previous equation, a new decision function can be obtained as following equation 3.6.

$$f(x) = sgn(\sum_{i=1}^{m} \alpha_i y_i k(x, x_i) + b) \tag{3.6}$$

Now to calculate threshold value b, implying to KKT conditions **[Schölkopf et al. 2002]** to the Eq., $\alpha_i > 0$ following formulation can be obtained

$$\sum_{i=1}^{m} \alpha_i \, y_i k(x_j, x_i) + b = y_j \tag{3.7}$$

So, the threshold value can be obtained as following equation 3.8.

$$b = y_j - \sum_{i=1}^{m} \alpha_i \, y_i k(x_j, x_i) \tag{3.8}$$

There are some popular forms of Kernel functions. Some of them are given by following equation 3.9, 3.10, 3.11 and 3.12 **[Giroti et al. 2018]**

Polynomial Kernel Classifier with a degree of d

$$k(x, x_i) = < x, x_i >^{d} \tag{3.9}$$

Gaussian Kernel

$$k(x, x_i) = exp\left(\frac{-\|x - x_i\|^2}{2\sigma^2}\right) \tag{3.10}$$

Radial basis function classifier with Gaussian Kernel of width c>0

$$k(x, x_i) = exp\left(\frac{-\|x - x_i\|^2}{c}\right) \tag{3.11}$$

Sigmoid kernel

$$k(x, x_i) = \tanh(\mathcal{B} < x, x_i > + \mathcal{C}) \tag{3.12}$$

Where $\mathcal{B} > 0$ and $\mathcal{C} \in \mathbb{R}$

3.1.3 Multi Class Support Vector Machines (SVMs)

The support vector machine is inherently a binary classifier where the class levels can take only two values: 1 and -1. However, many real-world problems, we found more than two classes. How to effectively extend the multi-class classification is still an on-going research issue. But several methods have been proposed where typically we construct a multi-class classifier by combining several binary classifiers.

There are two min strategies widely used for extending a binary classifier to be employed for a multi-class application: i) One-against-one and ii) One-against-all

In one-against-one strategy, the multi class problem constructs k(k-)/2 binary classification problems, where k is the number of classes. Now considering the data from one class against the data from another class as a binary classification task a k(k-1)/2 classifiers are designed for each problem.

In one-against-all approach, it constructs k SVM model where k is the number of classes. K binary classifiers are designed considering the data of one class against all other remaining data from other classes as a binary classification problem.

Beside these two strategies, there is another approach called Direct Acyclic Graph Support Vector Machine (DAGSVM) which is inherently a multi-class classifier. Its training phase is the same as one-against-one method by solving k(k-1)/2 binary SVM. However, in one-against-one approach (K (K-1)/2), its needs to be designed more binary classifiers comparing with one-against-all (K). The earliest used implementation for SVM multi-class classification is probably the one-against-all method.

Since the SVM classification problem is a multi-class support vector machine problem; it can be formulated by using One-Against-All method. In this method it constructs *k* SVM models where *k* is the number of class; The *i*th SVM is trained with all of the examples in the *i*th class with positive labels, and all other examples with negative labels.
Thus, given *l* training data $(x_1, y_1), \dots\dots, (x_l, y_l)$,

Where, $x_i \in R^n$, $i = 1, \dots., l$ and $y_i \in \{1, \dots\dots, k\}$ is the class of x_i, the *i*th SVM soled the following problem in equation 3.13.

$$\min_{w^i, b^i, \xi^i} \quad \frac{1}{2}(w^i)^T w^i + C \sum_{j=1}^{l} \xi_j^i \tag{3.13}$$

Subject to, $\left(w^i\right)^T \varphi(x_j) + b^i \geq 1 - \xi_j^i \qquad if y_j = i,$

$$\left(w^i\right)^T \varphi\left(x_j\right) + b^i \leq -1 + \xi_j^i \qquad if y_j \neq i,$$

$$\xi_j^i \geq 0, \quad j = 1, \ldots, l,$$

Where the training data x_i are plotted to a high dimensional space by function φ and C is the penalty parameter.

Minimizing $\left(\frac{1}{2}\right)(w^i)^T w^i$ means that we would like to maximized $2/\left\|w^i\right\|$, the margin between two groups of data.

When data are not linear separable, there is a penalty term $C \sum_{j=1}^{l} \xi_j^i$ which can be reduced the number of training error. The basic concept behind SVM is to search for a balance between the regularization term $\left(\frac{1}{2}\right)(w^i)^T w^i$ and the training error.

After solving (n), there are *k* decision functions in equation 3.14:

$$\begin{gathered}(w^l)^T \varphi(x) + b^l \\ \cdot \\ \cdot \\ \cdot \\ (w^k)^T \varphi(x) + b^k\end{gathered} \tag{3.14}$$

We say *x* is in the class which has the largest value of the decision function in equation 3.15.

Class of

$$x = \arg \max_{i=1,\ldots k} ((w^i)^T \varphi(x) + b^i) \tag{3.15}$$

Practically, it is used to solve the dual problem of (n) whose number of variables is the same as the number of data in (n). Hence *k*, *l*-variable quadratic programming problems are solved.

Figure 3.3 represents the visual of 5-flod cross validation to estimate the expected error leave-one-out method can be used. The leave-one-out method is defined by leaving out one of the training examples and use remaining for training the model and using the left-out training set for testing the model. This procedure is repeated until all training examples are used to test the accuracy of the model. This procedure averages the error generated from all texting set.

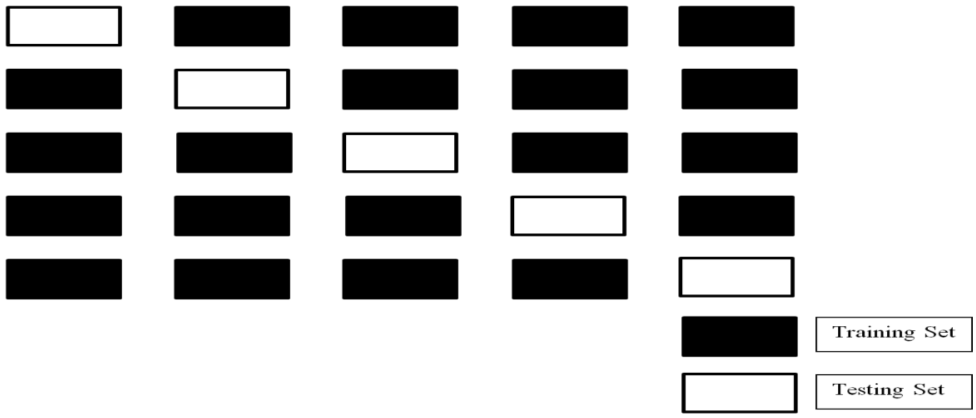

Figure 3.3: 5-Fold Cross-validation of Training Data Set

3.2 Support Vector Regression (SVR)

SVR is an extension of SVM because is follows almost all principles of SVM classification. A new type of loss function called ε^{-}insensitivity los function introduced by Vapnik used to perform regression in the high dimension feature space. The loss function defines the degree of penalty when the estimated value deviates from the real value. This ε^{-} insensitive function defines a tube: inside the tube, there is no penalty for deviation; while outside of the tube, a penalty occurs for any deviation. e defines the size of this tube which is used to balance the accuracy of approximation and the computational complexity. The model complexity can also be reduced by minimizing$\|w\|^2$. This can be described by introducing slack variables ξ_i and ξ_i^* where $i = 1, \ldots\ldots, n$ to measure the deviation of training sample outside ε^{-}sensitive zone. The objective function and constraints for SVR are in equation 3.16.

$$\text{Min } \frac{1}{2}\ \|w\|^2 + C \sum_{i=1}^{n} (\ \xi_i + \xi_i^*) \tag{3.16}$$

Subject to; $f(x_i) - y_i \le \varepsilon + \xi_i$

$$y_i - f(x_i) \le \varepsilon + \xi_i^*$$

$$\xi_i \ge 0,\ \xi_i^* \ge 0.\ i = 1,2,3, \ldots\ldots, n$$

This optimization problem can transform into the dual problem by including a dual set of Lagrange multiplier; which is in equation 3.17.

$$\max W(\alpha, \alpha^*) = -C\sum_{i=1}^{n}\left(\alpha^i + \alpha_i^*\right) + \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)y_i - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)\langle x_i, x_j\rangle \tag{3.17}$$

Subject to,

$$\sum_{i=1}^{n}(\,\alpha_i + \alpha_i^*) = 0$$

$\alpha_i^*, \alpha_i \in [0, C]$, $i = 1,2,3, \ldots\ldots, n$

In dual problem, kernel function $K\langle x_i, x_j\rangle$ is used to substitute $\langle x_i, x_j\rangle$. The desired regression function is then represented in equation 3.18.

$$f(x) = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)K\langle x_i, x_j\rangle + b \tag{3.18}$$

SVM generalization performance depends on a good setting of kernel parameters C, ϵ and kernel parameters.

The following equation is used to evaluate the performance of the proposed model. Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), relative Root Mean Squared Error (rRMSE) and Mean Squared Error (MSE) are used. Formulas of these evaluation measures are shown in the following equations 3.19, 3.20, 3.21 and 3.22 respectively.

$$\text{MAPE} = 100\,\frac{\sum_{i=1}^{n}\left|\frac{A-P}{A}\right|}{n} \tag{3.19}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|A-P|}{A} \tag{3.20}$$

$$\text{rRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{A-P}{A}\right)^2} \tag{3.21}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(A-P)^2 \tag{3.22}$$

Here, A stand for actual value and P stand for predicted value of machine and n is the number of samples. To calculate all these measures, a 5-fold cross validation method can be used.

3.3 Multivariate Linear Regression (MLR)

Regression is a statistical measurement-based predicting method for estimating the relationships among variables. It’s also used to evaluate the impacts of independent variables on dependent variables. Its outcome is based on the given input.

Linear regression is the simplest technique use to predict the relationship between a sclar dependent variable y and one explanatory or independent variables which is represented by X.

In many applications, there are more than one explanatory or independent variable that influences the response and this process is called multiple linear regressions. Multiple regression method is used for obtaining the relation between several independent variables and a dependent variable. The multiple linear regression model is just an extension of the simple linear regression model and represented as following equation 3.23.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \tag{3.23}$$

Where

y = an observed value of the response variable for a particular observation in the population

β_0 = the constant term

β_k = the coefficient for the kth explanatory/independent variable

x_k = a value of the kth explanatory variable for a particular observation

ε = the residual for the particular observation in the population

For the research work, predicting the reliability of machine is a multiple linear regression problem with 5 (k=5) independent or explanatory variable and one response or dependent variable y. the number of observations is n. So, the problem can be formulated by the following equation 3.24.

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_k x_{ik} + \varepsilon_i \qquad i = 1, \dots\dots.., n \tag{3.24}$$

Let, $y = \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix}$ $X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ . & . & . & & . \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$

$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_n \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix}$

Where, $y_{(n\times k)}$ = Matrix of response variable

$x_{(n\times k)}$ = Matrix of predictor variables

$\beta_{(n\times k)}$ = Matrix of coefficient

$\varepsilon_{(n\times k)}$ = Matrix of error

Solving the equation 3.24 mentioned above is based on least square method so that least square data fitting is a data fitting that estimates the coefficients (β) by minimizing the total value of the second power of deviations (E). In other words, least square data fitting is the same as the answer that minimizes the numerical product of $E'E=\Sigma E^2$. In order to solve the equation 3.24, it can be done as in following equation 3.25, 3.26.

$$x'.y = (x'.x)\beta \tag{3.25}$$

$$\beta = (x'.x)/x'.y \tag{3.26}$$

Three different criteria are used in order to evaluate the effectiveness of model and its ability to make precise prediction. These are Coefficient of Determination (R^2), Root Mean Square Error (RMSE) and Maximum Error are presented in equation 3.27, 3.28, 3.29 respectively.

$$R^2 = 1 - \frac{\sum (y_{obs} - y_{pre})^2}{\sum {y_{obs}}^2 - \frac{\sum {y_{pre}}^2}{n}} \tag{3.27}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_{pre} - y_{obs})^2}{n}} \quad (3.28)$$

$$Max\ (E) = max|y_{pre} - y_{obs}| \quad (3.29)$$

Where y_{pre} is predicted machine reliability and y_{obs} is observed machine reliability and n is the number of data.

3.4 Feature (Subset) Selection Method

There are three types of feature subset selection approaches: filters, wrappers, and embedded approaches which perform the features selection process as an integral part of a machine learning (ML) algorithm

Feature selection method generally consists of four steps shown in Figure 3.4 described below

(a) Generate candidate subset: The original feature set contains n number of features, the total number of competing candidate subsets to be generated is 2^n, which is a huge number even for medium-sized n. Subset generation is a search procedure that produces candidate feature subsets for evaluation based on a certain search strategy. The search strategy is broadly classified as complete (e,g. Breadth first search, Branch & bound, beam search, best first), heuristic (forward selection, backward selection, forward and backward selection), and random search (Las Vegas algorithm (LVW), genetic algorithm (GA), Random generation plus sequential selection (RGSS), simulated annealing (SA)).

(b) Subset evaluation function to evaluate the subset generated in the previous step (generate candidate subset) by using filter or wrapper approach. Filter and Wrapper approach differ only in the way in which they evaluate a subset of features. The filter approach is independent of the learning induction algorithm. Wrapper strategies for feature selection use an induction algorithm to estimate the merit of feature subsets. Wrappers often achieve better results than filters due to the fact that they are tuned to the specific interaction between an induction algorithm and its training data.

(c) Stopping Condition: Since the number of subsets can be enormous, some sort of stopping criterion is necessary. Stopping criteria may be based on a generation procedure/ evaluation function. Stopping criteria based on generation procedure include:

- Whether a predefined number of features are selected
- Whether a predefined number of iterations reached.

Stopping criteria based on an evaluation function can be:

- Whether addition (or deletion) of any feature does not produce a better subset
- Whether an optimal subset according to some evaluation function is obtained.

(d) Validation procedure: To check whether the feature subset selected is valid. Usually the result of original feature set is compared with the feature selected by filters/wrappers as input to some induction algorithm using artificial/real-world datasets. Another approach for validation is to use different feature selection algorithm to obtain relevant features and then compare the results by using classifiers on each relevant attribute subset.

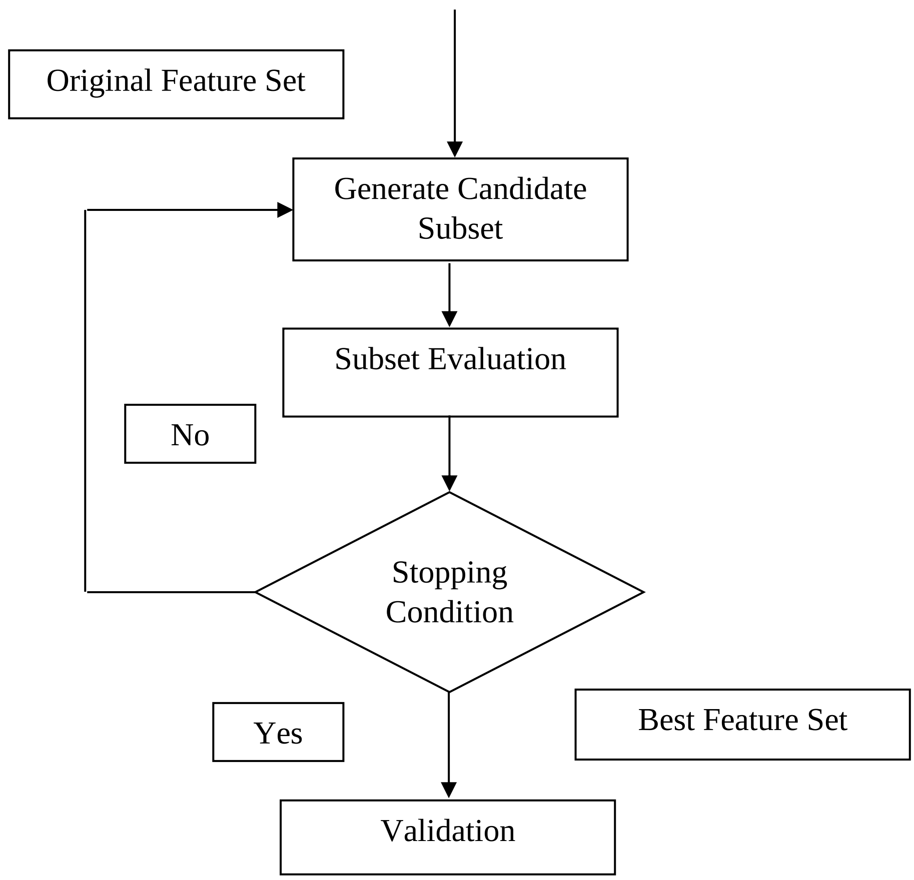

Figure 3.4: Steps for feature selection

3.4.1 The Filter Approach for Feature Selection

The filter approach actually precedes the actual classification process. The filter approach is independent of the learning induction algorithm, computationally simple fast and scalable. Using filter method, feature selection is done once and then can be provided as input to different classifiers. Various feature ranking and feature selection techniques have been proposed such as Correlation-based Feature Selection (CFS), Principal Component Analysis (PCA), Gain Ratio (GR), Chi-square Feature Evaluation, Markov blanket filter. Such filters are often used with forward selection, backward elimination, other methods. The authors have used decision tree consist of original feature set selection, feature subset selection and induction algorithms as filter approach to provide the relevant features as input to neural network classifier.

3.4.2 The Wrapper Approach for Feature Selection

Wrapper model approach shown in Figure 3.5 uses the method of classification itself to measure the importance of features set; hence the feature selected depends on the classifier model used. Wrapper methods generally result in better performance than filter methods because the feature selection process is optimized for the classification algorithm to be used. However, wrapper methods are too expensive for large dimensional database in terms of computational complexity and time since each feature set considered must be evaluated with the classifier algorithm used.

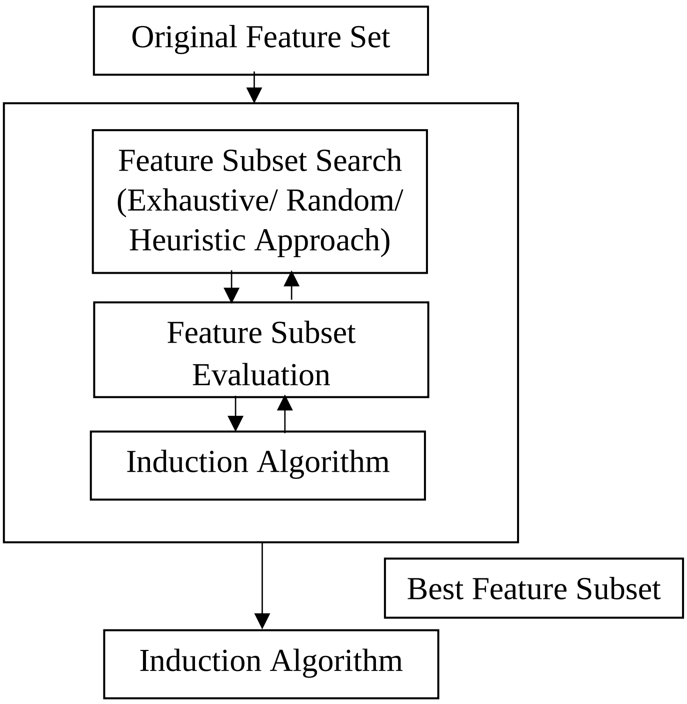

Figure 3.5: Wrapper approach for feature selection

CHAPTER 4: MODEL FORMULATION

Many residential and industrial electromechanical systems suffer from inevitable failure due to complex degradation processes and environmental condition. These unpredicted and unexpected failures may cause several consequences, including massive production losses, high repair costs and safety issues for personnel and environment. In Bangladesh, most of the cases in residential systems prefer corrective maintenance due to lack of maintenance knowledge and the inertia of advance expenses. Industrial system manages and tries to adopt some maintenance policies which are limited to inspection, cleaning and some replacement without proper plan and guidance. As a result, the benefit of planned maintenance is still out of reach due to a proper maintenance model.

Many maintenance models have been developed that helps to decrease the unexpected failure and reduce high operational cost. Among them some are effective, some are efficient and most of them are case specific. In case of traditional preventive maintenance policies, it is efficient than failure-based maintenance, but not effective, because it doesn't consider the systems reliability, as a result the maintenance actions is performed either too early or too late. On the other hand, in condition-based maintenance, it is more effective but not efficient, because the initial investment is too high which is feasible only for important, expensive and time sensitive machineries.

Considering this research gap, there is a scope to improve the traditional preventive maintenance model by improving its effectiveness. Without initial investment, the effectiveness can be improved by considering the systems reliability before predicting any maintenance requirements. The system reliability highly depends on various external and internal parameters of the system. This research work has considered Diesel Generator and Boiler as the system to develop the model.

4.1 Problem Formulation and Model Development

The purpose of this research work is to develop a planned preventive maintenance (PPM) model that can forecast the maintenance requirements more accurately and cost effectively. To develop this model machine reliability has been considered and the reliability value has been determined by analyzing qualitative and quantitative data. Various external and internal parameters of the machines will be selected based on the expert opinions and the relevant literature of that particular machine. The data will be collected based on these parameters to prepare the dataset. Machine learning algorithm will be used to formulate the problem.

A sequential framework has been designed which is shown in Figure 4.1 to develop and implement the models. The framework consists of four steps describe below:

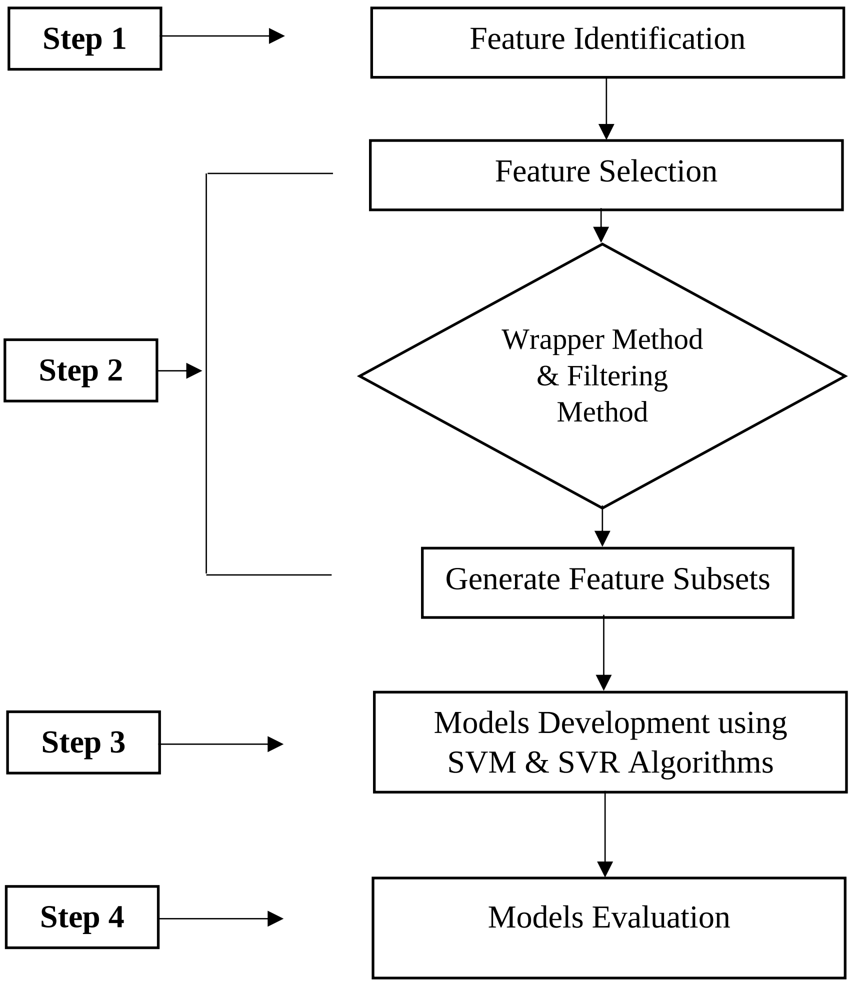

Figure 4.1: Flow chart of the current research

4.2 Feature Identification and Selection

To develop the model, following features are identified and considered based on the experts' opinions and reviewing the literatures.

i. Mean Time Between Failure (MTBF): The average period between system breakdowns.

ii. Mean Time to Repair (MTTR): The amount of time required to repair a system and restore it to fully functionality.

iii. Downtime (DT): The periods when a system is unavailable

iv. Machine Age (MA): The Completed amount of time of a system from its installation.

v. Machine Room Environment (MRE): The overall surrounding conditions round the machine where it is located.

vi. Average Operating Time (AOT): The average working period of a system.

vii. Manufacturer (M): The Company that produces the machines or system.

viii. Periodic Maintenance Practice (PMP): The types of maintenance actions are performed and its frequency.

ix. Alternative Machine (AM): The availability of backup or alternative source to maintain the services.

The above predictive features which have been identified earlier does not have same impact with its output variables. So, the impacts of each features on output variables are significant for model development. The selection of best feature subsets which is more significant is important to improve the model accuracy and interpretability of the model. The following two methods has been used for feature selection.

In this method the feature subsets are generated based on heuristics method (Forward Selection, Backward Selection, Forward & Backward Selection). This model uses the method of correlation-based feature selection to measure the importance of feature sets. The correlations of each feature with its output variable are determined by using data analysis tool pack of Microsoft Excel. Therefore, the features are ranked based on their correlation value from higher to lower. Finally, Backward elimination process are followed to generate probable feature sets.

In this method the feature subsets are generated based on Single Factor ANOVA test. Microsoft Excel data analysis tool pack is used to perform the single factor ANOVA test and determined the F and p-value for each feature. The insignificant features are identified based on their p-value. Avoiding the insignificant features, the desired feature subset has been obtained.

4.3 Model Development and Evaluation

Two separate models have been developed based on the selected feature sets. The first model is maintenance schedule prediction which is a classification problem and have been developed by using Support Vector Machine (SVM) algorithm. The second model is machine reliability prediction model development which is a regression problem and has been developed by using Support Vector Regression (SVR) algorithm.

4.3.1 Maintenance Schedule Prediction Model Development with Support Vector Machines (SVM)

The predetermined parameters has been used as an input variable which is called predictor variables are symbolized as followed. This predictor variable has been used to develop the maintenance schedule prediction model using SVM algorithm.

List of input variables:

Mean Time Between Failure (MTBF) (x_1)

Mean Time Between Repair (MTBR) (x_2)

Downtime (DT) (x_3)

Machine Age (MA) (x_4)

Machine Room Environment (MRE) (x_5)

Average Operating Time (AOT) (x_6)

Manufacturer (M) (x_7)

Periodic Maintenance Practice (PMP) (x_8)

Alternative Machine (AM) (x_9)

Let's *X* be the input feature vector. So,

$$X = \begin{bmatrix} \mathrm{x}_1 \\ \mathrm{x}_2 \\ \mathrm{x}_3 \\ \vdots \\ \mathrm{x}_n \end{bmatrix}$$

Where, *n = 9*

And *W* be the weight vector. So,

$$W = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \cdot \\ \cdot \\ \cdot \\ w_n \end{bmatrix}$$

Where, *n = 9*

The output variable which is called dependent variable y has five classes which is shown in Table 4.1 that indicates the planned preventive maintenance (PPM) periodicity is a multi-class support vector machine problem.

Table 4.1. Maintenance Class and Description

Maintenance Type	Description
Class-I	Monthly A-Check, Half-yearly B-Check and Yearly C-Check
Class-II	Daily A-Check, after running 300 hours B-Check and After running 840 hours C-Check
Class-III	Weekly A-Check, after running 300 hours B-Check and Yearly C-Check
Class-IV	Daily A-Check, after running 300 hours B-Check and Yearly C-Check
Class-V	Weekly A-Check, after running 300 hours, B-Check and after running 840 hours C-Check

Since the SVM classification problem is a multi-class support vector machine problem; it can be formulated by using One-Against-All method. In this method it constructs k SVM models where k is the number of class; in this classification problem we found 5 different classes. The ith SVM is trained with all of the examples in the ith class with positive labels, and all other examples with negative labels.

Thus, given l training data $(x_1, y_1), \ldots\ldots, (x_l, y_l)$,

Where, $x_i \in R^n$, $i = 1, \ldots., l$ and $y_i \in \{1, \ldots\ldots, k\}$ is the class of x_i, the ith SVM soled the following problem in equation 4.1.

$$\min_{w^i, b^i, \xi^i} \quad \frac{1}{2}(w^i)^T w^i + C \sum_{j=1}^{l} \xi_j^i \qquad (4.1)$$

Subject to, $(w^i)^T \varphi(x_j) + b^i \geq 1 - \xi_j^i \qquad if\, y_j = i,$

$$(w^i)^T \varphi(x_j) + b^i \leq -1 + \xi_j^i \qquad if\, y_j \neq i,$$

$$\xi_j^i \geq 0, \quad j = 1, \ldots., l,$$

Where the training data x_i are plotted to a high dimensional space by function φ and C is the penalty parameter.

Minimizing $\left(\frac{1}{2}\right)(w^i)^T w^i$ means that we would like to maximized $2/\|w^i\|$, the margin between two groups of data.

When data are not linear separable, there is a penalty term $C\sum_{j=1}^{l}\xi_j^i$ which can be reduced the number of training error. The basic concept behind SVM is to search for a balance between the regularization term $\left(\frac{1}{2}\right)(w^i)^T w^i$ and the training error.

After solving (n), there are *k* decision functions:

$$(w^l)^T\varphi(x) + b^l$$

$$\vdots$$

$$(w^k)^T\varphi(x) + b^k$$

We say x is in the class which has the largest value of the decision function in equation 4.2

$$\text{Class of } x = \arg\max_{i=1,\ldots k}((w^i)^T\varphi(x) + b^i) \tag{4.2}$$

Practically, we solve the dual problem of (n) whose number of variables is the same as the number of data in (n). Hence *k*, *l*-variable quadratic programming problems are solved.

The whole model has been formulated using MATLAB 2018a software package.

The maintenance schedule prediction model will be selected based on the accuracy level. The model that has the highest accuracy level among all models will be selected for further analysis and evaluation. The selected model will be evaluated by analyzing Confusion Matrix Parallel Coordination Plot and receiver operating characteristic curve or ROC Curve.

4.3.2 Machine Reliability Prediction Model Development with Support Vector Regression (SVR)

SVR is an extension of SVM because is follows almost all principles of SVM classification. A new type of loss function called ε^{-}insensitivity los function introduced by Vapnik used to perform regression in the high dimension feature space. The loss function defines the degree of penalty when the estimated value deviates from the real value. This ε^{-} insensitive function defines a tube: inside the tube, there is no penalty for deviation; while outside of the tube, a penalty occurs for any deviation. e defines the size of this tube which is used to balance the accuracy of approximation and the computational complexity. The model complexity can also be reduced by minimizing$\|w\|^2$. This can be described by introducing slack variables ξ_i and ξ_i^* where $i = 1, \ldots\ldots, n$ to measure the deviation of training sample outside ε^{-}sensitive zone. The objective function and constraints for SVR are in equation 4.3.

$$\text{Min } \frac{1}{2}\ \|w\|^2 + C\sum_{i=1}^{n}(\ \xi_i + \xi_i^*) \tag{4.3}$$

Subject to; $f(x_i) - y_i \le \varepsilon + \xi_i$

$y_i - f(x_i) \le \varepsilon + \xi_i^*$

$\xi_i \ge 0,\ \xi_i^* \ge 0.\ i = 1,2,3, \ldots., n$

This optimization problem can transform into the dual problem by including a dual set of Lagrange multiplier; which is in equation 4.4.

$$\max\ W(\alpha, \alpha^*) = -C\sum_{i=1}^{n}\left(\alpha^i + \alpha_i^*\right) + \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)y_i - \frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)\langle x_i, x_j\rangle \tag{4.4}$$

Subject to,

$$\sum_{i=1}^{n}(\ \alpha_i + \alpha_i^*) = 0$$

$\alpha_i^*, \alpha_i \in [0, C],\ i = 1,2,3, \ldots\ldots, n$

In dual problem, kernel function $K\langle x_i, x_j\rangle$ is used to substitute $\langle x_i, x_j\rangle$. The desired regression function is then in equation 4.5.

$$f(x) = \sum_{i=1}^{n}(\alpha_i^* - \alpha_i)K\langle x_i, x_j\rangle + b \tag{4.5}$$

SVM generalization performance depends on a good setting of kernel parameters C, ϵ and kernel parameters.

The following equation is used to evaluate the performance of the proposed model. Mean Absolute Error (MAE), relative Root Mean Squared Error (rRMSE) and Mean Squared Error

(MSE) are used. Formulas of these evaluation measures are shown in the following equations 4.6, 4.7 and 4.8 respectively.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|A-P|}{A} \tag{4.6}$$

$$\mathrm{rRMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{A-P}{A}\right)^2} \tag{4.7}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(A-P)^2 \tag{4.8}$$

Here, A stand for actual value and P stand for predicted reliability of machine and n is the number of samples. To calculate all these measures, a 5-fold cross validation method is used. The software packages MATLAB 2018a has been used in this study for the development of SVR model.

CHAPTER 5: MODEL IMPLEMENTATION

In order to implement the model, a real-life example has been considered as a case study in the context of Bangladesh. Diesel Generator (DG) is a complex electromechanical system widely used in residential and industrial building as a primary or secondary power supply units has been considered to implement the model. The research work has been performed on the basis of collected data from various residential and commercial office building of Dhaka city and some industrial building in the surrounding of Dhaka city. Although the needs, uses and maintenance practices of Diesel Generator (DG) are quite different for this different category projects however there are some growing needs and scope to improve the current practices for all the projects. Proper maintenance plan can enhance the expected productivity and services by ensuring the safety and environmental hazards.

In this research work, both qualitative and quantitative data has been collected by considering the subjective and objective manner of the machines. This data has been used to determine the machine condition by calculating its reliability value. These reliability values actually determine the maintenance requirements to develop the planned preventive maintenance schedule. So, the parameters selection and data collection are significantly important to develop the model. Expert opinions and literature on this particular machine have been used to select the parameters. Several projects were visited to collect the data. Quantitative data were acquired from the machine reading and record keeping registers. Project engineers, machine operators, building managers and experts were asking for qualitative data.

5.1 Maintenance Classification Model Formulation for Diesel Generator

In this work, maintenance schedule prediction involved two steps process, where the first step is to classify the maintenance requirements. For this purpose, a maintenance classification model has been developed by implementing Support Vector Machine (SVM). To develop the maintenance classification model, following features shown in Table 5.3 has been identified. Therefore, Support Vector Machine (SVM) has been trained using the dataset in Table 5.2 which contains historical data collected from the field survey. The features in Table 5.3 have been used to train SVM so that it can predict the maintenance requirements of any future machines. The standard checklist for Diesel Generator (DG) and the frequency of maintenance are given below.

A-Check:

Frequency of A-Check: Daily, Weekly or Monthly

Checklist for A-Check

1. Oil level, Coolant level, Gauge valve level

2. Radiator water level, Battery Voltages

3. Blower condition, All connection tight or not

4. Self-start, Voltage in all three phases, pf, Speed Governor

4. Check for any leakage

5. Cleanliness of DG set

B-Check:

Frequency of B-Check: After Running 300 Hours or After 6 Months; which become earlier

Checklist for B-Check

1. Check/ Change lube oil filter, Fuel filter

2. Check/ Change Air filter, Bypass filter

3. Check/ Change Water Separator,

4. Change the oil

5. Radiator cleaning with chemical

C-Check

Frequency of B-Check: After Running 840-1500 Hours or After 1Year; which become earlier

Checklist for B-Check

1. All steps of B Check

2. Check/ Change of Air filter (If needed)

3. Check/Change of other filters mentioned in B Check

4. Change of gasket of chambers

5. Radiator Cleaning

6. Check pollution level

Based on the above checklist and maintenance frequency, experts classify the maintenance requirements into five classes which is given below Table 5.1.

Table 5.1: Maintenance Classes for Maintenance Schedule Prediction Model Development

Class	**Dataset Output (Y)**	**Definition**
Class-I	1	Monthly A-Check, Half-yearly B-Check, Yearly C-Check
Class-II	2	Daily A-Check, After running 300 hours; B-Check, After running 840 hours; C-Check
Class-III	3	Weekly A-Check, After running 300 hours; B-Check, Yearly C-Check
Class-IV	4	Daily A-Check, After running 300 hours; B-Check, Yearly C-Check
Class-V	5	Weekly A-Check, After running 300 hours; B-Check, After running 840 hours, C-Check

For identifying a classification model which can predict the maintenance requirements based on the above-mentioned features different SVM models are checked by using MATLAB toolbox. Those models are given below.

i. Linear SVM- Uses linear Kernel

ii. Quadratic SVM- Uses quadratic Kernel

iii. Cubic SVM- Uses cubic Kernel

iv. Fine Gaussian SVM- Uses Gaussian Kernel with Kernel scale (√(No of features))⁄4

v. Medium Gaussian SVM- Uses Gaussian Kernel with Kernel scale √(No of features)

vi. Coarse Gaussian- Uses Gaussian Kernel with Kernel scale (√(No of features))*4

Table 5.2: Training Dataset Sample for Maintenance Schedule Prediction Model Development

Machine	**Features (X)**									**Output (Y)**
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	
	X1	X2	X3	X4	X5	X6	X7	X8	X9	MP
DG-1	9M	3	L	5	VP	30	1	VP	A	1
DG-2	5M	11	H	10	VP	105	2	VP	A	2
DG-3	4M	13	H	15	VP	115	3	P	NA	2
DG-4	10M	14	M	2	G	40	1	VP	A	1
DG-5	5M	11	H	13	VP	120	1	M	NA	2
DG-6	7M	7	M	7	P	50	4	P	A	3
DG-7	9M	2	L	2	M	35	2	VP	A	1
DG-8	4M	12	H	17	P	130	2	P	NA	2
DG-9	6M	8	M	8	P	70	3	VP	A	3
DG-10	9M	1	L	1	G	25	2	P	A	1
DG-11	2M	14	M	9	P	50	1	VP	NA	4
DG-12	11M	4	M	1	VG	45	3	M	A	1
DG-13	4M	15	H	9	VP	80	1	VP	NA	2
DG-14	9M	5	L	4	VP	30	2	VP	A	1
DG-15	9M	5	L	2	G	45	1	M	NA	1
DG-16	5M	20	H	10	VP	105	2	P	NA	2
DG-17	10M	4	L	3	M	20	1	M	A	1
DG-18	8M	9	H	5	P	75	3	VP	A	5
DG-19	6M	6	M	9	P	70	1	VP	A	3
DG-20	11M	5	L	5	VG	20	1	M	NA	1

Machine	Features (X)									Output (Y)
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	MP
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-21	1M	15	L	2	M	35	3	VP	NA	4
DG-22	6M	6	M	9	P	65	2	VP	A	3
DG-23	11M	3	M	2	G	40	1	M	A	1
DG-24	5M	12	H	13	VP	105	5	VP	NA	2
DG-25	9M	3	L	5	VG	30	1	VP	A	1
DG-26	10M	2	L	5	M	35	1	M	A	1
DG-27	4M	11	H	12	VP	110	3	VP	NA	2
DG-28	11M	2	L	3	M	35	1	VP	A	1
DG-29	7M	10	H	7	P	60	4	VP	NA	3
DG-30	8M	9	M	6	P	80	2	VP	A	5
DG-31	4M	13	H	20	VP	135	1	P	A	2
DG-32	7M	8	L	6	P	50	2	M	NA	3
DG-33	5M	12	H	5	VP	115	1	M	A	2
DG-34	11M	3	L	2	VG	25	1	VP	A	1
DG-35	4M	14	H	25	P	125	5	P	NA	2
DG-36	7M	6	M	7	P	65	1	G	A	3
DG-37	2	7	H	10	P	120	2	M	NA	5
DG-38	9M	4	L	2	G	40	1	M	A	1
DG-39	4M	9	H	11	P	90	4	VP	A	2
DG-40	7M	10	M	8	P	60	3	P	A	3

5.3 Feature Selection and Model Development for Diesel Generator

Table 5.2 represents the training dataset for all features. In this stage the subsets of above features will be selected using wrapper method and filtering method.

In Wrapper Method the feature subsets has been generated based on heuristics method (Forward Selection, Backward Selection, Forward & Backward Selection). This model has been used the method of correlation-based feature selection to measure the importance of feature sets. The correlations of each feature with its output variable are determined by using data analysis tool pack of Microsoft Excel. The negative sign represents the negative correlation while the positive sign represents positive correlation of the features with its output variable. The higher the values are, the higher the correlation of the features with the output variable. Therefore, the features has been ranked based on their correlation value from higher to lower.

Table 5.3: Correlation of Features with Output Variables

Features	**Correlation (Person's r)**	**Ranking**
Mean Time Between Failure (MTBF)	-0.58	1
Mean Time to Repair (MTTR)	0.43	2
Downtime (DT)	0.37	5
Machine Age (MA)	0.30	7
Machine Room Environment (MRE)	-0.39	4
Average Operating Time (AOT)	0.40	3
Manufacturer (M)	0.12	9
Periodic Maintenance Practice (PMP)	-0.23	8
Alternative Machine (AM)	-0.32	6

The following subsets has been selected using Backward Elimination process with the help of the ranking determined in Table 5.3.

Subset -1: {X1, X2, X3, X4, X5, X6, X7, X8, X9}
Subset -2: {X1, X2, X3, X4, X5, X6, X8, X9}
Subset -3: {X1, X2, X3, X4, X5, X6, X9}
Subset -4: {X1, X2, X3, X5, X6, X9}
Subset -5: {X1, X2, X3, X5, X6}
Subset -6: {X1, X2, X5, X6}
Subset -7: {X1, X2, X6}
Subset -8: {X1, X2}

In Table 5.4 the accuracy of different SVM models are given according to their subset. Five-fold cross validation has been used to determine the accuracy of the models. For all subset**,** six SVM models are checked to find the model which has the highest accuracy level.

Table 5.4: Prediction Accuracy Level of a Model Based on Wrapper Method

Subset	Model No	SVM Type	Accuracy (%)
Subset-1	Model-1	Linear SVM	86.3%
	Model-2	Quadratic SVM	83.8%
	Model-3	Cubic SVM	85%
	Model-4	Fine Gaussian SVM	16.3%
	Model-5	Medium Gaussian SVM	86.3%
	Model-6	Coarse Gaussian	77.5%
Subset-2	Model-7	Linear SVM	86.3%
	Model-8	Quadratic SVM	85%
	Model-9	Cubic SVM	86.3%
	Model-10	Fine Gaussian SVM	18.8%
	Model-11	Medium Gaussian SVM	85%
	Model-12	Coarse Gaussian	86.3%
Subset-3	Model-13	Linear SVM	87.5%
	Model-14	Quadratic SVM	87.5%
	Model-15	Cubic SVM	88.8%
	Model-16	Fine Gaussian SVM	41.3%
	Model-17	Medium Gaussian SVM	87.5%
	Model-18	Coarse Gaussian	87.5%
Subset-4	Model-19	Linear SVM	87.5%
	Model-20	Quadratic SVM	87.5%
	Model-21	Cubic SVM	88.8%
	Model-22	Fine Gaussian SVM	51.2%
	Model-23	Medium Gaussian SVM	87.5%
	Model-24	Coarse Gaussian	86.3%
	Model-25	Linear SVM	87.5%
	Model-26	Quadratic SVM	87.5%

Subset-5	Model-27	Cubic SVM	88.8%
	Model-28	Fine Gaussian SVM	55%
	Model-29	Medium Gaussian SVM	87.5%
	Model-30	Coarse Gaussian	87.5%
Subset-6	Model-31	Linear SVM	88.8%
	Model-32	Quadratic SVM	90. 0%
	Model-33	Cubic SVM	92. 5%
	Model-34	Fine Gaussian SVM	91.3%
	Model-35	Medium Gaussian SVM	88.8%
	Model-36	Coarse Gaussian	88.8%
Subset-7	Model-37	Linear SVM	90.0%
	Model-38	Quadratic SVM	91. 0%
	Model-39	Cubic SVM	91. 0%
	Model-40	Fine Gaussian SVM	67.5%
	Model-41	Medium Gaussian SVM	90.0%
	Model-42	Coarse Gaussian	88.8%
Subset-8	Model-43	Linear SVM	90.0%
	Model-44	Quadratic SVM	91. 0%
	Model-45	Cubic SVM	90.0%
	Model-46	Fine Gaussian SVM	81.3%
	Model-47	Medium Gaussian SVM	88.8%
	Model-48	Coarse Gaussian	86.3%

Table 5.5 has drawn a comparison among the subsets based on their accuracy level for all model. From this table it can be visualized the highest accuracy for each subset. Table 5.6 illustrates a comparison among the best accuracy level model (s) developed is each subset. Finally, it has been observed in Table 5.6 that subset-6, model-33 has the best accuracy level 92.5 % among all subsets determined in Wrapper method.

Table 5.5: Comparison Among Subset's Based on Accuracy Level

SVM Type	Subset-1	Subset-2	Subset-3	Subset-4	Subset-5	Subset-6	Subset-7	Subset-8
	Model: (1-6)	Model: (7-12)	Model: (13-18)	Model: (19-24)	Model: (25-30)	Model: (31-36)	Model: (37-42)	Model: (43-48)
Linear SVM	86.3%	86.3%	87.5%	87.5%	87.5%	88.8%	90.0%	90.0%
Quadratic SVM	83.8%	85%	87.5%	87.5%	87.5%	90. 0%	91. 0%	91. 0%
Cubic SVM	85%	86.3%	88.8%	88.8%	88.8%	92. 5%	91. 0%	90.0%
Fine Gaussian SVM	16.3%	18.8%	41.3%	51.2%	55%	91.3%	67.5%	81.3%
Medium Gaussian SVM	86.3%	85%	87.5%	87.5%	87.5%	88.8%	90.0%	88.8%
Coarse Gaussian	77.5%	86.3%	87.5%	86.3%	87.5%	88.8%	88.8%	86.3%

Table 5.6: Comparison Among Models Based on Best Accuracy Level of Each Subset

Subset	Highest Accuracy Model	Accuracy
Subset-1	Model-1 & Model-5	86.3%
Subset-2	Model-7, Model-9 & Model-12	86.3%
Subset-3	Model-15	88.8%
Subset-4	Model-21	88.8%
Subset-5	Model-27	88.8%
Subset-6	Model-33	92. 5%
Subset-7	Model-38, Model-39	91. 0%
Subset-8	Model-44	91. 0%

In Filtering Method the feature subsets are generated based on single factor ANOVA test. Microsoft Excel data analysis tool pack is used to perform the single factor ANOVA test and determined the F and p-value for each feature. Table 5.7 illustrates the results with decision where three features are not significant as their p-values is more than 0.05. Avoiding these three insignificant features, the desired feature subset is given below.

Table 5.7: Feature Selection Using Filtering Method

Feature	**ANOVA**		**Decision**
	F	p-value	
Mean Time Between Failure (MTBF)	225.41	0	Significant
Mean Time to Repair (MTTR)	122.38	0	Significant
Downtime (DT)	19.94	0	Significant
Machine Age (MA)	83.23	0	Significant
Machine Room Environment (MRE)	2.62	0.1072	Not Significant
Average Operating Time (AOT)	268.02	0	Significant
Manufacturer (M)	1.62	0.2048	Not Significant
Periodic Maintenance Practice (PMP)	2.66	0.1048	Not Significant
Alternative Machine (AM)	114.04	0	Significant

Selected Features' Set {X1, X2, X3, X4, X6, X9}

In Table 5.8 The prediction accuracy of different SVM models are given according to the selected subset. Five-fold cross validation has been used to determine the accuracy of the models. Six SVM models are checked to find the model which has the highest accuracy level. Finally, it has been observed in Table 5.8 that Model-49, 51 and 53 has the highest accuracy level 88.8%.

Table 5.8: Prediction Accuracy Level of a Model Based on Filtering Method

Feature Set	Model No	SVM Type	Accuracy (%)
{X1, X2, X3, X4, X5, X6, X9}	Model-49	Linear SVM	88.8%
	Model-50	Quadratic SVM	87.5%
	Model-51	Cubic SVM	88.8%
	Model-52	Fine Gaussian SVM	55.5%
	Model-53	Medium Gaussian SVM	88.8%
	Model-54	Coarse Gaussian	87.5%

5. 4 Model Selection and Evaluation

In this section the best model has been selected based on their accuracy level. In Wrapper method there were eight subsets among them subset 6 selected based on the highest accuracy level of its Model-33. On the other hand, in a single subset determined in Filtering method which has six different models.

Table 5.9: Comparison Between Wrapper Method & Filtering Method

SVM Type	Accuracy (%)	
	Wrapper Method	Filtering Method
	{X1, X2, X5, X6}	{X1, X2, X3, X4, X6, X9}
Linear SVM	88.8%	88.8%
Quadratic SVM	90. 0%	87.5%
Cubic SVM	92. 5%	88.8%
Fine Gaussian SVM	91.3%	55.5%
Medium Gaussian SVM	88.8%	88.8%
Coarse Gaussian	88.8%	87.5%

Table 5.9 illustrate a comparison between Wrapper method & Filtering method with its different models. Finally, it is observed that Model-33 in Wrapper method has the best accuracy level 92.5% with a training time of 0.50145 Seconds at a speed of 2200 observations per second in Figure 5.1.

Figure 5.1: Accuracy of Selected Model-33 from Subset-6

As Model-33 has best accuracy level, the model is studied further in different perspectives. The following Table 5.10 has shown class wise prediction accuracy and error level which is drawn from the confusion matrix of the selected Model-33.

Table 5.10: Class Wise Accuracy and Error Level of the Selected Model

Class	True Positive Rate	False Negative Rate
Class-I	94%	6%
Class-II	96%	4%
Class-III	93%	7%
Class-IV	50%	50%
Class-V	67%	33%

Figure 5.2 represent the confusion matrix of the selected Model-33, where it can be seen that for class one, 30 sample fallen in true class as it is predicted and only 2 sample fallen incorrectly which should fall in class 3, so that the sample 94% correctly belonging in class one. Similarly, for class two and three the accuracy is 96% and 93% respectively. But class four belonging only 2 samples where the true positive rate is 50%. Finally, class five represents 67% true positive rate with 33% false negative rate.

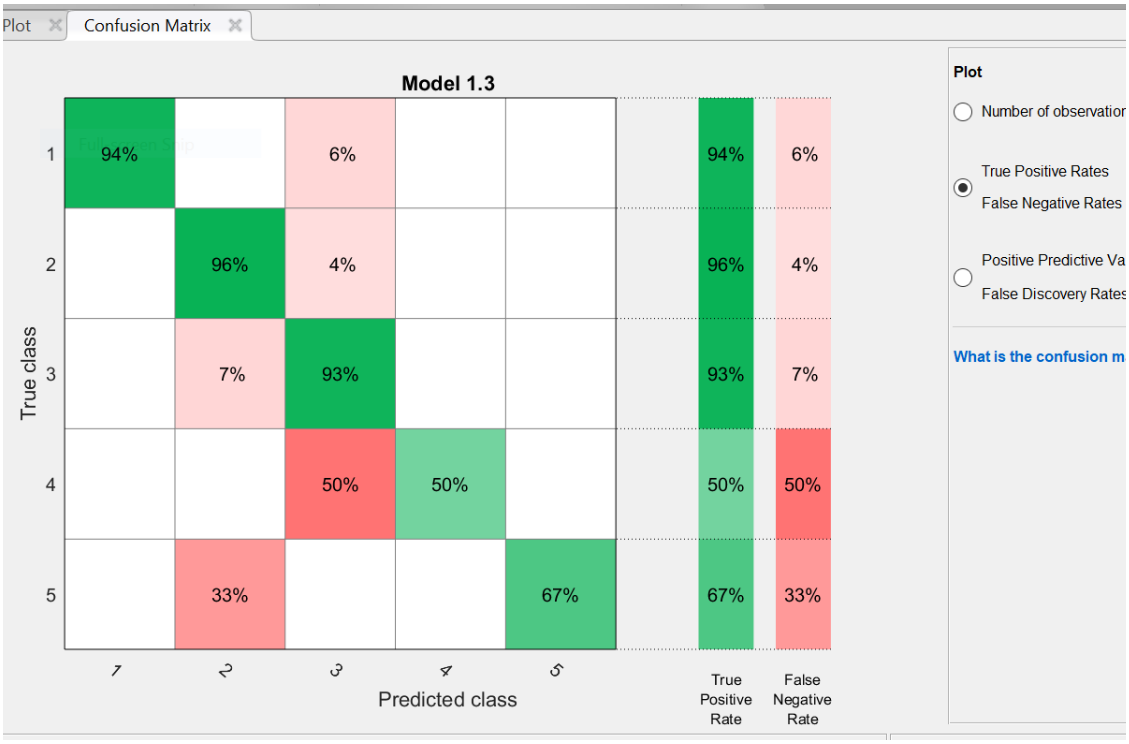

Figure 5.2: Confusion Matrix for Selected Model-33

In Figure 5.3, Figure 5.4, Figure 5.5 and Figure 5.6, scatter plots are given which reflects the correlation between different feature which is very important for overall model accuracy.

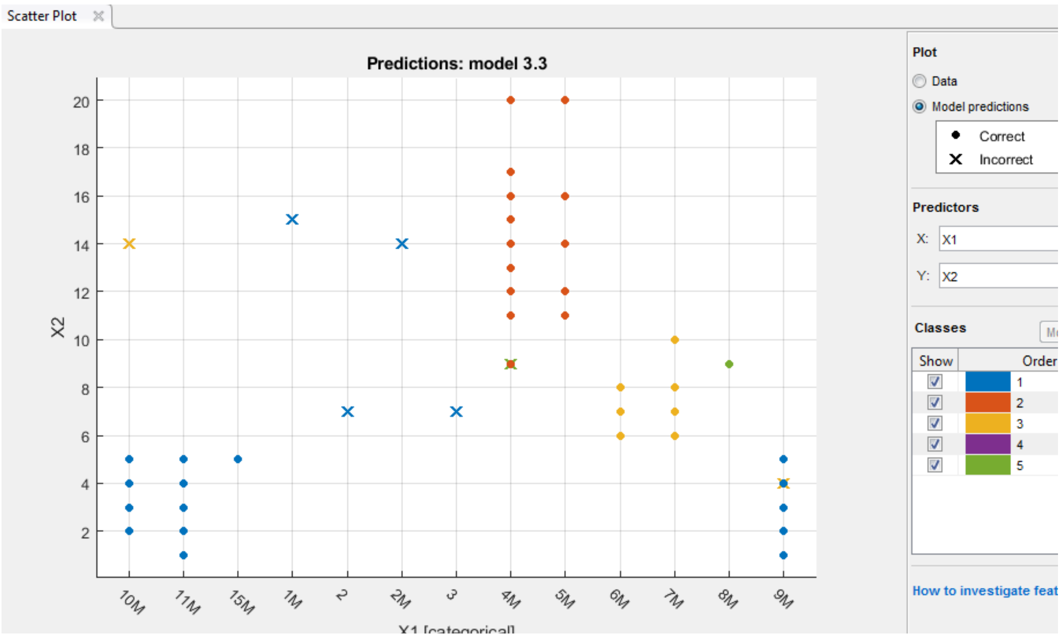

Figure 5.3: Scatter Plot for MTBF versus MTTR

In Figure 5.3, when data points are presented with respect to Mean Time between Failure (MTBF) and Mean Time to Repair (MTTR) they are easily separable. So, it is clear that these two features have a greater impact on maintenance prediction.

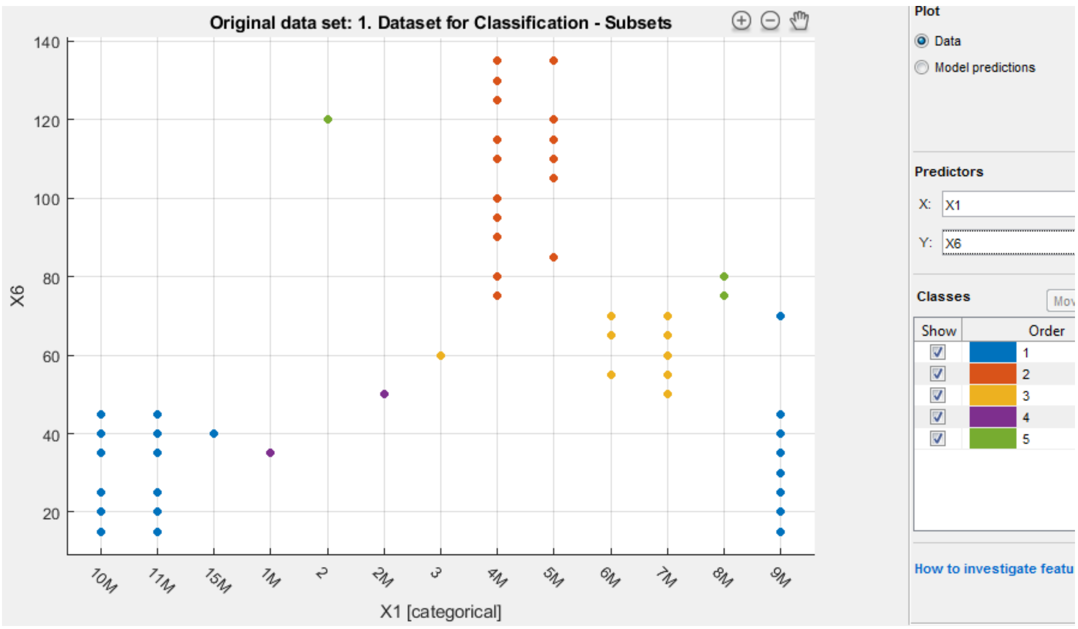

Figure 5.4: Scatter Plot for MTBF versus AOT

From Figure 5.4 it can be observed that when data points are presented with respect to Mean Time between Failure (MTBF) and Average Operating Time (AOT), they are also easily spreadable and has a positive impact on output.

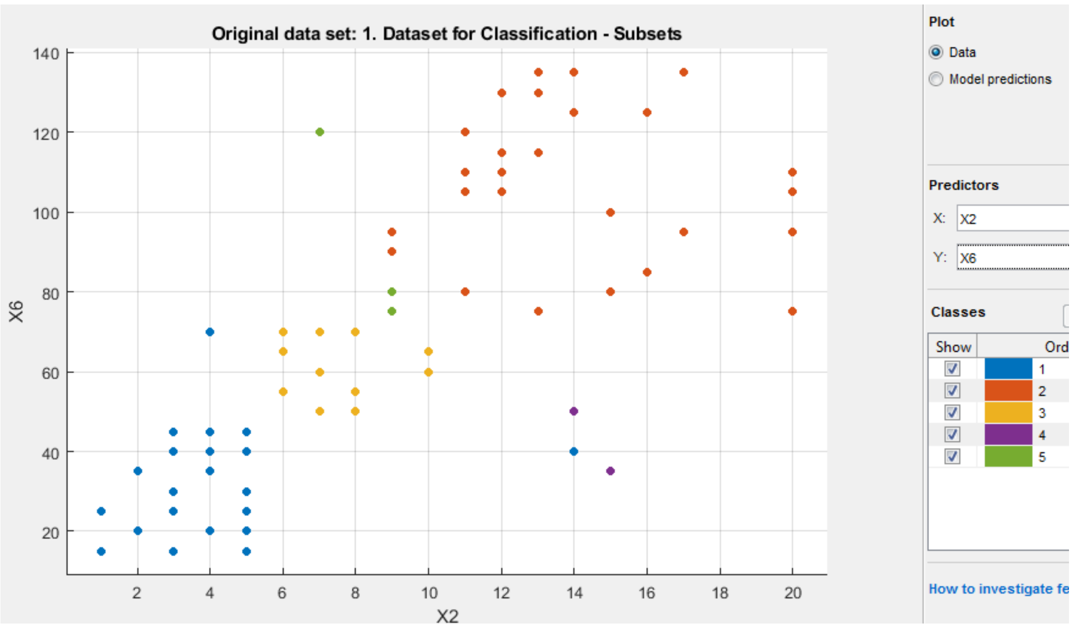

Figure 5.5: Scatter Plot for MTTR versus AOT

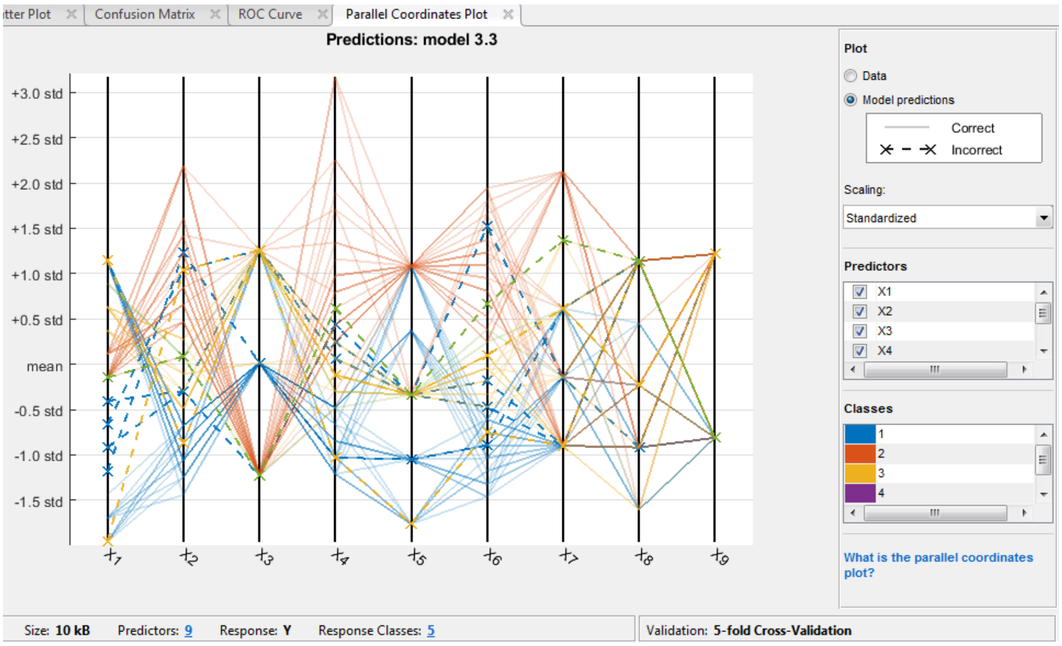

Figure 5.6: Parallel Coordinate Plot for all Features

Figure 5.5 illustrates that when data points are plot in a two-dimensional space with respect to Mean Time to Repair (MTTR) and Average Operating Time (AOT), These features can separate the classes in a better way.

Figure 5.6 demonstrates the parallel plot of all features which reflects how each feature has an impact on the classification model. In Figure 5.6 blue lines, red lines, yellow lines and orange lines respectively represent four different maintenance prediction classes. The areas between the features, where these lines are easily separable have a greater impact on the classification model compared to other features. All parameters are standardized h

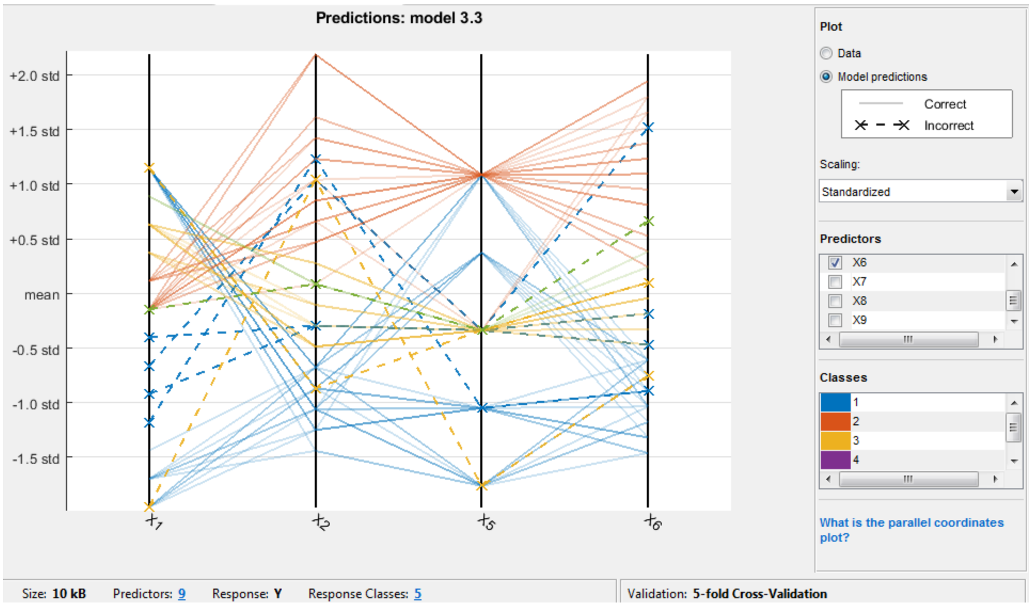

Figure 5.7: Parallel Coordinate Plot for Selected Model-33

In Figure 5.7 the parallel plot in Figure 5.6 is modified by keeping only four features those are Mean Time between Failure (MTBF), Mean Time to Repair (MTTR), Machine Room Environment (MRE) and Average Operating Time (AOT) which was selected in Subset-6 can separate the classes in a better way.

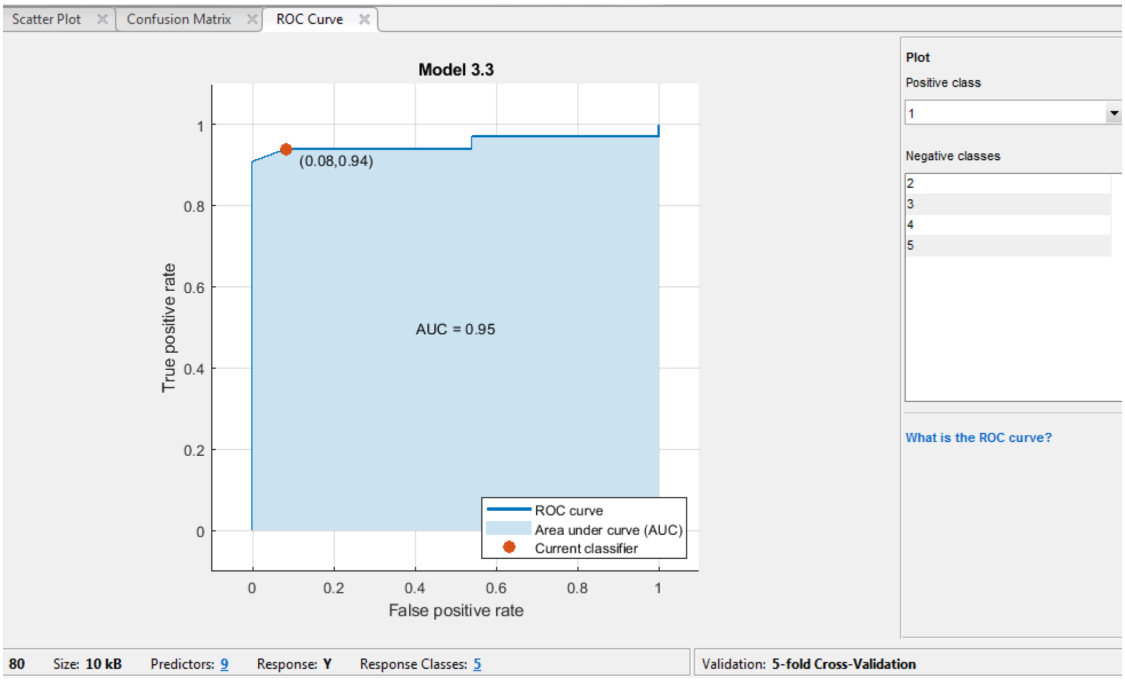

Figure 5.8: ROC Curve for Selected Model-33

Figure 5.8 Represent the ROC curve to quantify the performance of the classifier. The illustration shows that the curve hugs the upper left corner of the plot means a very good classification capability of the model. From the figure the true positive rate is 0.99 whereas the false positive rate is only 0.08. The Area under the Curve (AUC) is 0.95 that cover the maximum area of the curve which also represent the better performance of the model.

5.4 Machine Reliability Calculation of Diesel Generator through SVR

Machine reliability is the availability of the machine for a specific period of time is a significant thing to know for maintenance management. The probability of failure or breakdown of a machine depends on the reliability of the machine and it can be determined in different ways. Here a model is developed by using Support Vector Regression (SVR) to know the reliability value of the machine.

In this research, Machine reliability will be calculated considering the same parameters used in classification. A regression model is developed by implementing Support Vector Regression (SVR). To develop the regression model, Support Vector Regression (SVR) is trained using the historical data collected from the field survey. To develop the model a training dataset is used which is given as a sample in Table 5.11 and the complete dataset is given in Appendix B.

Table 5.11: Sample Dataset for Machine Reliability Prediction Model Development

Machine	Features (X)									Output (Y)
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	Machine Reliability
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-1	9M	3	L	5	VP	30	1	VP	A	0.76
DG-2	5M	11	H	10	VP	105	2	VP	A	0.59
DG-3	4M	13	H	15	VP	115	3	P	NA	0.54
DG-4	10M	14	M	2	G	40	1	VP	A	0.88
DG-5	5M	11	H	13	VP	120	1	M	NA	0.48
DG-6	7M	7	M	7	P	50	4	P	A	0.55
DG-7	9M	2	L	2	M	35	2	VP	A	0.74
DG-8	4M	12	H	17	P	130	2	P	NA	0.45
DG-9	6M	8	M	8	P	70	3	VP	A	0.56
DG-10	9M	1	L	1	G	25	2	P	A	0.76
DG-11	2M	14	M	9	P	50	1	VP	NA	0.67
DG-12	11M	4	M	1	VG	45	3	M	A	0.76
DG-13	4M	15	H	9	VP	80	1	VP	NA	0.62
DG-14	9M	5	L	4	VP	30	2	VP	A	0.77
DG-15	9M	5	L	2	G	45	1	M	NA	0.81
DG-16	5M	20	H	10	VP	105	2	P	NA	0.58
DG-17	10M	4	L	3	M	20	1	M	A	0.74
DG-18	8M	9	H	5	P	75	3	VP	A	0.55
DG-19	6M	6	M	9	P	70	1	VP	A	0.60
DG-20	11M	5	L	5	VG	20	1	M	NA	0.76
DG-21	1M	15	L	2	M	35	3	VP	NA	0.73
DG-22	6M	6	M	9	P	65	2	VP	A	0.62
DG-23	11M	3	M	2	G	40	1	M	A	0.77
DG-24	5M	12	H	13	VP	105	5	VP	NA	0.52
DG-25	9M	3	L	5	VG	30	1	VP	A	0.87
DG-26	10M	2	L	5	M	35	1	M	A	0.73

5.4.1 Feature Selection and Model Development for Regression

In this stage the subsets of above features has been selected using Wrapper method and Filtering method.

In Wrapper method the feature subsets has been generated based on heuristics method (Forward Selection, Backward Selection, Forward & Backward Selection). This model has been used the method of correlation-based feature selection to measure the importance of feature sets. The correlations of each feature with its output variable are determined by using data analysis tool pack of Microsoft excel. Therefore, the features has been ranked based on their correlation value from higher to lower.

Table 5.12: Correlation of Features with Output Variable

Features	**Correlation (Person's r)**	**Ranking**
Men Time Between Failure (MTBF)	0.71	4
Mean Time to Repair (MTTR)	-0.67	6
Downtime (DT)	-0.78	2
Machine Age (MA)	-0.75	3
Machine Room Environment (MRE)	0.69	5
Average Operating Time (AOT)	-0.82	1
Manufacturer (M)	-0.52	7
Periodic Maintenance Practice (PMP)	0.27	9
Alternative Machine (AM)	0.38	8

The following subsets has been selected using Backward Elimination process with the help of the correlation values of the features determined in the Table 5.12

Subset -1: {X1, X2, X3, X4, X5, X6, X7, X8, X9}

Subset -2: {X1, X2, X3, X4, X5, X6, X7, X9}

Subset -3: {X1, X2, X3, X4, X5, X6, X7}

Subset -4: {X1, X2, X3, X4, X5, X6}

Subset -5: {X1, X3, X4, X5, X6}

Subset -6: {X1, X3, X4, X6}

Subset -7: {X3, X4, X6}

Subset -8**:** {X3, X6}

Table 5.13: Prediction Accuracy Level of a Model Based on Wrapper Method

Subset	Model No	SVM Type	MAE	R squared
Subset-1	Model-1	Linear SVM	0.049129	0.69
	Model-2	Quadratic SVM	0.052711	0.64
	Model-3	Cubic SVM	0.057799	0.58
	Model-4	Fine Gaussian SVM	0.097849	0.07
	Model-5	Medium Gaussian SVM	0.048901	0.68
	Model-6	Coarse Gaussian	0.047874	0.71
Subset-2	Model-7	Linear SVM	0.049739	0.70
	Model-8	Quadratic SVM	0.056578	0.60
	Model-9	Cubic SVM	0.064677	0.37
	Model-10	Fine Gaussian SVM	0.08882	0.20
	Model-11	Medium Gaussian SVM	0.047962	0.71
	Model-12	Coarse Gaussian	0.047495	0.72
Subset-3	Model-13	Linear SVM	0.048437	0.71
	Model-14	Quadratic SVM	0.053679	0.66
	Model-15	Cubic SVM	0.061253	0.55
	Model-16	Fine Gaussian SVM	0.087982	0.21
	Model-17	Medium Gaussian SVM	0.047404	0.71
	Model-18	Coarse Gaussian	0.047147	0.72
Subset-4	Model-19	Linear SVM	0.054736	0.64
	Model-20	Quadratic SVM	0.057737	0.63
	Model-21	Cubic SVM	0.07308	0.37
	Model-22	Fine Gaussian SVM	0.081908	0.29
	Model-23	Medium Gaussian SVM	0.058341	0.61
	Model-24	Coarse Gaussian	0.053877	0.67
Subset-5	Model-25	Linear SVM	0.056505	0.62
	Model-26	Quadratic SVM	0.0597	0.60
	Model-27	Cubic SVM	0.096682	-1.27
	Model-28	Fine Gaussian SVM	0.075384	0.34
	Model-29	Medium Gaussian SVM	0.056526	0.64
	Model-30	Coarse Gaussian	0.05238	0.69

Subset	Model No	SVM Type	MAE	R squared
Subset-6	Model-31	Linear SVM	0.059029	0.61
	Model-32	Quadratic SVM	0.068886	0.48
	Model-33	Cubic SVM	0.12	-3.22
	Model-34	Fine Gaussian SVM	0.066269	0.47
	Model-35	Medium Gaussian SVM	0.057951	0.62
	Model-36	Coarse Gaussian	0.053516	0.68
Subset-7	Model-37	Linear SVM	0.053756	0.66
	Model-38	Quadratic SVM	0.056682	0.55
	Model-39	Cubic SVM	0.055249	0.62
	Model-40	Fine Gaussian SVM	0.055317	0.62
	Model-41	Medium Gaussian SVM	0.052763	0.67
	Model-42	Coarse Gaussian	0.052388	0.68
Subset-8	Model-43	Linear SVM	0.052749	0.66
	Model-44	Quadratic SVM	0.051662	0.66
	Model-45	Cubic SVM	0.056045	0.45
	Model-46	Fine Gaussian SVM	0.051226	0.65
	Model-47	Medium Gaussian SVM	0.048997	0.69
	Model-48	Coarse Gaussian	0.052964	0.67

In Table 5.13 represents the machine reliability based on Mean Absolute Error (MAE) and R squared value for different SVM models for different subsets. Five-fold cross validation has been used to determine the accuracy of the models. For all subset**,** six SVM models has been checked to find the model which has the highest accuracy level.

Mode- l2 from subset-2 and Model -18 from Subset-3 represents highest R Squared value 0.72 among all the models. Therefore, these two models are independently selected from Wrapper method for further comparison with the best model which will be selected from Filtering method.

In Filtering Method single factor ANOVA test has been performed by using Microsoft excel data analysis tool pack. Table 5.14 illustrates the results with decision where one feature is not significant as the its p-values is more than 0.05. Avoiding this insignificant feature, the desired feature subset is given below.

Selected Features' Set {X1, X2, X3, X4, X5, X6, X7, X8}

Table 5.14: Feature Selection Using Filtering Method

Feature	ANOVA		Decision
	F	p-value	
Mean Time Between Failure (MTBF)	407.02	0	Significant
Mean Time to Repair (MTTR)	181.80	0	Significant
Downtime (DT)	148.08	0	Significant
Machine Age (MA)	129.84	0	Significant
Machine Room Environment (MRE)	118.70	0	Significant
Average Operating Time (AOT)	279.14	0	Significant
Manufacturer (M)	107.09	0	Significant
Periodic Maintenance Practice (PMP)	137.58	0	Significant
Alternative Machine (AM)	0.8179	0.3671	Not Significant

In Table 5.15 demonstrate the Mean Absolute Error (MAE) and R squared value for different SVM models. The higher the R square values the higher the machine reliability. Five-fold cross validation has been used to determine the accuracy of the models.

Table 5.15: Prediction Accuracy Level of a Model Based on Filtering Method

Feature Set	Model No	SVM Type	MAE	R squared
{X1, X2, X3, X4, X5, X6, X7, X8}	Model-49	Linear SVM	0.048868	0.69
	Model-50	Quadratic SVM	0.051393	0.67
	Model-51	Cubic SVM	0.058005	0.58
	Model-52	Fine Gaussian SVM	0.096651	0.08
	Model-53	Medium Gaussian SVM	0.048602	0.69
	Model-54	Coarse Gaussian	0.047077	0.72

Mode- 54 represents highest R Squared value 0.72 among all the models generated in Filtering method. Therefore, this model is selected to compare with previously selected two model in Wrapper method.

5.4.2 Model Selection and Evaluation

Table 5.15 drawn a comparison among three different model with a view to select the best one. First two model were selected from Wrapper method and the last model was selected from Filtering method. These entire three models were chosen based on their higher R square value. Table 5.16 illustrates the same R square value for all the three models. Therefore, these three models have different Mean Absolute Error (MAE) value. Model-54 has the lowest MAE value with the same R square value. Hence Model-54 has selected for further analysis.

Table 5.16: Comparison Between Wrapper Method & Filtering Method

Model	**SVM Type**	**Wrapper Method**				**Filtering Method**			
		Subset-2 Subset-3				{X1, X2, X3, X4, X5, X6, X7, X8}			
		RMSE	MSE	MAE	R squared	RMSE	MSE	MAE	R squared
Model -12	Coarse Gaussian	0.0588	0.0034	0.047495	0.72	-	-	-	-
Model -18	Coarse Gaussian	0.0587	0.0034	0.047147	0.72	-	-	-	-
Model -54	Coarse Gaussian	-	-	-	-	0.0628	0.0039	0.047077	0.72

Figure 5.9 illustrates the result for selected model-54 where the R-Square value is 0.72 and the Mean Absolute Error is 0.047077 which is lower than other two models in Wrapper method, therefore selected for further analysis. The figure also shows the Residual Plot for true response where a trend is clearly visualized.

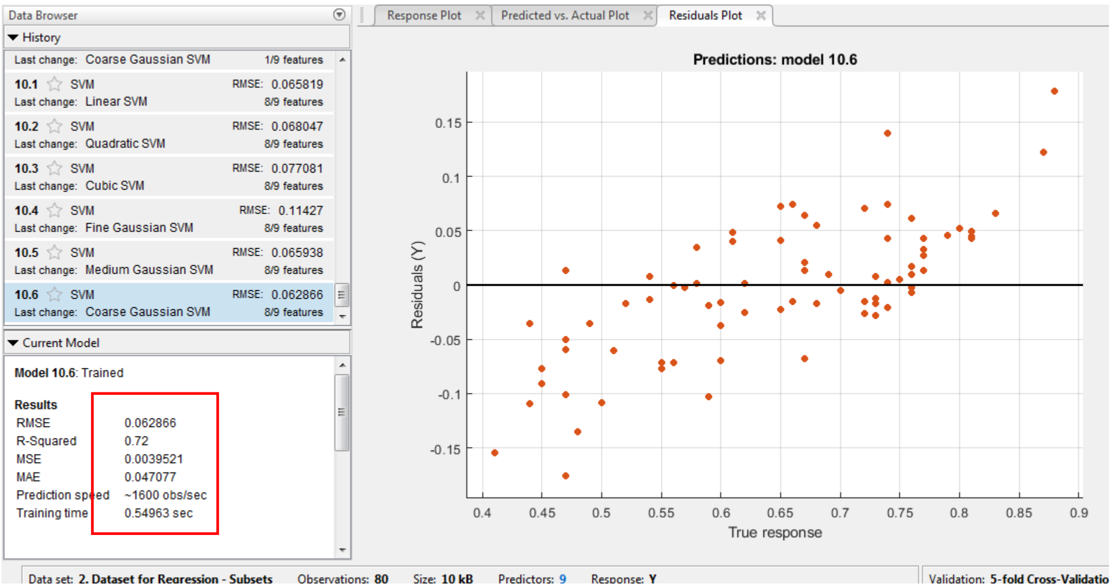

Figure 5.9: Selected Model-54 from Filtering Method

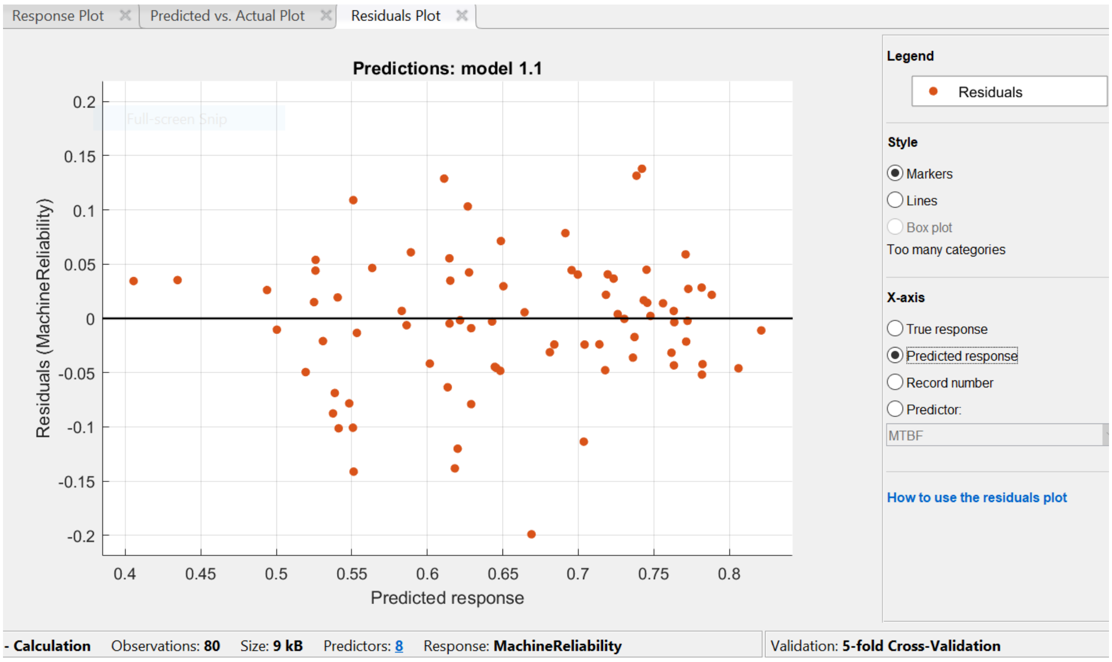

Figure 5.10: Residual Plot for Selected Model-54

In Figure 5.10 it can be seen that the residuals and predicted response are shown in a residual plot. The points in this plot are randomly dispersed around the horizontal axis. Most of the residual values are close to zero means the greater accuracy of the prediction. The points are pretty systematically distributed and slightly cluster towards the middle of the plot. Since the residual plot shows a fairly random pattern, indicating a decent fit to the data for a linear model.

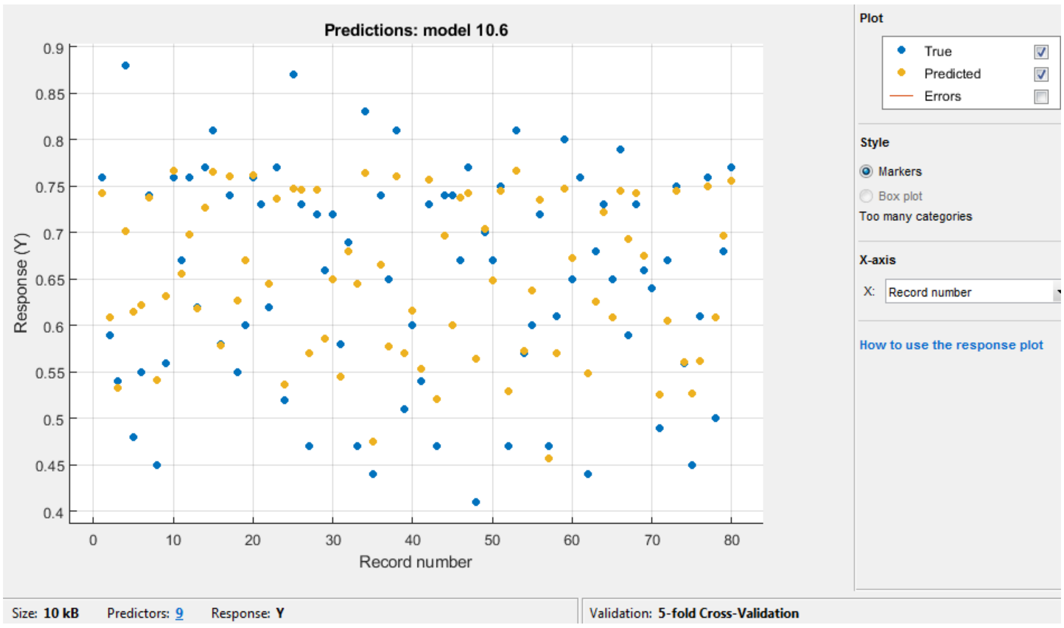

Figure 5.11: Response Plot for Selected Model-54

Figure 5.11 illustrate the response plot where the X axis represent the sample number and Y axis represent the reliability value. The difference between the true values and predicted values are clearly visible by the blue and yellow dots respectively. Most of the cases the position of this two-color points are very close and sometimes it's overlapped each other. So, the response plots demonstrate the greater accuracy of the model.

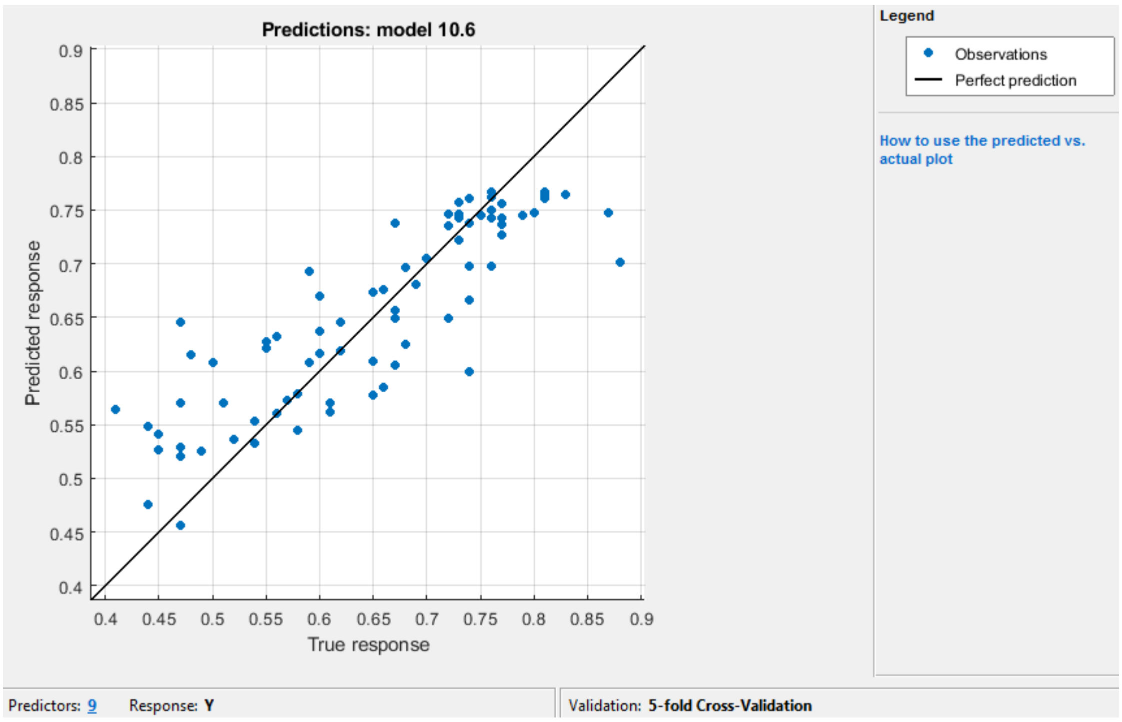

Figure 5.12: Predicted vs Actual for Selected Model-54

Figure 5.12 shows the relationship between true response and predicted response. Some points are fallen over the line where some are very close to the line and Most of the points are around the line. Very little amounts are out of the line. So, the overall true response and predicted response is very close therefore the model good enough for linear regression.

5.5 Machine Reliability Calculation for Diesel Generator through MLR

Regression is a statistical measurement based predicting method for estimating the relationships among variables. In this problem there are nine explanatory variables with one response variable. These multiple explanatory variables influence response variables since it is a multi variate linear regression problem. Table 5.17 represents the impact of explanatory variables on response variable by performing the regression in Microsoft office excel data analysis tool pack. Only four features among all the nine features seems significant based on their p-value.

Table 5.17: Feature Selection Using p- value

Feature	p-value	Decision
Mean Time Between Failure (MTBF)	0.00	Significant
Mean Time to Repair (MTTR)	0.1386	Not Significant
Downtime (DT)	0.1958	Not Significant
Machine Age (MA)	0.1142	Not Significant
Machine Room Environment (MRE)	0.0498	Significant
Average Operating Time (AOT)	0.0418	Significant
Manufacturer (M)	0.0892	Not Significant
Periodic Maintenance Practice (PMP)	0.00	Significant
Alternative Machine (AM)	0.2444	Not Significant

Table 5.18 illustrate the performance of the model with three different features set. Regression are performed for all these three different features set by using Microsoft Office Excel Data Analysis Tool pack and their results are shown in Table. It is very clear from the Table 5.18 that the previously selected features set in ANOVA test produces better result compare to other feature sets. This is the validation that the regression model is developed with the right feature set.

Table 5.18: Performance Measure for three Different Feature Sets

Performance Parameter	All Features	Significant Features	Selected Features (ANOVA test)
Multiple R	0.8959	0.8455	0.8959
R Square	0.72	0.7148	0.72
Adjusted R Square	0.7772	0.6996	0.7802
Standard Error	0.0555	0.0645	0.0551

In Table 5.17 the second column considered all the nine features where the third column has considered only four significant features from Table 5.16. The final column has been selected from filtering method by performing ANOVA test. From this comparison a same r square value with different standard error is found and the final column represents the better results because of its lower standard error.

Table 5.19 represents a comparison between Support Vector Regression (SVR) and Multivariate Linier Regression (MLR). Same dataset has been used for performing and measuring this result. Support Vector Regression (SVR) has been performed by using the software MATLAB 2018a where as Microsoft office excel data analysis tool pack has been used for Multivariate Linier Regression (MLR). The r square value for both SVR and MLR is same but the Mean Absolute Error (MAE) is lower for SVR. Therefore, SVR performing better than MLR.

Table 5.19: Comparison between SVR and MLR

Performance Parameter	SVR	MLR
R-SQUARE	0.72	0.72
MSE	0.0039	0.7802
MAE	0.0470	0.0551

5.6 Maintenance Schedule Prediction Model Development for Boiler

A well and organized planned preventive maintenance program avoid unnecessary downtime and costly repairs. In case of hazardous machineries like boiler it is also essential for safety reasons. The reason behind that Boiler is an enclosed vessel in which the water is heated and that provides a means for combustion and transfers heat to water until it becomes hot water or steam. The hot water or steam under pressure is then usable for transferring the heat to a process. When water is boiled into steam its volume increases about 1,600 times, producing a force that is almost as explosive as gunpowder. This causes the boiler to be extremely dangerous equipment and should be treated carefully. Therefore, it is essential to develop a well-planned maintenance program to avoid unnecessary accident, downtime and costly repair. In Bangladesh boiler is widely used in various industries for several purposes. But the maintenance and inspection of boiler is always invisible due to lack of awareness and ignorance. Several other reasons are also associated with it such as lack of skilled manpower and others. For this reason's boiler explosion is a common means in industrial factories specially the textiles and garments. The explosion can be reduced by developing and maintaining an appropriate maintenance plan. The detection of maintenance requirement can easily be identified by developing a model using Machine Learning technique.

5.6.1 Maintenance Classification Model Formulation

Maintenance classification model for boiler can be formulated by using machine learning technique which is Support Vector Machine (SVM). To develop the model it is needed to identify the predictor variables and response variable which is identified as follows.

5.6.1.1 Data Collection, Description and Model Formulation

To develop the maintenance classification model for boiler the data was collected from the site survey. Both qualitative and quantitively data were collected during the survey according to the predefined predictor variables which is also called features variable. The quantitively data were found from the respective gauges and meter or display of the of the boiler and the qualitative data were found from the expert's opinions and literature.

The data were collected based on the predefined predictor variables and response variables. After getting the data it is organized according to the following means.

The aim of this work is to develop of maintenance schedule prediction model for boiler to predict the maintenance requirement more accurately. So, it is essential to identify the features which is highly correlated with the maintenance of the boiler. The explosion of boiler depends on the variation of following parameters. So, the proper operation and maintenance of this parameters can reduce the explosion and ensure the efficiency of the boiler.

Operating Load: Operating load is the varying load of a boiler which is defined based on the required output over its maximum capacity. It's generally represents in percentage and may varies from hour to hour. When a boiler operates at its maximum rated capacity, it is referred to as maximum load.

Steam Pressure: Steam pressure is the generated pressure when the water is heated to produce the required steam. Every boiler belongs a maximum steam pressure where a maximum operating pressure is fixed in safety valve.

Water Level: The level of water in a steam boiler is another key parameter that must be carefully controlled, to ensure good quality steam is produced safely, efficiently and at the correct pressure. Normally boilers are designed in such a way that normal water level (NWL) provides sufficient space for steam velocity. Therefor normal operating water level (NOWL) is the desired condition during the operation of a boiler. However, the lack of supervision and discontinuity of water supply might cause water level go under low level point and cause explosions.

Boiler Pressure: Boiler pressure is such an important parameter that must maintain in a range and despite that that boiler can stop working due to the pressure is too low or too high. The ideal pressure for your boiler is usually between 1 and 2 bars. If the boiler pressure is below 1, that means low pressure. This could be because the lost water from the heating system. If the pressure gauge goes above 2, that usually means it's too high that may need to bleed the radiator.

Water Temperature: Water temperature is also significant to generate quality steam as well as ensuring safely operation of boiler by avoiding overheating. The average setting for a hot water boiler is 180°F. This provides the appropriate level needed for most cold weather temperatures. If the temperature setting is manual, the higher limit of temperature is 210°F and its suggested to remain down at 190°F. Once a boiler starts to go over 212°F, it may face serious problems.

The response variables of this model are the maintenance classes which is formed by the compilation of maintenance checklist and action plan. The checklist is developed based on the literature search and expert opinions. The checklist of boiler maintenance is given below which will be used to formulate the maintenance classes.

Daily Checklist

1. Blow Down
2. Water Controller Test
3. Softener Regeneration
4. Gauge Glass Test
5. Clean All Boiler Room
6. Water Pressure
7. Gas Press Check/ Fuel
8. Steam Pressure Check
9. Feed Water Tank Level Check

Weekly Checklist

1. Check Firing Rate Control
2. Check Air Modulation Motor Controls
3. Check Air Damper
4. Check Water level switch
5. Check Air Pressure Switch
6. Check Ignition Systems
7. Check Fuel Valves Pilot Flame & Main Flame
8. Check Combustion Satiety Controls Flame Failure Sensor
9. Check Flame Signal Strength
10. Check Feed Pump
11. Check Feed Pump Flow Meter
12. Check Gas Pressure

Monthly Checklist

1. Check Gas Pressure Valve
2. Check Steam Pressure Switch
3. Check Safety Pressure Switch
4. Check Gauge Glass
5. Check Blow Down Valve
6. Check Water level Sensor
7. Check Main Steam Stop Valve
8. Check Header Steam Pressure
9. Check Fuel, Vent, Stack, or Outlet damper
10. Check Combustion Air
11. Check Low Draft Fan, Air Pressure & Damper
12. Check High & Low Gas Pressure Interlocks
13. Check Low Oil Pressure Interlocks
14. Check Air Compressor Oil Level
15. Check Feed Water Modulation Valves
16. Check Low Cutoff Controller
17. Check High Cutoff Controller

Yearly Checklist

1. Boiler De- Scaling
2. Tube Carbon Cleaning
3. Fire Sail Change or Repair
4. Safety Valve Pressure testing or repair
5. Economizer De- Scaling
6. Boiler Automation Controlling system & Electronics system
7. Pre-Heater
8. Super Heater
9. Condenser
10. Condenser water pump
11. Feed water pump
12. Feed water tank
13. Chemical treatment process

14. Main Steam stop valve
15. Boiler Blow Down Valve
16. Burner
17. FD Fan or Blower
18. Boiler inside inspection

Two types of action plan are performed for boiler maintenance. The first action plan is to increase the boiler pressure when the boiler pressure suddenly dropped due to water leak causes no what water or water lost from the system and another reason is bleeding the radiator which occur when radiator fails to heat up the water. In that case action plan, one is performed to increase the boiler pressure.

Action Plan 1:

1. Check around the pipes
2. Check boiler for water or damp patches
3. Check radiators
4. Repressurize Boiler
    - Turn off the boiler and allow it to cool.
    - Find the filling loop, or attach it to boiler.
    - Make sure the boiler pressure gauge is seen while using the filling loop.
    - Open the valves on both sides to let water into your system.
    - Wait until the pressure gauge reaches 1.5 bar then close each valve, one after the other.
    - Switch the boiler back on and, if need, press the reset button.
    - Don’t forget to remove the filling loop if it’s an attachment. Be careful in case there’s any water left in it.
5. If the pressure drops again
6. Look for gas register engineer

The second action plan is to reduce the boiler pressure when the boiler pressure increased due to recently added water to the system for low pressure. The following action are suggested to reduce the boiler pressure to fix it to its desired level.

Action Plan 2:

1. Bleed the radiator which is lets some of the water out
2. If bleeding radiators doesn't work, one of the boiler parts could be faulty
3. Contact a Gas Safe Registered engineer to investigate
4. Check for temperature increase issues and fix it

Now the maintenance class can be drawn by compiling the above maintenance checklist and action plans. The Table 5.20 contains the maintenance class for the maintenance schedule prediction model development.

Table 5.20: Maintenance Class for Maintenance Schedule Prediction

Sl. No.	Maintenance Class	Description
1	Class 1	Daily Check Weekly Check Monthly Check Yearly Check
2	Class 2	Class 1 & Action Plan 1
3	Class 3	Class 1 & Action Plan 2

Maintenance schedule prediction model for boiler is formulated by using Support Vector Machine (SVM) algorithm. To formulate the model four predictor variables and three response variables has been identified. The data were collected from the field survey and finally a data set is prepared to formulate the model. Table 5.21 represents the dataset for this model development.

Table 5.21: Sample Dataset for Maintenance Schedule Prediction Model Development

Machine	Features (X)					Output (Y)
Boiler	Operating Load (%)	Maintenance Class	Water Level	Boiler Pressure (Bar)	Water Temperature (F)	Maintenance Class
	X1	X2	X3	X4	X5	
Obs. 01	50	4.5	0	1.2	170	3
Obs. 02	70	6.5	0	1.4	175	1
Obs. 03	75	6.8	1	1.4	190	2
Obs. 04	55	5.2	2	1.9	180	3
Obs. 05	85	8.3	0	1.4	185	1
Obs. 06	80	7.8	0	1.5	180	1
Obs. 07	70	6.2	0	1.3	170	1
Obs. 08	45	4.1	2	2.2	165	3
Obs. 09	40	3.7	2	1.3	160	3
Obs. 10	65	5.8	0	1.4	175	1
Obs. 11	85	8.9	1	0.7	205	2
Obs. 12	45	4	0	1.9	185	3
Obs. 13	85	7.8	0	1	190	1
Obs. 14	45	3.8	0	1.8	155	3
Obs. 15	90	8.5	0	0.9	195	2
Obs. 16	85	7.8	0	1.4	180	1
Obs. 17	50	4.8	1	0.8	160	3
Obs. 18	75	6.5	0	1.9	170	1
Obs. 19	70	6.4	1	1.5	190	2
Obs. 20	80	7.4	1	1	195	2

The dataset in Table 5.21 contains five predictor variables which is defined earlier. Water level is the most significant parameters that can affect the other parameters directly and indirectly. In this dataset for Water Level, 0 represent the normal operating water level (NOWL), where 2 represent above the normal operating water level (NOWL) and 1 represent below the normal operating water level (NOWL).

Now the correlation of each predictor variables with the response variable are identified by using data analysis tool pack of Microsoft excel which is shown in Table 5.22. The negative sign represents the negative correlation while the positive sign represents positive correlation of the features with its output variable. The higher the values are, the higher the correlation of the features with the output variable. Therefore, the features are ranked based on their correlation value from higher to lower.

Table 5.22: Correlation of Predictor Variable with Response Variables

Features	**Correlation (Person's r)**	**Ranking**
Operating Load (%)	-0.58	2
Steam Pressure (Bar)	-0.54	3
Water Level	0.70	1
Boiler Pressure (Bar)	0.18	4
Water Temperature (F)	-0.11	5

Table 5.22 represents that all the features are significantly correlated with its response variable therefore considered for model development.

For identifying a prediction model which is can predict the maintenance requirements based on the above-mentioned features different SVM models are checked by using MATLAB toolbox 2018 a. Those models are:

i. Linear SVM- Uses linear Kernel
ii. Quadratic SVM- Uses quadratic Kernel
iii. Cubic SVM- Uses cubic Kernel
iv. Fine Gaussian SVM- Uses Gaussian Kernel with Kernel scale
v. Medium Gaussian SVM- Uses Gaussian Kernel with Kernel scale
vi. Coarse Gaussian- Uses Gaussian Kernel with Kernel scale

Table 5.23: Accuracy of the Developed Model

Feature Set	Model No	SVM Type	Accuracy (%)
{X1, X2, X3, X4, X5}	Model-01	Linear SVM	86.0%
	Model-02	Quadratic SVM	80.0%
	Model-03	Cubic SVM	84.0%
	Model-04	Fine Gaussian SVM	62.0%
	Model-05	Medium Gaussian SVM	82.0%
	Model-06	Coarse Gaussian	70.0%

5.6.1.2 Model Selection and Evaluation

Table 5.23 has shown the accuracy level of the developed model considering the predetermined significant feature sets. From all this developed model, Model 01 in Linear SVM represents the highest accuracy level which is 86.0%. Therefore, this model is considered for further analysis.

Figure 5.13: Accuracy Level of the Selected Model 01

Table 5.24: Class Wise Accuracy and Error Level of the Selected Model

Class	True Positive Rate	False Negative Rate
Class-I	93%	7%
Class-II	81%	19%
Class-III	86%	14%

The Table 5.24 has shown class wise prediction accuracy and error level which is drawn from the confusion matrix of the selected Model 01. Now the confusion matrix for Model-01 which is given in Figure 5.14 where it can be seen that for class one, 14 sample fallen in true class as it is predicted and only 1 sample fallen incorrectly which should fall in class 1, so that the sample 93% correctly belonging in class one. Similarly, for class two and three the accuracy is 81% and 86% respectively.

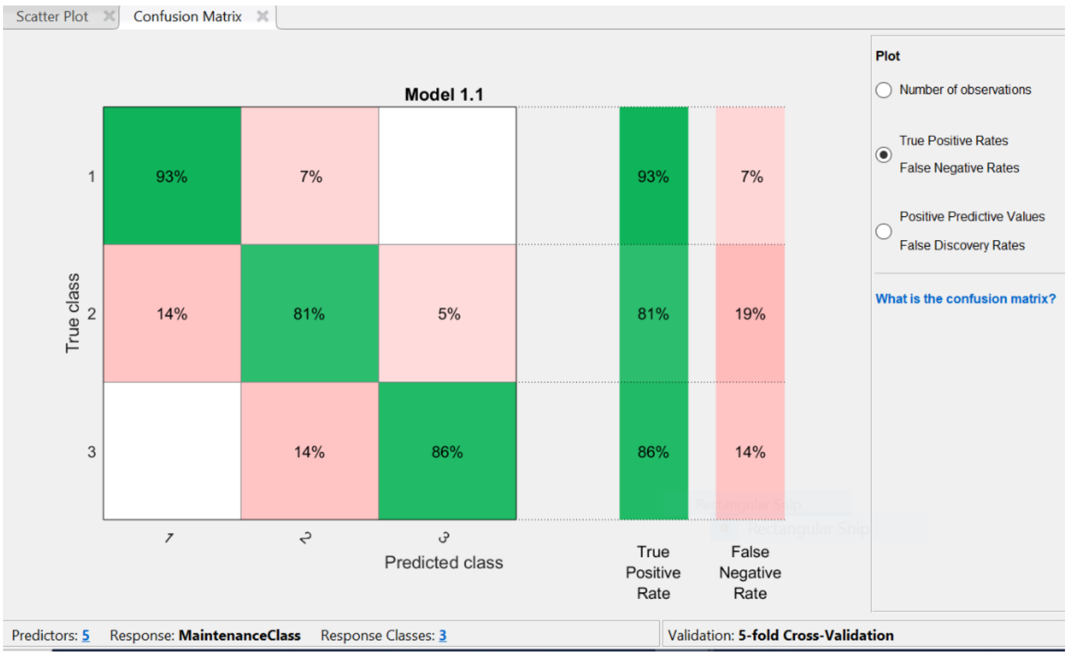

Figure 5.14: Confusion Matrix of the Selected Model 01

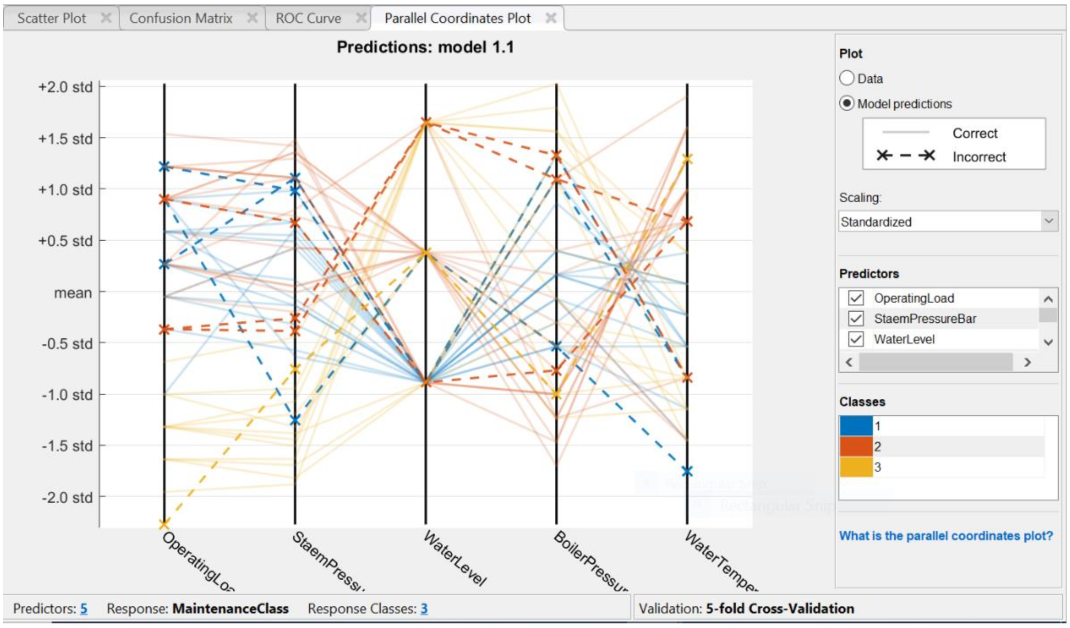

Figure 5.15: Parallel Coordinate Plot for all Features

Figure 5.15 demonstrates the parallel plot of all features which reflects how each feature has an impact on the classification model. In Figure 5.15 blue lines, red lines and yellow lines respectively represents three different maintenance prediction classes. The areas between the features, where these lines are easily separable have a greater impact on the classification model compared to other features.

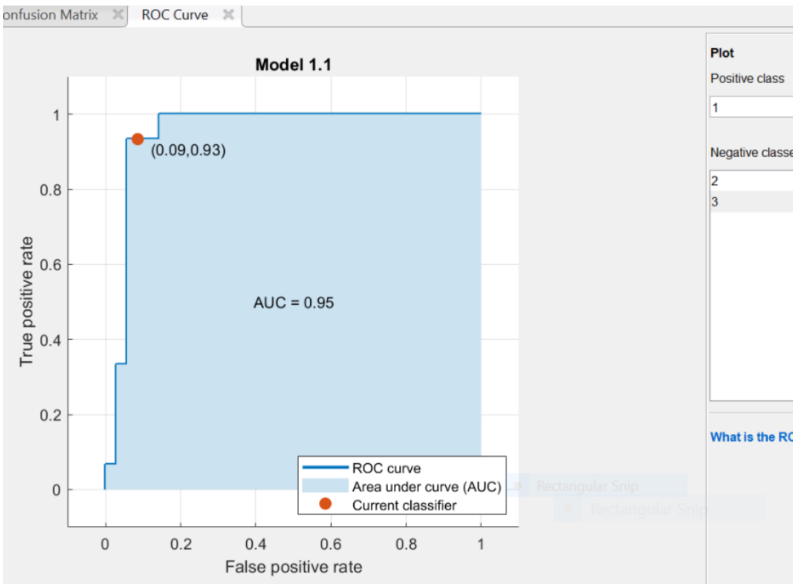

Figure 5.16: ROC Curve of the Selected Model 01

Table 5.25: ROC Curve Comparison for Three Classes

Particulars	Class 1	Class 2	Class 3
True Positive Rate	0.93	0.81	0.86
False Positive Rate	0.09	0.10	0.03
AUC	0.95	0.89	0.98

Figure 5.16 Represents the ROC curve to quantify the performance of the classifier. The illustration shows that the curve hugs the upper left corner of the plot means a very good classification capability of the model. Table 5.25 has drawn the data from the ROC curve for all the three classes, now for class 1, 2 and 3 the true positive rate is 0.93, 0.81 and 0.86 respectively whereas the false positive rate is only 0.09, 0.10, and 0.03 respectively. The Area under the Curve (AUC) is also 0.95, 0.89 and 0.98 that cover the maximum area of the curve which also represents the better performance of the model.

CHAPTER 6: RESULT AND DISCUSSION

6.1. Maintenance Schedule Prediction through Support Vector Machine (SVM)

The Machine Learning (ML) approach for predicting the maintenance requirements by using an effective preventive maintenance model is implemented for Diesel Generator (DG) and Boiler in Chapter 5. Analyzing various external and internal parameters, an effective Planned Preventive Maintenance (PPM) model needed to be developed that can forecast the maintenance requirements in advance.

Several subjective and objective features were identified initially. Then the features have been screened based on their impact on the output to develop and select the best model. Two structured method were followed to select the best features set. The first one is Wrapper method that generate eight different subsets by using backward elimination process based on the correlation of features to its output. The second one is Filtering method that select a feature set by eliminating insignificant features through ANOVA test. Each individual feature set generate six different model with different accuracy level when it run in MATLAB 2018a. Therefore, a total of 54 different models have been generated which is shown in Table 6.1.

Table 6.1: Comparison Among all Developed Models

SVM Type	**Wrapper Method**								**Filt. Met.**
	Sub. 1	**Sub. 2**	**Sub. 3**	**Sub. 4**	**Sub. 5**	**Sub. 6**	**Sub. 7**	**Sub. 8**	
	Mod.: (1-6)	Mod.: (7-12)	Mod.: (13-18)	Mod.: (19-24)	Mod.: (25-30)	Model: (31-36)	Mod.: (37-42)	Mod.: (43-48)	Model: (49 - 54)
Linear SVM	86.3%	86.3%	87.5%	87.5%	87.5%	88.8%	90.0%	90.0%	88.8%
Quadratic SVM	83.8%	85%	87.5%	87.5%	87.5%	90. 0%	91. 0%	91. 0%	87.5%
Cubic SVM	85%	86.3%	88.8%	88.8%	88.8%	92. 5%	91. 0%	90.0%	88.8%
Fine Gaussian SVM	16.3%	18.8%	41.3%	51.2%	55%	91.3%	67.5%	81.3%	55.5%
Medium Gaussian SVM	86.3%	85%	87.5%	87.5%	87.5%	88.8%	90.0%	88.8%	88.8%
Coarse Gaussian	77.5%	86.3%	87.5%	86.3%	87.5%	88.8%	88.8%	86.3%	87.5%

Table 6.2 is a simplified representation of the previous table. In this table the best models are picked for each subset based on their accuracy level. Model-33 form subset-6 in Wrapper Method, represent the highest accuracy level 92.5% among all the models. Therefore, this model has been selected and various analysis were performed through Scatter Plot, Confusion Matrix, Parallel Coordination Plot and ROC Curve.

Table 6.2: Prediction Accuracy Comparison for Selected Models

Method	Subset	Highest Accuracy Model	Accuracy
Wrapper Method	Subset-1	Model-1 & Model-5	86.3%
	Subset-2	Model-7, Model-9 & Model-12	86.3%
	Subset-3	Model-15	88.8%
	Subset-4	Model-21	88.8%
	Subset-5	Model-27	88.8%
	Subset-6	Model-33	92. 5%
	Subset-7	Model-38, Model-39	91. 0%
	Subset-8	Model-44	91. 0%
Filtering Method		Model-49, Model-51 & Model-53	88.8%

6.2. Machine Reliability Prediction through Regression

Machine reliability consideration is an important factor for accurate prediction of maintenance requirements. The machine reliability highly depends on various external and internal parameters of the machine which is considered in previous section for maintenance prediction model development. In this section the machine reliability value is developed by using Support Vector Regression (SVR). Finally, a comparison is drawn between the Support Vector Regression (SVR) and Multivariate Linier Regression (MLR) to understand the model effectiveness.

In Support Vector Regression (SVR), two structural methods has been followed to develop and select the best model as performed in previous section for maintenance prediction model development. The Wrapper method used backward elimination process to generate eight subsets and each subset has six different models when it run in MATLAB 2018a. Filtering method generated only one subset by eliminating the insignificant feature and these also have six different models.

Table 6.3 illustrates a comparison among different models to select the best one. In Wrapper method the best model is selected form each subset to compare with the model represented from Filtering method. Model -12 and Model-18 from Wrapper method has highest R square value which is same with the Model-54 from Filtering method. These three models have same R square value 0.72 but different Mean Absolute Error (MAE) value. Model-54 from Filtering method represent the lowest Mean Absolute Error (MAE) therefore selected as the best model. Various analysis has been performed through Residual Plot, Response Plot and Predicted Response vs True Response.

Table 6.3: Comparison Among Probable Reliability Prediction Models

Method	Subset	SVM Type	Model	MAE	R squared
Wrapper Method	Subset-1	Model-6	Coarse Gaussian	0.047874	0.71
	Subset-2	Model-12	Coarse Gaussian	0.047495	0.72
	Subset-3	Model-18	Coarse Gaussian	0.047147	0.72
	Subset-4	Model-24	Coarse Gaussian	0.053877	0.67
	Subset-5	Model-30	Coarse Gaussian	0.05238	0.69
	Subset-6	Model-36	Coarse Gaussian	0.053516	0.68
	Subset-7	Model-42	Coarse Gaussian	0.052388	0.68
	Subset-8	Model-47	Medium Gaussian SVM	0.048997	0.69
Filtering Method		Model-54	Coarse Gaussian	0.047077	0.72

Table 6.4 represents a comparison between Support Vector Regression (SVR) and Multivariate Linier Regression (MLR). Same dataset has been used for performing and measuring this result. Support Vector Regression (SVR) has been performed by using the software MATLAB 2018a where as Microsoft office excel data analysis tool pack is used for Multivariate Linier Regression (MLR). The r square value for SVR and MLR is same. But the Mean Absolute Error (MAE) is lower for SVR than MLR. Therefore, SVR is preforming better than MLR.

Table 6.4: Comparison between SVR and MLR

Performance Parameter	SVR	MLR
R-SQUARE	0.72	0.72
MSE	0.0039	0.7802
MAE	0.0470	0.0551

CHAPTER 7: CONCLUSIONS AND FUTURE WORK

7.1 Conclusion

This research presents a noteworthy contribution to develop a classification model by using Support Vector Machine (SVM) algorithms to predict the preventive maintenance schedule. Machine reliability is an important factor that deals with various objective and subjective measures and it becomes quite difficult to address through existing maintenance practices, whereas, this factor plays an important role for predicting maintenance schedule.

i. The proposed classification model developed by implementing SVM is capable to address the limitation of existing models in a cost-effective manner, which is a major contribution of this work.
ii. Feature selection is one of the unique and major steps that has followed in this research. All subjective and objective features do not have same impacts on its output, therefore defining the correlation of features with its output and selecting the best combination is important to improve the model accuracy.
iii. The machine reliability values are determined by using a regression model which is developed by using Support Vector Regression (SVR). The same features which are used in classification model are also used for regression model development. Here the predictor variable or output is the reliability of the machine.
iv. Finally, the model which is developed by using SVR are compared with Multivariate Linear Regression (MLR) for validation and it is very clear that the Regression model is developed with the right feature set.

7.2 Future Work

For future research work, this thesis can be developed by incorporating the following considerations.

i. In this research, to develop the Planned Preventive Maintenance (PPM) model, nine different subjective and objective features were selected from expert opinion and literature. In future some other features like Machine Availability, Maintainability, Cost of Failure etc. can be included to improve the model accuracy.
ii. For feature selection in Filtering method and Wrapper method, ANOVA test and Correlation-based Feature Selection (CFS) technique were used respectively. In

future most robust technique like Principal Component Analysis (PCA) and Markov Blanket Filter can be used for Filtering method. For Wrapper method Genetic Algorithm and Radian Basis Function (RBF) can be used.

iii. The research is carried out with a moderate number of sample size. A bigger sample would probably enhance more the quality and reliability of the research.

REFERENCES

Ahmad, R., & Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & industrial engineering*, *63*(1), 135-149.

Aiserman, M. A., Braverman, E. M., & Rozonoer, L. I. (1964). Theoretical foundations of the potential function method in pattern recognition. *Avtomat. i Telemeh*, *25*(6), 917-936.

Alaswad, S., & Xiang, Y. (2017). A review on condition-based maintenance optimization models for stochastically deteriorating system. *Reliability Engineering & System Safety*, *157*, 54-63.

Alrabghi, A., & Tiwari, A. (2015). State of the art in simulation-based optimisation for maintenance systems. *Computers & Industrial Engineering*, *82*, 167-182.

Arisoy, Y. M., & Özel, T. (2015). Machine learning based predictive modeling of machining induced microhardness and grain size in Ti–6Al–4V alloy. *Materials and Manufacturing Processes*, *30*(4), 425-433.

Badıa, F. G., Berrade, M. D., & Campos, C. A. (2002). Optimal inspection and preventive maintenance of units with revealed and unrevealed failures. *Reliability Engineering & System Safety*, *78*(2), 157-163.

Baptista, M., Sankararaman, S., de Medeiros, I. P., Nascimento Jr, C., Prendinger, H., & Henriques, E. M. (2018). Forecasting fault events for predictive maintenance using data-driven techniques and ARMA modeling. *Computers & Industrial Engineering*, *115*, 41-53.

Barker, C. T., & Newby, M. J. (2009). Optimal non-periodic inspection for a multivariate degradation model. *Reliability Engineering & System Safety*, *94*(1), 33-43.

Benkedjouh, T., Medjaher, K., Zerhouni, N., & Rechak, S. (2015). Health assessment and life prediction of cutting tools based on support vector regression. *Journal of Intelligent Manufacturing*, *26*(2), 213-223.

Ben Mabrouk, A., Chelbi, A., & Radhoui, M. (2016). Optimal imperfect preventive maintenance policy for equipment leased during successive periods. *International Journal of Production Research*, *54*(17), 5095-5110.

Berge, C. (1957). Two theorems in graph theory. *Proceedings of the National Academy of Sciences of the United States of America*, *43*(9), 842.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. springer.

Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory* (pp. 144-152).

Caballé, N. C., Castro, I. T., Pérez, C. J., & Lanza-Gutiérrez, J. M. (2015). A condition-based maintenance of a dependent degradation-threshold-shock model in a system with multiple degradation processes. *Reliability Engineering & System Safety*, *134*, 98-109.

Cho, S., Asfour, S., Onar, A., & Kaundinya, N. (2005). Tool breakage detection using support vector machine learning in a milling process. *International Journal of Machine Tools and Manufacture*, *45*(3), 241-249.

Cocconcelli, M., Capelli, L., Cavalaglio Camargo Molano, J., & Borghi, D. (2018). Development of a methodology for condition-based maintenance in a large-scale application field. *Machines*, *6*(2), 17.

Coria, V. H., Maximov, S., Rivas-Dávalos, F., Melchor, C. L., & Guardado, J. L. (2015). Analytical method for optimization of maintenance policy based on available system failure data. *Reliability Engineering & System Safety*, *135*, 55-63.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273-297.

Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent data analysis*, *1*(3), 131-156.

Das, S. (2001, June). Filters, wrappers and a boosting-based hybrid for feature selection. In *Icml* (Vol. 1, pp. 74-81).

De Jonge, B., Teunter, R., & Tinga, T. (2017). The influence of practical factors on the benefits of condition-based maintenance over time-based maintenance. *Reliability engineering & system safety*, *158*, 21-30.

Dekker, R. (1996). Applications of maintenance optimization models: a review and analysis. *Reliability engineering & system safety*, *51*(3), 229-240.

Do, P., Voisin, A., Levrat, E., & Iung, B. (2015). A proactive condition-based maintenance strategy with both perfect and imperfect maintenance actions. *Reliability Engineering & System Safety*, *133*, 22-32.

Drucker, H., Burges, C. J., Kaufman, L., Smola, A., & Vapnik, V. (1996). Support vector regression machines. *Advances in neural information processing systems*, *9*, 155-161.

Dy, J. G., & Brodley, C. E. (2004). Feature selection for unsupervised learning. *Journal of machine learning research*, *5*(Aug), 845-889.

Elangovan, M., Devasenapati, S. B., Sakthivel, N. R., & Ramachandran, K. I. (2011). Evaluation of expert system for condition monitoring of a single point cutting tool using principle component analysis and decision tree algorithm. *Expert Systems with Applications*, *38*(4), 4450-4459.

El-Ferik, S., & Ben-Daya, M. (2006). Age-based hybrid model for imperfect preventive maintenance. *IIE Transactions*, *38*(4), 365-375.

Estévez-Pérez, G., & Vilar, J. A. (2013). Functional ANOVA starting from discrete data: an application to air quality data. *Environmental and ecological statistics*, *20*(3), 495-517.

Ferreira, R. J., de Almeida, A. T., & Cavalcante, C. A. (2009). A multi-criteria decision model to determine inspection intervals of condition monitoring based on delay time analysis. *Reliability Engineering & System Safety*, *94*(5), 905-912.

Gao, R., Wang, L., Teti, R., Dornfeld, D., Kumara, S., Mori, M., & Helu, M. (2015). Cloud-enabled prognosis for manufacturing. *CIRP annals*, *64*(2), 749-772.

Garg, A., & Deshmukh, S. G. (2006). Maintenance management: literature review and directions. *Journal of quality in maintenance engineering*.

Graham, R. L., Lawler, E. L., Lenstra, J. K., & Kan, A. R. (1979). Optimization and approximation in deterministic sequencing and scheduling: a survey. In *Annals of discrete mathematics* (Vol. 5, pp. 287-326). Elsevier.

Giroti, G., Nakhate, T., Laddha, M., & Sarve, P. M. (2018). Person Identification through Voice using MFCC and Multi-class SVM. *International Research Journal of Engineering and Technology (IRJET)*, 690-693.

Grall, A., Bérenguer, C., & Dieulle, L. (2002). A condition-based maintenance policy for stochastically deteriorating systems. *Reliability Engineering & System Safety*, *76*(2), 167-180

Grall, A., Dieulle, L., Bérenguer, C., & Roussignol, M. (2002). Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE transactions on reliability*, *51*(2), 141-150.

Janecek, A., Gansterer, W., Demel, M., & Ecker, G. (2008, September). On the relationship between feature selection and classification accuracy. In *New challenges for feature selection in data mining and knowledge discovery* (pp. 90-105).

Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing*, *20*(7), 1483-1510.

Jiaa, C. L., & Dornfeld, D. A. (1998). A self-organizing approach to the prediction and detection of tool wear. *ISA transactions*, *37*(4), 239-255.

Jiang, X., Yuan, Y., & Liu, X. (2013). Bayesian inference method for stochastic damage accumulation modeling. *Reliability Engineering & System Safety*, *111*, 126-138.

Karegowda, A. G., Jayaram, M. A., & Manjunath, A. S. (2010). Feature subset selection problem using wrapper approach in supervised learning. *International journal of Computer applications*, *1*(7), 13-17.

Kohavi, R., & John, G. H. (1997). Wrappers for feature subset selection. *Artificial intelligence*, *97*(1-2), 273-324.

Lenz, B., & Barak, B. (2013, January). Data mining and support vector regression machine learning in semiconductor manufacturing to improve virtual metrology. In *2013 46th Hawaii International Conference on System Sciences* (pp. 3447-3456). IEEE.

Liao, H., Elsayed, E. A., & Chan, L. Y. (2006). Maintenance of continuously monitored degrading systems. *European Journal of Operational Research*, *175*(2), 821-835.

Liao, W., Pan, E., & Xi, L. (2010). Preventive maintenance scheduling for repairable system with deterioration. *Journal of Intelligent Manufacturing*, *21*(6), 875-884.

Li, H., Parikh, D., He, Q., Qian, B., Li, Z., Fang, D., & Hampapur, A. (2014). Improving rail network velocity: A machine learning approach to predictive maintenance. *Transportation Research Part C: Emerging Technologies*, *45*, 17-26.

Lin, D., Zuo, M. J., & Yam, R. C. (2001). Sequential imperfect preventive maintenance models with two categories of failure modes. *Naval Research Logistics (NRL)*, *48*(2), 172-183.

Lu, S., Shen, S., Huang, J., Dong, M., Lu, J., & Li, W. (2018). Feature selection of laser-induced breakdown spectroscopy data for steel aging estimation. *Spectrochimica Acta Part B: Atomic Spectroscopy*, *150*, 49-58.

Mabrouk, A. B., Chelbi, A., & Radhoui, M. (2016). Optimal imperfect maintenance strategy for leased equipment. *International Journal of Production Economics*, *178*, 57-64.

Paprocka, I. (2019). The model of maintenance planning and production scheduling for maximising robustness. *International Journal of Production Research*, *57*(14), 4480-4501.

Peng, H., Feng, Q., & Coit, D. W. (2010). Reliability and maintenance modeling for systems subject to multiple dependent competing failure processes. *IIE transactions*, *43*(1), 12-22.

Rafiee, K., Feng, Q., & Coit, D. W. (2015). Condition-based maintenance for repairable deteriorating systems subject to a generalized mixed shock model. *IEEE Transactions on Reliability*, *64*(4), 1164-1174.

Roth, V., & Lange, T. (2003). Feature selection in clustering problems. *Advances in neural information processing systems*, *16*, 473-480.

Sadeghian, O., Oshnoei, A., Nikkhah, S., & Mohammadi-Ivatloo, B. (2019). Multi-objective optimisation of generation maintenance scheduling in restructured power systems based on global criterion method. *IET Smart Grid*, *2*(2), 203-213.

Schölkopf, B., Smola, A. J., & Bach, F. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.

Schwabacher, M., & Goebel, K. (2007, November). A Survey of Artificial Intelligence for Prognostics. In *AAAI Fall Symposium: Artificial Intelligence for Prognostics* (pp. 108-115).

Sharma, S., Cui, Y., He, Q., Mohammadi, R., & Li, Z. (2018). Data-driven optimization of railway maintenance for track geometry. *Transportation Research Part C: Emerging Technologies*, *90*, 34-58.

Shaw, R. G., & Mitchell-Olds, T. (1993). ANOVA for unbalanced data: an overview. *Ecology*, *74*(6), 1638-1645.

Sherwin, D. (2000). A review of overall models for maintenance management. *Journal of quality in maintenance engineering*.

Sheu, S. H., & Chang, C. C. (2009). An extended periodic imperfect preventive maintenance model with age-dependent failure type. *IEEE Transactions on Reliability*, *58*(2), 397-405.

Shi, D., & Gindy, N. N. (2007). Tool wear predictive model based on least squares support vector machines. *Mechanical Systems and Signal Processing*, *21*(4), 1799-1814.

Shin, J. H., & Jun, H. B. (2015). On condition-based maintenance policy. *Journal of Computational Design and Engineering*, *2*(2), 119-127.

Sick, B. (2002). On-line and indirect tool wear monitoring in turning with artificial neural networks: a review of more than a decade of research. *Mechanical systems and signal processing*, *16*(4), 487-546.

Song, S., Coit, D. W., & Feng, Q. (2014). Reliability for systems of degrading components with distinct component shock sets. *Reliability Engineering & System Safety*, *132*, 115-124.

Susto, G. A., Schirru, A., Pampuri, S., McLoone, S., & Beghi, A. (2014). Machine learning for predictive maintenance: A multiple classifier approach. *IEEE Transactions on Industrial Informatics*, *11*(3), 812-820.

Ullah, I., Yang, F., Khan, R., Liu, L., Yang, H., Gao, B., & Sun, K. (2017). Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach. *Energies*, *10*(12), 1987.

Valdez-Flores, C., & Feldman, R. M. (1989). A survey of preventive maintenance models for stochastically deteriorating single-unit systems. *Naval Research Logistics (NRL)*, *36*(4), 419-446.

Wang, X., Li, L., & Xie, M. (2019). Optimal preventive maintenance strategy for leased equipment under successive usage-based contracts. *International Journal of Production Research*, *57*(18), 5705-5724.

Wu, D., Jennings, C., Terpenny, J., Gao, R. X., & Kumara, S. (2017). A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. *Journal of Manufacturing Science and Engineering*, *139*(7).

Xiao, L., Song, S., Chen, X., & Coit, D. W. (2016). Joint optimization of production scheduling and machine group preventive maintenance. *Reliability Engineering & System Safety*, *146*, 68-78.

Yang, C. H., Chuang, L. Y., & Yang, C. H. (2010). IG-GA: a hybrid filter/wrapper method for feature selection of microarray data. *Journal of Medical and Biological Engineering*, *30*(1), 23-28.

Yang, L., Ma, X., Peng, R., Zhai, Q., & Zhao, Y. (2017). A preventive maintenance policy based on dependent two-stage deterioration and external shocks. *Reliability Engineering & System Safety*, *160*, 201-211.

Ye, Z. S., Xie, M., & Tang, L. C. (2013). Reliability evaluation of hard disk drive failures based on counting processes. *Reliability Engineering & System Safety*, *109*, 110-118.

Zahedi, Z., Salim, A., Yusriski, R., & Haris, H. (2019). Optimization of an integrated batch production and maintenance scheduling on flow shop with two machines. *International Journal of Industrial Engineering Computations*, *10*(2), 225-238.

Appendix: A

Training Dataset for Maintenance Schedule Prediction Model Development:

Machine	**Features (X)**									**Output (Y)**
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	MP
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-1	9M	3	L	5	VP	30	1	VP	A	1
DG-2	5M	11	H	10	VP	105	2	VP	A	2
DG-3	4M	13	H	15	VP	115	3	P	NA	2
DG-4	10M	14	M	2	G	40	1	VP	A	1
DG-5	5M	11	H	13	VP	120	1	M	NA	2
DG-6	7M	7	M	7	P	50	4	P	A	3
DG-7	9M	2	L	2	M	35	2	VP	A	1
DG-8	4M	12	H	17	P	130	2	P	NA	2
DG-9	6M	8	M	8	P	70	3	VP	A	3
DG-10	9M	1	L	1	G	25	2	P	A	1
DG-11	2M	14	M	9	P	50	1	VP	NA	4
DG-12	11M	4	M	1	VG	45	3	M	A	1
DG-13	4M	15	H	9	VP	80	1	VP	NA	2
DG-14	9M	5	L	4	VP	30	2	VP	A	1
DG-15	9M	5	L	2	G	45	1	M	NA	1
DG-16	5M	20	H	10	VP	105	2	P	NA	2
DG-17	10M	4	L	3	M	20	1	M	A	1
DG-18	8M	9	H	5	P	75	3	VP	A	5
DG-19	6M	6	M	9	P	70	1	VP	A	3
DG-20	11M	5	L	5	VG	20	1	M	NA	1
DG-21	1M	15	L	2	M	35	3	VP	NA	4
DG-22	6M	6	M	9	P	65	2	VP	A	3
DG-23	11M	3	M	2	G	40	1	M	A	1
DG-24	5M	12	H	13	VP	105	5	VP	NA	2
DG-25	9M	3	L	5	VG	30	1	VP	A	1
DG-26	10M	2	L	5	M	35	1	M	A	1
DG-27	4M	11	H	12	VP	110	3	VP	NA	2
DG-28	11M	2	L	3	M	35	1	VP	A	1
DG-29	7M	10	H	7	P	60	4	VP	NA	3
DG-30	8M	9	M	6	P	80	2	VP	A	5
DG-31	4M	13	H	20	VP	135	1	P	A	2
DG-32	7M	8	L	6	P	50	2	M	NA	3
DG-33	5M	12	H	5	VP	115	1	M	A	2
DG-34	11M	3	L	2	VG	25	1	VP	A	1
DG-35	4M	14	H	25	P	125	5	P	NA	2
DG-36	7M	6	M	7	P	65	1	G	A	3
DG-37	2	7	H	10	P	120	2	M	NA	5
DG-38	9M	4	L	2	G	40	1	M	A	1
DG-39	4M	9	H	11	P	90	4	VP	A	2
DG-40	7M	10	M	8	P	60	3	P	A	3

Machine	Features (X)									Output (Y)
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	MP
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-41	5M	16	H	9	VP	85	5	M	NA	2
DG-42	9M	5	L	3	M	15	1	P	A	1
DG-43	5M	20	H	11	VP	110	5	VP	NA	2
DG-44	9M	5	L	2	VP	45	3	P	A	1
DG-45	4M	20	H	10	VP	95	1	VG	NA	2
DG-46	9M	3	L	4	M	30	2	VP	A	1
DG-47	9M	4	L	5	VG	40	1	M	A	1
DG-48	4M	15	H	12	VP	100	3	VP	NA	2
DG-49	10M	5	L	5	VG	45	3	M	A	1
DG-50	6M	7	M	9	P	70	1	VP	A	3
DG-51	9M	4	L	5	VG	35	2	G	A	1
DG-52	4M	13	M	17	VP	75	5	M	NA	2
DG-53	10M	5	L	1	G	25	1	VG	A	1
DG-54	4M	11	H	6	VP	80	5	G	NA	2
DG-55	9M	4	M	7	P	70	3	P	NA	1
DG-56	9M	3	M	4	G	45	1	P	A	1
DG-57	5M	14	H	25	VP	135	5	M	A	2
DG-58	4M	15	H	18	VP	100	1	P	NA	2
DG-59	10M	4	L	1	M	20	2	G	A	1
DG-60	7M	6	M	7	P	65	1	P	NA	3
DG-61	15M	5	L	2	VP	40	3	M	A	1
DG-62	4M	17	H	14	VP	135	2	VP	NA	2
DG-63	7M	10	M	7	P	65	3	G	NA	3
DG-64	9M	4	M	2	M	45	1	P	A	1
DG-65	4M	9	M	12	VP	95	1	VP	NA	2
DG-66	9M	3	L	5	G	30	2	G	A	1
DG-67	7M	6	M	6	P	70	1	G	NA	3
DG-68	9M	2	L	1	VP	35	2	P	A	1
DG-69	6M	8	L	8	P	55	2	M	NA	3
DG-70	3	7	M	8	P	60	1	VP	A	3
DG-71	4M	13	H	10	VP	130	5	M	A	2
DG-72	5M	16	M	9	VP	85	2	VP	NA	2
DG-73	9M	2	L	2	M	20	3	VG	A	1
DG-74	4M	16	H	13	VP	125	2	VP	A	2
DG-75	4M	20	M	15	VP	75	5	VP	A	2
DG-76	5M	12	M	20	VP	110	1	M	NA	2
DG-77	10M	3	L	3	M	15	2	VG	A	1
DG-78	4M	17	H	9	VP	95	1	P	A	2
DG-79	7M	6	M	6	P	55	1	M	NA	3
DG-80	11M	1	L	1	G	15	3	G	A	1

Appendix: B

Training Dataset for Machine Reliability Prediction Model Development:

Machine	Features (X)									Output (Y)
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	Machine Reliability
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-1	9M	3	L	5	VP	30	1	VP	A	0.76
DG-2	5M	11	H	10	VP	105	2	VP	A	0.59
DG-3	4M	13	H	15	VP	115	3	P	NA	0.54
DG-4	10M	14	M	2	G	40	1	VP	A	0.88
DG-5	5M	11	H	13	VP	120	1	M	NA	0.48
DG-6	7M	7	M	7	P	50	4	P	A	0.55
DG-7	9M	2	L	2	M	35	2	VP	A	0.74
DG-8	4M	12	H	17	P	130	2	P	NA	0.45
DG-9	6M	8	M	8	P	70	3	VP	A	0.56
DG-10	9M	1	L	1	G	25	2	P	A	0.76
DG-11	2M	14	M	9	P	50	1	VP	NA	0.67
DG-12	11M	4	M	1	VG	45	3	M	A	0.76
DG-13	4M	15	H	9	VP	80	1	VP	NA	0.62
DG-14	9M	5	L	4	VP	30	2	VP	A	0.77
DG-15	9M	5	L	2	G	45	1	M	NA	0.81
DG-16	5M	20	H	10	VP	105	2	P	NA	0.58
DG-17	10M	4	L	3	M	20	1	M	A	0.74
DG-18	8M	9	H	5	P	75	3	VP	A	0.55
DG-19	6M	6	M	9	P	70	1	VP	A	0.60
DG-20	11M	5	L	5	VG	20	1	M	NA	0.76
DG-21	1M	15	L	2	M	35	3	VP	NA	0.73
DG-22	6M	6	M	9	P	65	2	VP	A	0.62
DG-23	11M	3	M	2	G	40	1	M	A	0.77
DG-24	5M	12	H	13	VP	105	5	VP	NA	0.52
DG-25	9M	3	L	5	VG	30	1	VP	A	0.87
DG-26	10M	2	L	5	M	35	1	M	A	0.73
DG-27	4M	11	H	12	VP	110	3	VP	NA	0.47
DG-28	11M	2	L	3	M	35	1	VP	A	0.72
DG-29	7M	10	H	7	P	60	4	VP	NA	0.66
DG-30	8M	9	M	6	P	80	2	VP	A	0.72
DG-31	4M	13	H	20	VP	135	1	P	A	0.58
DG-32	7M	8	L	6	P	50	2	M	NA	0.69
DG-33	5M	12	H	5	VP	115	1	M	A	0.47
DG-34	11M	3	L	2	VG	25	1	VP	A	0.83
DG-35	4M	14	H	25	P	125	5	P	NA	0.44
DG-36	7M	6	M	7	P	65	1	G	A	0.74
DG-37	2	7	H	10	P	120	2	M	NA	0.65
DG-38	9M	4	L	2	G	40	1	M	A	0.81
DG-39	4M	9	H	11	P	90	4	VP	A	0.51
DG-40	7M	10	M	8	P	60	3	P	A	0.60

Machine	Features (X)									Output (Y)
DG	MTBF	MTTR	DT	MA	MRE	AOT	M	PMP	AM	Machine Reliability
	X1	X2	X3	X4	X5	X6	X7	X8	X9	
DG-41	5M	16	H	9	VP	85	5	M	NA	0.54
DG-42	9M	5	L	3	M	15	1	P	A	0.73
DG-43	5M	20	H	11	VP	110	5	VP	NA	0.47
DG-44	9M	5	L	2	VP	45	3	P	A	0.74
DG-45	4M	20	H	10	VP	95	1	VG	NA	0.74
DG-46	9M	3	L	4	M	30	2	VP	A	0.67
DG-47	9M	4	L	5	VG	40	1	M	A	0.77
DG-48	4M	15	H	12	VP	100	3	VP	NA	0.41
DG-49	10M	5	L	5	VG	45	3	M	A	0.70
DG-50	6M	7	M	9	P	70	1	VP	A	0.67
DG-51	9M	4	L	5	VG	35	2	G	A	0.75
DG-52	4M	13	M	17	VP	75	5	M	NA	0.47
DG-53	10M	5	L	1	G	25	1	VG	A	0.81
DG-54	4M	11	H	6	VP	80	5	G	NA	0.57
DG-55	9M	4	M	7	P	70	3	P	NA	0.60
DG-56	9M	3	M	4	G	45	1	P	A	0.72
DG-57	5M	14	H	25	VP	135	5	M	A	0.47
DG-58	4M	15	H	18	VP	100	1	P	NA	0.61
DG-59	10M	4	L	1	M	20	2	G	A	0.80
DG-60	7M	6	M	7	P	65	1	P	NA	0.65
DG-61	15M	5	L	2	VP	40	3	M	A	0.76
DG-62	4M	17	H	14	VP	135	2	VP	NA	0.44
DG-63	7M	10	M	7	P	65	3	G	NA	0.68
DG-64	9M	4	M	2	M	45	1	P	A	0.73
DG-65	4M	9	M	12	VP	95	1	VP	NA	0.65
DG-66	9M	3	L	5	G	30	2	G	A	0.79
DG-67	7M	6	M	6	P	70	1	G	NA	0.59
DG-68	9M	2	L	1	VP	35	2	P	A	0.73
DG-69	6M	8	L	8	P	55	2	M	NA	0.66
DG-70	3	7	M	8	P	60	1	VP	A	0.64
DG-71	4M	13	H	10	VP	130	5	M	A	0.49
DG-72	5M	16	M	9	VP	85	2	VP	NA	0.67
DG-73	9M	2	L	2	M	20	3	VG	A	0.75
DG-74	4M	16	H	13	VP	125	2	VP	A	0.56
DG-75	4M	20	M	15	VP	75	5	VP	A	0.45
DG-76	5M	12	M	20	VP	110	1	M	NA	0.61
DG-77	10M	3	L	3	M	15	2	VG	A	0.76
DG-78	4M	17	H	9	VP	95	1	P	A	0.50
DG-79	7M	6	M	6	P	55	1	M	NA	0.68
DG-80	11M	1	L	1	G	15	3	G	A	0.77

Appendix: C

Feature Details:

Features	Feature Details	Feature Type	Representation
X1	Men Time Between Failure (MTBF)	Quantitative	Months
X2	Mean Time to Repair (MTTR)	Quantitative	Hours
X3	Downtime (DT)	Qualitative	Low (L) Medium (M) High (H)
X4	Machine Age (MA)	Quantitative	Years
X5	Machine Room Environment (MRE)	Qualitative	Very Poor (VP) Poor (P) Moderate (M) Good (G) Very Good (VG)
X6	Average Operating Time (AOT)	Quantitative	Hours Per Month
X7	Manufacturer (M)	Qualitative	1. Very Good (VG) 2. Good (G) 3. Moderate (M) 4. Poor (P) 5. Very Poor (VP)
X8	Periodic Maintenance Practice (PMP)	Qualitative	Very Poor (VP) Poor (P) Moderate (M) Good (G) Very Good (VG)
X9	Alternative Machine (AM)	Qualitative	Available (A) Not Available (NA)

Appendix: D

Training Dataset for Maintenance Schedule Prediction Model Development for Boiler

Machine	**Features (X)**					**Output (Y)**
Boiler	Operating Load (%)	Steam Pressure (Bar)	Water Level	Boiler Pressure (Bar)	Water Temperature (F)	Maintenance Class
	X1	X2	X3	X4	X5	
Obs. 01	50	5	2	2	175	3
Obs. 02	70	6.5	0	1.4	175	1
Obs. 03	90	8.5	0	0.9	200	2
Obs. 04	60	6	2	2	185	3
Obs. 05	75	6.8	1	1.4	190	2
Obs. 06	55	5.2	2	1.9	180	3
Obs. 07	85	8.3	0	1.4	185	1
Obs. 08	65	6.3	2	1.8	190	3
Obs. 09	80	7.8	0	1.5	180	1
Obs. 10	70	6.2	0	1.3	170	1
Obs. 11	45	4.1	2	2.2	165	3
Obs. 12	40	3.7	2	1.3	160	3
Obs. 13	90	8.8	1	0.8	200	2
Obs. 14	65	5.8	0	1.4	175	1
Obs. 15	65	6.1	2	1.9	165	3
Obs. 16	85	8.9	1	0.7	205	2
Obs. 17	45	4	0	1.9	185	3
Obs. 18	85	7.8	0	1	190	1
Obs. 19	45	3.8	0	1.8	155	3
Obs. 20	90	8.5	0	0.9	195	2
Obs. 21	75	8	2	0.8	200	2
Obs. 22	85	7.8	0	1.4	180	1
Obs. 23	80	7.5	0	1.7	175	1
Obs. 24	50	4.8	1	0.8	160	3
Obs. 25	75	6.7	1	1.2	155	2

Machine	Features (X)					Output (Y)
Boiler	Operating Load (%)	Steam Pressure (Bar)	Water Level	Boiler Pressure (Bar)	Water Temperature (F)	Maintenance Class
	X1	X2	X3	X4	X5	
Obs. 26	90	8.5	1	0.9	195	2
Obs. 27	75	6.5	0	1.9	170	1
Obs. 28	55	4.6	1	1.9	165	3
Obs. 29	70	6.4	1	1.5	190	2
Obs. 30	80	7.6	0	1.3	155	1
Obs. 31	80	7.4	1	1	195	2
Obs. 32	85	7.9	1	1.1	195	2
Obs. 33	50	4.5	0	1.2	170	3
Obs. 34	85	8.9	1	0.8	165	2
Obs. 35	50	4.4	2	2.1	160	3
Obs. 36	35	5.5	1	0.9	200	2
Obs. 37	75	9.1	0	0.8	210	2
Obs. 38	95	9	0	0.9	205	2
Obs. 39	55	7.7	0	1.1	160	1
Obs. 40	50	4.3	2	0.9	165	3
Obs. 41	75	6.9	0	1.1	170	1
Obs. 42	85	4.7	1	1.1	150	2
Obs. 43	70	7.4	1	0.9	195	2
Obs. 44	80	6.6	0	1.2	190	1
Obs. 45	75	8.5	0	0.9	195	2
Obs. 46	70	7.4	0	1.8	170	1
Obs. 47	75	6.8	1	1.3	155	2
Obs. 48	85	5.7	0	1.5	180	1
Obs. 49	85	8.5	0	0.7	205	2
Obs. 50	90	8.3	0	0.8	165	2