ADDIS ABABA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
FACULTY OF INFORMATICS
DEPARTMENT OF INFORMATION SCIENCE

Gas Turbine Engine Performance Classifier: Artificial Neural Network Approach

By

Haileleoul Gudeta

A thesis submitted to

the School of Graduate Studies of Addis Ababa University

in partial fulfillment of the requirements for the Degree of Master

of Science in Information Science

July 2004

ADDIS ABABA UNIVERSITY
FACULTY OF INFORMATICS
DEPARTMENT OF COMPUTER SCIENCE

DISTRIBUTED CLIENT'S PROFILE MANAGEMENT SYSTEM FOR PERVASIVE SYSTEMS

BY

HAILELEUL TEGENE MEKURIA

Name and Signature of Members of the Examining Board

Dr. Dida Midekso	Chairman, Examining Board	
Dr. Dawit Bekele	Advisor	
Dr. Kumuda Raimond	External Examiner	
Dr. Mulugeta Libsie	Local Examiner	

ACKNOWLEDGEMENT

I would like to express my sincere thank to my advisor **W/o Rahel Bekle**, for constructive comments and encouragement she gave me on my work. I am also grateful to **Ato Mesfin Getachew** for his dedicated and unreserved support during the entire course of this research.

I am immensely indebted to my beloved parents for giving me unconditional care, love and support throughout my life. The encouragement, affection and advice and support I got from my sisters and brothers is also immeasurable on any scale and imprinted deep in my heart.

I am also grateful to my classmates for their dedication and effort in developing team sprit during difficult times we encountered by sharing ideas and resources.

At last, I would like to extend my sincere thank and express special regards to all the people around me who have helped me during the last two years of my study. Especially the ones with open mind and positive attitude. ***Many many many....thanks to all of you***.

TABLE OF CONTENT

.... i
ACKNOWLEDGEMENT ii
TABLE OF CONTENT.......... iii
LIST OF TABLES iv
LIST OF FIGURES ....

CHAPTER ONE 1
1.1 Background 6
1.2 Statement of The Problem And Its Importance 10
1.3 Objectives 10
1.3.1 General Objective 10
1.3.2 Specific Objectives.......... 11
1.4 Research Methodology.......... 13
1.5 Scope and limitation 13
1.6 Organization Of The Thesis. ....

CHPATER TWO
GAS TURBINE ENGNIE PERFORMANCE ASSESSEMENT
.... 14
2.1 Introduction.......... 14
2.2 Gas Turbine Engine 17
2.3 Gas Turbine Engine Performance 19
2.4 Evaluating Gas Turbine Engine Performance. 22
2.5 Research Areas Using Gas Turbine Engines Data 22
2.5.1 Engine monitoring systems (EMS),.......... 23
2.5.2 Automated Fault Diagnosis 24
2.5.3 Neural Network Based Gas Turbine Engine Diagnosis ....

CHPATER THREE
LITERATURE REVIEW 25
3.1 Data Mining 27
3.1.1 Data mining functionalities.......... 29
3.1.2 Neural Network As A Data Mining Technique 30
3.2 Neural Networks ....

3.2.1 Introduction 30
3.2.2 Applications Of Neural Network 33
3.3.4 Basic Structures Of Neural Network 34
3.2.4 Classification, (Types) Of Neural Network 39
3.2.4.1 Neural Network Connection 40
3.2.4.2 Learning Methods Of Neural Network 40

CHPATER FOUR **44**
EXPERIMENTATION 44
4.1 Introduction 44
4.2 Data Collection 44
4.3 Data Preprocessing 48
4.3.1 Data Cleaning, ... 49
4.3.2 Data Transformation, ... 50
4.3.3 Data Reduction, ... 52
4.4 Modeling Technique 54
4.5 Model Test Design 55
4.6 Building the model 55
4.6.1 Defining network architecture 55
4.6.2 Data partitioning for model training, ... 56
4.6.3 Training the model 57
4.6.4 Evaluating the Model. 60
4.6.4.1 Using confusion matrix, ... 60
4.6.4.2 Cross Validation. 62

CHPATER FIVE **64**
CONCLUSION AND RECOMMENDATION **64**
5.1 Conclusion 64
5.2 Recommendations 68

REFERENCES **71**
DECLARATION **78**

LIST OF TABLES

Table 1 : Attributes Selected and Collected 46
Table 2 : Collected Data Class Distribution 48
Table 3 : Nominal Attributes Transformation 51
Table 4 : Ordinal Attributes Transformation 52
Table 5 : Attributes Selected to Build Model 54
Table 6 : Partitions of data set for training ..... **Error! Bookmark not defined.**
Table 7 : Initial Classificaton Accuracy 58
Table 8 : Classificaton Accuracy with 10 neurons at hidden layer and different learning rate 59
Table 9 : Classificatoin Accuracy with 11 neurons at hidden layer and different learning rate 59
Table 10 : Classificaton accuracy with 12 neurons at hidden layer and differentlearning rate 59
Table 11 : Confusion mateix for the best model selected 61
Table 12 : Sensititity and specificity of the two classes 62
Table 13 : Classificaton accuracy of the model with ten fold cross validation 63

LIST OF APPENDICES

Appendix 1 - Engine Specifications (PT6A-27) 75

Appendix 2 - Log of Engine Test 76

Appendix 3 - Engine Performance Curve (sample Form)............... 77

ABSTRACT

Conventional method of assessing the performance of gas turbine engine involves analyzing different engine parameters manually and comparing them with their respective acceptable limits in engines maintenance manuals. The method takes lengthy and complicated processes that demand personnel with many years of experience and allows subjective judgment of the personnel involved in evaluation process. Technological advances in design and constructions of gas turbine engines adds more engine parameters thereby setting more hurdles on the process of engines performance evaluation.

This paper reports on the finding of a research that had the objective to build a model that classify the performance, either accepting or rejecting, of PT6A-27 model turboprop gas turbine engine. The engines considered were those undergoing evaluation for performance after they went through repair or overhaul.

The data used to build the mode first passed through different data preprocessing and analyzing techniques. The model employed neural network built using backpropagation algorithm on a neural network toolbox found in MATLAB 6.5.

The model built classifies gas turbine engines by their performances into their respective classes. The classification accuracy found was encouraging for the model to be adapted in real problem. The outcome of this research can also put a corner stone for further researches in using data mining technique for gas turbine engine maintenances.

CHAPTER ONE
INTRODUCTION

1.1 Background

Currently commercial airlines are struggling hard to keep their publicized timetables and schedules. Many factors can attribute for their difficulties [11]. One among many is failure to sustain their operational support machineries functioning properly [25]. Reasons for machineries technical failure include overlooking or paying little attention to machineries operating instructions, absence of adequate training for personnel working on machines and most importantly due to high-tech complexity of the systems [12]. The problem can easily infringe safety standard of airlines, not to mention its huge potential to leave dark mark on airlines reputations. If the failure of the machinery happens to be the part of the aircraft itself, for example its engine, the issue will gain substantially huge magnitude.

Engine is one of the important and critical components of an aircraft for its operation [9]. Aircraft engines provide thrust, or propulsive force that enable aircrafts to move in a forward direction [35] both on ground and in air. As one of major aircraft component, the health of the engine has to be closely monitored while it is operating. However monitoring its health, safe operation, is not restricted only while it is on aircraft, but aircrafts engine health monitoring also takes place during test after it under goes repair or overhaul in maintenance facilities.

Normally, maintenance technicians who took training do perform the maintenance and repair tasks on aircraft engines. The training includes adherence to principles of good workmanship plus a consistent practice of following the instructions provided in the engines manufacturer's manuals [27]. As a consequence current practices on aircraft engines maintenance and operation exclusively adopts instructions and procedures entirely stated in manuals. However, in practice manuals cannot provide exhaustive solutions for all technical problems encountered during engine maintenances [35]. In many instances, the experience of maintenance personnel on the system plays major role to make logical approach to identify and rectify technical problems. But experience will have effect only for engine systems with minimum or no complexity. However current aircraft engines are equipped with high-tech sophisticated apparatuses that are highly complex to analyze their behavior through experience [25].

In addition to high-tech sophistication, the over all operation of aircraft engines is governed by many variables [12]. These variables include ambient air temperature, ambient air pressure, oil pressure, oil temperature, fuel temperature, rotational speed (rpm) of rotors, temperatures at different parts inside the engine, pressures at different parts of the engine …etc. [12].

During the entire operation of aircraft engines these variables are continuously watched and closely monitored. But the number of variables to be observed and their complex relationship inhibit effective monitoring. For this reason engine manufacturers recommend very few numbers of variables to be monitored continuously assuming any problem within the engine would be reflected on those variables. Since the information about the engine is not adequately observed with those selected variables alone, some engine parts that are critical for safe operation are periodically inspected, repaired or replaced on regular bases. The engine itself is periodically inspected, and replaced as well for repair for the same reasons.

Taking periodic maintenance actions on aircraft engines and aircraft engine components have two important disadvantages. First, the operation of the engine has to be interrupted during the entire period, which is very expensive in the case of airline industry. Second, during periodic components removal, the component may not need replacement at all. It might still be in good shape to give service. However, it is a standard procedure for components removed from engine to be sent to maintenance facilities at least to be checked, or repaired as required. This has a potential to incur maintenance cost unnecessarily.

However effective analysis of engine variables can provide early information on the over all condition of the engine and it will also give early warning when there is a need for components replacement or repair there by minimizes or even avoids the above-mentioned problems. As stated previously however, analyzing and interpreting engine variables is difficult for human being due to its size and complexity.

However data mining, which is a defined as an automatic process of discovering patterns in data [27] can be applied to gain good understanding on condition of the engine, by analyzing large number of engine variables involved. If adequate data regarding engine operation is collected, data mining can help to monitor engine operation effectively by identifying and analyzing different patterns.

Data mining uses different techniques to discover interesting, patterns within data [19,34]. One of the techniques it adapts is machine learning, which is the field of study that is concerned with the question on how to construct computer program that automatically improve with experience using algorithms [32].

Machine learning algorithms have proven to be of great practical value in variety of application domains. For example, they are useful in data mining problems where large database may contain valuable implicit regularities that can be discovered automatically [6]. Machine learning are also highly effective in places where humans might not have the knowledge needed to develop effective algorithms (e.g. face recognition). And also they are suitable in domains where the program must dynamically adapt to the changing conditions like manufacturing and controlling process [32]. Taking these things into considerations the applications of machine learning technology appear to hold great promise for enhancing the effectiveness and of technical maintenance practices in general and engine maintenance practices in particular [15].

Machine learning adapts different types of learning techniques, one of which is artificial neural network or neural network. Artificial neural network provides a general, practical method for learning real, discrete, and vector valued functions from examples [27]. Compared to different conventional methods, artificial neural network is the most suitable machine-learning scheme for real world applications of considerable complexity [32].

Among different learning techniques, artificial neural network's important advantage for machine learning is its the ability to process data that are too complex for conventional technologies and to problems that do not have an algorithmic solution [28]. In general, because of their abstraction, from the biological brain, artificial neural network are well suited for problems that people are good at solving, but for which computers are not [22].. Artificial neural network is also used and give highly promising results on many technical researches of this kind [24, 42, 3].

Normally technical maintenance on gas turbine engine is executed by consistent and strict reference to operating and maintenance instructions, which are usually given by the engine's manufacturer. However due to the growing number of engine input and output variables, this method becomes very time consuming and often does not lead to optimal results [27]. Taking longer time during engine maintenance will cost substantial time and money.

As one of an important technical maintenance activity for gas turbine engines, a performance evaluation involves large number of variables with complex relationship. This research uses neural network model to analyze these variables. The research output affirms the potential use of data mining technique for gas turbine engine performance evaluation and it insights further researches to use data mining for other gas turbine engines maintenance activities.

1.2 Statement of The Problem And Its Importance

Gas turbine engines are widely used for powering aircrafts. The quality of maintenance on gas turbine engines is crucially important since it will be installed on aircrafts that carry the life and property of many people. Therefore the maintenance and repair performed on gas turbine engines must be checked for its rightness by evaluating engine's performance.

The existing performance evaluation practices on gas turbine engines usually employ diagnostic procedure, performed manually, which is error prone [15] and also takes considerable time to reach on conclusion. Automating the system will resolve these two problems.

The other problem in the existing gas turbine engine performance evaluation is its requirement to have experienced maintenance personnel. It is always customary for the maintenance personnel to integrate, analyze and categorize the values of the entire observed parameters that characterize the performance condition of the engine. This process is largely based on personnel experience [34]. Automating the system will relax this uncompromising condition.

Therefore the justification to conduct the proposed research is stated as follows.

- The customary way of determining the performance of gas turbine engine is mostly performed by analyzing the test data manually, which is always time taking and error prone. Automating the system will enable the same task to be executed within short period of time with minimum or no effect on accuracy of results.

- Conventional assessment on the condition of the engine is purely heuristic that demands maintenance personnel many years of work experience. However automating the system will assist to perform the task by less experienced personals with little or no effect on accuracy of results.

- Conventional method of evaluating gas turbine engine performance involves approximation of some engine parameters. These approximation are subjective to the individual who is doing the evaluation task. This may give a chance for different results if different people do the evaluation task.

- The research will use engine performance data collected on single model engine for over 10 years. The use of this data for model development will raise maintenance personnel awareness on data storage practices.

The manual analysis of the engine performance is one of the leading problems, which costs considerable time and money, especially in the field of aviation business. Normally the manual analysis is required to be done offline. It requires the engine run for test activity to be finished first before the analysis of the data is started. Manual collection of data on chart, feeding the data collected to computer program, and finally plotting the graph using standardized data takes considerable amount of time.

The other problem associated with this kind of practices is the need of personnel with many years of work experience to execute the tasks, especially to take the final decision on the performance of the engine.

The collection and storage of data is not used for no other than deciding the performance of that particular engine. However, these engine performances related data has a potential to understand more and provide information on other engines of the same model.

This research attempted to address these problems, by automating engine performance evaluation using neural network models. The model provides excellent classification assistance with the help of actual engines test performance signatures. In addition to facilitating quick engine performance evaluation, the research output also avoids the need for personnel with many years of work experience. The model developed also enables less experienced personnel to do the job with equal competency. The research has also taken steps in utilizing engine test data that was stored for no purpose.

The objective of this research includes building a classifying model to gas turbine engines with reference of their performance. There are different techniques and algorithms used for classification purposes. Among many, these include decision tree and neural network [8]. This research uses neural network due to the following reasons.

Neural networks have potential to approximate complex-non-linear mapping in an environment with many numbers of attributes [32]. The nature of the data this research uses has the same property. The noise tolerance nature and its ability to classify patterns with noise make neural network potential choice for many applications [16].

While using neural networks to model, priori assumptions about the distributions of the data, or the form of interactions between factors are not necessary. This fact is very important because it enables the model to learn more as the training of the model progress [21]. Neural networks can also be updated with fresh data, making them useful for dynamic environment. This makes the model made of neural network to be maintainable and usable in a dynamic environment. The other advantage of neural network is its potential to be highly automated [27]. This is an important point especially to pursue further work on this research topic.

Considering the above-mentioned points in mind, conducting a research to develop a model using neural network that can classify gas turbine engine performance with better classifying accuracy independent of the personnel involved will have considerable importance.

1.3 Objectives

1.3.1 General Objective

The general objective of this research is to explore the potential applicability of data mining technology in developing a model that can support gas turbine engines performance evaluation.

1.3.2 Specific Objectives

In order to accomplish the above-indicated general objective, the research work has carried out the following specific objectives.

- Reviewing literatures on application of data mining for classification
- Identification and collection of gas turbine engine performance data.
- Preparing and preprocessing the data
- Design, build and train models to classify gas turbine engine performance
- Test and evaluate models

- Analyze the outcome of the research

1.4 Research Methodology

To conduct the research the following methods were followed

1.4.1 Literature Review

To have strong base on data mining and artificial neural network, books, journal articles and research papers from different sources including the Internet were reviewed.

1.4.2 Data collection

The source of data for the research was obtained from Ethiopian Airlines Technical Service. The data source contains gas turbines engines test records that were collected and to evaluate engines performance and latter stored for no other reason. It is a standard practice to test and evaluate the performance of gas turbine engine after repair or maintenance is done on engines. Therefore the data contains all the parameters collected for evaluation and the evaluation result, i.e. weather the performance was accepted or rejected. Since the life history and the type of repair performed on the engine has influence on engine performance during test, additional data was collected from the same source. Most of the data are numerical records but it also includes few categorical data. More than 600 gas turbine engine test records were collected for this research.

1.4.3 Data analysis and preparation

The collected data was analyzed to remove data that were irrelevant for the objective of the research. In this step, noisy, irrelevant and redundant data were removed to reduce the data size and make the data analysis very efficient. Finally subset of the input variables that has significant weight for the research were identified and latter converted in a format that would be appropriate to build and train the model.

1.4.4 Training and building the model

Analyzed and preprocessed data were used to build the model. The model was built using artificial neural network using MATLAB 6.5. The choice of neural network was made mainly due to the type of data used for the research, which is a real time data having many variables with complex and non-linear relationship. The optimal architecture and other relevant neural network parameters are determined and the neural net was designed. Since the gas turbine engine test data used for this research contains the collected test data with the out come of the evaluation, the model was built using supervised learning method.

1.4.5 Evaluation of the model

After the model were built and trained, the model was evaluated for its classification accuracy. Confusion matrix and ten fold cross validation used to evaluate the correct classification of the model and assesses the outcome of the research.

1.5 Scope and limitation

The scope of this research is to assess merits or quality of the potential applicability of data mining technology to support gas turbine engines performance evaluation. Even if the outcome of the research can be adapted to different models of gas turbine engines, this research is limited to assess the possible applicability of data mining in classifying the performance of one particular model gas turbine engine with which Ethiopian Airlines is currently operating.

The shortage of time to conduct the research and difficulty in obtaining required software for the research hampered significantly the progress of the research. Moreover the out come of the research is restricted to the data that was collected to conduct this research.

1.6 Organization Of The Thesis.

The thesis is organized into 5 chapters. The first chapter presents introduction of the research. The second chapter gives an overview on how performance of gas turbine engine is evaluated and the different researches areas using gas turbine engine data. The third chapter deals with data mining technology and its relationship with the current research problem. The same chapter also tries to give insight on neural network. The fourth chapter contains the step-by-step tasks executed while performing the experiment. The final chapter is dedicated for the conclusion and recommendations based on the findings of the research.

CHAPTER TWO

GAS TURBINE ENGINE PERFORMANCE ASSESSMENT

2.1 Introduction

This chapter starts with introducing on basic principles and operation of gas turbine engine. It continues explaining the meaning of gas turbine engine performance. In subsequent section discussion on performance evaluation for the engine model considered in this research is made. In the last sections of the chapter, different research areas for better utilization of gas turbine engine data is discussed.

2.2 Gas Turbine Engine

History of airplane is only 100 years old. However its development within this period is so dramatic that results in shrinking our world [9]. The distance that used to require months journey is now covered with few hours due to technological developments of airplanes, and mainly due to technological advance in design and construction of aircraft engines [38].

Leonardo Da Vinci discovered the basic theory and operation of gas turbines in the 15th century. But it was in 1904 gas turbine was first proposed for aircraft application [9,20]. The proposal was not immediately followed by implementations mainly due to lack of appropriate material for manufacturing gas turbines for aircraft engines due to extreme thermal and mechanical stress it is required to withstand during operation

[38]. But progress in scientific study on materials metallurgy and arm race during Second World War, manufacturing gas turbine for aircraft engines application begin in mid 1930s.

Basically two types of engines used to power aircrafts, namely reciprocating engines and gas turbine engines. Both types of engines are considered internal combustion engines, because burning fuels to release thermal energy that will be later transformed into mechanical energy takes place inside them [37]. However, the process of energy conversion inside each type of engines is entirely different.

In reciprocating engines, the thermal energy is converted to mechanical energy in the form of reciprocating motion. The reciprocating movement then changed to rotational movement using the internal mechanical arrangement and feed to the system the engine is attached with [37]. On the other hand in the case of gas turbine engines, released thermal energy accelerate gas (air) to strike (hit) and spin gas turbine, which is found inside the engine. Actually the name of the engine itself is derived from the gas turbine found inside the engine [38].

All reciprocating engines drive propeller to develop thrust when they are used for aircrafts. But gas turbine engines develop same thrust using any one of the following two methods. The first method develops thrust by ejecting accelerating gas (air) at a very high speed. These types of gas turbine engines are also called pure jet engines. Newton's second and third laws of motion are the most applicable reasons for development of thrust by pure jet engines [9]. On the second method, the extracted thermal energy from burning fuel drives turbines, which are coupled with propellers. These types of gas turbine engines develop thrust similar to reciprocating engines by driving propellers [20]. These types of gas turbine engines are called turboprop engines.

Currently gas turbine engine is the major type of engine used to power aircrafts [38]. As compared to its counterpart, reciprocating engine, it is relatively lighter in weight and has capability to operate for long period with less risk on technical failure [9]. When compared with a reciprocating engine having equivalent weight, gas turbine engines can develop more than ten times thrust a reciprocating engine is capable of. These are the main reasons for gas turbine engine wide application in commercial and military aircrafts [9].

Regardless of its type and application, gas turbine engines consists of five major sections namely inlet, compressor, combustion, turbine and exhaust section.

Air is admitted to the engine using inlet section. The way air is flowing inside the engine is an important factor for the over all operation of the engine. Therefore extreme care is taken during design, manufacture and repair of air inlet section to assure disturbance free airflow inside the engine. Because airflow disturbance has significant influence in the over all performance of the engine [38].

After the air is entered inside the engine, it is compressed inside the compressor section. Air is compressed mainly to increase the static pressure of the air that enables it to accelerate it at a very high speed when exposed to high temperature inside combustion section. After air leaves the compressor section it is mixed with fuel and burns inside combustion section. The heat released inside the combustion section accelerates the air from combustion section and leave the engine through exhaust after driving the gas turbine found inside the turbine section [9, 20, 37,38].

Each section has different types and features depending on the intended use of the engine. Other than for aircraft applications, gas turbine engines nowadays are also used to power large sailing vessel, oil field pumps armed tanks, and also electrical generators in remote areas [12]. Discussing on the types and different features of the engine is skipped since it is not relevant to the research. But since the research is concerned mainly on the performance of gas turbine engine, the following section introduces gas turbine engines performance.

2.3 Gas Turbine Engine Performance

After an engine is manufactured, repaired or overhauled, it is a standard procedure to test and check proper operation of the engine. The purpose of testing the engine is to make sure that the engine will produce the expected and required power or thrust within allowable engine parameter limits and also to ensure smooth operation of the engine [29]. The test that all engines are required to take before giving service is called ***Engine Performance*** check.

Testing the performance of gas turbine engine is accomplished in an electronic instruments equipped test cell [9, 29]. During engine test, these instruments generate vast amount of numeric and symbolic data, by sensors installed in various parts of the engine to make sure that various systems of the engine operates properly [23]. The engine performance evaluation is made depending on the data collected from these instruments.

The engine manufacturers develop and issues detail procedures and instructions on how to test engines performance with. Procedures and instructions, which are recommended by the engine manufacturer, must be followed precisely to ensure correct information is obtained regarding the performance of the engine [9, 29]. Usually these instructions are found in Engine Overhaul Manuals that are prepared and issued by engine manufacturers.

Specific engine performance checks, instructions and the procedures to be followed vary depending on the type of engine and the type of repair it had gone though [9,29]. Introducing the different types of gas turbine engine performance checks is not relevant issue to this research. However since the research uses the engine test data on specific model engine, the overall test procedure of that particular model is revised in the following section. Introduction of engine performance evaluation for this particular model is limited to the extent relevant to the research work.

2.4 Evaluating Gas Turbine Engine Performance.

The model of the gas turbine engine that is used on this research is PT6A-27, which is a turboprop gas turbine engine widely used to power DHC-6 aircraft. Detail specification of the engine is attached in the appendix-1.

As already stated the performance of gas turbine engine is done during engine test. To monitor proper function and the condition of the engine as it runs, different sensors have to be installed on different parts of the engine to generate electrical signals. The signals that need to be sensed while the engine is under test include, rotational speed, air pressures, air temperatures, oil temperature, oil pressure fuel temperature, and fuel flow rate, [29].

Therefore the first thing to be done during engine test is to install those sensors and connect them with their respective electronic display equipments. All personnel in charge of running the engine will monitor these readings that are displayed on electronic instruments. The next step is to start running the engine. The persons who are in charge of testing the engine should be very much aware of all the procedures and instructions required during test [9, 29]. The detail procedures and instructions for this task is explicitly stated in engine overhaul manual.

These test procedures generate massive amount of data on instruments [35]. During the process of engine performance check, instruments readings will be recorded on a table specifically prepared for this purpose by engine manufacturer. The table is attached in the appendix-2.

After all the required data are collected, the engine will be stopped, and processing of collected data will start. These collected readings first processed with a computer program called Data Reduction Computer Program that is supplied by the engine manufacturer.

The computer program simply converts the collected readings into standard atmospheric conditions. This is mainly because atmospheric temperature and atmospheric pressures highly affect the overall operation of the engine. Especially the compressor rotational speed, compressor discharge air pressure, air temperature inside turbine section and many others are dependent on atmospheric conditions. Therefore collected data must be adjusted with equivalent values when sensed at standard or sea level condition. i.e. at 29.92 “Hg atmospheric pressure with temperature of 15^0C (59^0 F). The Data Reduction Program enables any engine to be evaluated for its performance regardless of the atmospheric conditions by converting the engine parameters to their equivalent values when sensed at sea level conditions.

The output data from the Data Reduction Program is used to plot graphs with strict instructions given on the engine overhaul manual. The graph plotting is done using parameters that are recommended by engine manufacturer to be important in showing the performance state of the engine. For PT6A-27 engines, the three parameters that are used for same purpose are compressor rotor speed, air temperature inside turbine section and the fuel flow readings. These readings are plotted on a graph specifically prepared for this purpose. The sample graph used for plotting engine performance is attached on appendix-3. Engine performance graph is plotted manually that depicts the performance of the engine.

Since the graph is plotted manually, and involves approximation of different values the outcome of the engine performance evaluation is not free from subjective judgment for a certain extent. Maximum care and caution is taken during plotting, but the bias cannot be eliminated completely.

The plotted graph approximately determines the rotational speed (RPM) of the compressor, the air temperature inside the combustion chamber (T_t5) and the engine's specific fuel consumptions as the engine develops around 680 SHP (shaft horse power) at standard or sea level conditions. These determined values are compared with the limits specified in engine overhaul manual either to accept or reject the performance of the engine.

if they are all found within the acceptable ranges, the engine is assumed to perform well and the test will be accepted. But if anyone or two or all of these three parameters are found out of the specified limits, the performance of the engine will be rejected.

2.5 Research Areas Using Gas Turbine Engines Data

Currently researches are underway to optimize the utilization of gas turbine engine especially in the field of aviation. Some of these researches with their objective are summarized as follows.

2.5.1 Engine monitoring systems (EMS),

Engine monitoring system, EMS, is a system that is used to examine the normal and smooth operation of the engine, with the help of sensors installed on different parts of the engine [35]. The system requires periodic collection of data from sensors as the engine operates and analyses collected data to tell the condition of the engine. For the case of EMS data analysis is performed manually.

Naturally gas turbine engines constitute a complex system, requiring adequate monitoring to ensure operational safety. Important parameter indicators display engine performance through vital information such as rotational speed, engine pressure ratio, exhausts gas temperatures, etc [35]. Any of these parameters can serve as an early indicator to prevent costly component damage and/or catastrophic failure, and thus help reduce the number of incidents and the cost of maintenance [35]. But identifying the exact parameter that has a potential to give more information than what it actually indicates is still a research question under EMS.

To address the practical problems for Engine Monitoring System (EMS) in commercial aircrafts, researches have been searching for better feature extraction in selecting important engine parameters with the purpose of providing a reliable means to monitor engine failures [35]. The current state of practice of engine monitoring system (EMS) is flooded with problems that are mainly concentrated on false alarms [35].

2.5.2 Automated Fault Diagnosis

The conventional means of achieving automated diagnosis is by establishing a library of faults, based on field experience and manufacturer data, and using this knowledge to build a system to identify the potential failure sources and recommend corrective action. Nowadays there are many commercially available software packages that are implemented to achieve this task. However, the reliability of these packages depends greatly upon the accuracy of faults identified by experts. Years of accumulated knowledge are typically necessary to establish all the necessary rules for engine diagnostics [35] and researches are underway to develop an improved automated fault diagnosis system on many engine models. But still, there is no a single package that give 100% support for engine fault diagnoses [35].

2.5.3 Neural Network Based Gas Turbine Engine Diagnosis

Research is undergoing to build a model to traditional methods of gas turbine engine diagnosis using neural network approach. Neural network utilizes its property of non-linear modeling technique to develop a diagnosis model for many sophisticated machines. A neural network model can also provide a general tool for classifying different types of faults. The main advantage of neural networks in this line is their ability to learn the faulty and normal operating signatures of the engine from actual test data and help with the reliable classification of faults in engines, without requiring detailed system models. In regard to this, neural networks are trained using deliberately entered faulty data.

According to [42], neural network has been applied for engines diagnostic monitoring and control that includes,

- Interpreting sensory data to extract new information.
- Detecting specific signatures from sensors, in order to detect and identify faulty condition.
- Achieving the non-linear mapping necessary for efficient engine modeling and the implementation of advanced engine control strategy.

CHAPTER THREE

REVIEW OF RELATED LITERATURE

3.1 Data Mining

Ability of human beings to generate and collect data has been increasing rapidly in the last several decades. Attributing factors for this increase include the wide spread use of cheep and powerful computers, computerization of many businesses, and technological advances in data collection tools [8]. As a result organizations in the late 1990's typically have large stores of data [40]. This fast-growing tremendous amount of data, collected and stored in large and numerous databases, has far exceeded human ability for comprehension without powerful tools. This explosive growth in stored data has generated an urgent need for new techniques and automated tools that can intelligently assist in transforming the vast amounts of data into useful information and knowledge [2,8].

The abundance of data, coupled with the need for powerful data analysis tools, has been described as data rich but information poor situation [8]. To bridge the gap between data rich and information poor conditions, different techniques are being developed. One among many is data mining [2].

Data mining is the task of discovering interesting and key patterns from large amounts of data where the data can be stored in databases, data warehouses, or other information repositories [8, 40].

Data mining is defined as the process of automated extraction of valid, previously unknown, comprehensible and actionable information from large databases and to make crucial business decisions [34, 6, 15]. Data mining is also popularly referred to as knowledge discovery in database, KDD [8]. Even if data mining is defined differently on different books, its idea revolves around finding information that is hidden in large data, which is new and difficult to comprehend by human being.

There have existed many traditional data analyzing tools and techniques [8] to deduce information from data. However in such tools, the user puts a hypothesis about specific data relationships and then the tools would help in verifying or disproving the hypothesis. In these traditional analysis tools, the discovery of important information out of data is highly dependent on the question posed by the tool user. This poses a serious limitation to traditional tools [8] as compared to data mining. In addition, when the numbers of variables to be analyzed become many it would be difficult to identify the important variables [36]. However this does not mean data mining has entirely replaced all other traditional techniques.

Data mining is a multidisciplinary field of study, drawing work from areas including database technology, artificial intelligence, machine learning, neural networks, statistics, pattern recognition, knowledge-based systems, knowledge acquisition, information retrieval, high performance computing and data visualization [8].

The discovered knowledge using data mining can be applied to decision-making, process control, information management, and query processing. Therefore data mining is considered as one of the most important interdisciplinary fields of study in the information industry [8].

Regardless of the problem area it tries to solve, the general process involved during the data mining activity include defining the problem, select the data, preparing the data, mining the data, and later deploying the model which will assist to take business action [5]. But it does not mean that the above steps are always carried out in specific order. Data mining is an iterative process [10].

3.1.1 Data mining functionalities

Using the above mentioned general data mining activities, different functionalities of data mining are used to specify different kinds of patterns within data [8]. These functionalities include,

- Concept/class description: characterization and discrimination
- Classification and prediction
- Association analysis
- Cluster analysis

- Outlier analysis
- Evolution analysis

In concept/class description, classes and concepts are described in summarized, concise and yet precise terms. The concept description can be done through data characterization, by summarizing the data of the class under study, data discrimination, by comparing the target class with one or a set of comparative classes, or by both data characterization and discrimination [8].

Classification is the process of finding a set of models (or functions) that describe the distinguishing data classes or concepts where as prediction is the process that predicts values rather than class labels. Data classification is a two-step process. In the first step, a model is built describing a predefined set of data classes or concepts. In the second step, the model is used for classification [5]. Prediction is usually used in the case where the predicted values are numerical data and is often specifically referred to as prediction [5].

Association analysis is the discovery of association rules showing attribute-value conditions that occur frequently together in a given set of data. Association analysis is widely used for market basket or transaction data analysis [15].

Cluster analysis is a process that uses cluster or group objects based on the principle of maximizing the entraclass similarities and minimizing the interclass similarity. Cluster analysis can help to formulate common rules for objectives found in one cluster [8]

The objects that don't comply with the general behavior or model of the data are studied in outlier analysis. In different application like fraud detection, the rare events can be more interesting than the more regularly occurring ones [8]

Data evaluation analysis is the last functionalities of data mining that describes regularities or trends for objects whose behavior changes over time. Although this may include characterization, discrimination, association, classification, or clustering of time-related data, distinct features of such analysis include time-series data analysis, sequence or periodicity pattern matching and dissimilarity based data analysis [8].

3.1.2 Neural Network As A Data Mining Technique

There are various techniques that are used to extract and discover important relationships within data. Some of the technique are neural networks, decision trees, genetic algorithms and nearest neighbor method [5, 39]. These techniques use statistical and machine learning methods to search databases for patterns that describe relationships in the data or predict future values or behavior [17].

As it will be discussed in the next section, neural networks are well suited for data mining tasks due to their ability to model complex multi-dimensional data [41]. They also provide a general, practical method for learning real-valued, discrete-valued, and vector-valued functions. Neural network learning is robust to errors in the training data and has been successfully applied to problems such as interpreting visual scenes, speech recognition, and learning robot control strategies [27].

The advantages that neural network offer over traditional statistics technique is that the model does not have to be explicitly defined before the experiments begin. There is no need to have preconceived idea about the model. Neural network can grasp the relevant data to develop the model, whereas to derive a statistical model, prior knowledge of the relationships between the factors under investigations is required [6].

3.2 Neural Networks

3.2.1 Introduction

Ever since computers were invented, human beings have wondered whether they might be made to learn. Yet human being does not yet know how to make computers learn as nearly as human brain does. But different attempts are made to make computers learn by imitating the method human brain functions [21, 27]. Computer scientists figured out that by simulating the actual human brain network into computers, they could make computers solve problems like human brains do [13].

Neural network is an artificial representation of the human brain that tries to simulate its learning process [31, 21]. Therefore the appropriate terminology for neural networks is artificial neural networks. But it is customary to leave the artificial and write only neural network [23, 22].

The neural network is composed of a large number of highly interconnected processing elements called neurons working in parallel and learns from its own experience [32]. Each artificial neuron is based on a simplified model of the neurons found in the human brain. Information processing in neural network is similar to the way a human brain does.

Neural network is a set of connected input/output units where each connection has weight associated with it. During the learning phase, the network learns by adjusting the weights so as to be able to predict the correct output for the input samples. Neural network learning is also referred to as connectionist learning due to the connections between units [8].

Advantages of neural networks include their high tolerance to noisy data as well as their ability to classify patterns on which they have not been trained. In addition, several algorithms have recently been developed for the extraction of rules from trained neural networks. These factors contribute towards the usefulness of neural networks for classification in data mining [23].

Neural networks differ greatly from traditional computers in both form and functions. While neural networks use a large number of simple processors to do their calculations, traditional computers greatly use one or a few extremely complex processing units [4].

Neural networks also differ greatly from traditional computers in the way they are programmed. Rather than using programs that are written as a series of instructions, as do all traditional computers, neural networks are "taught" with a limited set of training examples. The network is then able to "learn" from the initial examples to respond to information sets that it has never encountered before [1].

Traditional computers on the other hand use sequential processing, and perform faster than brains in computational problems but they are inefficient in problems like pattern recognition of knowledge processing. The reason is, traditional computers use a cognitive approach to problem solving, i.e. the way problem is to be solved must be known and stated in small unambiguous instructions [32].

However, neural networks and conventional algorithmic computers are not in competing with one another. There are some tasks more suited to algorithmic approach for conventional algorithmic computers and there are other tasks that are more suited to neural networks that require parallel processing [1]. In some areas cooperation of both may be required. Therefore one is not absolutely superior over the other.

3.2.2 Applications Of Neural Network

Inspired by biological nervous system, neural network is being used to solve a wide variety of complex scientific, engineering, and business problems. Potential application of neural network is mainly for problems with no known algorithms [41] or problems' analytical description is too complicated for computer processing. Commercial applications of neural network include investment portfolio trading, data mining, process control, noise suppression, data compression, and speech recognition. Neural network are ideally suited for such problems because, similar with their biological counterparts, a neural network can learn, and therefore can be trained to find solutions, realize patterns, classify data, and forecast future events [1, 23, 32].

Typically, neural networks may be employed in cases where example data is available to train the neural network, which sufficiently covers the underlying problem domain [4, 32, 23]. Since computational power has been increasing rapidly, neural networks have become a very competitive way of modeling nonlinear systems on the basis of measured input/output data [23, 32].

Neural networks are applicable in virtually every situation in which a relationship between the predicator variables (independent inputs) and predicated variables (dependent, outputs) exists, even when that relationship is very complex and not easy to articulate. Few representative examples of problems to which neural network analysis has been applied successfully are [22, 28, 32]

- Detection of medical phenomena
- Stock market prediction
- Credit assignment
- Monitoring the condition of machinery
- Engine management:
- Hand written character recognition,
- Speech recognition,
- Finger print recognition
- Neural networks that detect hypertension and heart abnormalities

3.3.4 Basic Structures Of Neural Network

Artificial neural network is approximately modeled on the human brain [23], which is the biological neural network [4]. Therefore discussing the basic structure of biological neurons in human brain first is of help to better understand artificial neural network.

The human brain is composed of approximately 100 billion neurons [4]. Each of these neurons is connected with up to 200000 other neurons [23]. The extensive number of neuron and their high degree of interconnection are important reasons for human brain to be capable of making vast number of calculations in a short amount of time. Neuron activity is typically excited or inhibited through these connections. The fastest neuron switching times are known to be on the order of 10^{-3} seconds, quite slow compared to computer switching speeds of 10^{-10} seconds [26].

All natural neurons have four basic components, which are dendrites, somia, axon and synapses.

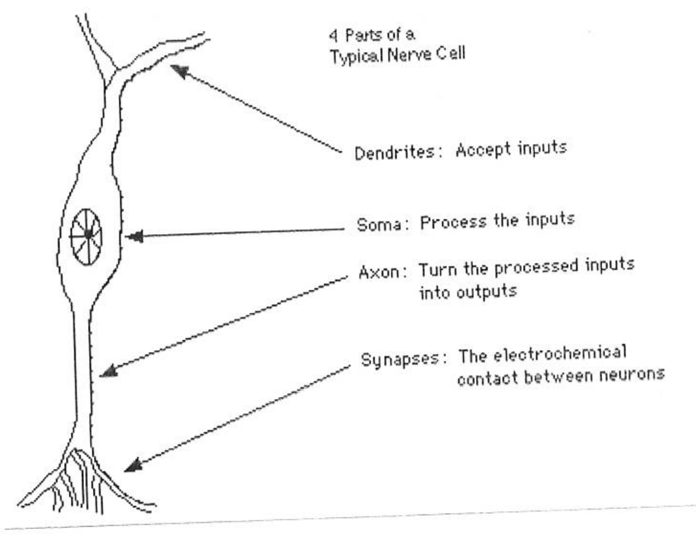


Figure 1 : The Basics of Biological Neuron [22]

A typical neuron collects signals from others through a host of fine structure called dendrites [32]. Information from other neurons, in the form of electrical impulse, enters the dendrites at connection points called synapses. A central cell body of a neuron, which is somia, accepts incoming information in the form of electrical signal and processes it accordingly [4]. Then finally the neuron sends out spikes of electrical activity through a long, thin strand known as axon, which splits into thousands of branches [23]. At the end of each branch, a structure called synapse converts the activity from the axon into electrical effects that inhibit or excite activity in the connected neurons. Learning in neuron occurs by changing the effectiveness of the synapses so that the influence of one neuron on another changes [32].

The basic units of artificial neural networks simulate the four basic functions of natural neurons [23]. In rough analogy, artificial neural network neurons are built out of a densely interconnected set of simple units called neuron [26, 16].

For a given neuron, inputs are summed using an adder weighted by respective synapse of the neuron. The result of the adder is directed to the activation function (transfer function). The purpose of the activation function is to produce an output depending on the neuron inputs sum result it receives from the adder[11] there by it limits the amplitude of the output. A neuron fires signals only if the sum of input signal exceeds its threshold [13]. With these arrangements each unit takes a number of real-valued inputs and produces a single real-valued output. The following figure illustrates this concept.

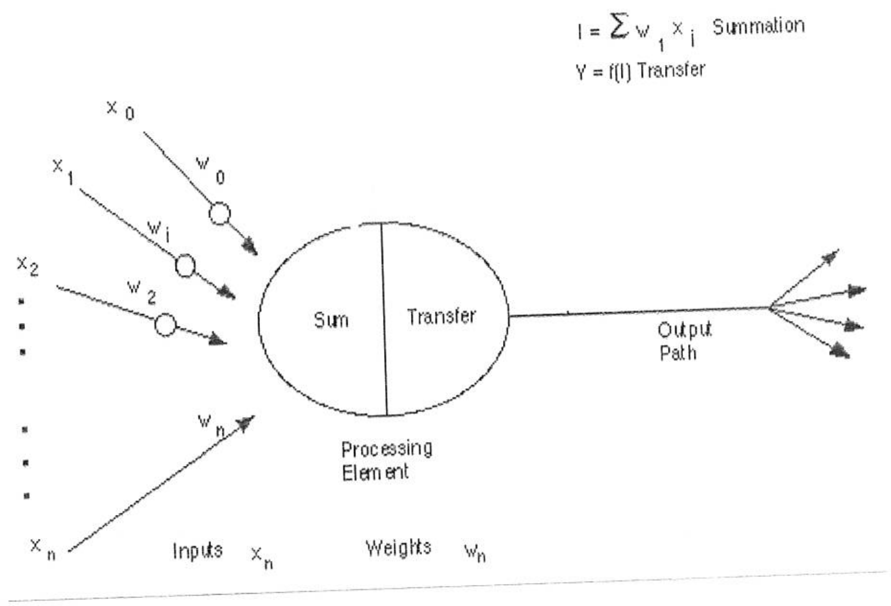


Figure 2 : The Basics of Artificial Neuron [22]

As it is shown from the above figure a neuron contains an adder and transfer function inside. There are many types of transfer functions used. The three most commonly used transfer functions are threshold transfer function, piecewise-linear transfer function and log-sigmod transfer function. Log-sigmod transfer function is by far the most common form of transfer function used in the construction of artificial neural networks. Log-sigmoid transfer function takes the input, which may have any value between plus and minus infinity, and squashes into a range zero to one. It is commonly used n backpropagation networks [11], which will be discussed in the next section.

The formula for log-sigmoid (sigmoid) transfer function is:

$$v = \frac{1}{(1 + e^{-s})^{-1}}$$

Where S= sum of the inputs to the neuron and

V= value of the neuron

A single neuron by itself cannot do much in problems of pattern recognition. in fact the real power of neural networks comes when neurons are interconnected to form a multilayer structure [11, 21]. The commonest type of multilayered artificial neural network consists of three layers, namely input layer, hidden layer and output layer [4]. The input layer takes external signals and transmits signals to the neurons in the hidden layer. Similarly neurons in hidden layer takes signals from input layer and transmit to output layer [30]. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. Similarly the behavior of the output units depends on the activity of the hidden units and the weights between the hidden and output units [32]. The connection of the neurons and the rule used to train the neural network are crucially important factors to determine the learning process of the network.

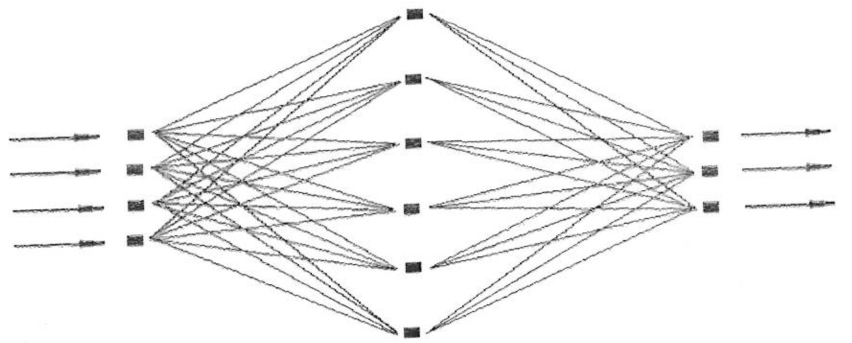


Figure 3 : Muliti layer Structure of Artificial Neural Network [4]

In many cases, the problem specification helps to define the structure of the network. Usually the number of network inputs is equal to the number of problem inputs, and the number of neurons in the output layer depends on the number of problem outputs [11]. The number of hidden layers is arbitrary, although in practice, only one is used. The weighted outputs of the last hidden layer are inputs to the units making up the output layer [8]. Using this arrangements it is possible for the multilayer network to represent highly nonlinear decision surface that are much more expressive than the linear decision surfaces of single units [27].

3.2.4 Classification, (Types) Of Neural Network

The connections between neurons and the learning rule of the artificial neural network are the two means to describe the classification of artificial neural networks [22]. Let us see these two important points in more detail.

3.2.4.1 Neural Network Connection

The flow of data within the network is determined by the connections of neurons [6]. With this respect neural networks are generally classified as feedforward and recurrent networks.

In ***feedforward neural network***, information flows strictly in one direction only, from input layers to output layers. Feedforward artificial neural network tends to be straightforward networks that associate inputs with outputs [32]. Feedforward networks are used in situations where all the information to bear on a problem can be presented at once [2].

A ***recurrent neural network*** that is also called feedback neural networks distinguishes itself from a feedforward neural network in that it has at least one feedback loop [16]. Feedback networks can have signals traveling in both directions by introducing loops in the network [32]. Such networks are used in situations where we need the neural network to somehow store a record of prior inputs and factor them in with the current data to produce an output. Feedback networks are very powerful and can get extremely complicated [32].

3.2.4.2 Learning Methods Of Neural Network

Neural networks can also be categorized based on the learning methods they use. The learning methods of a neural network are nothing but the way weights are adjusted as the neural network gains experience or learns [22]. Learning corresponds to choosing the appropriate weight value for each edge in a neural network [27]. The learning rule of neural network is broadly categorized as supervised and unsupervised learning.

In **supervised learning** of a neural network, both the inputs and the outputs are provided [18] to the network. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights, which control the network [16]. This process occurs over and over as the weights are continually adjusted. The set of data, which enables the training, is called the "training set." During the training of a network the same set of data is processed many times as the connection weights are continuously refined [18].

There are many laws (algorithms) used to implement adjusting the weights during training for supervised learning. The most common technique is backward-error propagation, more commonly known as back-propagation [18], which was first proposed in 1980 [8].

Backpropagation learns by iteratively processing set of training samples, comparing the networks prediction for each sample with the actual known class label. For each training sample, the weights are modified so as to minimize the mean square error between the network's prediction and the actual class. These modifications are made in the "backward" direction, i.e. from the output layer through each hidden layer down to the first hidden layer (hence the name backpropagation). Although it is not guaranteed, in general the weights will eventually converge, and the learning process stops [1].

The ***Backpropagation*** algorithm is the most commonly used ANN learning technique. It is appropriate for problems with the following characteristics [8].

- Instances are represented by many attribute-value pairs
- The target function output may be discrete-valued, real-valued, or a vector of several real- or discrete-valued attributes.
- The training examples may contain errors
- Long training times are acceptable
- Fast evaluation of the learning target function may be required.
- The ability of humans to understand the learned target function is not important.

The weight update loop in backpropagation may be iterated thousands of times in a typical application. A variety of termination conditions can be used to halt the procedure. One may choose to halt after a fixed number of iterations through the loop [1, 27], or once the error on the training examples falls below some threshold, or once the error on a separate validation set of examples meets some criterion [18,8,27]. The choice of termination criterion is an important one because too few iterations can fail to reduce error sufficiently, and too many can lead to over fitting the training data [27].

The other type of training is called **unsupervised training**. In unsupervised training, the network is provided with inputs but not with desired outputs [27, 16]. The system itself must then decide what features it will use to group the input data. This is often referred to as self-organization or adaptation [27, 16, 18].

CHAPTER FOUR
EXPERIMENTATION

4.1 Introduction

This chapter presents detailed steps taken when the experiment of the research was conducted. The discussion includes the data collection, data preprocessing, model building and model testing for its accuracy.

4.2 Data Collection

As mentioned earlier, the objective of this research is to explore the potential applicability of data mining technology in developing a model that can support gas turbine engines performance evaluation.

An important requirement to any data mining activity is the data itself. The data this research has made use of is obtained from Ethiopian Airlines Technical Service. As one of the major gas turbine engines operator in the country, Ethiopian Airlines keeps records of all the engines it is operating. Different gas turbine engines model documentations were explored for this research.

The history of each engine starting from date of manufacture was found in engine's documentations. The history includes different maintenance tasks like periodic maintenances, engine disassembly, engine parts inspections, modifications, repair, assembly, and performance evaluation check records the engine had during entire duration of its service. Since the research is specifically focused on engines performance evaluation, data that was concerned with performance evaluation were collected from engine documentation.

The data size and data consistency of its recording system in the documentations were the main criteria used to select the model of gas turbine engine for this research. The PT6A-27 model gas turbine engine, which is used to power DHC-6 aircraft, was found to satisfy these criteria.

The data that was identified for this research was available in a manual format. Therefore it had to be converted into an electronic format, since it was compulsory step to exercise data mining activity on data. Direct entering of data from engine performance check records into computer was difficult. Therefore a table was prepared to collect all the required data. The attributes that were collected are shown in the following table.

No	Code	Description	Values
1	Date	The date the engine was evaluated	Date
2	Test Cell	Test Cell No.	Unique value (categorical)
3	Engine Model	Engine model	Unique Value (Categorical)
4	Engine Work Order	Engine identification In maintenance floor	Unique Value
5	Engine Serial Number	Engine identification number issued by manufacturer	Unique value
6	SB 1457 or SB 1479 or SB 1290 incorporated?	Modification incorporation codes	2 Unique Categorical Values (Yes/No)
7	1st Stage Vane Class	Compressor turbine vane flow area	Number
8	2nd Stage Vane Class	Power turbine vane flow area	Number
9	Fuel S.G	Specific gravity of fuel	Number
10	TSG	SG Fuel sample temperature	Number
11	LHV	Latent heat of vaporization	Number
12	SG BASE	Turbine Meter Calibration Standard	Number
13	Test Condition (Cycle)	Engine power settings during data collection.	4 categorical nominal unique values
14	Baro	Ambient Barometric Pressure	Number
15	PSN	Test Cell Static Pressure	Number
16	Ng	Gas Generator (compressor) Speed	Number
17	Nf	Power Turbine Speed	Number
18	Torque	Pressure	Number
19	Wf	Fuel flow	Number
20	Tf	Fuel Temperature	Number
21	Tt_1 (Avg)	Ambient air temperature	Number
22	T_t5	Air temperature at engine station 5	Number
23	PS3	Compressor discharge air pressure	Number
24	$T_{t7}1$	Engine Exhaust Nozzle Temperature #1	Number
25	$T_{t7}2$	Engine Exhaust Nozzle Temperature #2	Number
26	$T_{t7}3$	Engine Exhaust Nozzle Temperature #3	Number
27	$T_{t7}4$	Engine Exhaust Nozzle Temperature #4	Number
28	Total Time	Time duration the engine operated since new	Number
29	Time Since Overhaul	Time duration the engine operated since last overhaul	Number
30	Total Cycles	Total cycles the engine operated since new	Number
31	Cycles since Overhaul	Total cycles the engine operated since last overhaul	Number
32	Type of repair	Type of repair before performance evaluation test	2 unique categorical Values
33	Evaluation Result	The outcome of evaluation.	2 unique categorical values

Table 1 : Attributes Selected and Collected

Attributes from item-1 to item-27 in table-1 were taken from *Log of Engine Test*, a table that is prepared by engine manufacturer to collect data during engine run for performance evaluation. *Log of Engine Test* table is attached at the appendix-2. Out of all the attributes recorded on *Log of Engine Test* table, this research uses only the ones that are required by Data Reduction Computer Program. (See section 2.4 for Data Reduction Computer Program).

The last six attributes in table-1 are not found in *Log of Engine Test table*. However the first four of these six attributes, that are "Total Time" (item-28), "Time Since Overhaul" (time-29), "Total Cycle" (time 30), and "Cycles Since Overhaul" (item-31) tell duration (age) of the engine on service since manufacture and last overhaul. The last two attributes that are "Type of Repair" (item-32) and "Evaluation Result" (item-33) contain information regarding the type of maintenance the engine had undergone just before the performance evaluation test and the outcome of the evaluation respectively. According to domain expert openion, these data have influence on the performance of the engine during evaluation. On some case the test performance evaluation acceptance values may depend on these values. Therefore they were included as the part of the data collection.

After the data was collected from PT6A-27 engines documentations with the help of the above table, the next step was to enter it into computer. These manual entering was a lengthy, painstaking and methodical process. Since wrong data have erroneous effect on the model to be built, extreme care was taken. After the data conversion into electronic format was done the data is crosschecked with the documents it was extracted from. The domain experts were also asked for assistance to go through the data if unusual data entry was made.

The data is entered in Excel format to take advantage of easier data manipulation and compatible interaction with MATLAB 6.5, which is used as a platform for building the model.

The class distribution of the entire collected records is shown in the following table.

Evaluation Result	Number of Records	Percentage
Accepted	362	60
Rejected	240	40
Total	602	100 %

Table 2 : Collected Data Class Distribution

4.3 Data Preprocessing

The data that was collected and converted into an electronic format passed through important steps of preprocessing as cleaning, integration, transformation and reduction. Preprocessing the data enhanced the excellence of the data thereby helping to improve the precision and effectiveness of the preceding data mining process [27]. The tasks executed under each data preprocessing steps are discussed as follows. However some tasks under data preprocessing were already done during the data collection process.

For example the data acquired regarding engine performance evaluation and the engine life history were located at different places within engine documentation. These two data regarding one engine had to be integrated from different sources in one record during data collection. Therefore task of data integration was done during data collection. However each of the remaining tasks under data preprocessing are described as follows.

4.3.1 Data Cleaning,

The reliability on output of data mining procedure highly depends on cleanliness of the data [8]. Data cleaning includes filling in missing values, smoothening noisy data, identifying or removing outliers and removing inconsistencies [5, 8]. With regard to this, the data cleaning activities done is discussed as follows.

Originally 602 engine performance evaluation records were collected. However a total of 22 records removed due to their missing values that account more than 50% of the total values considered. The missing attributes were also very important for determining the engine test performance class, as suggested by domain expert. The cause of missing values was due to missing papers from engines documentations. Therefore decision was made to exclude those records form the list.

Originally the use of Log of Engine Test to collect engine performance data for evaluation greatly assured the data to be consistent. As indicated on table-1, these data include engine and atmospheric parameters like atmospheric pressure, atmospheric air temperature, rotational speed, fuel flow, fuel temperature, air pressure ... etc. The instruments from which these readings were taken are periodically checked for proper operation. This also helps further to enforce consistent way of data collection and minimizes greatly the introduction of noisy data. Yet seven records were found to have PS3 reading (item-23 in table-1) extremely low for normal functioning engine at any condition. PS3 is the engine compressor discharge air pressure. The value is crosschecked with the original documents and the problem was identified to be from data recording itself. The problem is consulted with the domain expert and those values were replaced with the average values of the class each record belongs in.

4.3.2 Data Transformation,

Data transformation is also an important task under data preprocessing. It transforms data or consolidates into forms appropriate for data mining [8]. Data transformation involves data smoothening data aggregation, data generalization, data normalization and new attribute construction. The collected and cleaned data for this research constitute attributes with different scale of measurements. Most of the attributes in table-1 have numerical values however the data also contains two attributes “Type of Repair” (item-32) and “Evaluation Result” (item-33) that takes nominal categorical values and an attribute "Test Condition" (item 13) that takes four ordinal categorical values.

Since the technique selected to conduct the research was a neural network the entire attributes transformations should meet the neural network inputs criteria. All the inputs to neural network should be in numerical format [11] much preferably in a normalized format. This research adapts min-max normalizations for all numerical data attributes in the range between 0 and 1.

Attributes with out numerical values, were transformed in a format recognizable by neural network. For data attributes that contains nominal categorical values, with two values, one bit is assigned having 0 and 1 for each values. These assignments were considered for attributes "Type of Repair" (item-32) and "Evaluation Results" (item-33) in table-1. Respective assignments are shown in the following table.

Field Name with Actual Value	**Transformed Value**
Type of Repair = "Repair"	0
Type of Repair = "Overhaul"	1
Evaluation Result = "Rejected"	0
Evaluation Result = "Passed"	1

Table 3 : Nominal Attributes Transformation

The field "Test Condition" (item-13 in table-1) contains numerical values representing ordinal categorical values having four possibilities. Two binary digit numbers are used to represent these numerical values. Their respective transformed format is shown in the following table.

Field Name With Actual Values	Transformed Representation
Test Condition = "1"	00
Test Condition = "2"	01
Test Condition = "3"	10
Test Condition = "4"	11

Table 4 : Ordinal Attributes Transformation

4.3.3 Data Reduction,

A technique of data reduction is applied on data to obtain a reduced representation of the data set that is much smaller in volume, yet closely maintains the integrity of the original data [5, 8]. Data mining on the reduced data set produces the same (or almost the same) analytical results with original data set. Therefore during the process of data reduction care was taken not to remove data that affect the data mining analytical results.

Among strategies for data reduction such as data cube aggregation, dimension reduction, data compression, numerosity reduction, discretization and concept hierarchy generation [5], the strategy selected for this research is dimension reduction since the other strategies are more applicable to data warehouse.

Dimensional reduction is the process of identification and removal of irrelevant, weakly relevant or redundant attributes from the data. For this research dimension reduction was done in three steps. On the first step, relevancies of the attributes for outcome of the result were considered and a total of five attributes removed. Attributes that were identified as irrelevant for this research were “Date” (item-1), “Test Cell” (item-2), “Engine Model” (item-3), “Engine Work Order” (item-4), and “Engine Serial Number” (item-5) that are all found in table-1. On the second step of data reduction, attributes that had constant values throughout the data were identified and removed. These include “LHV” (item-11) and “PSN” (item-15).

In the third step of data reduction, correlation of all attributes computed and analyzed. Since one of the important characteristics of neural network is its potential to represent the non-linear relationship among inputs, attribute pairs having high correlation value replaced by anyone of the two. Decision on which attribute to select was made in consultation with domain experts.

After the above removal of attributes, the total number of attributes left was 12 including the class label. They are listed in the table shown below.

No	Name of Attribute	Descriptions
1	Fuel SG	Specific Gravity of Fuel Used
2	Baro	Atmospheric Pressure
3	Test Condition	Test Condition
4	Ng	Compressor Turbine RPM
5	Nf	Power Turbine RPM
6	Torque	Torque Output
7	Wf	Fuel Flow
8	TT5	Air Temperature at station 5
9	TT1	Atmospheric Temperature
10	Work Done	Type of Work Done
11	TT	Total Time Engine Give Service Since New
12	Evaluation Result	Engine Performance Result

Table 5 : Attributes Selected to Build Model

4.4 Modeling Technique

This research used the best known and the simplest learning algorithm for a neural network, which is ***backpropagation*** that is efficiently applied in supervised method of learning. Back propagation learning algorithm was discussed in section 3.3.6.2.

The platform used to design, develop and train the model was MATLAB 6.5 that has a neural network toolbox. The Neural Network Toolbox in MATLAB 6.5 provides tools for the design, implementation, visualization, and simulation of neural networks. It also allows importing potentially large and complex data sets and provides comprehensive support for many proven network paradigms. Because neural networks require intensive matrix computations, MTLAB 6.5 provides a natural framework for rapidly implementing neural networks and for studying their behavior and applications.

4.5 Model Test Design

Estimating classifier accuracy helps to evaluate how accurately the classifier will label the data on which the classifier has not been trained [8] and also to compare different classifiers. There are different methods used to assess classifier accuracy. Holdout and cross-validation are two common ones based on randomly sampled partitions of the given data [8]. And this research will use holdout method for evaluating classification accuracy during training.

4.6 Building the model

4.6.1 Defining network architecture

The architecture of a neural network includes a description of how many layers a network has, the number of neurons and transfer functions on each layers [11]. Usually neural networks use three layers architecture, i.e. input layer, hidden layer and output layers [7]. In this study and the same is adapted to the model intended to build.

Usually the number of neurons on input and output layers is defined by the problem to be solved. The number of neurons in input layer is equal to the number of input variables to the model, and the number of neurons on output layer is equal to the number of output variables [28]. Therefore this research will use a model with twelve neurons on input layer and only one neuron at output layer.

For the case of neurons on hidden layer, there is no clear rule as to the "best" number. Most of the time trial-and-error method is employed, starting with a small number and additional neurons are gradually added until some performance goal is achieved [7]. In other cases, starting point for the number of neurons in hidden layer is taken equal to half the sum of the number of input and output units [28]. Since its number is critical to the performance of the model, many models with different number of neurons will be built and checked for classification accuracy until one is found with best performance.

Since the output of the model is constrained between zero and one, Logsig transfer function will be used on both hidden and output layers of the model. An advantage of Logsig transfer function is that, it has good performance within this constrained range [11].

4.6.2 Data partitioning for model training,

When neural networks are modeled using neural network toolbox in MATLAB 6.5, first the data that will be used for training, validation and test has to be spitted exclusively. First the entire data is partitioned into two. The first set contains 75% (435 records) of the data that will b used to train the model. The remaining 25% (145 records) are reserved to test the classification accuracy of the model. The data, which is used for training, is further divided into two. It's 75% (327 records) is used for training whi8le the remaining 25% (108 records) are used for validating the models.

4.6.3 Training the model

In neural network, training to find optimal network design is a trial and error process [1], therefore other than the neural net parameters mentioned in section 4.6.1, the training function used is a gradient descent with momentum. At an early stage of training the model, the learning rate 0.01 learning rate was used.

In neural network, training is carried out by comparing the neural network output and actual value to adjust the connection weight depending on magnitude the error. Learning rate determines how big a change must be made on the connection weight adjustments of the network towards the correct value.

On an initial stage of training, the model managed to correctly classify only 34% of the test records. Neural networks training using MATLAB provide graphical illustration on the progress of the training. The graphical illustrations observed shows that the training and the validation errors decrease very slowly. The possible reasons for slow rate of error decrease can be small number of neurons at the hidden layer and very low learning rate used. Therefore learning rate is increased until significant improvement on performance was observed.

At learning rate of 0.1, performance of the model starts to improve significantly. Therefore learning rate of 0.1 was used to train many models with different number of neurons on hidden layers. The number of correct classification with their respective number of neuron is summarized in the following table.

At learning rate 0.1

# of neurons on Hidden layer	1	7	9	11	13	15	17	19	21	23	25	30
Classification Accuracy(%)	64.9	71.3	77.2	77.4	73.6	72.4	77.8	82.2	78.4	78.6	79.2	78

Table 6 : Initial Classificaton Accuracy

The above table shows, the better classification accuracy was achieved, when the hidden neuron was 19. With this reference, the immediate neighbors of 19 and 19 itself selected. Classification accuracy of these three architecture was checked by varying the learning rate.

Therefore three models with 18, 19 and 20 neurons in hidden layer were built and checked by varying the learning rate with 0.1 intervals. Their respective classification accuracy with different learning rate is summarized in the tables shown below. The range of learning rate shown for each model in the following tables is limited to values with better classification accuracy.

 # Of neurons on Hidden layer = 18

Learning Rate	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Classification Accuracy(%)	81.9	87.14	91.2	94.4	94.9	93.9	96.45	96.8	97.04

Table 7 : Classificaton Accuracy with 18 neurons at hidden layer and different learning rate

 # Of neurons on Hidden layer =19

Learning Rate	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Classification Accuracy(%)	82	87.39	93.79	95.4	96.35	95.94	97.24	97	97.28

Table 8 : Classificatoin Accuracy with 19 neurons at hidden layer and different learning rate

 # Of neurons on Hidden layer = 20

Learning Rate	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Classification Accuracy(%)	81.28	88.87	96.5	93.21	96.13	96.05	96.4	97.72	97.32

Table 6 : Classificaton accuracy with 20 neurons at hidden layer and differentlearning rate

The above tabulated experimental results show that a the model that has the best classification accuracy is the one that has 20 neurons on the hidden layer that was trained with learning rate of 0.8. Its classification accuracy is 97.72%.

4.6.4 Evaluating the Model.

The model that was identified in the previous steps was evaluated using confusion matrix and cross validation methods.

4.6.4.1 Confusion matrix,

The classification accuracy of the neural model in classifying the performance evaluation of gas turbine engine was also evaluated by using new records that were not used during model training. For this purpose, a total of 12 records of gas turbine engine performance data from both classes were acquired in which seven of these new records belongs in "Accepted" class and the remaining five records belongs in "Rejected" class.

These selected data was given to the model with best classification accuracy identified in section 4.6.3, and respective predicted classes are shown in confusion matrix in table-1. For problems of classification, confusion matrices describe more on the test results showing the actual versus predicted class values. It not only shows how well the model classifies but also presents the details necessary to pinpoint where things might have gone wrong [15] during classification.

		Actual		
		Accepted	Rejected	Total
Predicted	Accepted	6	0	6
	Rejected	1	5	6
	Total	7	5	12

Table 10 : Confusion mateix for the best model selected

The confusion matrix shown above indicates the model misclassifies only one record out of 12. Close inspection was made on the record misclassified by the model. The result found affirms the competency of the model for real world application. The misclassified engine performance had a marginal performance. The decision to accept its' performance was made due to shortage of spare engine in the maintenance facility. This property is important especially because, the model can be designed to give more information on the state of the performance in addition to just the acceptance or rejection of the engine performance.

Other means of assessing the classification accuracy of a classifying model is by determining the sensitivity and specificity of the classification. This method is an alternate way of illustrating the classification accuracy of a classifier model as shown below.

	Sensitivity	Specificity
Accepted	85.7%	100%
Rejected	100%	85.7%

Table 7 : Sensititity and specificity of the two classes

4.6.4.2 Cross Validation.

The other method to evaluate the classification accuracy of the model is to use cross validation. This method is usually used on cases where there is limited data set for training and testing. The data set is divided into a number of equal size divisions. Then, each division is used for the test data, and all the others are used for training the model. The classification accuracy determined from each division is averaged to get the overall classification accuracy [15]. The standard way of predicting the error rate of a learning technique given a single, fixed sample of data is to use stratified ten fold cross-validation [15]. This research will also use this method for evaluating the model classification accuracy. The classification accuracy for each partition is shown in the following table.

Partition #	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Classification Accuracy (%)	95.51	97.93	99.31	100	98.97	98.97	93.8	95.1	99.66	97.24

Table 8 : Classificaton accuracy of the model with ten fold cross validation

From the above table the average classification accuracy from the ten fold cross validation is 97.65%, which is almost equal to the best model identified in section 4.6.3.

From the experiment finding it is clear that the models memorize better with more number of neurons on hidden layer trained with higher learning rate. This finding will be used as a good input for further researches on same or related fields of study.

CHPATER FIVE

CONCLUSION AND RECOMMENDATION

5.1 Conclusion

Ability of human beings to generate data has been increasing rapidly in the last several decades. This enormous amount of data, collected and stored in large and numerous databases, has gone beyond human ability for understanding without powerful tools. This explosive growth in stored data has generated an urgent need for new techniques and automated tools that can intelligently assist in transforming the vast amounts of data into useful information and knowledge [2,8]. One of the potential solutions for this problem is data mining.

The use of data mining technology has progressively become popular and proven to be pertinent for many sectors such as stock market prediction, credit assignment, biomedical and DNA data analysis, financial data analysis and also in monitoring the conditions of machineries. Particularly in the case of monitoring the conditions of machineries, data mining tools can be instrumental in reducing maintenance costs.

The objective of this research undertaking was to assess the potential of data mining technique in classifying performance of gas turbine engines by developing a model that could help in predicting the acceptance of gas turbine engine performance.

The methodology used to conduct the research followed three basic steps, namely data collection, data preparation and model building/testing. However the iterative nature of data mining activity made steps not to be always followed in strict order. In some instances, there were needs to go back and forth between the different steps.

A data set totaling 602 gas turbines performance evaluation test records with their respective results was collected to build a neural network model in supervised method of learning. In order to identify a model that can classify the performance of gas turbine engines, several models were built and checked for their classification accuracy.

The model with the best classification accuracy is found to be the one that has eleven neurons in hidden layer and trained with 0.7 learning rate. The model has twelve input nodes taking eleven attributes and one out put node. The input variables are "TSG" (item-10), "Test Condition" (item-13), "Baro" (item-14), "Ng" (item-16), "Nf" (item-17), "Torque" (item-18), "Wf" (item-19), "TT1" (item-21), "TT5" (item-22), "TT" (item-28) and "Work Done" (item-32). The model has 97.72% classification accuracy.

After the model selection was made, the model was supplied with twelve new records and the accuracy of classification was evaluated. The evaluation was made using two methods, confusion matrix and ten-fold cross validation method. From the confusion matrix, it was learnt that the model correctly classifies eleven records out of twelve. One accepted record was misclassified as rejected. Further investigation on misclassified record revealed it had marginal performance. As an additional means to assess the classification accuracy, sensitivity and specificity were also computed.

During the entire course of the research, reference, recommendations and sentiments given by domain experts were monumental. The domain expert played major role for identification of attributes “TT” (item-28) and “Work Done” (item-32). The research later proved their importance as shown on the classification accuracy.

However it is far from truth to assume data mining technology can replace experts who have many years of working experience in the area. After all with out domain experts involvement, it was practically impossible event to do researches in developing and updating models of this type. Generally speaking, the principal ingredient required to develop a model using data mining technology is the domain expert. The encouraging result found from this research was mainly by unreserved contribution of the domain experts. Therefore a model of this type only supports but by no means replaces the domain experts

The encouraging results found in this research indicate that data mining application is really a technology that can be considered to assist gas turbine engines performance evaluation not only for Ethiopian Airlines, but also to all other, organizations and corporations that operate gas turbine engines.

The results found in this research work have demonstrated the potential applicability of data mining technology to classify acceptance of gas turbines engines performance. It has shown that reliable assessment of gas turbine engine performance can be done using only some of the engine parameters collected during performance evaluation check.

If automating the system is made, quick assessment on engine's performance is possible even when the engine running. The encouraging results obtained from neural network also indicate that neural networks can be used to analyze inputs from sensors mounted on different parts of the engine. It also indicates neural networks can be trained to differentiate acceptable gas turbine engine performance.

The encouraging classification accuracy of the model found in this research can insight more researchers to be done for adapting models of this type for different models gas turbine engines maintenance practices.

5.2 Recommendations

This research work has disclosed the potential applicability of data mining technology to classify gas turbine engines based on their performance. Based on the conclusion of this research work, the researcher would like to make the following recommendations especially in relation to the possible application of data mining technology in supporting gas turbine engines maintenance activities.

The researcher would also like to note that this research, as an academic exercise, should only be considered as a preliminary effort to assess the applicability of data mining technology for gas turbine engines performance assessment. Accordingly, the findings of this research undertaking can fairly be considered as a contribution towards more in depth and comprehensive study in the application of data mining for gas turbine engines maintenance activities.

Some of the following recommendations relate to points for consideration in a similar future research work.

- **Conducting Researches Of Same Kind On Different Models Of Gas Turbine Engines**

Even if this research work was only on one model gas turbine engine, namely PT6A-27, it is the considered opinion of the researcher that researches of these kind can be conducted on other models of gas turbine engines to develop a fully functional prototypes that can assist aviation maintenance personnel in assessing the performance on engines they operate.

- **Developing Diagnosis System**

This research can be used as a corner stone to explore the possibility of using data mining technology to develop gas turbine engines diagnosis system. Instead of following library of faults to identify and rectify technical problems on gas turbine engines, signatures of engines behavior on different types of faults, with the help of engine parameters can be taught for the model that will specifically identify when the problem arises.

- **Studying The Applicability Of Other Data Mining Techniques For Same Purpose.**

The neural network technique considered here is only one of the different data mining techniques. The experiment with this technique has shown an encouraging result. But different techniques such as decision tree could be checked on the similar problem area for their potential.

- **Building A Model That Can Do More Than Classification**

This research addresses the classification of gas turbine engines performance. It either accepts or rejects the performance. However even if engines fail in same class they may not have the same degree of acceptance or rejection. Further work on the same research area can help to develop models that can differentiate this kind of situations.

- **Early Warning On Deterioration Of Engine Performance**

This research is worked on engines that are under performance evaluation test. However research of the same kind can be done on engines that are giving service. for example gas turbine engines that are mounted on aircraft are always monitored for normal operation using different sensors. Until any one or more of engine parameters are observed exceeding the limit specified by the engine manufacturer, the engine keeps on working until its due time is reached for change or removal. However, even if all parameters are within limit specified by the engine manufacturer, there could be a chance for the engine to deteriorate continuously through time. Early warning on the condition of the engine can be found by studying the trend of different engine parameters simultaneously using data mining technique.

REFERENCES

[1] Artificial Neural Network Technology. http://www.dacs.dtic.mil/techs/neural/neural_ToC.html

[2] Askale Worku (2001): Possible Applications of Data Mining Technology in Supporting Loan Disbursement Activity at Dashen bank S.C, Addis Ababa University, Addis Ababa

[3] Bahlmann, C., & Heidemann, G. & Ritter, H.: Artificial Neural Network for Automated Quality Control of Textile Seams http://ftp.informatik.uni-freiburg.de/papers/lmb/ba_he_ri_pr99.pdf

[4] Bernander, O. (2003): Neural Network, Microsoft Encarta Encyclopedia 2003 1993-2003 Microsoft Corporation.

[5] Berry, Micheal J.A. & Linoff, G., (1997): Data Mining Techniques, For Marketing Sales And Customer Support, New York: John Wiley And Sons Inc, Canada.

[6] Bigus, Joseph P (1996): Data Mining With Neural Networks, Solving Business Problems From Application To Development To Decision Support,' New York: McGraw Hill,

[7] Boger, Z. & Weber, R.: Finding An Optimal Artificial Neural Network Topology In Real-Life Modeling, http://www.dr-rene-weber.de/files/nc2000.paper1403-109.pdf

[8] Bruce, M. (1996): Defining data mining, Miller Freeman, Inc http://www.dbmsmag.com/9608d53.html

[9] Bent, R. & L. McKinley, (1985): Aircrafts Powerplants, Aviation Technology Series, USA

[10] David, S. (2001): Data Mining Blunders Exposed, Database programming and design magazine http://www.db2mag.com/db_area/archives/2001/q2/miners.shtml

[11] Demuth, H. & Beale, M.: Neural Network Toolbox User Guide, Version 3.0

[12] Escobar, J. (2004): Gas Turbine Engine Troubleshooting Tips, Aircraft Maintenance Technology, Vol(15) No(15), USA

[13] Khuu, H. & Lee, H. & Tsai, J. Machine Learning with Neural Networks and Support Vector Machines, http://www.cs.wisc.edu/~hiep/Sources/Articles/svms_nns_research.pdf

[14] Frieze, M. & M.Ennet, C. & Walker, C., Influence of Missing Values on Artificial Neural Network Performance, http://www.sce.carleton.ca/faculty/frize/MIRG_2001/Ennett_medinfo2001.pdf

[15] H.Whitten, I. & Frank, E. (2000): Data Mining : Practical Machine Learning Tools and Techniques with Java Implementations, Academic Press, USA

[16] Haykin, S. (1994): Neural Networks, A Comprehensive Foundation, Prentice Hall, Canada

[17] Herb, E., (1997): Predicting customer behavior with Neural Network". Database programming and design magazine http://grwy.ha.cn/sweetheart/9701eds2.html

[18] http://www.dacs.dtic.mil/techs/neural/neural_ToC.html

[19] Han, J., & Kamber M. (2001):Data Mining, Concepts and Techniques, Academic Press, USA

[20] Jane, J. (1975): Gas Turbine Theory and Jet Propulsion, India

[21] Jachen, F. (1999): Neural Network Overview http://rfhs8012.fh-regensburg.de/~saf39122/jfroehl/diplom/e-1-text.html

[22] Jeannette, L. (1994) : Introduction to Neural Networks, Design, Theory, and Application, California Scientific Software Press, Nevada City

[23] Klerfors, D., Artificial Neural Network, what are they? How do they work? In what areas are they used? http://hem.hj.se/~de96klda/NeuralNetworks.htm

[24] Koehl, M. : On-Board Engine Life Usage Monitoring By Real Time Computation. http://www.enginemonitoring.com/publications/icof2001.pdf

[25] Lambardo, D., (2000): Turbine Engines Data Transformation, Maintenance With Diagnostics, Aircraft Maintenance, V(19) Num(10), USA

[26] Letourneau, S. & Famili, F., Data Mining to Predict Aircraft Component Replacement http://www.site.uottawa.ca/~stan/papers/1999/componentFailureIEEE.pdf

[27] Mitchell, T., (1997): Machine Learning, McGraw-Hill, Singapore

[28] Neural Network, ST Neural Networks from StatSoft, http://www.statsoftinc.com/textbook/stneunet.html

[29] Pratt & Whitney Canada Overhaul Manual, Manual Part No 3013243

[30] Price, J. & J.Simoneau, M., Neural Networks Provide Solutions to Real-World Problems, http://www.mathworks.nl/company/newsletters/news_notes/sum98/sum98net.html

[31] Roelof K Brouwer Peng, PhD , An Integer Recurrent Artificial Neural Network for Classifying Feature Vectors. http://www.dice.ucl.ac.be/proceedings/esannpdf/es1999-18.pdf

[32] Stergiou, C., & Siganos, D.,: Neural Networks http://www.doc.ic.ac.uk/ñd/surprise_96/journal/vol4/csll/report.html

[33] Stich, T. & K.Spoerre, J. & Velasco, T., (2000) : The Application Of Artificial Neural Networks To Monitoring And Control Of An Induction Hardening Process, Journal of Industrial Technology, Vol-16, No-1, Nov 1999-Jan 2000.

[34] Thomas, C., & Begg, C. & Strachan, A. (1999): Database systems, A Practical Approach to Design, Implement and Management, New Jersey: Addison-Wesley

[35] Turney, P. & Halasz, M.: Contextual Normalization Applied to Aircraft Gas Turbine Engine Diagnosis
http://arxiv.org/ftp/cs/papers/0212/0212031.pdf

[36] Two Crows Corporation. (1999): Introduction To Data Mining And Knowledge Discovery
http://www.twocross.com/

[37] US Department of Transportation , Federal Aviation Administration, (1976) Airframe & Powerplant Mechanics (Powerplant Handbook),

[38] Vincent, E. (1950): The Theory and Design of Gas Turbine and Jet Engines, USA

[39] Weisoing, L., (1999): Data mining on the internet
http://www.trident.mcs.kent.edu/~javed/DL/surveys/IAD99s-datamining/

[40] Z Solutions, Neural Network and Data Mining,
http://www.zsoultions.com/sowhy.htm

[41] Solutions (1999): Neural Networks And Data Mining.
http://www.zsolutions.com/sowhy.htm

[42] R.J.Howlett, M.M.de Zoysa, S.D. Walters, Monitoring Internal Combustion Engines by Neural Network Based Virtual Sensing,
http://rjhowlett.complexnet.co.uk/papers/2000-2.pdf

Appendix 1 - Engine Specifications (PT6A-27)

Engine Specifications (PT6A-27)

OPERATING CONDITION	ESHP	SHP	PROPELLER SPEED (rpm)	JET THRUST (lbs)	SPECIFIC FUEL CONSUMPTION (lb/eshp/hr at 15°C (59°F)
Takeoff and Max. Continuous	715	680 (1)	2200 (3)	90	0.602
Max. Climb and Max. Cruise	652	620 (2)	2200	80	0.612

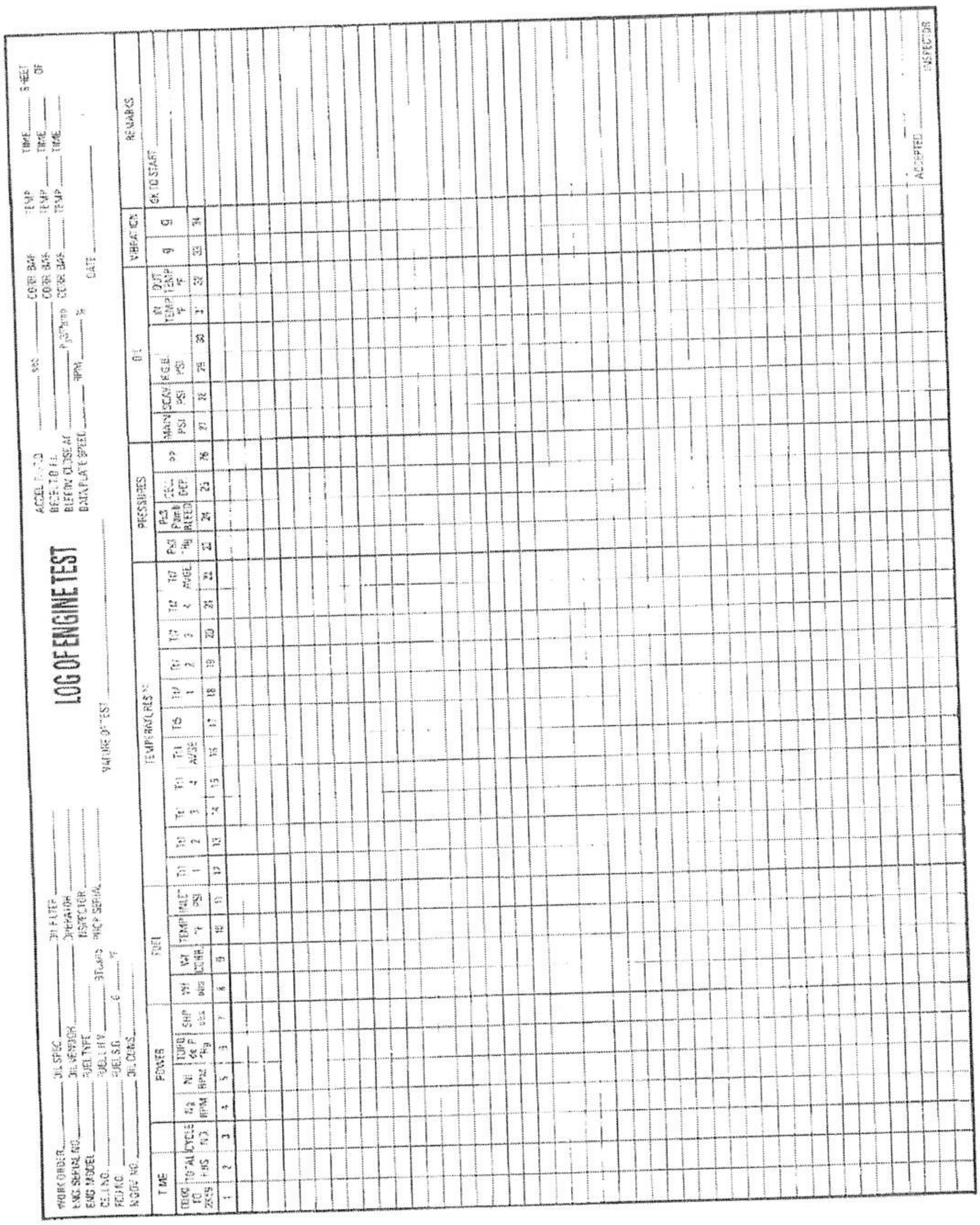
(1) Available to +21.7°C (+71°F) at 2200 rpm propeller speed.

(2) Available to +20.6°C (+69°F) at 2200 rpm propeller speed.

(3) Corresponding Rotor Speeds: Gas Generator - 38100 rpm (max). Power Turbine - 33000 rpm.

Engine type	Free Turbine
Type of combustion chamber	Annular
Compression ratio	7.0:1
Propeller shaft rotation (looking FWD)	Clockwise
Propeller shaft configuration	Flanged
Propeller shaft gear ratio	0.0668:1
Engine diameter (combustion case at room temperature)	19.0 in. (482.6 mm)
Engine length at room temperature	62.0 in. (1574 mm)
Dry weight (incl. standard equipment)	314.0 lb. (142.4 kg)
Oil consumption, maximum average	0.2 lb/hr (0.091 kg/hr)

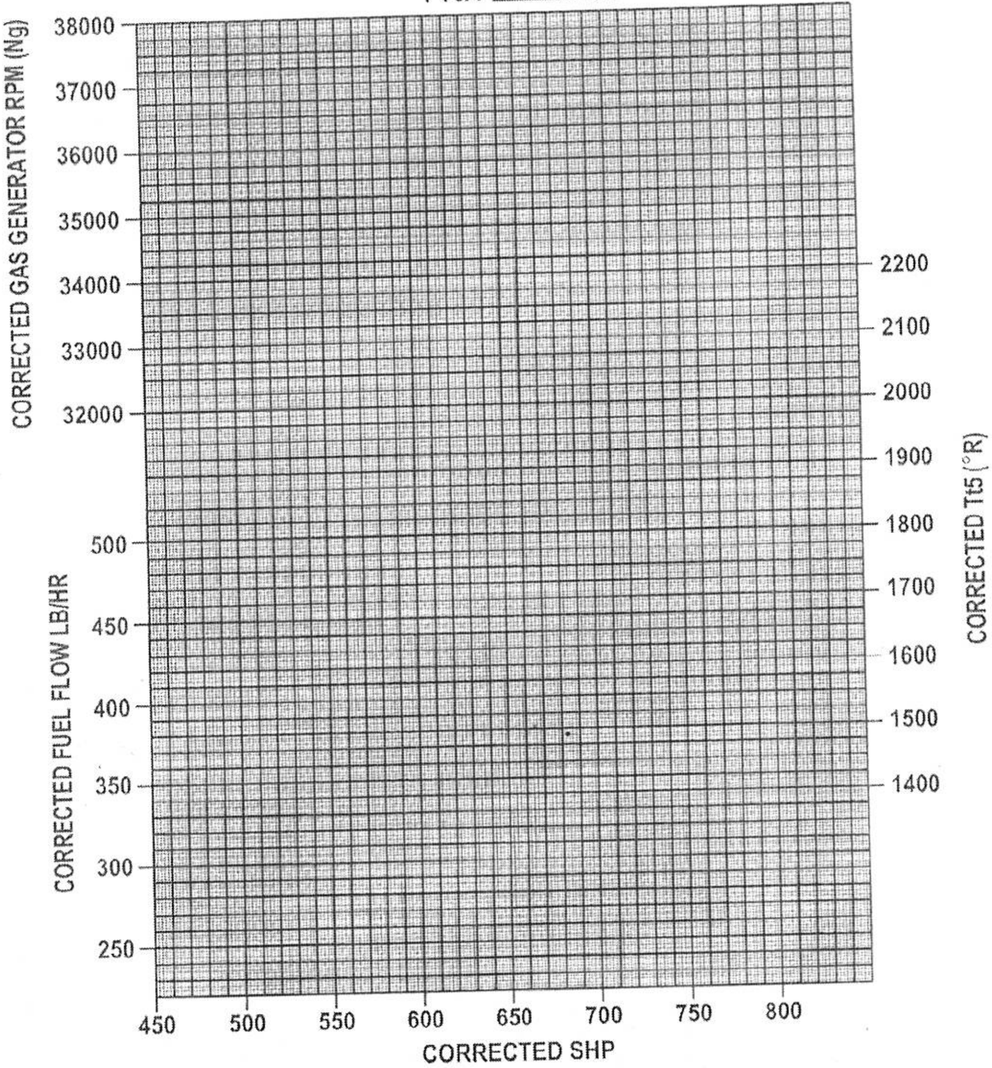
Appendix 2 - Log of Engine Test


LOG OF ENGINE TEST

Log of Engine Test - Sample Form

APPENDIX - 3
Engine Performance Curve (sample Form)

PT6A-________ S/N PC-E____________



CORRECTED ENGINE PERFORMANCE

RATING	SHP	Ng	SFC	Tt5°R
TAKE-OFF				
DATA PLATE SPEED		RPM %		
MAX. Ng WITH TRIM STOP		RPM AT T1 °F		

OPERATOR ____________

INSPECTOR ____________

DATE ____________

Engine Performance Curve - Sample Form

DECLARATION

This thesis is my original work, has not been presented for a degree in any other university and all sources of material used for the thesis have been duly acknowledged.

Haileleoul Gudeta

THE THESIS HAS BEEN SUBMITTED FOR EXAMINATION WITH OUR APPROVAL AS UNIVERSITY ADVISORS

W/ro Rahel Bekle