Intelligent Condition Assessment of Power Transformer Based on Data Mining Techniques

by

Monsef Tahir

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In recent years, the trade-off between quality and cost of power system components has become a matter of interest for many utilities. The widespread use of costly electricity networks either in residential or industrial areas has encouraged service providers to find a proper strategy that will minimize the overall life-cycle cost while keeping components in good working condition. The power transformer, which represents approximately 60% of the overall cost of the network, is ranked as one of the most important and expensive components. However, the transformer's sudden failure puts the system in a serious or critical condition which in most cases causes catastrophic loss to both utilities and customers. Significant attention has been given to monitoring and diagnostic techniques that observe any abnormal behaviour, assess the transformer's condition, and therefore minimize the probability of unplanned outage. Yet, applying many various monitoring tests is not always applicable due to the following factors: some tests require the unit to be taken out from service for testing, insufficient availability of man power, and significant cost of applying all the tests. Thus, there is a vital demand for an intelligent method of minimizing the number of monitoring tests without losing much information about the transformer's actual condition.

In this research, data mining techniques have been employed to evaluate the transformer's state through intelligent selection criteria that determines the optimal number of monitoring tests in cost-effectiveness. Feature selection technique based on ranker search method has been used to rank the monitoring tests (features) in a priority sequence from their individual evaluation, and to select the most inductive tests that provide the most information about the unit's condition. When the measured data from monitoring tests is collected and prepared, a diagnostic technique is applied to assess the condition of the transformer. In this regard, Support Vector Machine (SVM) has been utilized to perform this task due to its robust classification accuracy. SVM is first applied to the full number of tests, and then the number of monitoring tests is reduced by one after each classification process using the feature selection algorithm. The selected number of monitoring tests has shown the best possible accuracy the classifier can reach over the whole number of tests. Radial Basis Function (RBF) classifier has been used in the classification process for results comparison purposes. This proposed work contributes towards finding an intelligent method of evaluating the transformer state as well as minimizing the number of tests without losing much information about the unit's actual condition. Therefore, this method facilitates deciding a wise course of action regarding the transformer: either maintain, repair, or replace.

Acknowledgements

I would like to take this opportunity to express my deep gratitude to my supervisor Dr.Ramadan Elshatshat. During my master's, he guided me wisely and precisely to put me on the right path. I greatly got benefited from his experience and positive comments.

I also would like to thank my friends in Waterloo specifically the Libyan community for offering help and support all the times. And special thanks to the closest friends Elfaitoy, Nabil and Haithm.

I would like to thank all UW staff and members.

Finally, I would like to thank my family, relatives and all my friends in Libya for their support and encourage.

Dedication

This thesis is dedicated to my lovely parents who are raising their hands every day for asking Allah to help and protect me.

To: my lovely wife for her patience and support all the times.

To: my brothers and sisters for encourage and advice.

Table of Contents

List of Figures

List of Tables

Chapter 1
Introduction

1.1 General

In recent years, high investment in electricity networks has encouraged power utilities to find optimal management of capacity while optimizing the cost of the current components over their lifespans [1]. The power transformer represents approximately 60% of the overall cost of the network, and is designed to operate safely between 25 and 30 years. The power transformer is ranked as one of the most important and expensive components in the electricity sector [2, 3]. Recently, more attention has been paid to the life cycle management and condition monitoring techniques of power transformers because of their significant contribution in minimizing forced outage risks, reducing cost, and extending the nominal end of life. Approximately 27 monitoring methods to assess transformer condition are presented in the literature [3]. Furthermore, in many cases the capital cost of transformer unplanned outage costs millions of dollars [4]. The gradual deterioration in a transformer occurs for many reasons, specifically for in-service aged transformers. Overloading, lack of maintenance, design problems, environment temperature, and other factors speed up the deterioration process; therefore, condition monitoring and assessment procedures aim to track behaviour of components and detect incipient faults early; hence, reduced maintenance costs and proper health assessment can be achieved.

The wide diversity in monitoring and diagnostic techniques greatly help in the evaluation of power transformer status, but at the same time a lot of effort, expense, and skilled people are needed in order to achieve these tasks. Moreover, the time required for the unit to be out of service for off-line monitoring tests causes serious compensation challenges for service providers. This thesis presents an intelligent model based on optimization technique to properly monitor and diagnose the power transformer. Furthermore, the proposed model helps in finding an efficient yet reliable limited number of monitoring techniques to address transformer condition and hence minimize the challenges. The primary goal of this thesis is to help decision makers apply a cost-effective plan for an optimally selected number of monitoring methods to achieve proper condition assessment.

1.2 Motivation

In the past decades many and various monitoring methods have been developed to help utilities make the right decisions when performing maintenance, repairs, or replacement scenarios. The evaluation process of power transformer must be conducted based on measured and reliable information while resolving any problems that have occurred during its service. However, many challenges impede the application of many monitoring tests every time. Availability of skilled people, time demands (especially for major and off-line tests), electricity interruption in the absence of spare transformers, and the high costs incurred to perform these tests constitute the greatest obstacles to service providers. As the number of monitoring tests increases, the challenges increase. Therefore, the need for a technique that overcomes these challenges to compromise between the reliable assessment of transformer condition and the application of many monitoring tests has recently become a matter of interest. In this thesis, an intelligent model based on tests minimization technique has been proposed to both technically and economically assess transformer condition.

1.3 Thesis Objectives

The main contribution of this research is presenting a reliable and efficient model that evaluates transformer condition both economically and technically based on a specific number of monitoring tests which have been intelligently selected. The proposed model selects the most efficient and useful tests that provide as much information about transformer state as possible without going through all the monitoring tests. This work contributes towards solving many of the previously mentioned issues in order to meet utilities' satisfaction. Furthermore, this is an efficient model for decision makers to make wise decisions when it comes to maintenance, repair, replacement, or disposal actions. This study's objective is further sub-divided:

1. To investigate different classification techniques that might be used to evaluate transformer condition based on the available test data;
2. To develop feature selection algorithm that optimally selects the most inductive type of tests;
3. To develop intelligent classifier model that assigns each transformer to its proper class among five different classes based on measured data; and
4. To validate the efficiency of the proposed classifier and compare it with the most commonly used estimation techniques.

1.4 Methodology

The proposed methodology of this thesis is based on minimizing the number of monitoring methods by employing feature selection techniques using ranker search method. After preprocessing the data, the full set of tests is ranked first in a priority sequence by the algorithm that computes the information gain value of each individual test of the data set. After that, the classifier is applied to the full data set, and then the tests are reduced one at a time by eliminating the least inductive test and applying the classifier again. The best performance achieved by the classifier is the optimal number of tests that will be used to assess transformer condition. The output of the classifier will be the actual condition of the tested transformer. This model will help decision makers in making the right decision when it comes to planning actions. The following flowchart in Figure 1.1 explains the process of the methodology.

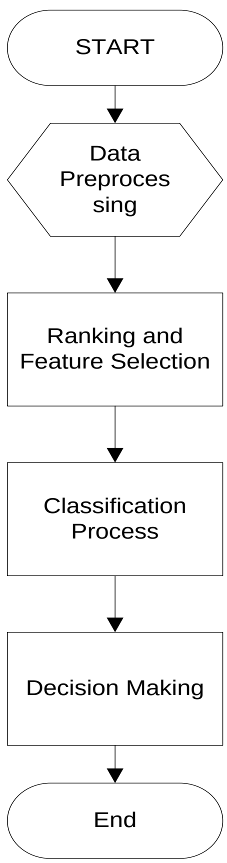

Figure 1.1 Thesis Methodology

1.5 Thesis Outline

This thesis includes six chapters following the introduction that have been organized as follows. Chapter two is an overview of transformer asset management in terms of maintenance and risk assessment, along with a survey on the most effective monitoring and diagnostic techniques. Chapter three explains the diagnostic techniques used in this research. Support vector machine (SVM) and radial basis function (RBF) as classifiers are investigated and explained in terms of operation and algorithm structure. Moreover, feature selection technique in specific “InfoGainAttributeEval” is considered for use in feature (test) reduction process. Chapter four introduces the most common set of monitoring tests used in this thesis. In chapter five, the model (classifier) has been developed to assess transformer condition based on measured data from in-service transformers. Chapter six presents the conclusion which summarizes the work in this thesis and its main contribution in evaluating transformer condition.

Chapter 2
Literature Review and Background

2.1 Power Transformer Asset Management

In recent years, the trade-off between quality and cost of power system components has become the main concern for many utilities. Due to widespread electricity networks and their high cost in both residential and industrial areas, a proper strategy is needed to minimize overall cost while keeping the component in good condition until it reaches its designed end-of-life or beyond. Hence, over the past 15 to 20 years asset management approaches have been applied on electricity sectors. Asset management can be defined as a series of management, engineering, financial, and economic procedures applied to an existing asset in order to reach a satisfactory or acceptable level of service in the most cost effective way. Effective asset management begins from study, design, construction, operation, maintenance, replacement, and disposal, as shown in Figure 2.1. In order for asset management practices to minimize the overall life cycle cost, a comprehensive program should be applied from the beginning through to the end of an asset's life so that a full return will be achieved from the asset service.

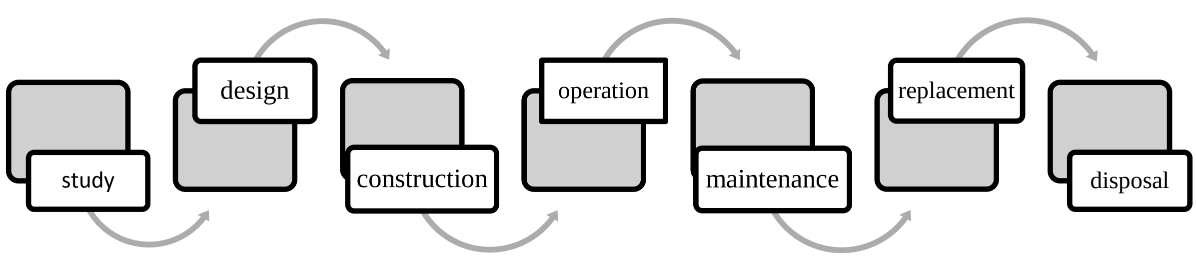

Figure 2.1 An effective asset management practice

In a power system network, the power transformer is ranked as one of the most important and expensive components. Its main function is to transform power from one circuit to another. Among all power system devices, power transformer is a reliable element which may function for more than 40 years, with some even working for 60 years. However, the sudden failure of the power transformer places the system into serious or critical condition which in most cases costs both utilities and customers catastrophic losses; moreover, the time for repairing or replacing such a component takes days to weeks and in some cases more than a month. This time increases the economic loss and risk on the system, specifically if the load is loaded to another transformer. Therefore, a successful asset management strategy can be applied to keep a component in good condition while being in service on order to avoid sudden interruption in electricity. Figure 2.2 represents a complete asset management approach which consists of two main sections: one for life assessment and the other one for decision options. Economic based action alternatives represent the main actions conducted on a transformer while being in service for full economic benefits.

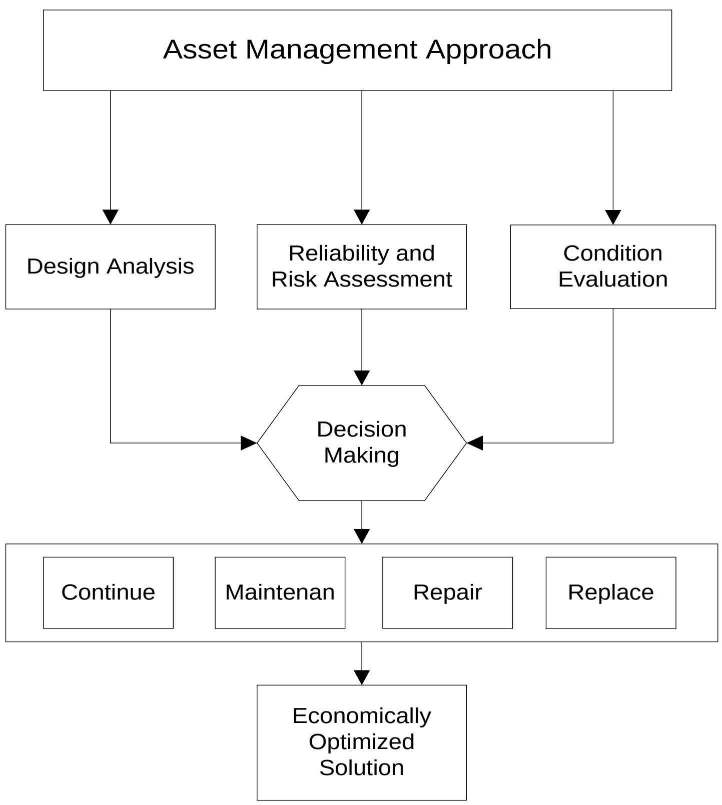

Figure 2.2 Power transformer asset management approach

2.1.1 Maintenance

In order to maintain good condition and implement satisfactory function, power transformer, like any other power system component, needs regular maintenance to facilitate successful operation. Satisfactory service from this equipment can be achieved if regular inspection procedures are applied and corrective measures are taken when necessary [5]. On the other hand, lack of attention and regular maintenance reduces the reliability and the chance of unplanned outage becomes more likely. There are many benefits of performing regular maintenance on power transformer to ensure satisfactory outcomes during the service life, which include:

- Extending the end-of-life;
- Increasing reliability and reducing unplanned outages;
- Minimizing risks;
- Minimizing overall cost (life cycle cost); and
- Making reliable decisions when considering repair and replacement options.

To ensure that all of the abovementioned advantages are met, three types of maintenances are proposed as shown in Figure 2.3:

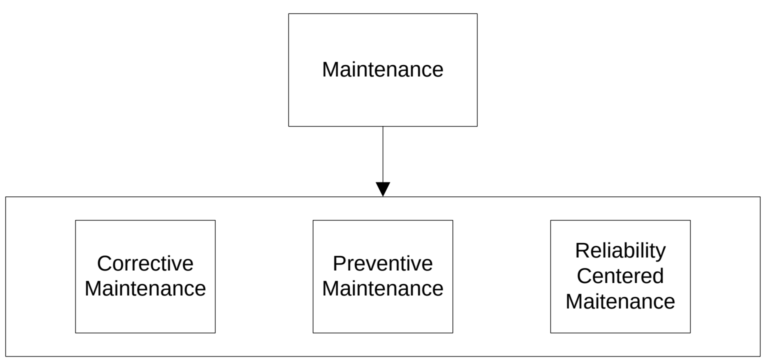

Figure 2.3 Types of maintenance

2.1.1.1 Corrective Maintenance

Corrective maintenance has been used for a long time. This type of maintenance is only performed when a failure or problem occurs to the component. Currently, this kind of maintenance is considered as a poor type of maintenance if conducted alone since it only takes place after the occurrence of the problem or unplanned outage. The advantages and disadvantages of corrective maintenance are listed and compared with the other types of maintenance in Table 2.1.

2.1.1.2 Preventive Maintenance

Preventive maintenance is planned maintenance that is conducted on a regular basis to keep transformers working continuously at their highest levels as well as minimize the probability of failure. Moreover, preventive maintenance is a very useful practice for detecting incipient faults that may lead to catastrophic risks if not repaired. This type of maintenance requires certain efforts in terms of condition monitoring and diagnostic techniques, and the time span between them depends on the type of preventive maintenance.

Two types of preventive maintenance are presented: time-based maintenance (TBM) and condition-based maintenance (CBM). TBM is a very common practice conducted regularly by utilities when performing testing and monitoring actions and the time span between two practices is constant. The disadvantages of TBM are the amount of work required for testing and inspection as well as the cost of planned outage. The time span between tests plays an important role in minimizing risks if it is short, and contributes to unplanned outage if it is too long.

In contrast, CBM does not rely on the time span between inspections. Time for conducting CBM varies from one transformer to another based on its condition - whether it is a severe and urgent case to perform maintenance or not. CBM is preferred and is widely used by utilities even more so than TBM because it avoids some drawbacks including manpower and cost. To perform this type of maintenance, much attention must be given to monitoring, inspection, and diagnostic techniques for better condition assessment. Experts or maintenance planners then choose the appropriate time and action to conduct maintenance based on transformer condition. CBM can be classified into two types: on-line where sensors are used to take measurements including partial discharge, winding movement, and temperature; and off-line such as tan (δ), recovery voltage, and gas in oil.

2.1.1.3 Reliability Centred Maintenance (RCM)

Reliability centered maintenance (RCM) strategy can be considered a type of preventive maintenance, yet it is a qualitative rather than quantitative approach of maintenance as presented previously. By definition, RCM is a combination of tools and methods that help utilities identify the minimum set of preventive maintenance tasks that are prerequisite to address serious component failures without compromising service reliability [6]. RCM can be considered the most intelligent type of maintenance because it recognizes that each component in the system differs from other components in term of design, location, and loading conditions specifically in the case of power transformer. Therefore, RCM assigns maintenance level and priority based on the overall reliability and degree of risk of the transformer as well as the facility.

Table 2.1: Advantages and disadvantages of maintenance types

Maintenance type	Advantages	Disadvantages
Corrective maintenance	1. Low cost 2. Less staff	1. Increased cost due to unplanned downtime 2. Increased labor cost, especially if overtime is needed 3. Inefficient use of staff resources
Preventive maintenance	1. Cost effective in many capital-intensive processes 2. Increased component life cycle 3. Energy savings 4. Reduces equipment or process failure	1. Catastrophic failures still likely to occur 2. Labor intensive 3. Includes performance of unneeded maintenance
Reliability Centred Maintenance (RCM)	1. Most efficient maintenance program 2. Lowers costs by eliminating unnecessary maintenance 3. Minimizes frequency of overhauls 4. Increases component reliability	1. Can have significant startup cost, training, and equipment 2. Savings potential not readily seen by management

In conclusion, maintenance strategy is an essential procedure conducted by utilities to improve the overall efficiency of the facility and minimize life cycle cost of the component. Moreover, reliability and minimization of risks can be improved if the three types of maintenances are applied properly. Table 2.1 lists the advantages and disadvantages of each type of maintenance [7].

2.1.2 Transformer Failure Modes and Reliability Assessment

The investment in electricity sectors due to increased demand is continuing to grow. As a result, any unplanned shutdown in electricity may cause a catastrophic loss for utilities and customers. The study of reliability and failure modes of power system components helps in better understanding certain problems so as to avoid any negative consequences following these failures. Power transformer, as an essential part of power grid, has two modes of failure: firstly, constant failure rate that cannot be improved by maintenance such as design problems, human mistakes, or weather conditions, and that are constant throughout transformer lifetime; secondly, failure modes that change over time while transformer is in service. Such failure modes can be cured by maintenance such as insulation degradation, high temperature, partial discharge, bushings, load tap changers, bad contacts, and others. Condition monitoring techniques and maintenance procedures are the main keys in the early detection of incipient faults or any other problems and as a result prevent any risks or failures which may cause transformer to break down. Table 2.2 and Figure 2.4 list the major failures and the monitoring detection techniques as well as the failure statistics of the defective components [8, 9].

Table 2.2: Failure modes and condition monitoring techniques

Failure Type	**Condition Monitoring Technique**
Cellulose insulation degradation	Degree of Polymerization, Fluid analysis
Oil decomposition	DGA analysis, Fluid analysis
LTC failure	DGA analysis, Internal inspection
Partial Discharge	DGA, (acoustic and electric signal testing)
Bushing failure	Power factor test, Visual inspection
Short turns	Resistance test, Winding ratio test
Heat exchange, devices failure	Thermography, Function test, Vibration test

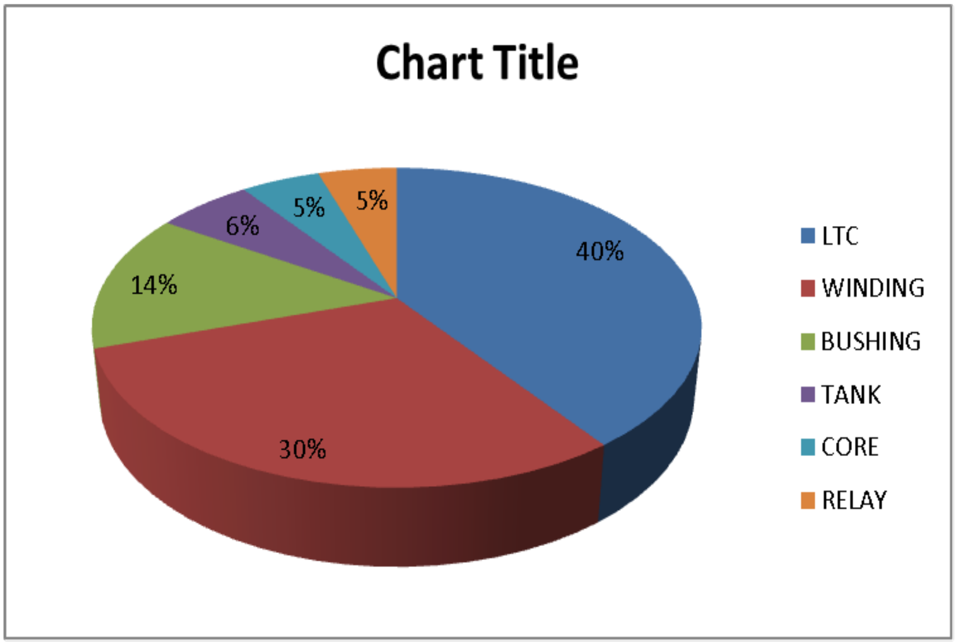

Figure 2.4 Failure statistics of the defective components

2.1.3 Monitoring Techniques

Condition monitoring and assessment of power transformer, which is the first step in an effective asset management, are gaining significant attention from utilities. Monitoring techniques provide useful information to the operation and planning for teams to keep tracking their assets. Monitoring power transformer is achievable through many techniques, either on-line or off-line depending on the type of the used technique. Such techniques utilized for monitoring power transformer can be classified into electrical, chemical, and mechanical practices. These practices help utilities monitor their assets and observe any abnormal changes in the obtained measured data. In addition, the benefit of monitoring power transformer creates a historical record of measured data which is used in defining the date and type of maintenance as well as assessing the condition of a transformer. Understanding the current condition of a transformer helps stakeholders to estimate the investment and set up a plan. Moreover, performing such a strategy is useful specifically in the case of aged transformers where their condition can be assessed precisely, hence avoiding any future possible shortcomings.

In the literature, many papers have been published in the field of power transformer condition assessment. Over the past twenty years, numerous methodologies and tools have been presented in this area in order to discover the proper way to evaluate the state of a component based on special techniques or historical data collected from this component. The complexity in construction, design, operation, and environment factors of power transformer have made the accurate assessment process be very difficult. However, the lifespan and health of power transformer during its service mainly depends on its insulation system and specifically the insulating paper; therefore, considerable attention has been given to this aspect in terms of diagnosis and assessment. Traditionally, chemical and electrical techniques are widely used in insulating paper condition assessment, yet a review of all these techniques in one paper is very rare. In [10, 11] a review has been created of chemical diagnostic methods and their interpretations schemes. [12] Polarization measurement techniques, including furan analysis by high performance liquid chromatography (HPLC) for analyzing cellulose ageing have been implemented. Also, the DP measurement for cellulose mechanical strength and its relation to cellulose ageing phenomena was also investigated in this paper. [13] has demonstrated the use of spectroscopy and multivariate statistical analysis to accurately measure the water and oil content in insulating paper. In [14], new on-line and off-line monitoring techniques in insulation system have been presented. Other papers have examined the thermal effect on insulating paper as in [15, 16]. However, condition of power transformer does not only depend on its insulating system whereas other parts such as windings, tap changers, cooling system, tank, and bushing play an important role in transformer failure. Assessing the condition of each from the previous mentioned accessories can lead to a full and accurate assessment. However, a variety of monitoring and diagnostic techniques have been used to assess the state of each component. In [4] a comprehensive survey was made on electrical, chemical, and physical tests. Their interpretations have been presented and a methodology has been developed to use data acquisition derived from condition monitoring and standard diagnosis for rehabilitation purposes of transformers. [17] has proposed three types of tests. The first type is an on-line diagnostic technique which includes physical, thermal, and insulating oil assessment. The second type is an off-line diagnostic technique represented by electrical tests. The third type is a more advanced diagnostic technique, and is essentially a frequency response analysis and partial test. An approach called Health Index (HI) has been proposed in [1, 2] where the majority of thermal, mechanical, electrical, and chemical diagnostic tests are presented. HI approach can be simply defined in [2] as a practical and reliable tool that combines the results of operating observations, field inspections, and site and laboratory testing into an objective and quantitative index, thereby providing the overall health of the asset. 24 techniques are used to assess a transformer condition via HI in this paper whereas in [1] three more tests were added to the 24 used in [2] which include lose factor, conductivity factor, and Polarization index to increase the reliability of the assessment, and then special software was used to predict HI.

2.1.4 Application of Artificial Intelligence Techniques into Condition Assessment

Artificial intelligence (AI) techniques have been used widely in many different fields including medicine diagnosis, weather forecasting, and control. In terms of transformer condition assessment, AI tools have recently been broadly used to translate monitoring and maintenance data into a health index or a condition reflecting the efficiency of a component. In the literature many diagnostic tools including neural networks, support vector machines, expert systems, fuzzy logic, and various softwares have been utilized to assess transformer condition based on type and size of the used data. In [18] a proposed methodology to assess insulating oil condition of power transformer is conducted which utilizes ANN for the application of extracting chromatographic information using physicochemical analysis. An expert system is proposed in [19] to diagnose the fault in power transformer based on fuzzy logic and artificial neural network using data from dissolved gas analysis test. A comprehensive assessment model based on hierarchical fuzzy theory is proposed to evaluate transformer condition in [20]. The transformer accessories in this model are assessed individually by related monitoring parameters and then the overall condition is predicted by combining individual conditions. An integrated fuzzy logic and evidential decision making model is presented in [21] to assess power transformer condition. DGA, oil, and electrical tests are established to facilitate the assessment model. A statistical learning technique has proposed supporting the use of vector machine algorithm for assessing field transformer condition based on DGA data [22]. The algorithm used in this paper has demonstrated its ability to quantify the insulation condition based on polarization and depolarization current (PDC) measurement. In order to obtain better classification results, an optimized SVM parameters via Particle Swarm optimization PSO algorithm has been proposed in [23]. Again, DGA data is used where transformer condition is divided into five conditions as excellent, good, normal, attention, and fault.

The drawbacks of the used techniques are the number or the diversity measured data. Most of the mentioned papers rely on DGA, oil test, or temperature measurements for assessing transformer condition. However, for full and reliable assessment, more monitoring information including power factor, cooling systems, LTC, and winding should be used as an input data to some machine learning tools to assure the achievement of confident assessment.

2.2 Discussion

The reliable assessment of power transformer greatly depends on the amount of available information. While in service, a transformer is exposed to many abnormal operating conditions. As described earlier, applying different types of maintenance prevents transformers from early deterioration and keeps transformer in good working condition until reaching end of life. Furthermore, studying the failure rate and risks associated with the component helps in preventing expected faults or sudden breakdown by taking correct preventative action. However, one of the most significant actions that can be taken to keep transformer in good health is to monitor and diagnose each associated part. Hence, collecting data is the first step in understanding the current state of a transformer so that a decision or judgment can be taken based on the interpretation of the gathered data. Nevertheless, in the field of monitoring techniques many challenges are faced by utilities, and in some cases it is impossible to perform all the tests. Some of these challenges can be summarized by expense, lack of skilled workers, types of tests, time demands, and value of test. Therefore, there is clearly a need for a model to evaluate the condition based on historical operating information.

Chapter 3
Artificial Intelligence Techniques

3.1 Introduction

Monitoring methods of power system components helps utilities in identifying the condition of their assets through measured data. However, this measured data is a kind of raw data that needs to be prepared and processed before it can be used. Therefore, statistical analysis methods are used to preprocess data by removing information that is not useful or by filling in missing points. In the case of power transformers, many monitoring methods result in sets of information that need to be translated into values that represent their condition. Hence, artificial intelligence technique, which is defined as the capability of a machine to imitate intelligent human behaviour, has been used in this thesis and is represented here by support vector machine (SVM) to translate raw data into a condition or health index. Section 3.2 summarizes the basic structure and algorithms used in SVM as well as the training process for better performance, whereas section 3.3 briefly introduces radial basis function machine (RBF) for the purpose of comparison with SVM. Finally, future selection from data mining technique has been utilized to reduce the number of monitoring tests by selecting the most inductive attributes which significantly affect the health index.

3.2 Support Vector Machine

SVMs have been developed recently [24, 25]. Originally, SVMs worked out for linear two-class classification with margin, meaning the minimal distance from the separating hyperplane to the closest data points. SVM learning machine seeks an optimal separating hyperplane where the margin is maximal. An important and unique feature of this approach is that the solution is based only on those data points which are at the margin. These points are called support vectors. The linear SVM can be extended to nonlinear when the problem is first transformed into a feature space using a set of nonlinear basis functions. In the feature space - which can be very high dimensional - the data points can be separated linearly. A key advantage of the SVM is that it is not necessary to implement this transformation or to determine the separating hyperplane in the possibly very-high dimensional feature space; instead, a kernel representation can be used, where the solution is written as a weighted sum of the values of certain kernel function evaluated at the support vectors [26].

Consider the training data (x_i, y_i) $i = 1,2 \ldots \ldots \ldots .., n,$ $y_i \in \{+1,1\}$ as shown in Figure 3.1.

Where H is the classification line of two types of sample, H1, H2 respectively, are all in the smallest distance from the classification line, and parallel to the H line. For optimal classification, the distance between H1 and H2 must be maximized [23].

In order to obtain the largest distance, that is $\frac{\|w\|^2}{2}$ is the smallest. For all x_1, x_2 there is:

$y_i(w.x_i + b) - 1 \geq 0$ for i=1,2 ……….., n …………..(1.1)

Solving the optimal hyperplane problem is ultimately a quadratic programming problem:

Min $\emptyset(w, b) = \frac{1}{2}\|w\|^2$

S.t $y_i(w.x_i + b) - 1 \geq 0,$ $i = 1,2 \ldots \ldots \ldots .., n$ ……………. (1.2)

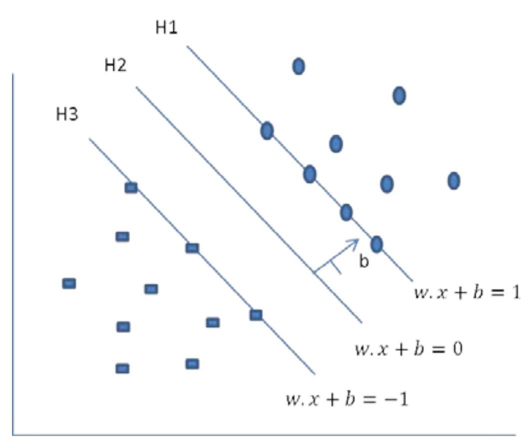

Figure 3.1 Hyperplane line

Some samples cannot be separated linearly; therefore, non-negative slack variables are introduced ξ_i. The constraint of hyperplane is changed as:

$y_i(w.x_i + b) - 1 + \xi_i \geq 0,$ $i = 1,2 \ldots \ldots \ldots .., n$ …………………………..(2.3)

If $0 < \xi_i < 1$, the samples were correctly classified, otherwise the samples were misclassified. Generalized optimal separating surface becomes:

$$Min\ \emptyset(w, b) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \quad \ldots\ldots(2.4)$$

The first step is to make the classification of an interval as big as possible; the second step is to make errors as small as possible. Among it, C is an adjustable parameter [2].

The optimization problem in (2.4) can be solved easily in its dual formulation. By introducing Lagrange multipliers, the original constrained problem is converted to an unconstrained problem.

$$L(w, b, \xi, a) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\eta_i\,\xi_i - \sum_{i=1}^{n}a_i[y_i(w.x_i + b) - 1 + \xi_i] \quad \ldots\ldots(2.5)$$

a_i are the Lagrange multipliers for all training sample data. In solving (2.5), the optimal classification function can be obtained as:

$$Min\ F(a) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}a_i a_j y_i y_j (x_i, x_j) - \sum_{i=0}^{n}a_i$$

$$S.t \quad \sum_{i=0}^{n}a_i\ y_i = 0 \qquad i = 1,2, \ldots \ldots \ldots.., n$$

$$0 \le a_i \le C, \qquad i = 1,2, \ldots \ldots.., n \quad \ldots\ldots (2.6)$$

The focus of the problem is a linearly constrained convex quadratic programming. The solution is unique. Most of them have the solution of zero. The non-zero solution for support vector is recorded as a^*.

$$f(x) = sgn\{\sum_{i=1}^{m}a_i^*\ y_i(x_i, x) + b^*\} \quad \ldots\ldots (2.7)$$

m is the number of the support vectors; sgn(x) is the sign function [26].

It is evident that formula (2.5) only contains the inner product operation of unspecified sample and training samples support vector. To solve linear classification problems in feature space, the inner product operation in space must be known. Non-linear problems can be transformed into a high-dimensional linear problem. Kernel functions are the keys to be used to achieve this transformation.

3.2.1 Kernel

The Kernel Trick can be applied to any algorithm that is solely dependent on dot products between two vectors. It has been demonstrated that Kernel can be implemented to develop nonlinear generalization of any algorithm that can be cast in terms of dot products. The basic theoretical concept for building nonlinear support vector machine is to first map the data points in the original space or the input space to the feature space (higher dimensional space), then a linear classification function is determined in the new space. The transformation of the data is performed using the mapping $x \rightarrow \Phi(x)$, where x is a vector of independent variables. The Kernel dot product between two points i and j in the feature space takes the form of $\Phi(x_i).\Phi(x_j) = k(x_i x_j)$.

Table 3.1 shows some Kernels that can be used in the case of SVM to perform dot product:

Table 3.1. Types of Kernel functions

Function	**Form**
Linear	$k(x, x') = \langle x.x' \rangle + c$
Polynomial	$k(x, x') = (\langle x.x' \rangle + 1)^d$
Gaussian Radial Basis Function	$k(x, x') = \exp(-\frac{\|x - x'\|^2}{2\sigma^2})$
Exponential Radial Basis Function	$k(x, x') = \exp(-\frac{\|x - x'\|}{2\sigma^2})$
Multi-Layer Perceptron	$k(x, x') = \tanh(\sigma\langle x.x' \rangle + \gamma)$

3.3 Radial Basis Function

Radial basis function neural networks (RBFNN) are well suited for solving function approximation and pattern classification problems due to their simple topological structure and ability to reveal how learning proceeds in an explicit manner [24]. An RBFN is primarily composed of three layers, including input layer, hidden layer, and output layer. The input layer is composed of input data, while the hidden layer transforms the data from the input space to the hidden space using a non-linear function. The output layer, which is linear, yields the response of the network. The argument of the activation function of each hidden unit in an RBFN computes the Euclidean distance x between the input vector and the center of that unit. The network has a better accuracy only when the early stop technique is appropriately used.

The values of RBF depend only on the distance from the origin, so $\varphi(x) = (\|x\|)$ or r on the distance from other point c, where c is called center, hence $\varphi(x, c) = (\|x - c\|)$. Therefore, any function φ satisfies $\varphi(x) = (\|x\|)$ is a nonlinear function which has many types, as shown in Table 3.2, and ⟦ ⟧ is the Euclidean Norm [25].

The popular Gaussian function is the most common type of RBF used in the hidden layers. There are three parameters used in radial function center c_i, shape φ, and distance measure $r = \|x - c_i\|$.

Table 3.2 Types of radial basis functions

Function	**Form**
Gaussian	$\varphi(r) = \exp(-\beta r^2) \quad \beta > 0$
Multiquadric	$\varphi(r) = (\beta^2 + r^2)^{\frac{1}{2}} \quad \beta > 0$
Polyharmonic Spline	$\varphi(r) = r^k \ln(r), k = 2,4,6$
Thin Plate Spline	$\varphi(r) = r^2 \ln(r)$

3.3.1 Radial Basis Function Network

The basic RBFN has an input layer with input nodes, a hidden layer of RBF nodes, and an output layer with linear nodes. Figure 3.2 shows the architecture of radial basis function neural network.

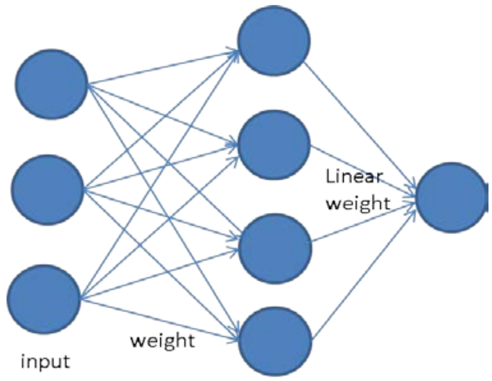

Figure 3.2 RBFN structure

RBFN classification function is introduced as:

$y_j(x) = \sum_{i=1}^{n} w_{ij}\varphi(\|x - c_i\|)$ ………………………..(2.8)

y_j represents the activity of the output node j, $\varphi(\|x - c_i\|)$ is the activity of the middle node i with a radial basis function centre on the vector x_i, x is the actual input vector, and w_{ij} are the weights between the hidden layers and output layers [25]. As previously stated, the most popular type of radial basis function network used in classification is the Gaussian function.

$\varphi(\|x - c_i\|) = exp\left(-\frac{\|x-c_i\|^2}{2\sigma^2}\right)$ $\sigma > 0$ …………………. (2.9)

Where σ is the Gaussian's width. The parameters of the radial basis function neural network that must be determined include weights, spreads, and centers [25].

3.3.2 Learning Strategy

In the neural network, it is usually assumed that the number of basis functions is less than input data. Therefore, determining the number of neurons is important for the construction of an RBFN. In addition, the placement of neurons and the calculation of the adjusted parameters within the neural system are also essential. Two approaches have been presented for the learning process, including hybrid learning algorithm and supervised learning algorithm, as proposed in [27].

3.4 Feature Selection

Variable selection, attribute selection, dimensionality reduction, or what is called feature selection is the process of selecting a small number of highly predictive features out of a large set of candidate attributes that might be strongly irrelevant or redundant. This selection plays a fundamental role in pattern recognition, data mining, and more general machine learning tasks [28]. In other words, the least number of attributes that most contribute to accuracy will be selected by eliminating the unrelated or unimportant ones.

Feature selection plays an important role in preprocessing data, specifically when tens or hundreds of thousands of variables are available. There are many advantages for applying such techniques, which are summarized as follows:

- Reducing training and execution time
- Defining the course of dimensionality to improve prediction performance
- Facilitating data visualization and data understanding
- Reducing measurement and storage requirements

There are two approaches for applying feature selection in order to avoid the course of dimensionality.

1. **Forward selection**: In this method the algorithm first begins the selection with no variables, and then starts adding them one by one. The error is calculated at each step. In this process the algorithm selects the most effective attributes that significantly reduce the error until no observable reduction occurs in the error.
2. **Backward selection**: In contrast with the first technique, the algorithm begins with all the variables and eliminates them one by one. At each step the algorithm eliminates the feature which reduces the error.

Feature selection technique has been implemented using WEKA software, which has many types of selection algorithms and either supervised or unsupervised techniques. The algorithm used in this research is called **InfoGainAttributeEval** with Ranker search method. This algorithm evaluates attributes based on information gain, followed by ranking based on their evaluation.

The classifier InfoGainAttribEval is a tree classifier using InfoGain in every parent (level) to decide about which attribute to be used. A decision tree can be constructed top-down using the information gain in the following way [29]:

- begin at the root node
- determine the attribute with the highest information gain which is not already used as an ancestor node
- add a child node for each possible value of that attribute
- attach all examples to the child node where the attribute values of the examples are identical to the attribute value attached to the node
- if all examples attached to the child node can be classified uniquely add that classification to that node and mark it as leaf node
- go back to step two if there are unused attributes left, otherwise add the classification of most of the examples attached to the child node

3.5 Discussion

SVM has demonstrated its leading robustness in the field of classification and regression problems, compared to other machine learning techniques such as artificial neural network and radial basis function. SVM is superior in always finding the global rather than local solution. However, selecting the correct value of each parameter is a considerable challenge for any programmer because the classification process depends on these parameters to be successful. Some methodologies have been proposed to optimize these values to assure better performance, such as using a hybrid system integrated SVM with genetic algorithm (GA).

Chapter 4
Monitoring Techniques

4.1 Preamble

Life cycle management and condition monitoring techniques of power transformer have recently been gaining considerable attention because of their direct impact on minimizing unplanned outage risks, cost, and extending the nominal end of life. The gradual deterioration in a transformer occurs for many reasons, specifically for in-service aged transformers. Overloading, lack of maintenance, design problems, environment temperature, and other factors speed up the deterioration process; therefore, condition monitoring procedures aim to track component behavior and detect incipient faults early. This monitoring is directly reflected on maintenance costs and proper health assessment. In recent years many various monitoring techniques have been developed to help utilities make the right decisions when performing maintenance, repairs, or replacement scenarios. Due to the complex construction of a power transformer, many monitoring tests have been developed to assess each specific part. For instance, dissolved gas analysis, oil quality, degree of polymerization, and furan content have been widely used to assess insulation systems including insulating paper and oil. Other electrical tests have been used for assessing winding conditions including AC and DC insulation tests, bushing, excitation current, winding resistance, turns ration, and load tests. Finally, thermal tests such as infrared thermography and hot spot temperature measurement have been applied to complete the monitoring process. All previous monitoring techniques are applied periodically on oil power transformers for full assessment. The results of these procedures are essential to evaluate the most cost-effective option when making operation, maintenance, repair, or replacement decisions.

4.2 Dissolved Gas Analysis (DGA)

Most in-service oil-immersed power transformers are exposed to abnormal operating conditions which cause minor or major deterioration in the insulating system, depending on whether the fault type is thermal or electrical. Electrical faults occur because of partial discharge, arcing, and low energy sparking; whereas thermal faults occur due to overloading, cooling problems, and overheating in the insulating material. Under abnormal operating conditions, small amounts of gases will be liberated from insulating oils or paper that are generally compostable when exposed to thermal or electrical stress. DGA test has been used widely to extract, detect, and quantify dissolved gases and their concentration in insulating fluid; therefore, incipient faults inside a transformer can be detected with their causative. The most important gases generated from transformer insulating material can be listed as follows: methane (CH4), ethane (C2H6), ethylene (CH4), acetylene (C2H2), and hydrogen (H2). These gases are classified as hydrocarbon gases, whereas carbon oxides are represented by carbon monoxide (CO) and carbon dioxide (CO2). In addition, two gases that may be generated but are considered non-fault gases are nitrogen (N2) and oxygen (O2) [12]. The fault type can be detected based on the concentration and rate of each gas over the acceptable limit provided by national standards such as Dorenenburg, Bureau of Reclamation, and IEEE, as shown in Table 4.1. In addition, a criteria for classifying transformer state to four conditions based on TDCG has been proposed in IEEE standard C57.104-1991 [30]. Over the past few years, many diagnostic techniques have been proposed to distinguish the fault type based on the level and concentration of the key gases.

Table 4.1. Recommended limits of dissolved gases

Gas	**Dorenenburg/Stritt**	**IEEE**	**Bureau of Reclamation**
Hydrogen	200	100	500
Methane	50	120	125
Ethane	35	65	75
Ethylene	80	50	175
Acetylene	5	35	7
Carbon Monoxide	500	350	750
TDCG		720	
Carbon Dioxide	6000	2500	10000

4.2.1 Key Gases Method

The interpretations of incipient faults in oil-immersed power transformer using key gas method helps in detecting the fault type based on rates of gases. Four general fault types and their relative proportions are listed as follows [30]:

A. **Thermal-Oil:** Decomposition samples include ethylene and methane with small amount of hydrogen and ethane.

B. **Thermal Cellulose:** High rate of carbon dioxide and carbon monoxide are developed from overheated cellulose.

C. **Electrical Corona**: Hydrogen and methane with small amount of ethane and ethylene may be presented due to low energy electrical discharge.

D. **Electrical Arcing**: High rate of hydrogen and acetylene are produced with small quantities of methane and ethylene. If the faults involve cellulose, carbon monoxide and carbon dioxide may be produced.

4.2.2 Dorenenburg Ratio Method

Dorenenburg methodology uses four gas ratios to indicate a fault type from the existence of three fault types. The four key gas ratios are R1=CH4/H2; R2=C2H2/C2H4; R3=C2H2/CH4; and R4=C2H6/C2H2. The first step is to compare these values with special concentration to identify possible problems or the existence of any faults. Next, limiting values that provide a suggested fault diagnosis are compared with these ratios in order to diagnose any existing faults [30].

4.2.3 Rogers Ratio Method

Rogers ratio method is similar to Dorenenburg method as it uses the same ratios procedure. This technique uses three ratios instead of four, which are R1=CH4H2; R2=C2H2/C2H4; and R3=C2H4/C2H6. The fault can be clarified by applying key gas method when the concentration of the gases exceeds the limits provided in Table 4 [30].

In addition to the previous mentioned techniques, other methods used to clarify fault type such as Duval's triangular [31], total combustible gas (TCG), and total dissolve combustible gas (TDCG) as stated in IEEE PC57.104 D11d standards [30].

4.3 Oil Quality

Insulating oil in the power transformer functions similarly to blood in the human body, in that useful information about the transformer's condition can be extracted from oil. In addition to its main function of insulating liquid between conductors and core, insulating oil is also used for cooling due to its excellent electrical insulating properties as well as its high boiling point. However, during its service, the power transformer is exposed to thermal and electrical stress which degrades oil; hence, some of its strength characteristics are lost. Therefore, physical, chemical, and electrical tests are used to verify oil quality and detect any faults occurring during the service. Table 4.2 lists many types of tests suggested by national standards, including ASTM and IEC. The most frequently used tests among this list are Color, interfacial tension, dielectric break down, power factor, acidity, and water content.

Table 4.2 Pertinent ASTM tests for mineral insulating oil

Type	**Test**	**ASTM standard/method**	**IEC**
Physical Tests	Color	ASTM D1500	ISO 2049
	Interfacial tension	ASTMD971	ISO 6295
	Visual examination	ASTM D1524	
Electrical Tests	Dielectric breakdown voltage	ASTM D877	IEC60156
	Dissipation factor	ASTM D924	IEC247
	Dielec breakdown impulse vol	ASTM D3300	
Chemical Tests	Gas content	ASTM D2945	
	Neutralization number	ASTM D664	IEC62021
	Water content	ASTM D1533	IEC 60814
	Furans in insulating liquids	ASTM D5837	

Reconditioning oil or replacing oil with new oil will eliminate most of the information that is used to assess the transformer's condition. Therefore, such tests are applied from time to time to create a database system that helps utilities understand the deterioration behavior of their transformers and make the right decision. Table 4.3 shows suggested test limits by voltage class for good oils to remain in continued service and more deteriorated values, as stated in [32].

Table 4.3 Suggested limits for continued use of in service oil

Test and method	**Voltage Class**		
	≤69 kV	>69 – <230 kV	230 kV and above
Dielectric strength Strength kV(1,2 mm gap)	23 40	28 47	30 50
Power factor 25 º C, 100 º C	0.5 5.0	0.5 5.0	0.5 5.0
Water content mg/kg maximum (ppm)	35	25	20
Interfacial tension mN/m minimum	25	30	32
Acidity mg KOH/g maximum	0.2	0.15	0.1

4.4 Furan Test

There are two types of insulating systems in power transformer: insulating oil and cellulosic paper (solid insulation). Insulating oil and cellulosic paper are degraded by the ageing process or by abnormal operating conditions such as overloading followed by high temperature. In the solid insulation, furanic compounds are generated from cellulosic paper and the generation process is accelerated with the presence of oxygen and moisture under higher operating temperatures. However, the concentration level of these compounds is used as a good indication for ageing or incipient fault conditions. Many papers in the literature have examined the use of furanic compounds to understand the ageing phenomena of cellulosic paper. For example, one study used high performance liquid chromatography (HPLC) technique for ageing analyzing [12]. Acceptable limits suggested by ASTM from IEEE are shown in[12].

4.5 Power Factor

Power factor (known as dissipation factor or tan delta) test is a common and practical test used to monitor and assess transformer and bushing conditions. Power factor is the ratio of resistive current component to the leakage current when the voltage is applied. This test is used widely by utilities as a routine test to evaluate capacitive insulation condition between windings and compartments [2]. According to [33], power factor less than 1% is considered to be good; power factor between 1% and 2% must be investigated; and if the power factor is over 2% an action is required that takes into consideration the historical change in power factor for full evaluation. This test is performed as follows: high-voltage to ground, high- to low-voltage winding, low-voltage winding to ground, high- to tertiary-voltage winding, low- to tertiary-voltage winding, and tertiary-voltage winding to ground insulation [33].

4.6 Infra-red Emission Testing

During its service, the power transformer is exposed to some abnormal operating conditions such as overloading, cooling system blockage, connection problems, and high environmental temperature (ambient temperature). Notably, when the transformer operates under higher temperatures, especially hot spot temperature oil decomposition, cellulosic paper deterioration and power losses are the outcomes. Therefore, on-line monitoring using thermal imaging is a very common and important practice to prevent unwanted operating conditions and protect insulating material, which mainly affects ageing factor and early breakdown [34]. The hot spot temperature has been recommended by IEC not to exceed 98 º C, as 6 º C deviation on the hottest temperature may double the ageing rate [35]. Infrared thermography testing is used to detect all the hot areas with any temperatures that are higher than the transformer's external surface. The hot areas emit infrared radiation which can be measured on the infrared spectra band of the electromagnetic spectrum [36]. Four colors are displayed in the thermal imaging: white, red, blue, and black. Hottest areas are displayed in white and red colors and colder areas are displayed in black and blue colors. Table 4.4 classifies temperature levels and their severity where temperature excess is the difference between normal temperature and the highest temperature [33].

Table 4.4 Classification of overheating severity

Classes	**Temperature Excess**
Attention	0-9 º C
Intermediate	10-20 º C
Serious	21-49 º C
Critical	>50 º C

4.7 Winding Resistance Test

In this test, the pure ohmic resistance of the transformer windings and tap changer are tested. This test is conducted on both low and high voltage windings to compare their deviations from manufacturer's name plate as well as the existence of the tap changer in each position. The deterioration in the ohmic resistance occurs for many reasons, including poor design, poor environment, overloading, and lack of maintenance. Therefore, this test is important for the purpose of calculating conductor losses ($I^2R\ component$) and for calculating winding temperature at the end of the temperature test cycle. A capable ohmmeter is required ranging from 10 ohm down to some fractions. The test's basic principle is to inject a stable DC current through the winding and then measure the voltage drop on this winding. During the test, the temperature should be recorded for future comparisons because winding resistance varies with oil temperature. Any deviation from the nameplate value greater than 5% indicates serious damage in the conductor [33]. A ranking method that classifies winding conditions into four classes based on the severity of the deviation is applied in [2].

4.8 Turns Ratio Test

It is known that the main function of the transformer is step-up or step-down voltage from one side to another depending on the number of turns in each coil. However, these coils are subject to mechanical and electrical stress that may cause shorts between turns of the same coil or open winding circuits. Therefore, turns ratio test is used to detect these problems which may indicate insulation breakdown between turns. As noted on transformer nameplate, a good ratio must be within 5% of the ratio of the rated voltage between windings. For new transformers, the ratio is expected to be 1% compared to the nameplate, and a ratio greater than 2% indicates severe deviation requiring immediate action. The direct relationship between the ratio of primary-to-secondary voltage must meet the ratio of number of turns in the primary-to-secondary windings as indicated by ($V1 / V2 = N1 / N2$) where V1=primary voltage, V2=secondary voltage, N1=number of primary turns, and N2=number of secondary turns.

4.9 Load Test

During its service, power transformers are exposed to different loading conditions. Power transformers are designed to provide satisfactory service for a specific period of time or beyond based on the rated nameplate output (KVA). However, in many cases, the demand exceeds the rated value for short or long periods of time which results in continuous hot spot conductor temperatures. Hence, the deterioration rate of insulating material (oil and cellulosic paper) will be increased. According to [37], transformer operation follows one of the following five loading conditions:

A. Normal Life Expectancy Loading: Continuous Load

This is a constant load subjected to the rated nameplate output KVA while the ambient temperature does not exceed 30 º C. In addition to the ambient temperature (30 º C), hot spot rise (15 º C), and average winding rise (65 º C), a continuous hot spot temperature of 110 º C is the result of this condition.

B. Normal Life Expectancy Loading: Cyclical Load

For this load, the hottest spot conductor temperature ranges higher and lower than the normal temperature of 110 º C at a cyclical load and normal constant ambient 30 º C where the load cycles up and down transformer rating.

C. Long Time Emergency Loading

This rare type of loading results from the long outage of some components in the system that cause transformer loading. Hence, the hottest spot conductor temperature may reach 140 C. Long time emergency loading has a severe impact on transformer life by increasing the aging process of the insulating system, and may have associated risks.

D. Short Time Emergency Loading

This rare condition of loading may occur two to three times during the transformer's life. It happens because of one or two events and may cause hottest spot conductor temperature to reach to 180 º C. The insulation life is accelerated significantly during this condition.

E. Planned Loading Beyond Nameplate: Normal Operation

In this loading, top oil temperature or hottest spot conductor temperature exceeds the suggested limits for normal life loading. This can be considered a normal loading accepted by users since it is more likely to occur without component outage or emergencies. The hottest spot conductor temperatures vary from 120 º to 130 º C.

Calculating the loss of insulation life in all cases is recommended since it is important for utilities loading policy. The recommended maximum temperatures provided by ANSI/IEEE C57.91-1995 and IEC St 354 are given in [37]. The monthly or yearly load peaks can be used to estimate load history condition based on the recommended values given by ASTM/IEEE and IEC.

4.10 Tap Changer Condition

Tap changer plays the primary role in changing the voltage ratio of power transformer to the secondary voltage within the acceptable limit. On-load tap changer (OLTC) selects the winding taps along the transformer winding for the appropriate ratio with the continuity of the load current. However, as with other transformer parts, OLTC is subject to many abnormal operating conditions which cause it to be the top reason for transformer failure. Major failure occurs due to insulation degradation and contact failure inside the devices. In addition to thermal and electrical stress, bad contacts cause hotspots in the tank which speed up the degradation rate of the insulating oil. There are many techniques used to assess OLTC condition including DGA, oil test, dynamic resistance measurements (DRM), number of operations, temperature, relay timing, and maintenance. However, applying the same interpretation of the previous test to all OLTC types is not correct because each type has its own sensitivity to degradation mechanisms. DGA and oil quality tests have been used as the most dependable techniques for detecting thermal and electrical faults in the OLTC. However, the interpretations of gas ratios and DGA limits differ from one standard to another. For example, IEC has different limits than IEEE. The general differences between these as well as normal gas concentration introduced by Weidman-ACTI and WECC gas limits are given in [2].

4.11 Bushing Condition

Bushings are one of the most important accessories of transformers. Statistical studies have shown that failure in bushing causes long term outage of transformers; therefore, a catastrophic economic loss may occur because of this condition. Some studies show that approximately 45% of transformer failures occur because of bushings, and 90% of failures are caused by moisture entering the bushing through leaky gaskets [38]. Transformer bushings come in many types, including air-insulated bushing, oil-filled bushing, oil-paper insulated bushing, and gas-insulated bushing. However, oil-paper is the most commonly used transformer bushing because of its characteristics in terms of good electrical and heat transfer properties. Like transformers, bushing exposed to electrical, mechanical, environmental, and thermal stress increases the deterioration rate of the bushing insulation. There are many monitoring procedures to assess bushing conditions such as power factor, capacitance, partial discharge, DGA, moisture, oil sampling, and thermovision. The normal and emergency limits for power factor, water content, DGA results, breakdown voltage, and temperature are provided in [39].

4.12 Cooling Equipment

One serious problem faced by transformers is the rise of temperature over the acceptable limits. Heat in the transformer is generated from the current flowing in the windings (copper losses), hysteresis, and eddy current losses. Proper heat dissipation is a very crucial action because the continuity of heat over the acceptable limit contributes largely to the degradation of insulating system. Therefore, the transformer’s cooling system dissipates heat in winding and insulating oil which saves the transformer from early breakdown. External cooling systems used in power transformers in order to accelerate the dissipation rate of heat include Oil Natural Air Natural (ONAN), Oil Natural Air Forced (ONAF), Oil Forced Air Forced (OFAF), Oil Forced Water Forced (OFWF), Oil Directed Air

Forced (ODAF), and Oil Directed Water Forced (ODWF). Cooling assessment of power transformer must be performed on all the cooling parts, including radiator, cooler gaskets, fans, and pumps.

4.13 Core-To-Ground Resistance Test

The transformer is intentionally grounded via one connection. Core-to-ground test or Megger test is used to detect any loss in the connection. Moreover, this test is also used to detect the presence of a spurious or undesired ground. If the resultant resistant is very low, then the intentional core ground is intact. In order to check for undesired core ground connection, the intentional ground has to be removed and megger between core the grounded transformer tank. If the resultant resistance is very high, the unintentional or undesired ground does not exist. Hot metal gases such as methane, ethylene, and ethane can be generated from unintentional core ground; therefore, megger test can be used as a supplement to DGA test [4]. Table 4.5 Core-to-Ground test classifies transformer conditions based on Core-to-Ground resistance test.

Table 4.5 Core-to-Ground test

Limits (MΩ)	**Diagnostics**
R> 1000 MΩ	new transformer
R>100 MΩ	service-aged transformer
R 10-100 MΩ	deterioration between core and ground
R<10 MΩ	serious deterioration

4.14 Frequency Response Analysis Test

During transportation operations, because of high faults, or even because of lightning, a displacement in the winding of a transformer can occur. Changes in resonance frequencies or magnitudes are linked to deviations of inductances or capacitances, which are defined by physical dimensions and materials of a transformer. Thus, changes in resonances are evidence for diagnosis of winding mechanical faults such as displacements and deformations [40]. The SFRA is a non-destructive method and can be considered a mechanical fault detector. Developing this method is based on the relationship between the displacement or deterioration in windings and the change in frequencies between the two windings due to changes in the capacitance and inductance of windings. In essence, the method consists of applying a low voltage sinusoidal signal made of a sweep of frequencies covering a range between 10Hz and 2MHz on one end of the transformer winding, and measuring the other end of the winding [41].

According to [40], different winding faults can be interpreted as illustrated in Table 4.6.

Table 4.6 Sub-band division of frequency responses

Frequency	**Failure Sensitivity**
<2 kHz	core deformation, open circuits, shorted turns, and residual magnetization
2-20 kHz	bulk winding movement between windings and clamping structure
20-400 kHz	deformation within the main or tap windings
400 kHz- about 1MHz	movement of the main and tap windings, ground impedance variations

4.15 Leakage Reactance Test

Sometimes called short circuit test, leakage reactance test is used to detect winding or core deformation caused by faults, lightning strikes, or during transportation operation. When deformation occurs, changes in the magnetic flux will be followed by changes in the measured leakage reactance, which may cause immediate failure to the transformer. Usually this test is performed in the field and compared with transformer nameplate information, previous tests, or transformers from the same design and rating. The factory test conditions, which are indicated on the nameplate, may be different from field test conditions in terms of load current which is not always available in the field [9]. Limits of results obtained from this test and their deviation from nameplate information are shown on Table 4.7.

Table 4.7 Leakage reactance test

Deviation from Nameplate %	**Condition**
$\Delta\%<1$	Good condition
$1< \Delta\% <3$	Minor deterioration
$3< \Delta\% <5$	Significant deterioration
$5< \Delta\%$	Severe deterioration

4.16 Excitation Current Test

Excitation current test is effective in identifying problems such as abnormal core ground, short circuit windings, open circuit windings, LTC problems, and poor electrical connections. However, when one of these problems occurs, a change in the effective reluctance of the magnetic circuit will occur, which also affects the current required to force a given flux through the core. This test is performed on the high voltage side in order to minimize the excitation current by keeping the secondary side open. The applied voltage must be high but not exceed the rating voltage for line-to-neutral on wye connection and line-to-line voltage on delta connection. The results from this test should be compared with previous tests or similar components by taking into consideration the exact applied voltage. The acceptable limits and the percentage of deviation from transformer nameplate or previous tests are discussed and scored in [4].

4.17 Maintenance

As mentioned in the literature review, maintenance in its three classes (corrective, preventive, and reliability centered) is an essential task that must be conducted regularly in addition to monitoring tests. Other diagnostic tests involved in the overall transformer condition can be conducted by internal and external visual inspection during maintenance, including oil leakage, gaskets, tank condition, and connectors. Maintenance history, as a good indicator of transformer condition, summarizes problems and actions that have been conducted on the unit. Ranking power transformer in terms of maintenance actions is proposed in [2].

4.18 Discussion

Condition monitoring of power transformers is a very important step in assessing conditions and analyzing risk of failure. Many monitoring methods have been proposed to prevent transformers from sudden failure as well as create databases consisting of information about each part of the unit in order to track any change in the measured data from the rating nameplate. However, monitoring methods vary in terms of the importance or weight factor. Some methods must be conducted when a transformer is off-line and others can be conducted when a transformer is on-line, depending on whether the type of test is electrical, chemical, or thermal. The challenges utilities may experience determine the time and type of test to be used. Many tests are impossible to perform in some cases, especially for off-line tests that may result in catastrophic economic loss to providers and customers. Another problem is that not all tests provide reliable information for making clear decisions. For example, in the case of insulating paper, it is very difficult to accurately estimate its condition and expected life time. The experts play the main role in deciding the type and time of the tests as well as the interpretation of the obtained data. However, it is clear that the most common and practical tests used by most utilities include DGA, oil quality test, power factor, infra-red, bushing condition, turns ration, and load tap changer. Other tests are used when more information is required, especially in the case of aged transformers.

Chapter 5
Proposed Intelligent Assessment Technique

5.1 Preamble

It is well established that more monitoring techniques leads to more available data, which results in the achievement of accurate health assessment. However, it is not always easy or even economically viable to perform many tests because of limited budget, availability of skilled workers, and time constraints. Collecting data from the previously mentioned techniques (see Chapter 4) is the first step in assessing the current state of a transformer; thus, a decision or judgment can be taken based on the interpretations of the measured data and their variation from the transformer nameplate. In fact, translating these measurements into a value or class helps in discovering the transformer's state without focusing on individual interpretations of each test. In this chapter, an SVM-based diagnostic test is used to assess the state of the transformer's condition. The preparation and preprocessing is first conducted, followed by feature (test) selection technique based on ranker search algorithm to select the most effective tests that best describe the transformer's condition. RBF classifier has been used as a second classifier for results comparison and to prove the validity of SVM model.

5.2 Data Preparation and Preprocessing

Collecting a large set of data from many different tests without preprocessing makes the interpretation process very difficult. However, in most cases the availability of all measured data of a single transformer is very rare and expensive if it is possible. Normally, either utilities perform the most important tests, or specified tests based on experts' experience provide the needed information. However, even for conducted tests there will be some missing data. Usually the missing data is lost during the collection process or when mistakes are made with the data entry. For instance, performing oil quality test includes measurement of breakdown voltage, interfacial tension, color, moisture, and neutralization number, but does not include dissipation factor and furan tests. Moreover, from the collected data, some transformers had undergone too few tests which made the preparation process more difficult. Therefore, statistical methods have been employed to fill in the missing data that is close to the real or measured values. These statistical techniques are mean substitution, regression substitution, and more advanced methods such as maximum likelihood. Also, transformers of the same size and class working under the same operating conditions could be taken as a reference to approximate the missing values.

In this research, for each individual test of the monitoring methods a ranked method is used to classify the raw data from 4 to 0 where 4 is excellent, 0 is fail, and 1, 2, and 3 are values in between. This ranking is made taking into consideration the interpretations and limitations provided by national standards such as ASMT/IEEE and ICE. The methods for ranking these values have been proposed in [2, 4]. For example, the DGA test has been ranked into five conditions based on the percentage of total dissolved compostable gases TDCG as stated in section 4.2, and this continues for the rest of the tests. After preprocessing the data, a diagnostic technique is needed to translate the information obtained from all the tests of one unit to a value which is represented by the Health Index HI. Heuristic techniques, specifically support vector machine for classification SVM, are used as a diagnostic technique to estimate the overall condition or HI of a transformer. SVM has been chosen due to its robust classification results compared with other heuristic tools such as neural network, as shown in section 5.3. SVM, as any machine learning tool, needs to be trained and tested in order to be used as a classifier; therefore, data from 70 power transformers with the condition of each transformer (label) is used in this research.

Training and testing the classifier has been conducted using k-fold cross-validation technique. K-fold cross-validation is a technique for assessing how the results of a classifier can be generalized to an independent sample. In k-fold cross-validation, the actual data is randomly divided equally into four sub-samples. From the four sub-samples, one sub-sample is used for testing the classifier (validation), and the remaining three sub-samples are used for training purposes. The cross-validation process is repeated four times, with each of the four sub-samples used exactly once as a validation data. The four outputs from the folds are then averaged to produce a single estimation. The advantage of this technique is that each sample or observation is used once as a training and testing sample [42]. The input data of the classifier are the features or attributes, and here it is the measured data (tests); whereas the output is the label which represents the condition (HI). After training, testing, and adjusting the parameters of the classifier, it can be applied to other transformers of the same size and class to estimate conditions.

5.3 Condition Assessment Using SVM

Due to its superb characteristics, support vector machine (SVM) has been chosen to assess the transformer condition from the available data. Compared to other classifiers such as ANN, RBF, and decision tree classifiers, SVM has the following significant advantages:

- It has a regularization parameter, which helps in avoiding over-fitting;
- It is defined by a convex optimization problem (no local minima);
- It uses a Kernel trick where expert knowledge can be built about the problem; and
- It works well with a small set of data for training.

In this research, the dimensionality of the measured data that is used in SVM is not large. In general, by assuming that the number of transformers used in the classifier is (M) and the number of tests of each transformer is (N), then a matrix with a dimension of M x N represents the whole dataset used in SVM and RBF for training and testing purposes. In this research, SVM is applied to 70 transformers and is represented by 19 features (M=70 and N=19). Moreover, in order to generalize and improve the classifier, the data is divided into a number of subsets using k-fold. The condition (label) of each unit has been provided with the measured data based on the experts' knowledge of transformers' ages and any problems that might have occurred during their service. Therefore, the transformer conditions are divided into five different categories as follows: A, B, C, D, and E, where A represents excellent condition or new transformer and E represents failed unit, as shown in Table 5.1.

Table 5.1 Transformer Condition Classification

Class	**Percentage%**	**Condition**
A	85-100%	Excellent/New
B	70-85%	Very Good
C	50-70%	Maintain/Repair
D	30-50%	Deteriorated
E	<30%	Failed

Up until now the methodology of ranking transformer conditions for three, four, or five classes is still a matter of discussion; however, such a ranking or classification strategy has been proposed in different publications as in [2, 17]. The percentages of each class for the whole set of the transformer conditions as received are shown Figure 5.1. As is evident from Fig 5.1, class A represents 24% (17 of 70 units) in contrast with class B which is indicated by 47% (33 of 70 units) holding the highest percentage. Class B has very high returns because the majority of the transformers are aged transformers that have been in service for more than 15 years. However, class B transformers are still in very good condition since many of them are very close to class A transformers, which indicates new transformers or transformers that have been in service for less than ten years and have never experienced abnormal operating conditions such as overloading for a long period of time, exposure to lightning, or internal faults. Class C transformers, which represent approximately 21% (15) of the whole dataset, are still in good condition but need more attention in terms of maintenance and repairs for some parts. In contrast, class D transformers are quite deteriorated units which are very rare to find in service. Only four Class D transformers of 70 represent 6% of the dataset. Finally, 1 of 70 units, or approximately 2%, represents a class E transformer that is already out of service and is only included for the purpose of classifying the performance of different transformers.

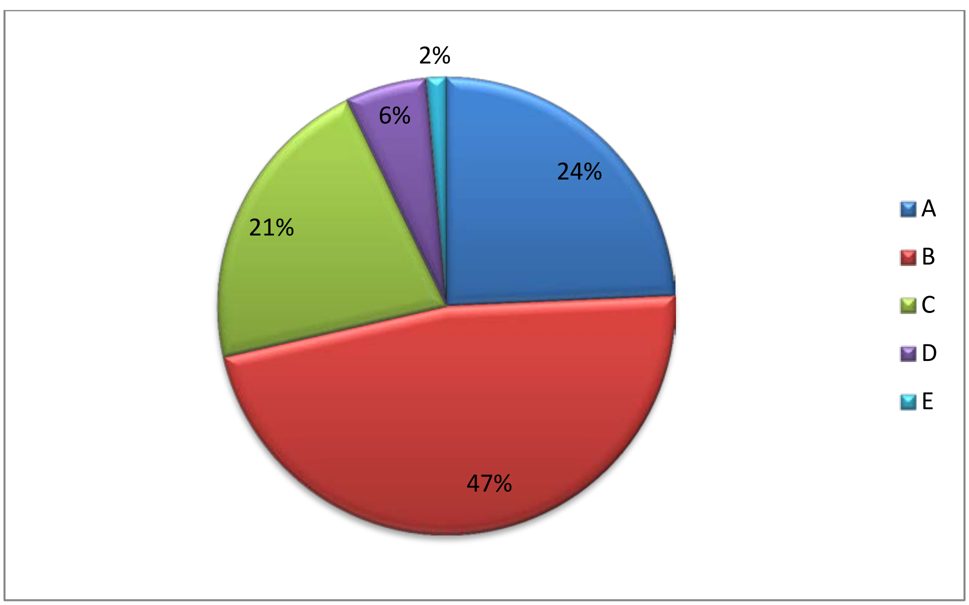

Figure 5.1 Transformer Classes

In order to assure proper performance for the classifier, SVM parameters should be adjusted correctly. For instance, if the regularization parameter or penalty factor C is too large, this will lead to a high penalty for non-separable points, and result in overfit. If the regularization parameter is too small, this may result in under fitting. In addition, the Kernel type and its parameters such as σ need to be chosen properly, as the correct selection for these parameters makes the model work properly on a new set of data. In this research, the values of SVM parameters are chosen based on trial and error strategy until the best performance is achieved. Other techniques might be used such as Particle Swarm Optimization (PSO) to optimize these parameters as reported in [23]. Table 5.2 shows the parameters used in the classifier and their values.

Table 5.2 SVM Parameters

Parameter	**Value**
C	2000
Epsilon	1.0E-12
Kernel	RBF
Tolerance parameter	0.001
K-fold cross validation	4

5.4 Feature Selection

Feature selection technique plays an important role in reducing the dimensionality of the data, or in other words, minimizing the number of tests that are required to give proper condition assessment. Hence, the InfoGainAttributeEval algorithm, which ranks attributes by their individual evaluation, has been used as a supervised technique with ranking methodology to rank features in priority sequence from most to least importance. The ranking strategy enables the user to perceive the most inductive attributes affecting the output (label). In order to make sure from the selection process, and the most effective features that will provide useful information about transformer condition, different selection algorithms have been examined and compared to InfoGainAttributeEval. Theses algorithms built in Weka software and designed to be used in the same way as the proposed algorithm. These algorithms are: InfoGainAttributeEval, sfcSubsetEval, consistencySubsetEval, gainRatioAtrributEval, svmAttributeEval which are explained in [43] . SVM classifier has been run on each algorithm to evaluate each algorithm's validity on the same number of tests. The precision rate is used as an evaluator to compare between them. From Figure 5.2 it is clearly shown that the proposed algorithm InfoGainAttributeEval has achieved the best performance over the rest of the selection algorithms.

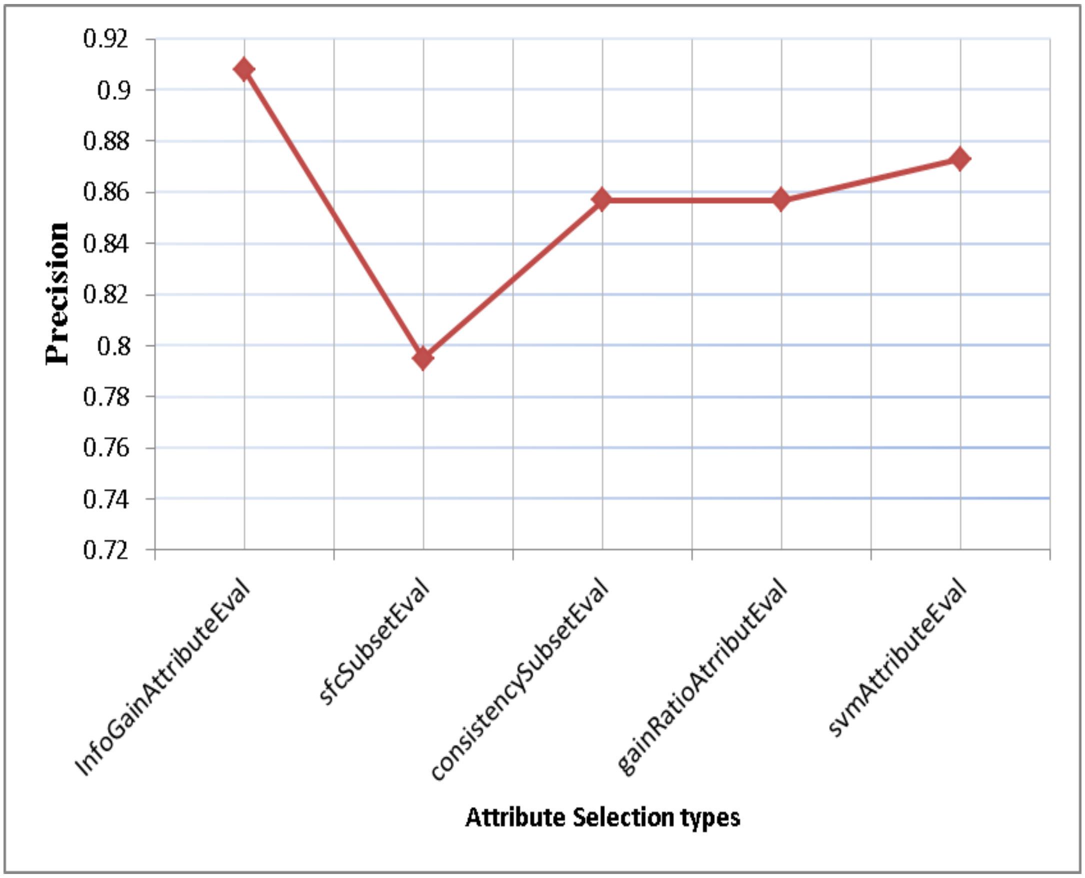

Figure 5.2 Feature selection algorithms comparison

The 19 tests used in this research have been ranked based on their individual evaluations using InfoGainAttributeEval algorithm as follows: DGA, infra-red, load history, bushing condition, power factor, DGA of LTC, leakage reactance, winding resistance, LTC oil quality, oil quality, turns ratio, cooling, furan, excitation current, core-to-ground, frequency response analysis, main tank condition, connectors, and gaskets. This ranking provides a clear indication of the validity of the feature selection technique since the most common and important test has been ranked first [2]. The physical meaning of DGA test being ranked first is of high importance, because this routine test provides information about the actual condition of a transformer. The concentration of the dissolved gases when performing DGA test reflects the severity and type of the fault occurring inside the transformer. While in-service, transformers are exposed to thermal and electrical stress which occasionally results in variation in DGA test conditions, as opposed to other stable tests such as core-to-ground or turns ratio. Therefore, the algorithm selects the most inductive tests that vary between transformers in their levels and affect the output (condition) as in the case of DGA, infra-red, load history, etc.

5.5 Classification Process

First of all, the whole data set of the seventy units and their labels are prepared and adjusted for use in the classifier. The entries are the 19 features (tests) and the labels (conditions) of each unit for the whole data set. The classification process begins with all the features and then reduces each feature step-by-step starting with the least inductive using feature selection techniques. At each step, the classifier is run to see the correctly classified instances and identify at which number of attributes the least percentage of error is achieved.

The flowchart in Figure 5.3 summarizes the feature selection and classification process and shows the classification of tests using SVM and RBF. First, data is collected by employing monitoring tests and feedback from experts, followed by the preparation and preprocessing of data as explained previously. When the data is ready for classification, the third step begins selecting and ranking the best tests (features) based on their individual evaluation. Next, the training and testing process using k-fold cross-validation is performed. After training the classifier and adjusting its parameters, the classification operation is conducted and the results are saved for comparison with new results. The second cycle starts by deleting the least important feature and then retrains the classifier with the new number of features. Each time, a new subset of data is generated using k-fold cross validation. This process continues until the number of tests reaches one which is indicated by k=1, meaning that the deletion process has been stopped. Thus, the accuracy of selected tests is printed out as the best results of the classifier.

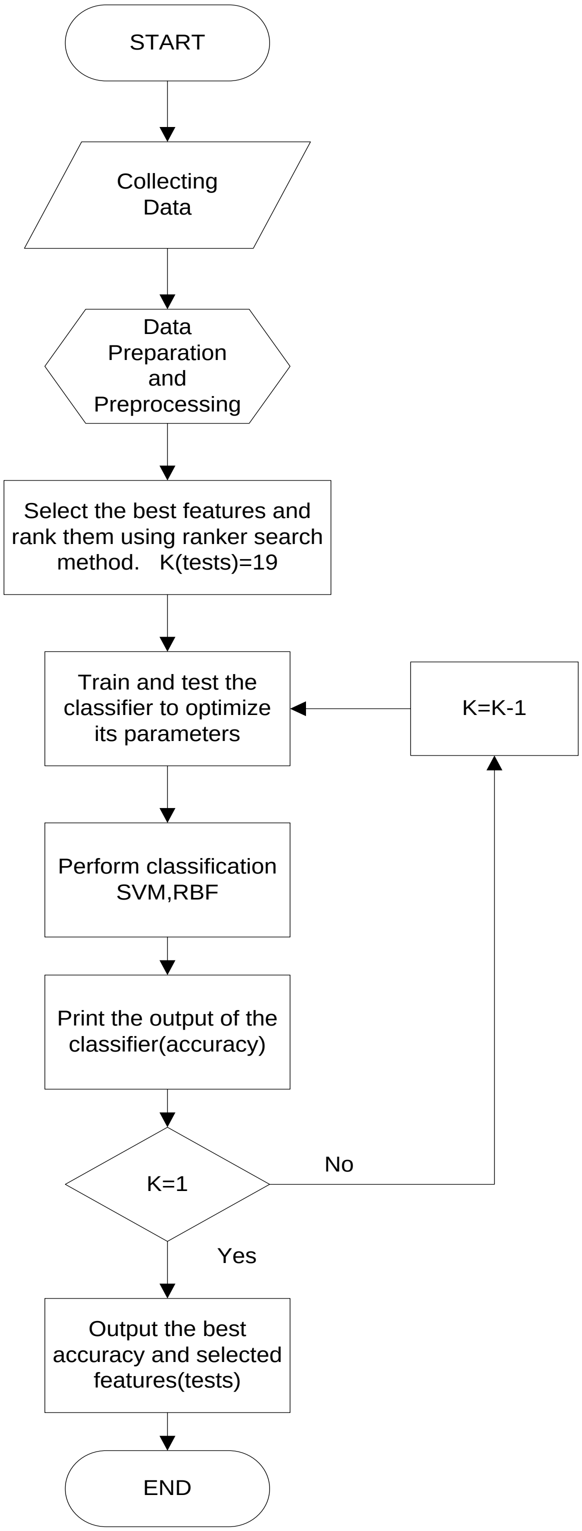

Figure 5.3 Classification process using SVM and RBF

5.6 Simulation Results

The SVM algorithm has been developed and tested in Matlab environment. The following figure shows the performance of SVM classifier over each number of features (tests), where each point over the number of tests represents the classifier accuracy at each cycle or loop as explained in the flowchart. Note that these results represent the performance of SVM for a combination of test data generated using k-fold. The trend is evident of the relationship between the accuracy of the classifier and the number of tests.

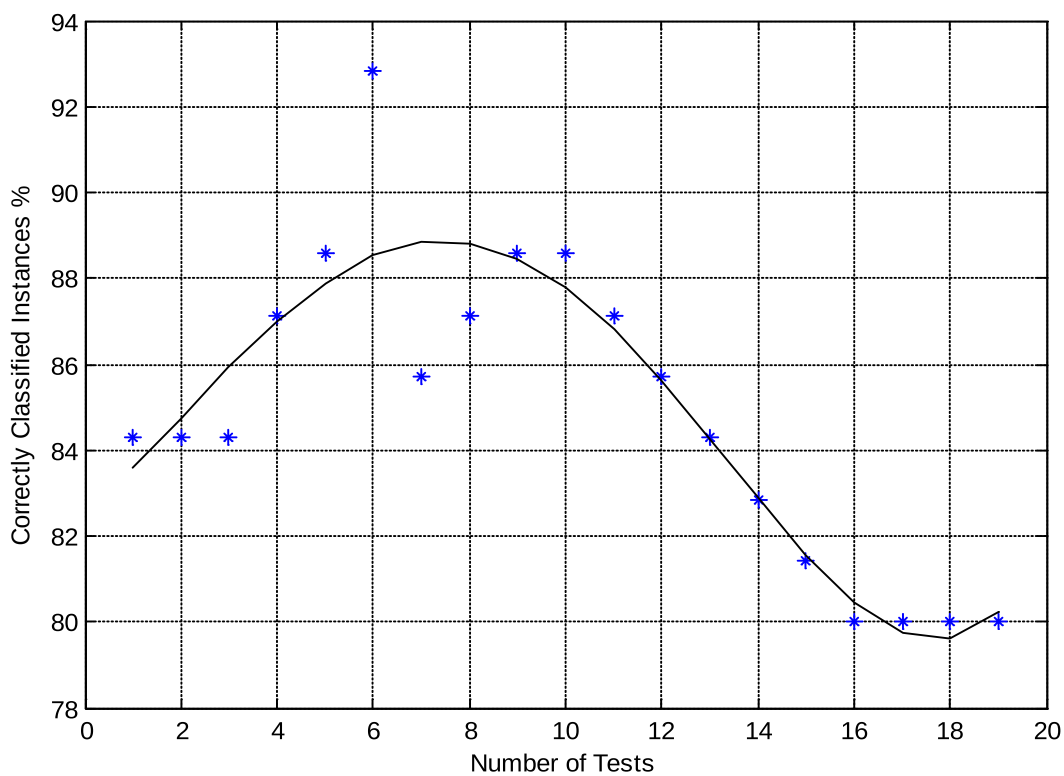

Figure 5.4 SVM Performance at each number of attributes

As shown in Figure 5.2, the x-axis represents the number of tests from one to 19 and the y-axis represents the correctly classified instances at each set of tests. It was observed that for the first three features the classifier performance achieved accuracy of 84.28%, which can be considered a good start for the classifier. However, as the number of attributes increased the classification process improved; hence, the correctly classified instances increased as illustrated at four and five attributes. At six features it was observed that the classifier provided the best performance (accuracy 92.85%), which indicates the best number and type of monitoring tests that can be utilized to indicate transformer condition. However, as the number of tests increased further, the classification accuracy became less, as in seven, nine, and ten features having the same accuracy. Finally, the classification accuracy decreased until it reached 16 features with accuracy of 80% and continued with the same performance until completion.

In data mining, when the dimensionality of the data increases, many techniques of data analysis and classification problems become more difficult. Moreover, a high number of attributes results in lower classification accuracy. On the other hand, when the number of the training data is very small, the created model in the case of a supervised learning technique will be less reliable [44]. Therefore, feature selection techniques have shown their capability in improving the classification performance when many features are used. The basic idea of feature subset selection is to eliminate the redundant or irrelevant attributes from the data as they can lead to a reduction of the classifier performance [45].

In order to have better visualization and to observe the trend of the data points, a curve fitting technique has been used to find the best curve that passes through these points. Curve fitting procedures have been implemented using Matlab software as a powerful technique for dealing with such a problem. This technique helps in identifying the margin and the required tests that should be used in the assessment process. It was observed that the classification accuracy improved as the number of tests decreased until feature nine, at which point the classification accuracy decreased. This trend can be interpreted as the non-inductive or irrelevant features being considered noisy information that reduces classification accuracy. Hence, utilizing the top ranked features obtained by the feature selection algorithm improves the classifier performance by ignoring the attributes of less information. More information regarding the classification accuracy is given by the classifier in Weka software at each test. In each run, full information is provided about the classifier including correctly/incorrectly classified instances, average error, accuracy, and most importantly the confusion matrix as explained in detail in the following tables and paragraphs. Table 5.3 displays one of the performance indices that provide the best results (six features). This case, which can be considered optimal in terms of selecting the best number of tests, gives utilities the type and number of monitoring tests to perform in order to assess transformer condition. Some terms shown in the tables are explained clearly in [46]. These measurements are useful for comparing classifiers. For example, Kappa statistic measures the agreement of prediction with the true class -1.0 signifying complete agreement.

Table 5.3 SVM Performance at Six Features

Classifier Performance	**Results**
Correctly classified instances	65 = 92.8571 %
Incorrectly classified instances	5 = 7.1429 %
Kappa statistic	0.8932
Mean absolute error	0.2446
Root mean squared error	0.3234
Relative absolute error	89.7888 %
Root relative squared error	88.1441 %
Total number of instances	70

Other terms shown in Table 5.4 such as True Position (TP), False Position (FP), Precision, and F-measure are explained further in [46]. These terms are critical in understanding the classification accuracy where the classified and misclassified examples are stated as a percentage for each individual class. The following are definitions of terms used in Table 5.4 [46]:

True Positive (TP) rate: the proportion of examples which were classified as class x among all examples which truly have class x

False Positive (FP) rate: the proportion of examples which were classified as class x but belong to a different class

Precision: the proportion of the examples which truly have class x among all those which were classified as class x

F-Measure: a combined measure for precision and recall defined as

$$2 * precision * Recall/(Precision + Recall)$$

Table 5.4 SVM Detailed Accuracy by Class

TP Rate	FP Rate	Precision	Recall	F-measure	Roc-Area	Class
1	0	1	1	1	1	**A**
0.939	0.054	0.939	0.939	0.939	0.949	**B**
0.867	0.036	0.867	0.867	0.867	0.96	**C**
1	0.015	0.8	1	0.889	0.992	**D**
0	0	0	0	0	0.435	**E**
0.929	0.034	0.917	0.929	0.922	0.959	**Weight Avg**

In the field of artificial intelligence, the success of a classifier is measured by the confusion matrix, which in some cases is called the contingency table. The confusion matrix is a specific table design which permits the visualization of classification accuracy. In this problem, the correctly classified and misclassified instances are shown in the confusion matrix in Table 5.5.

Table 5.5 SVM Confusion Matrix

a	b	c	d	e	Classified as
17	0	0	0	0	**a=A**
0	31	2	0	0	**b=B**
0	2	13	0	0	**c=C**
0	0	0	4	0	**d=D**
0	0	0	1	0	**e=E**

From the confusion matrix it is evident that the 17 class A transformers have been classified correctly and there are no misclassified instances. For the 33 class B transformers, there are 31 units correctly classified as B and two units classified as C**,** which indicates two misclassified instances. Transformers with class C that are represented by 15 units have two class B units while the other 13 units are classified correctly as C class. Finally, transformers with lower conditions such as C are classified correctly for four units and the one failed E class transformer is classified as D. Overall, it was observed that five of 70 transformers were incorrectly classified, which represents a 7.14% error rate. Therefore, the small number of misclassified instances can be considered a good indicator for the validity of the classifier performance.

5.7 Comparison With RBF

Radial Basis Function (RBF) classifier has been applied to the same data set following the same exact strategy as SVM in terms of data size and feature selection algorithm. RBF is used as a benchmark to validate the performance of the proposed algorithm.

It is clear from Figure 5.5 that the best performance of the classifier is given at ten, 14, and 15 tests with classification accuracy of approximately 90%. In addition, the procedures of applying SVM are followed. The performance of the classifier is improved by starting with 19 tests and then reducing the number of the tests step-by-step. The accuracy increases until it reaches its maximum at 14 and 15 features, at which point the classification accuracy starts decreasing. A sudden drop occurred at six features, in contrast with SVM where at six features the best performance was achieved. After this, the performance improves with small variation until it reaches the first attribute.

Comparing these results with SVM classifier gives a clear indication of each classifier's performance in terms of the minimum number of features that can be used to achieve better accuracy. Comparing these results with SVM classifier gives a clear indication of each classifier's performance in terms of the minimum number of features that can be used to achieve better accuracy.

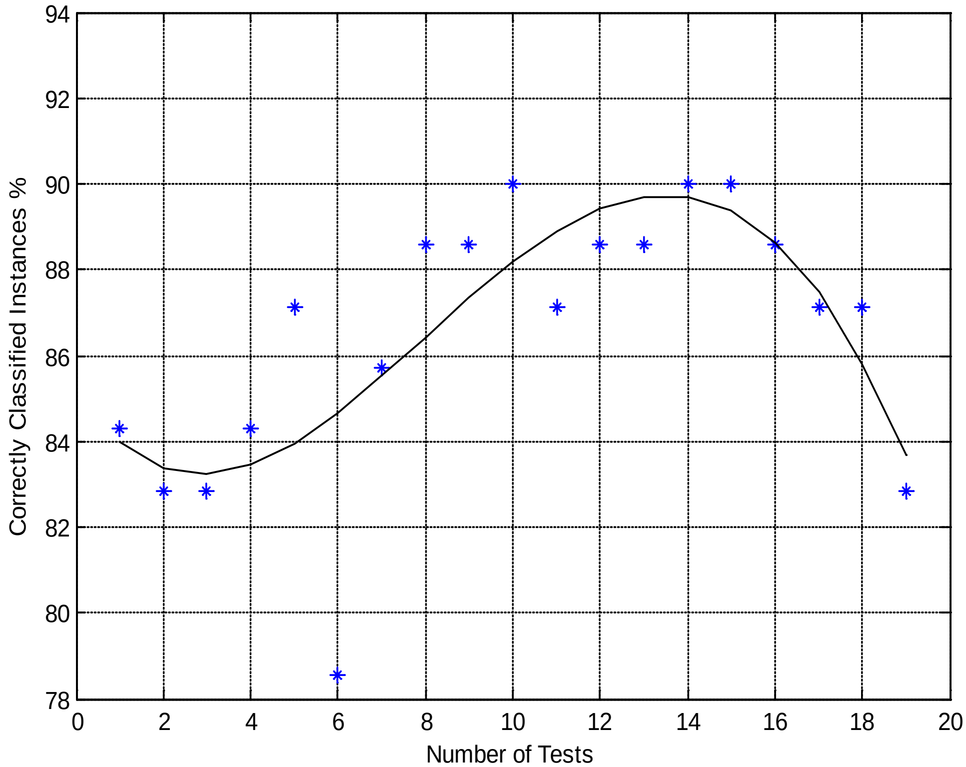

Figure 5.5 RBF Performance at each number of attributes

Table 5.6 RBF Performance at Ten Features

Correctly classified instances	63 = 90%
Incorrectly classified instances	7 = 10%
Kappa statistic	0.8498
Mean absolute error	0.0438
Root mean squared error	0.1947
Relative absolute error	16.0664 %
Root relative squared error	53.0805 %
Total number of instances	70

RBF performance parameters including TP Rate, FP Rate, Precision, Recall, and F-Measure are shown in detailed accuracy by class in Table 5.7. Comparing this table with SMV is very beneficial in distinguishing between the performances of the classifiers.

Table 5.7 Detailed Accuracy by Class RBF

TP Rate	**FP Rate**	**Precision**	**Recall**	**F - measure**	**Roc-Area**	**Class**
1	0	1	1	1	1	**A**
0.97	0.054	0.941	0.97	0.955	0.947	**B**
0.8	0.036	0.857	0.8	0.828	0.863	**C**
0.5	0.03	0.5	0.5	0.5	0.792	**D**
0	0.014	0	0	0	0.79	**E**
0.9	0.035	0.899	0.9	0.899	0.927	**Weight Avg**

As previously explained, the confusion matrix is most important. The classified and misclassified instances are shown in Table 5.8 in which transformers with condition A are classified correctly, one instance of B is misclassified, three instances of C are misclassified, two instances of D are misclassified, and finally one instance of E is misclassified. When comparing the confusion matrix from SVM and RBF, it was observed that both matrices correctly classified transformers with condition (class) A. In B classes, SVM misclassified two B class transformers as C class, whereas RBF misclassified one B class unit as C class. For C class, SVM demonstrated better performance by misclassifying two units as B class, while RBF misclassified three units. Moreover, SVM correctly classified all D class transformers and misclassified the only transformer with E class to D class, whereas RBF misclassified two instances of D class and the one instance of E class.

Table 5.8 RBF Confusion Matrix

a	**b**	**c**	**d**	**e**	**Classified as**
17	0	0	0	0	**a=A**
0	32	1	0	0	**b=B**
0	2	12	1	0	**c=C**
0	0	1	2	1	**d=D**
0	0	0	1	0	**e=E**

5.8 Results and Discussion

This chapter investigates the proper use of data mining techniques in the condition assessment of power transformers. Diagnostic techniques have been used to perform this task through two classifiers: Support Vector Machine (SVM) and Radial Basis Function (RBF). Firstly, the 19 monitoring tests have been ranked based on their individual evaluation using the ranker search from the attribute selection filter in Weka Software. Secondly, the features (tests) were reduced by one each time, and then the classifiers were executed on the new set of data to identify the classification accuracy. This strategy will allow utilities to minimize the number of tests to a level that properly indicates the transformer condition without losing much information. The performance of each classifier has been presented and compared.

In many cases, conducting all the monitoring tests on power transformers is not a cost-effective action for effective asset management. Therefore, a good method of meeting utilities satisfaction is to rank these tests by priority sequence in terms of their effectiveness in transformer condition, as well as select a limited number of these tests to improve classifier performance. SVM achieved better performance in predicting transformer condition with the first six tests when its accuracy reached 92.85%. This implies that new transformers can be classified reliably and the risk of misclassifying is very small. While RBF classifier has also achieved good performance, it is not comparable with SVM in terms of accuracy and number of tests. RBF achieved the best performance with the first ranked ten, 14, and 15 tests, in which classifier performance reached 90%. A comparison of the two classifiers in terms of their performance on the measured data is shown in Table 5.9.

Table 5.9 Performance Comparison (SVM and RBF)

Comparison Terms	**SVM**	**RBF**
Number of instances	70	70
Accuracy %	92.85%	90%
Number of tests at best performance	6	10, 14, 15
Correctly classified instances	65	63

Overall, applying the methodology of this research facilitates quick assessment of transformer condition and reliable decisions without performing all tests. There are many beneficial applications including the minimization of monitoring costs, time, and work that is required to perform these tests. This model will allow utilities to assess their transformers properly with a limited number of specific monitoring tests, and create reliable plans regarding proper transformer action.

Chapter 6
Conclusion

6.1 Thesis Conclusion

Power transformers are a vital and essential component in any electric network. Indeed, transformer cost varies between thousands and millions of dollars depending on the design and size of the unit. Therefore, any failure leads to unplanned outage or early breakdown before the designed lifespan expires, which costs utilities catastrophic loss. Hence, asset management practices have been widely utilized in the last decades to minimize overall life cycle cost and maintain cost-effectiveness. Such a practice ensures operating the system in an efficient and reliable way to meet the providers' and customers' satisfaction. Moreover, an effective asset management procedure needs many practices and work starts from design until disposal. In this research, condition assessment of power transformers, which is a part of asset management, has been investigated and a case study from in-service transformers has been conducted to develop a model that helps in identifying transformer condition from measured data.

A literature review is introduced which describes transformer importance and its contribution to the electricity sector. Various maintenance strategies are discussed as a part of asset management. In addition, a survey is conducted of monitoring and diagnostic techniques and of what has been done thus far in the area of transformer condition assessment. Common monitoring methods used to assess transformer state as well as the limits and variations from the unit nameplate interpreted based on national standards such as ASTM/IEEE and IEC have been investigated. Diagnostic techniques such as SVM and RBF as well as feature selection algorithms are explained for their employment in the assessment task.

Applying many monitoring techniques may be economically unreliable, require a lot of skilled work, and have many challenges and drawbacks from application. Some of these challenges can be summarized as electricity interruption in the case of off-line tests, time consumption, and high cost. Hence, the proposed model in this research employs data mining techniques to find a proper method of minimizing these challenges while proving reliability as a valid model. SVM classifier is used to predict transformer condition or health index among five classes A, B, C, D, and E from optimal to least optimal, respectively. The 19 tests (features) from the measured data have been prepared and processed to be used in the classification process. When the data are prepared and scaled, the feature selection based on ranker search method has been used to rank and select the most inductive tests (features) in a priority sequence. The algorithm ranked the tests based on their individual evaluation and their strong impact on the transformer’s condition. For instance, DGA, infra-red, and load history tests were selected first because they are the most common and important tests used by utilities to evaluate the state of the transformer. The full number of tests has been used to train and test the classifier (SVM) using k-fold cross-validation in which each feature (test) is used in the training and testing at least once, and then the tests have been reduced by eliminating the least inductive feature (test) after each classification process. Moreover, after each eliminating process, the classifier is retrained for the new number of tests. The number of tests for which the classifier gives best performance is selected as the recommended number of tests that reflect the transformer's actual condition. SVM reached its maximum performance (92.85%) at six tests, which are DGA, infra-red, load history, bushing condition, power factor, and DGA of LTC. In fact, these tests are the most common tests conducted by utilities and manufacturers because of the useful information they provide about the transformer's state. For instance, DGA provides information about faults inside a transformer and their causes. The thermography test identifies hot spot areas and overloading conditions. Loading history helps in determining how many times the unit has been operated under overloading conditions and their impact on the insulation system. Power factor is a routine test used to evaluate capacitive insulation condition between windings and compartments as well as assess bushing condition. Finally, DGA of LTC is used as the main test to assess tap changer condition. The criteria followed for selecting the specific number of tests prove the applicability of the model in predicting the transformer's actual condition with the least number of monitoring techniques. This process results in a cost-effective solution to assess the power transformers. The RBF model has been built in the same way and compared with SVM in terms of accuracy and number of tests. Although its classification accuracy was good (90%), this percentage was achieved at ten, 14, and 15 tests. However, RBF model is not comparable with SVM because of the high number of monitoring tests it needs to achieve this performance.

Considerable attention has been given to transformer condition assessment because of its direct impact on planning and budget as well as reliability and risk assessment. Furthermore, operating transformers in acceptable condition even after their designed lifespan has been considered to meet service providers' and customers’ satisfaction.

6.2 Future Work

A similar model using artificial intelligence tools could be designed to predict the transformer's future condition. Historical data from one unit represents different types of tests that could be utilized to train the model in order to use it on other units of the same size. Assessing the future condition of the transformer would provide utilities with more time to develop an effective plan and minimize the probability of unplanned outages through awareness of the deterioration rate or expected faults of the transformer.

References

[1] N. Dominelli, "Equipment health rating of power transformers," 2004, pp. 163-168.

[2] A. Jahromi, R. Piercy, S. Cress, and W. Fan, "An approach to power transformer asset management using health index," *Electrical Insulation Magazine, IEEE,* vol. 25, pp. 20-34, 2009.

[3] B. GORGAN, P. NOTINGHER, L. BADICU, and G. TANASESCU, "CALCULATION OF POWER TRANSFORMERS HEALTH INDEXES," *health,* vol. 602, p. 929481.

[4] Z. Xiang and E. Gockenbach, "Asset-Management of Transformers Based on Condition Monitoring and Standard Diagnosis [Feature Article]," *Electrical Insulation Magazine, IEEE,* vol. 24, pp. 26-40, 2008.

[5] "IEEE Recommended Practice for Installation, Application, Operation, and Maintenance of Dry-Type General Purpose Distribution and Power Transformers," *ANSI/IEEE Std C57.94-1982,* p. 3, 1982.

[6] J. Spacek, "Maintenance strategies of power equipments with a brief view to condition monitoring of power transformers," 2008, pp. 16-19.

[7] A. Pride, "Reliability Centered Maintenance (RCM)," *WBDG Electronic Resource [http://www. wbdg. org/resources/rcm. php],* 2008.

[8] J. McCalley, Y. Jiang, V. Honavar, J. Pathak, M. Kezunovic, S. Natti, C. Singh, and P. Jirutitijaroen, "Automated integration of condition monitoring with an optimized maintenance scheduler for circuit breakers and power transformers," *Power Systems Engineering Research Center,* 2006.

[9] X. Zhang and E. Gockenbach, "Asset-Management of Transformers Based on Condition Monitoring and Standard Diagnosis [Feature Article]," *Electrical Insulation Magazine, IEEE,* vol. 24, pp. 26-40, 2008.

[10] A. Emsley and G. Stevens, "Review of chemical indicators of degradation of cellulosic electrical paper insulation in oil-filled transformers," 1994, pp. 324-334.

[11] J. Singh, Y. R. Sood, and R. K. Jarial, "Condition monitoring of power transformers-bibliography survey," *Electrical Insulation Magazine, IEEE,* vol. 24, pp. 11-25, 2008.

[12] T. K. Saha, "Review of modern diagnostic techniques for assessing insulation condition in aged transformers," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 10, pp. 903-917, 2003.

[13] P. J. Baird, H. Herman, G. C. Stevens, and P. N. Jarman, "Spectroscopic measurement and analysis of water and oil in transformer insulating paper," *Dielectrics and Electrical Insulation, IEEE Transactions on,* vol. 13, pp. 293-308, 2006.

[14] B. H. Ward, "A survey of new techniques in insulation monitoring of power transformers," *IEEE Electrical Insulation Magazine,* vol. 17, pp. 16-23, 2001.

[15] D. Allan and C. Jones, "Thermal–oxidative stability and oil-paper partition coefficients of selected model furan compounds at practical temperatures–9 th International Symposium on high voltage engineering," 1995.

[16] A. M. Emsley and G. C. Stevens, "A reassessment of the low temperature thermal degradation of cellulose," in *Dielectric Materials, Measurements and Applications, 1992., Sixth International Conference on,* 1992, pp. 229-232.

[17] Y. Z. Yang Ghazali, M. A. Talib, and H. Ahmad Rosli, "TNB experience in condition assessment and life management of distribution power transformers," in *Electricity Distribution - Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*, 2009, pp. 1-4.
[18] F. R. Barbosa, O. M. Almeida, A. P. S. Braga, M. A. B. Amora, and S. J. M. Cartaxo, "Application of an artificial neural network in the use of physicochemical properties as a low cost proxy of power transformers DGA data," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 19, pp. 239-246, 2012.
[19] B. Németh, S. Laboncz, I. Kiss, and G. Csépes, "Transformer condition analyzing expert system using fuzzy neural system," 2010, pp. 1-5.
[20] W. Feng-Jiao, Z. Guan-Jun, W. Shi-Qiang, X. Hao, W. Da, and L. Min, "Research on condition assessment method of intelligent power transformer," in *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*, 2011, pp. 1-4.
[21] R. Liao, H. Zheng, S. Grzybowski, L. Yang, Y. Zhang, and Y. Liao, "An integrated decision-making model for condition assessment of power transformers using fuzzy approach and evidential reasoning," *Power Delivery, IEEE Transactions on*, vol. 26, pp. 1111-1118, 2011.
[22] M. Hui, T. K. Saha, and C. Ekanayake, "Statistical learning techniques and their applications for condition assessment of power transformer," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 19, pp. 481-489, 2012.
[23] L. Jinling and W. Mijia, "Condition assessment for power transformer based on improved particle swarm optimization and Support Vector Machine," in *Critical Infrastructure (CRIS), 2010 5th International Conference on*, 2010, pp. 1-6.
[24] P. Dash and S. Samantaray, "An accurate fault classification algorithm using a minimal radial basis function neural network," *Engineering intelligent systems*, vol. 4, pp. 205-210, 2004.
[25] W. Weihua, "License Plate Recognition Algorithm Based on Radial Basis Function Neural Networks," in *Intelligent Ubiquitous Computing and Education, 2009 International Symposium on*, 2009, pp. 38-41.
[26] H. G. Suykens J.A.K., *Advances in learning theory* vol. 190 NATO-ASI: IOS Press, 2003, 2003.
[27] G. F. Lin and L. H. Chen, "A non-linear rainfall-runoff model using radial basis function network," *Journal of Hydrology*, vol. 289, pp. 1-8, 2004.
[28] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *The Journal of Machine Learning Research*, vol. 3, pp. 1157-1182, 2003.
[29] T. M. Mitchell, "Machine learning," ed: McGraw-Hill New York:, 1997.
[30] "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers," *IEEE Std C57.104-1991*, p. 0_1, 1992.
[31] M. Duval, "Dissolved gas analysis: It can save your transformer," *Electrical Insulation Magazine, IEEE*, vol. 5, pp. 22-27, 1989.
[32] "IEEE Guide for Acceptance and Maintenance of Insulating Oil in Equipment," *IEEE Std C57.106-2006 (Revision of IEEE Std C57.106-2002)*, pp. 1-36, 2006.
[33] M. Wang, A. Vandermaar, and K. Srivastava, "Review of condition assessment of power transformers in service," *Electrical Insulation Magazine, IEEE*, vol. 18, pp. 12-25, 2002.
[34] G. Duke, "Predictive maintenance a case study in infrared thermography," *Electrical Maintenance*, pp. 11-12, 1998.
[35] I. Arifianto and B. Cahyono, "Power transformer cooling system optimization," in *Properties and Applications of Dielectric Materials, 2009. ICPADM 2009. IEEE 9th International Conference on the*, 2009, pp. 57-59.

[36] R. Yadav, S. Kumar, A. Venkatasami, A. M. Lobo, and A. M. Wagle, "Condition based maintenance of power transformer: A case study," in *Condition Monitoring and Diagnosis, 2008. CMD 2008. International Conference on*, 2008, pp. 502-504.
[37] M. A. Franchek and D. J. Woodcock, "Life-Cycle considerations of loading transformers above nameplate rating," in *65th Annual International Conference of Doble Clients*, 1998.
[38] A. Setayeshmehr, A. Akbari, H. Borsi, and E. Gockenbach, "On-line monitoring and diagnoses of power transformer bushings," *Dielectrics and Electrical Insulation, IEEE Transactions on*, vol. 13, pp. 608-615, 2006.
[39] K. Jan, "Operating Damages of Bushings in Power Transformers," *TRANSACTIONS ON ELECTRICAL ENGINEERING*, p. 89, 2012.
[40] A. Shintemirov, W. Tang, and Q. Wu, "Transformer winding condition assessment using frequency response analysis and evidential reasoning," *Electric Power Applications, IET*, vol. 4, pp. 198-212, 2010.
[41] M. De Nigris, R. Passaglia, R. Berti, L. Bergonzi, and R. Maggi, "Application of modern techniques for the condition assessment of power transformers," in *Cigré Session*, 2004, pp. A2-207.
[42] G. J. McLachlan, K. A. Do, and C. Ambroise, *Analyzing microarray gene expression data* vol. 422: Wiley-Interscience, 2004.
[43] http://www.cs.waikato.ac.nz/ml/weka/index.html.
[44] A. G. K. Janecek, W. N. Gansterer, M. Demel, and G. F. Ecker, "On the relationship between feature selection and classification accuracy," in *JMLR: Workshop and Conference Proceedings*, 2008, pp. 90-105.
[45] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artificial intelligence*, vol. 97, pp. 245-271, 1997.
[46] R. R. Bouckaert, E. Frank, M. Hall, R. Kirkby, P. Reutemann, A. Seewald, and D. Scuse, "WEKA Manual for Version 3-7-4," 2011.