ESTIMATING THE TRANSFORMER HEALTH INDEX USING ARTIFICIAL INTELLIGENCE TECHNIQUES

by

Alhaytham Y. Al Qudsi

A Thesis Presented to the Faculty of the
American University of Sharjah
College of Engineering
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in
Electrical Engineering

Sharjah, United Arab Emirates

June 2016

Approval Signatures

We, the undersigned, approve the Master's Thesis of Alhaytham Y. Al Qudsi.

Thesis Title: Estimating the Transformer Health Index Using Artificial Intelligence Techniques

Signature	**Date of Signature** (dd/mm/yyyy)
___________________________ Dr. Ayman El-Hag Associate Professor, Department of Electrical Engineering Thesis Advisor	_______________
___________________________ Dr. Mostafa F. Shaaban Assistant Professor, Department of Electrical Engineering Thesis Committee Member	_______________
___________________________ Dr. Michel Pasquier Associate Professor, Department of Computer Science and Engineering Thesis Committee Member	_______________
___________________________ Dr. Nasser Qaddoumi Head, Department of Electrical Engineering	_______________
___________________________ Dr. Mohamed El-Tarhuni Associate Dean, College of Engineering	_______________
___________________________ Dr. Leland Blank Dean, College of Engineering	_______________
___________________________ Dr. Khaled Assaleh Interim Vice Provost for Research and Graduate Studies	_______________

Acknowledgement

In the name of Allah, the Most Gracious, the Most Merciful: The best and hardest moments of my life occurred during the time I spent at AUS as a M.Sc. student. Writing this thesis was a dream that came true. The presented work is a result of all the graceful moments and great people that the Most Merciful allowed me to have in my life. I thank Allah Almighty for gracefully giving me the strength and support of finishing this work with the support of the surrounding people.

My greatest appreciation and acknowledgment go to the greatest man I have ever met, personally and academically. My university advisor, teacher, elder brother and mentor Dr.Ayman El-Hag. He stood with me and supported me up to the last minute of my study at AUS as a student. Having him as a mentor in my life is a great blessing from the Most Merciful. I will always owe him this work.

Sincere thanks is dedicated for the great team of university authority members. Dr. Nasser Qaddoumi, Dr.Mohamed El-Tarhuni and Dr.Khaled Assaleh for never considering the option of letting me go when they had all the reasons to do so.

I would like to thank my parents for always reminding me of how important earning this degree is, and how they always were by my side. My father who worked restlessly with all what he had for a prosperous future. My mother who raised me well to make sure that I will never give up on my dreams. Moreover, I would like to thank my brother, Mohammad and sister Leen for their support and motivation.

Finally, every great man is supported by a great wife. I would like to thank my wife Dana for giving me all the time and space to read and write. Moreover, special thanks is dedicated for my baby Mariam for neither giving me the time nor the space with her crying and screaming. But she always cheered me up with her laughs when I was in need for help.

For those who never gave up…

Abstract

Transformer Asset Management (TAM) is concerned with the strategic activities that monitor and manage the transformer asset in the power system. The outcomes of TAM aim at setting proper monitoring methods and maintenance plans, with minimal cost of time and money. Monitoring methods in the form of electrical, chemical and physical tests are conducted to assess the transformer operational condition. The main part, which is directly related to the ageing of the transformer, is the oil-paper insulation system. The standard practiced monitoring test methods used by TAM companies are considered highly effective and useful. However, a full feedback of the transformer's condition requires a number of monitoring tests to be conducted. Such an exercise is considered expensive and difficult to implement for some of the tests. Moreover, the individual conducted tests cannot provide a comprehensive understanding of the transformer condition based on a single factor. Thus, the concept of the Health Index (HI) was developed to accurately assess the transformer's condition and effective remnant age. The main components involved in the HI computation are related to the transformers' insulation condition, service record and design. Finding the transformer HI is normally done through using several industry computational methods. The drawback of these methods is the large number of tests required to achieve high level of condition assessment accuracy. Thus, alternative Artificially Intelligent (AI) methods should be used to design the HI model. AI methods, such as Artificial Neural Networks (ANN), can learn the pattern of the response output (HI), based on a given set of input (monitoring tests). The use of feature selection technique such as stepwise regression, can lead to an effective reduction of redundant tests in the presence of more significant ones. The presented work produces a general cost-effective AI based HI predictor model that can be used by different utility companies. Such a predictor would be able to produce a HI output value with a 95% prediction accuracy using only a subset of the required input features. Furthermore, the model can produce the same prediction accuracy with a predicted costly feature as one of the input features.

Search Terms: ***Transformer Asset Management, Health Index, Artificial Intelligence, Artificial Neural Network and Stepwise Regression.***

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AI	-	Artificial Intelligence
ANN	-	Artificial Neural Network
BDV	-	Breakdown Voltage
CA	-	Condition Assessment
CBM	-	Condition Based Maintenance
CM	-	Condition Monitoring
DDF	-	Dielectric Dissipation Factor
DGA	-	Dissolved Gas Analysis
DGAF	-	Dissolved Gas Analysis Factor
FF	-	Feed Forward
FFA	-	Furan Analysis
FRA	-	Frequency Response Analysis
FSVM	-	Fuzzy Support Vector Machine
HI	-	Health Index
LTC	-	Load Tap Changer
IR	-	Insulation Resistance
MLP	-	Multi-Layer Perceptron
OQA	-	Oil Quality Analysis
OQF	-	Oil Quality Factor
ppm	-	Particles Per Million
RCM	-	Reliability Centered Maintenance
SVM	-	Support Vector Machine
TAM	-	Transformer Asset Management
TBM	-	Time Based Maintenance

Chapter 1. Introduction to Transformer Asset Management

The growing demand of the electrical power grid forces the utility companies to take decisions that will ensure the continuous supply of power with high standards of network reliability. Providing such reliable power grids demands the continuous operation of the electrical equipment used in supplying the power. Transformers are one of the asset elements that may be subjected to operation failure due to internal faults which in turn will affect the whole power grid. Thus, transformers require special attention and careful supervision by utility companies to ensure the continuous operation.

1.1. Definition and Objectives

Transformer Asset Management (TAM) defines the set of strategic activities that are practiced by the utility companies in order to be fully aware of the operating conditions of the transformers. This will lead to make decisions related to the transformers in question to be replaced, refurbished or relocated to insure reliable power supply. The prime objective of asset management is to define the transformer life-cycle management strategy that will define monitoring priorities and provide a planned maintenance strategy for all transformers [1]. This objective takes into consideration the importance of reducing the maintenance cost and avoids accidental transformer operation failure. Moreover, accomplishing this objective requires the use of multiple transformer diagnostic models that are developed in order to evaluate the transformers operational conditions. Such diagnostic methods can be used to assess the lifetime and reliability of the transformers, and can further draw conclusions about the causes of transformer ageing and future operational failure. TAM strategy cycle defines a set of standard procedures to be followed in order to develop a complete understanding of the transformers' condition and to give an overall assessment of the required maintenance plan for a reliable supply of power [2].

1.2. Building a Risk-Assessment Database

Fully understanding the consequences of the transformers' failure raises the awareness of the utility companies of the involved risk associated with each transformer. This is highly important to define a priority list of energized transformers. Risk-Assessment studies like the FMEA (Failure Mode and Effect Analysis) are used to perform this task [1]. Building a Risk-Assessment database is associated with listing a set of all the possible failure causes of a transformer whether being internal or external. This helps identifying the most sensitive areas where time and money have to be invested for developing better problem-recognition models and remedy maintenance solutions.

The main reasons involved in the transformer operation failure include ageing, deterioration, or damage of different internal or external components of the transformer such as the insulation system, load tap changer, windings, tank and bushings. Factors leading to such damage can be age-based factors, such as the reduced dielectric strength of the insulation system due to insulation contamination. Other factors are mainly due to the electrical, mechanical and thermal stresses due to external short circuits, incipient faults, transient switching, lightning strikes and excessive overloading. Having these factors identified, the utility company can predict the probability of failure and remaining life-time using formulated probabilistic models.

1.3. Setting the Condition Monitoring and Assessment Strategy

Once the utility company builds a transformer risk-assessment database, condition monitoring (CM) and condition assessment (CA) are performed. Condition monitoring refers to the development of special methods for monitoring and acquiring the information of a certain parameter in the transformer [3]. An example would be taking oil samples of a transformer to use the information of particular substances composition for later analysis to determine the strength of the oil insulation system. Condition assessment, on the other hand, processes the acquired CM data to give an evaluation of the transformer's current performance and its predicted life-time. CA can either be done through standard diagnostic methods or novel techniques that deal with artificial intelligence.

1.4. Adapting Effective Maintenance Plans

Transformers that are poorly maintained have short life-time expectancy [3]. The maintenance plan practiced by the utility company is considered poor if some of the transformers are under high risk of operational failure with the utility being unaware of it. TAM requires utility companies to adapt one of three strategic maintenance plans which are namely the corrective, preventive and reliability-based maintenance plans.

1.4.1. **Corrective maintenance.** The corrective maintenance exercise is only conducted when an alarm is triggered or an unplanned outage occurs. This maintenance plan was the first adapted maintenance strategy in power systems. Corrective maintenance is considered a cost-saving maintenance approach in which regular inspections are not required, low manpower is needed and is only desirable when the transformer failure is due to easily replaceable accessories. However, if the operational failure was due to a severe internal damage in a main transformer component that could have been detected with proper monitoring, corrective maintenance would be considered as a failing strategy. Moreover, energizing the transformer would require a significant number of tests if the fault cause was un-identified.

1.4.2. **Preventive maintenance.** Preventive maintenance plans take into account the use of the CM and CA methods in order to fully assess the transformer condition. Preventive maintenance plans can be categorized into time-based maintenance (TBM) and condition-based maintenance (CBM) plans. TBM sets fixed time span intervals for the inspection and maintenance of the transformers. This strategy is followed by most of the utility companies. It increases the power supply reliability by preventing un-planned outage through early detection of possible operational threats. The trade-off in this maintenance plan is between the time-span interval setting and the maintenance cost. The shorter the time-span, the lower is the probability failure but with an expensive inspections and manpower. On the other hand, the longer the time-span, the more vulnerable the transformer is to failure due to incipient and external faults but with lower maintenance cost. CBM sets a maintenance plan on the basis of the produced outcomes and conclusions of the CA process and hence has a comprehensive understanding of the transformers' operational conditions. This helps the utility company to make an informative decision of the required time-span for the next maintenance exercise and what the components that have to be repaired are. This approach is an excellent preventive of operation failures due to the constant monitoring, with a cost-saving strategy of less man-power and detection of possible incipient faults at an early stage. The disadvantage in this plan is the need for continuous online condition monitoring of the transformer, and the required sophisticated CA methods that can correctly identify the appropriate maintenance time and type.

1.4.3. **Reliability centered maintenance.** With the risk-index associated with each transformer being well known, utility companies can take studied decisions on the maintenance requirements. RCM combines the knowledge of the risk-index with the outcomes of the CA to make a proper maintenance management plan amongst a population of in service transformers. This maintenance strategy is considered as the optimum choice that is subjected to the constraints of maintenance cost and operation reliability. Designing such smart systems requires a large database of CM acquired data and risk-assessment data for proper CA training.

Figure 1 illustrates the general TAM strategy and how the exercised maintenance plans are incorporated accordingly.

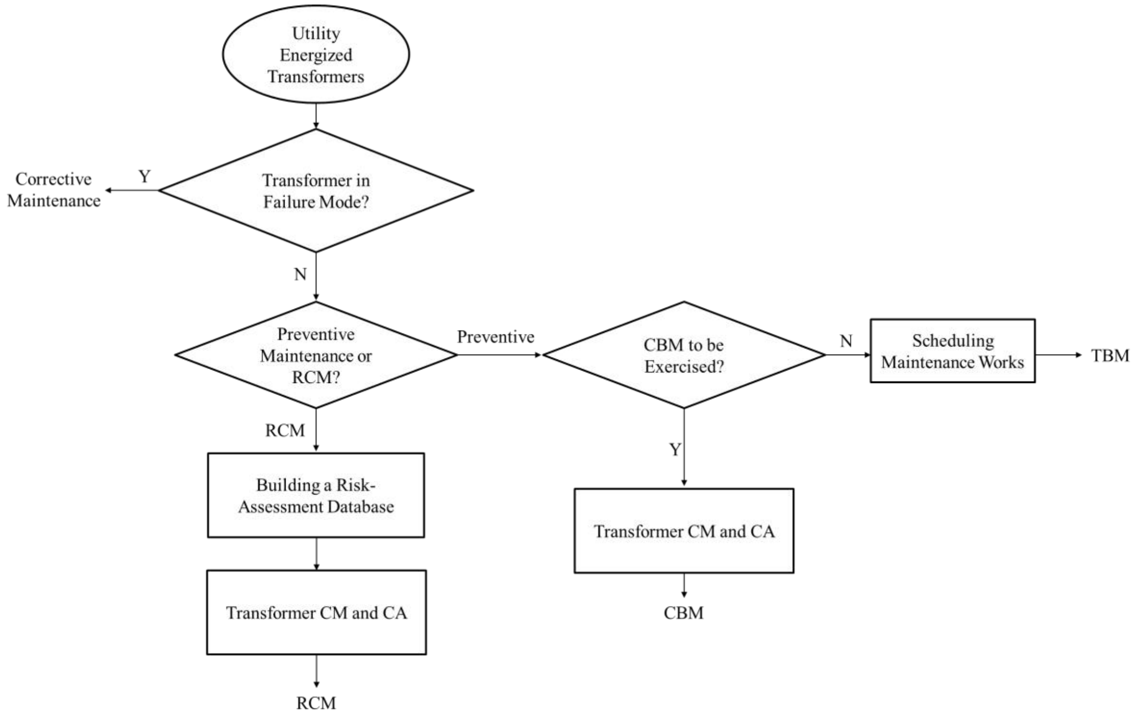

Figure 1: TAM strategy flowchart

Chapter 2. Background

2.1. Transformer Health Condition

The transformer is used to step the voltage level up or down depending on its function and location in the power system. Each transformer has a voltage and power rating that should not change during the entire duration of service. The factors that can facilitate a change in such parameters are related to the weakness and degraded health of the major transformer components. The main component associated with the transformer health is its insulation system. The performance of the transformer and its expected outcomes can only be maintained through the proper care and awareness of the strength of its insulation system. The strength of the insulation systems is measured through its physical, electrical and mechanical properties. Other transformer components and accessories, like tap changer, attribute to the overall health of the transformer, but are relatively of lower significance when compared to the insulation system.

2.1.1. **Oil-paper insulation system.** The insulation system of the transformer is composed of solid and liquid forms of insulation. The liquid or oil insulation of a transformer system plays a vital role in providing an insulating medium that will prevent the passage of electrical current between conductors of different potential levels. Moreover, the transformer oil acts as a cooling medium for the transformer. The oil is pumped through the transformer windings to absorb the dissipated heat due to the winding copper and core losses. The solid or paper insulation is mainly used to cover the transformer conductors and insulate the windings. The paper insulation is relatively more vulnerable than the oil insulation to excessive thermal and electrical stresses. A weakness in the paper insulation strength can result in creating conducting paths between the transformer windings. Moreover, paper pressboards are used in the transformer system for the isolation of high voltage parts that is filled with mineral oil.

2.1.1.1. *Oil insulation system.* The chemical composition of the mineral transformer oil is based on three carbonic structures. The first structure is paraffin which is a straight organic chain of Normal-Alkanes (N-Alkanes or waxes) [4]. High concentrations of N-Alkanes can increase the oil viscosity and prevent the free flow movement. They are associated with forming solid or sludge substances that can block the movement of oil within the transformer. The second structure is the naphthenic structure which is a cyclo-alkanes of good solvent properties. The third and most important structure that constitutes to the oil's thermal and electrical property is the aromatic structure. These can be in the form of monoaromatics or polyaromatics (PAC). PAC contributes to the dissolved gas property in the transformer oil.

The transformer oil is mineral insulating product which is produced by a set of refining processes on extracted crude oil. The refining processes include the fractional distillation, dewaxing, extraction and hydrogenation with the crude oil as a starting product. Fractional distillation of crude oil involves separation of the oil components on the basis of their different boiling points. The useful organic material produced by the distillation process is later subjected to a dewaxing solvent which is used to remove N-alkane compounds. The extraction process later will be used to remove the reactive polar molecules of the distillated fluid. Finally, a high temperature catalytic reaction occurs in the presence of hydrogen (hydrogenation) to chemically convert the aromatic and polar compounds in the extracted fluid to the useful organic mineral oil.

2.1.1.2. ***Paper insulating system.*** The solid insulation of the transformer system comprises of the paper and pressboard components. The paper material is made from Kraft un-bleached cellulose which is known for its mechanical and electrical strength. The cellulose fiber (extracted from softwood) is a polymer chain of D-anhydroglucopyranose units that are tied together through β-1,4-glycosidic bonds [5]. Production of Kraft paper is done by processing the softwood using the Sulphate method, followed by a series of extensive cleaning processes to remove the undesirable resins and mineral substances.

2.1.2. **Other transformer health components**. The electrical and physical properties of other transformer components should be taken into account during the CA process of the transformer [1], [2]. These include the electrical properties of the load tap changer (LTC), turns ratio, winding resistance, transformer impendence and capacitance, bushings, etc. Physical properties include the corrosive condition of the tank, efficiency of the cooling system, the mechanical strength of the gaskets, etc. Degradation of the electrical and physical properties of these components can reduce the expected lifetime and increase the transformer probability to fail.

2.2. Condition Monitoring and Assessment Procedure for Transformers

TAM defines a standard set of CM test procedures that should be followed in order to properly assess the current situation of the transformer, and get a good approximation of its remaining lifetime. These test procedures are used to acquire the transformer condition parameter data that is required by the CA methods. Related works of TAM indicate the initiative of developing new data-mining based methods to predict the condition parameters for reducing the maintenance procedure costs. This section will go through the traditional and data-mining based CM methods.

2.2.1. **Dissolved Gas Analysis (DGA).** One of the fundamental methods used in assessing the transformers' operating condition is the analysis of the dissolved gases in the insulating oil. DGA is considered a crucial tool for interpreting internal faults in the transformer. Dissolved gases are decomposed by-products of oil and cellulose paper insulation material as a result of the transformer internal electrical and thermal faults [6]. Incipient electrical faults such as partial discharge or intense arcing can cause the breakdown of the insulation material through ion bombardment (in partial discharge) or by the arcing thermal energy. Decomposition of oil material occurs due to the breaking of Carbon and Hydro-Carbon bonds. New produced molecular gaseous products dissolve in the oil solution. Paper decomposition, on the other hand, occurs by breaking the Glycosidic bonds in the cellulose polymer chain. The rate of thermal breakdown of the cellulose paper material depends on the temperature and volume of the Kraft paper material. Cellulose decomposition is accelerated due to factors like heat, moisture and oxygen [5]. The CM method for DGA requires having test samples of the transformer oil to find the composition of certain dissolved gases known as the key gas elements through the application of chromatographic separation. The key gases are namely Hydrogen (H_2), Methane (CH_4), Acetylene (C_2H_2), Ethylene (C_2H_4), Ethane (C_2H_6), Carbon Monoxide (CO) and Carbon Dioxide (CO_2) [7].

Interpretation of the dissolved gases using CA methods can be done by practiced standard diagnostic methods. The IEEE.C57.104 guide for DGA defines two CA methods for interpreting the internal transformer faults [7]. One method is the investigation of the dominant key gases in oil samples along with the total dissolved combustible gas (TDCG) concentration. The other method is through the calculation of certain dissolved gas ratios using either the Duval triangle, Roger or Doerenburg ratio. Based on the DGA findings, four main incipient faults can be interpreted which are mainly low level thermal (150-300°C), high level thermal (above 300°C), low intensity electrical discharge and high intensity arcing faults. For example, high levels of Hydrogen indicate the possibility of partial discharge fault inside the transformer winding. Calculating the ratio of Methane to Hydrogen can indicate the presence of high level thermal faults.

Based on the DGA findings, it has been concluded that the breakdown of the insulating oil will mainly result in release of Hydrogen, Methane, Acetylene, Ethylene and Ethane. On the other hand, the breakdown of cellulose will result in high amounts of released Carbon Dioxide and Carbon Monoxide gases [6]. Table 1 shows an example of one of the condition-based DGA method that is recommended by the IEEE.C57.104 standards. Based on the TDCG, a C1 condition indicates the safe operation of the transformer with least probability of failure threat. A C2 condition indicates an abnormal operating condition in which additional investigations are required if any key gas concentration exceeds the specified limit. C3 level indicates the immediate requirement for further investigation if any key gas concentration limit is exceeded. A C4 limit is reached when the transformer is in its worst condition and has to be immediately investigated [7].

TAM in DGA plays a vital role in defining the required action strategy that should be done based on the outcomes of the standard or data-mining based approaches. The utility can make a supported decision on scheduling the next sampling interval. Moreover, the utility can make early failure-preventive actions that can be either to remove the transformer or to be more cautious with its operation.

Table 1: Condition-based DGA [7]

Status	*Dissolved key gas concentration limits (µL/L (ppm)*							
	Hydrogen	*Methane*	*Acetylene*	*Ethylene*	*Ethane*	*Carbon Monoxide*	*Carbon Dioxide*	*TDCG*
C1	100	120	1	50	65	350	2500	720
C2	101-700	121-400	2-9	51-100	66-100	351-570	2500-4000	721-1920
C3	701-1800	401-1000	10-35	101-200	101-150	571-1400	4001-10,000	1921-4630
C4	≥1800	≥1000	≥35	≥200	≥150	≥1400	≥10000	≥4630

2.2.2. **Oil Quality Analysis (OQA).** The quality of the oil is mainly characterized by its insulation strength against any electrical or thermal stress. Verifying the oil quality is done through a series of electrical, physical and chemical tests conducted on oil test samples. Electrical tests include the dielectric strength (or breakdown voltage) and dielectric dissipation factor (DDF). Physical tests are conducted for measuring the oil's interfacial tension (IFT) and visual appearance (color). Chemical tests are conducted for measuring the acidity and water content of the transformer oil [8].

2.2.2.1. *Dielectric strength.* The measure of the oil's dielectric strength is an indication of the oil's capability to withstand the electrical stress caused by an electrical field without the breakdown of its insulation property. Dielectric strength is measured by means of the breakdown voltage value (BDV).

The CM practiced methods for measuring the BDV is through the application of either ASTM D877 or the D1816 standards. For new unused oil samples, the D877 standard subjects two front-flat cylindrical electrode disks (with 2.5mm separation gap) to an increasing high voltage stress inside the oil medium. Measurement of the BDV occurs when the oil insulation breakdowns and arcing occur. The D1816 standard follows the same experimental procedures except for the use of pre-used transformer oil, and applying the voltage stress on spherical electrodes (with 1-2mm separation gap) instead of flat ones. Other tests such as the breakdown impulse voltage tests are used to test the quality of oil insulation against transient conditions such as lightning or load switching [8]. Measuring the BDV is required to deduce the degree of contamination of the oil [9]. The presence of acidic compounds and free water significantly reduces the dielectric strength of the oil. Solid containments in the presence of high levels of dissolved water in the insulating oil can further reduce the BDV strength of the oil insulation.

2.2.2.2. *Acidity.* Mineral oil oxidation is accelerated in the presence of atmospheric oxygen and the copper element in the transformer winding. The oxidation process results in the production of acidic compounds that contaminate the oil. In the presence of water and other contaminants, acidic compounds can be very harmful to the oil dielectric property and may result in lowering the breakdown voltage values [9]. Moreover, acidic products can cause the corrosion of the metallic components such as the transformer windings or tank. Increasing acidity indicates the ageing of the transformer oil. CM methods for measuring acidity are simple and are based on basic chemistry-neutralization methods. The test standard used is the ASTM D947 for using potassium hydroxide in neutralizing 1g of the transformer oil sample [8].

2.2.2.3. ***Water content.*** The presence of water particles in the transformer oil can be produced as a by-product of the degraded insulation material (due to thermal or electrical stress) or from the atmosphere (related to weak ingress protection) [9]. Water is a main constituent that can reduce the dielectric strength and lower the BDV of transformer oil. It can be found in the form of free water particles (clouded oil) or dissolved in oil. Dissolved water is acceptable at very low concentrations. Increasing concentrations of dissolved water indicate a high water solubility value which is mainly caused at high oil temperature. High concentrations of water can enhance the formation of acidic products in the oxidation process. The followed CM test standard is the ASTM D1533 Karl Fischer method, which relies on basic titration procedures for detecting concentration of moisture in oil [8].

2.2.2.4. ***Interfacial tension (IFT).*** The strength of the tension force at the surface boundary between oil's organic compounds and other fluids is measured by means of IFT. Having high measurements of IFT is a good indication of the preserved chemical properties of mineral oil [9]. As the transformer ages with time, soluble polar molecules are formed due to factors related to oil acidity and oxidation. The organic oil compounds lose their non-polar property in the presence of such polar contaminants and result in the formation of new oxidized contaminants or sludge. Observation of sludge material and other solid contaminants in the transformer oil is an indication of low IFT values. The practiced CM test methods used in measuring IFT are ASTM D-971 and ASTM D-2285 standards. The D971 standard measures the amount of force required to break the interface between oil using metallic platinum rings, while the D2285 measures the required volume of a water drop that the oil can withstand before the tension at the surface breaks [8]. Measuring the IFT using these standards requires the use of specifically designed testing equipment. Novel CM methods (addressed later in this thesis) are being developed in order to reduce the cost of IFT measurement for assessing the condition of the transformer [10].

2.2.2.5. *Dielectric dissipation factor and insulation resistance*. The transformer oil acts as a medium for cooling and insulating material. When considered as an insulating medium between two points at the transformer's winding surface, the transformer oil is modeled as a capacitor. Under ideal conditions, the resistivity of the oil is considered infinitely high giving an exact 90° phase shift between the capacitive current and the voltage. In real situations, the oil resistivity value lowers with time, causing a heat dissipating resistive current to pass through the oil that reduces the phase difference to a value less than 90°. This is identified as the loss in dielectric strength of the transformer oil caused by a leakage current. The DDF is the trigonometric tan of angle difference between the 90° angle and the new phase angle [9]. The DDF can also be computed as a ratio of the real to reactive leakage current. The dissipated heat will hasten the breakdown of the oil insulation material. Low oil resistivity indicates the presence of polar contaminant substances, acidic oxides and water content [10]. Reduction in the oil resistivity and increasing DDF measurements are indications of weakening the dielectric strength of the transformer oil. The CM method for measuring the DDF is done through the ASTM D924 standards in which a current is passed through the oil test sample in specifically designed cells to measure the capacitive and resistive components by the application of calibrated capacitive and resistive bridge circuits [8].

2.2.2.6. *Color.* The transformer oil in a new condition is clear and transparent. Changes in the color occur with ageing due to the formation of contaminants, sludge, free water and other insoluble products [9]. This will cause darkening of the oil and increasing the intensity of color. ASTM D1500 CM test provides a variety of color samples that can indicate the degree of degradation in the transformer oil. Color tests are generally good indicatives of the precise condition of the transformer oil, but can only provide a general understanding of its status [8].

Having an understanding of the standard CM testing techniques, the utility can conduct CA studies on the oil quality to interpret the transformer's operational conditions. The IEEE Std. C57-106 code defines the CA method through a set of bounded limits for the condition parameters. Oil parameters, such as IFT or water content, should fall within the acceptable range defined by the standard. Based on the produced outcome of the CA method, the oil type will be classified as a class I, II or III. A class I oil type is considered in a satisfactory condition for further use.

Alternatively, oil is classified as a class II if the water content and BDV requirements are not met. Class II oil can be reconditioned to retain its insulation strength. Finally, class III oil type is characterized by low IFT measures, high DDF values and high acidity. Class III oil is a very poor oil that requires the utility to either conduct intense reconditioning or replacement [8]. The practiced CA method for oil quality gives a good understanding of the transformer oil insulation weakness points. However, it cannot give an overall assessment of the transformer insulation system.

Table 2 shows the common recommended limits for oil quality parameters based on the IEEE Std. C57-106 standards.

Table 2: Recommended oil quality limits [8]

CM Test	*Value for Voltage Class*		
	≤69kV	*Between 69kV and 230kV*	*≥230kV*
Minimum Dielectric Strength (kV) for 1mm and 2mm gap	23 40	28 47	30 50
Maximum DDF (at 25°C)	0.5	0.5	0.5
Minimum IFT (mN/m)	25	30	32
Maximum Acidity (mg KOH/g)	0.2	0.15	0.10
Maximum Water Content (ppm)	35	25	20

2.2.3. **Furan concentration and degree of polymerization.** Degradation of the transformers paper insulation mainly occurs due to thermal, chemical and electrical stress. This is due to overload and short circuits that can dissipate a massive amount of heat through the transformer winding. The transfer of heat from the windings can cause winding hot spots. Hot spots due to the non-uniform distribution of heat in the transformer windings can rapidly increase the degradation rate of the surrounding paper insulation. Degraded paper is characterized by the loss of the mechanical tensile strength and electrical insulation property. Kraft paper is normally used in the solid insulation system of the transformer winding. The Kraft paper material is made up of cellulose. Degradation of the cellulose material is accelerated through the elements of heat, water and acidity. This is done through a thermo-chemical breakdown reaction known as pyrolysis. Pyrolysis of the glucose units in the cellulose chain produces water and gaseous by-products which are namely Hydrogen, Methane, Carbon Dioxide and Carbon Monoxide [11], [12]. Another cellulose degraded by-product formed in the Pyrolysis of glucose is 2-Furfuraldehyde (2FAL) or Furan [5]. Testing the presence of Furan in oil is the main CM test method used to interpret the strength of the paper insulation. The ASTM D5837 standard indicates the detection of Furanic compounds through the use of High-Performance Liquid Chromatography (HLPC).

Another method for measuring the extent of solid insulation degradation is through the measurement of Degree of Polymerization (DP). DP is a way of expressing the number of B-Glucose macro-molecular units which are still attached to the cellulose chain for a given volume of insulation material. High DP measurement indicates strong tensile strength of the paper material and its insulating capabilities. DP measurement is done by dividing the number-average molecular weight of the glucose polymer to the molecular weight of a single glucose-monomer unit. The DP of new paper material is around 1300 units, which drops to 200 units for aged oil. Related work has shown the possibility of an existence of negative correlation between the DP measurement and the concentration of dissolved Furanic compounds [11]. Therefore, two CA methods can indicate the paper insulation condition based on the outcome of the CM tests. One is through the use of DGA to detect the presence of high levels of Carbon Dioxide and Carbon Monoxide or through the use of the key gas ratio CO_2 /CO. The second method would be through the analysis of Furan concentrations and DP units. Table 3 shows the extent of paper insulation degradation as a function of Furanic concentration and DP [2].

Table 3: Inference of degradation using DP and 2-FAL concentration [2]

2-FAL (ppm)	*DP*	*Extent of Degradation*
0-0.1	800-1200	Insignificant
0.1-0.5	700-550	Significant
1.0-2.0	550-450	Cause of Concern
>10	<300	End of Life

2.2.4. **Assessment of other transformer components.** As mentioned earlier, the health of a transformer is mainly based on the strength of its oil-paper insulation system. Nevertheless, the health of other components of the transformer should also be taken into consideration to have a complete diagnostic CA procedure. One of the main transformer components that is vulnerable to damage is the LTC. The LTC is used to allow the transformer to vary its output voltage level by varying the winding ratio, without any interruption of the current. The CM methods for assessing the LTC condition include obtaining the data for the LTC oil dissolved gases, oil quality and contact resistance of the tap changer. CA for the LTC can be done by analyzing the collected data through DGA or OQA [26]. Moreover, the cooling efficiency of the transformer can be tested with Infrared Thermography which can indicate cooling problems due to overheated components such as the transformer windings or bushing. Other measurements include the rated leakage reactance using frequency response analysis (FRA) CA method, which is an indication of the possible winding deformation [1]. In addition, Core-to-Ground insulation tests can be done to indicate any loose connection that can lead to core grounding. Additional tests such as those for the turn's ratio can indicate the insulation failure between the windings of the same coil. A winding resistance test can detect loose connections or broken conductor strands. Many other CM tests and CA outcomes can be produced from analyzing different components of the transformer.

Figure 2 summarizes the standard transformer CA methods which are mainly used in assessing the health of the transformer.

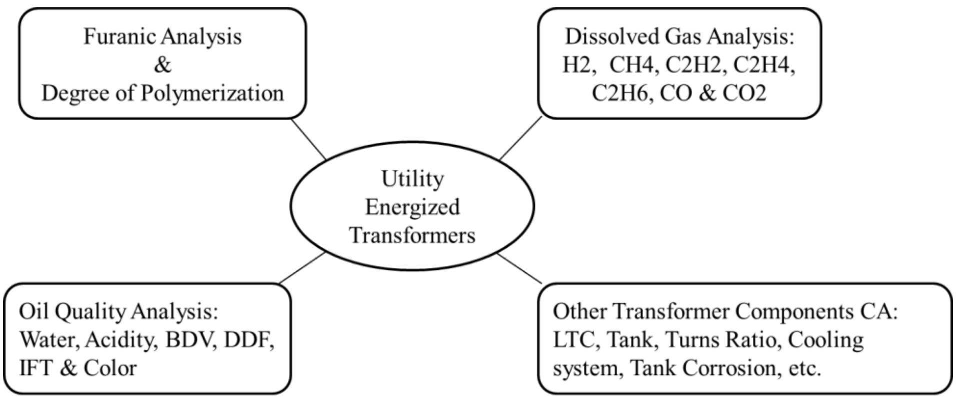

Figure 2: Standard transformer CA methods

2.3. Condition Monitoring and Assessment Using Artificial Intelligence (AI)

Proper data acquisition of the CM parameters can be done through any of the previously discussed standard techniques. Standards set by the IEEE or IEC codes are used for proper CA of the transformer using the acquired condition parameter data. However, some of the standard CM methods are considered expensive and hard to conduct. For example, measurements of the IFT number and Furanic concentration levels in the transformer oil are considered costly in terms of the required equipment and expertise. This was a strong initiative for researches to study alternative data-mining based techniques that can reduce the cost of CM maintenance applications.

2.3.1. **Dissolved Gas Analysis (DGA).** The practiced CM methods for acquiring the dissolved key gases composition in oil are based on chromatographic methods which are considered cheap and easily performed methods. AI methods have been used to predict the dissolved key gases composition as an alternative to the standard diagnostic CM methods. In [13], an ANN approach was used to predict the dissolved key gases in the transformer's oil using oil quality parameters which are acidity, BDV, water content, IFT, density and the power factor. Prior to using ANN for predicting any key gas, exhaustive feature selection techniques have been used to identify the oil quality parameters of highest statistical significance in predicting each key gas. For example, oil quality parameters of BDV, IFT and water content are considered to be the input features of highest statistical significance in predicting the concentration of Hydrogen. Based on 140 training and 51 testing oil samples of known dissolved key gas concentration, 96-100% of accuracy level is recorded for predicting the concentration of Hydrogen, Methane, Ethylene, Acetylene, Ethane, Carbon Dioxide and Carbon Monoxide in the transformers' oil.

The major contribution of AI-based research in the area of CA using DGA can be seen in developing new alternative incipient-fault interpretation methods with a higher level of decision accuracy. An approach has been used in [14], where a cascade of fuzzy logic models is used to predict the incipient transformer fault type. The input to the overall model is the concentration of the dissolved key gases. Each fuzzy model predicts the fault type based on a standard diagnostic method such as the Duval triangle or Doerenburg method. Based on the testing outcomes of each individual model, a weight is assigned to the output accordingly. The final interpretation of the incipient fault type is based on calculating the cumulative decision factor of all the individual fuzzy model outcomes. The overall model was successfully tested and verified against 70 transformer oil samples of known gas concentrations and associated fault types. In related works, an ANN approach has been used in [15]- [16] for interpreting the transformer incipient fault type using key gas elements as input features. In [16], the ANN model was used to classify the fault type for a given oil sample based on key gas ratios as input features (as per the IEC 60599 standards). Other works predict the fault type using Support Vector Machine (SVM) models based on the key gas ratio methods [17]. The first SVM model classified the transformer as faulty or not. Subsequent cascaded models classified the faulty transformers as one of the four internal transformer faults, i.e. low thermal, high thermal, low electrical discharge or high arcing fault.

2.3.2. **Oil Quality Analysis (OQA).** OQA is a CA method which relies on the composition of certain substances and chemicals in the transformer oil. However, researchers have done extensive studies to find AI alternatives for the standard diagnostic methods used for determining the oil quality tests. The methodology used by many researches was to utilize the easily conducted condition parameters for predicting the more expensive tests. In [18] a multi-stage ANN approach was used to predict the oil's BDV, IFT, water content and acidity using only the insulation resistance (IR) measurements. The idea was based on designing a series of cascaded feed-forward ANN stages with a back-propagation learning algorithm, in which each stage produced a particular oil feature output that was forwarded to the next network stage to predict another oil feature. The choice of the input features and corresponding output oil feature was based on the direct correlation of the features with each other. IR was used for predicting BDV and IFT values that were used along with oil color to predict the water content. A success rate of 84%, 95% and 75% was accomplished for predicting BDV, IFT and acidity respectively. A similar approach for predicting BDV and water content using IR was used in [19]. The prediction of BDV and water content was done through a single-hidden layer feed-forward ANN, with the success accuracy of 95% and 83% for BDV and water content prediction respectively. The objective in both [18] and [19] was to create a workable platform of assessing the transformer condition using only IR as an input feature, and thus reducing the CM maintenance cost. In [20], prediction of BDV, IFT and water content using IR data was done through the application of polynomial classifiers. A success rate of 84% and 93% was obtained for prediction of BDV and Interfacial tension respectively due to their high correlation with IR. Poor prediction of water content was obtained as a result of a temperature variation and solubility levels in the oil samples. The drawback in [19] and [20] was the limited number of available oil transformer samples.

A recent novel approach has been used for predicting IFT through the absorbance of light and using fuzzy logic [21]. An oil sample with degraded organic insulation compounds has the potential of absorbing light in the Ultra Violet-Visible wavelength spectrum. This property can be detected by the use of absorption spectroscopy which analyzes the absorption of light in any medium with respect to the change in wavelength. In [21], the wavelength of the incident light is changed from 200-1,100nm and the absorption spectral response of the light in different oil samples was observed. The lower the IFT number (poor quality), the higher is the concentration of the insulation degraded organic compounds. This results in a higher wavelength range of light absorption and higher energy absorbance peak. In the fuzzy logic approach, the absorbance peak and wavelength range were used as condition parameters for setting the If-Then rules to predict the IFT number as an output. All of the oil samples IFT numbers were previously known using standard CM methods (training and testing). The accuracy rate for predicting IFT for new 15 oil samples was 87%. ANN was also used with the same spectral inputs to yield an accuracy rate of 80% for the same test samples. Another application of fuzzy logic in AI-based CA method was published in [22]. The fuzzy-based approach was used to develop a cascade of fuzzy models that can predict a transformer oil criticality based on the given inputs. The models were built based on a wide range of CM diagnostic tests of transformers with different operational conditions. Interpretation of the estimated lifetime and transformer condition was done on each transformer oil sample by utility experts. One fuzzy model predicts the electrical quality of the oil based on the power factor and BDV. Another model predicts the physical quality of the oil using IFT and UV spectral analysis. The physical and electrical oil qualities are forwarded to a new fuzzy model to predict the overall oil criticality. The success of the method was verified with correctly identifying the weaknesses of the transformer with the possible internal faults.

2.3.3. **Furanic Content in Oil Analysis (FFA).** CM methods for data acquisition of the Furanic content in oil are considered relatively expensive and costly compared to the other tests. Furan data is normally acquired through outsourced companies that have the proper experimental facilities and equipment. Over the last decade, there have been multiple approaches to estimate Furan concentration through the application of AI. In [23] ANN was used to predict Furan content using other relatively easier measured condition parameters as input features. The approach was to selectively use the features of highest statistical significance in Furan prediction using step-wise regression. Selection of Carbon Dioxide, acidity and BDV for Furan prediction produced a mean accuracy precision rate (MAPE) of 85%. The study validated the correlation between Carbon Dioxide, acidity and BDV with Furan. The study proceeds with adding other features such as Carbon Monoxide which resulted in increasing the MAPE to 90%. Similarly in [24], Furan content in oil was classified into five standard grades using other oil condition parameters as input features. This was done using the decision tree algorithm, ANN, Support Vector Machine, KNN and Naïve Bayes. The classification rates were poor due to the imbalance of the number of sample in each grade class. Synthetic Minority Over-Sampling Technique (SMOTE) was used to solve the imbalance problem. In addition, the number of grade classes was reduced to three. These measures increased the recognition rate from 73%, 68% and 58% to 80%, 74% and 77% using decision tree, ANN and KNN respectively. In other work [25], Furan prediction was done through the application of fuzzy logic. Similar to what has been done with IFT, an approach of using the light spectral response of the oil samples due to Furan content was used. Furan is an organic by-product of paper degradation that absorbs light energy in the Ultra Violet-Visible wavelength range. A positive correlation was indicated between the Furan content and the maximum absorbance peak of light energy by the oil sample. Similarly, a positive correlation was also found between the Furan levels and the range of spectral wavelengths where the energy is absorbed.

2.4. Transformer Health Index as a CA Method

2.4.1. **Concept and objectives.** TAM is concerned with defining a set of strategies for properly managing and maintaining a population of transformers in service. This is normally done through the diagnostic CM methods, which can assess the operational conditions and the insulation strength of the transformers based on the CA techniques. The standard CA insulation techniques are namely the DGA, OQA and FFA. The strength of the other transformer condition parameters such as the LTC, winding resistance, leakage reactance, bushing etc., are also assessed through defined CA methods which are set by the IEEE and IEC codes. Each test parameter can highlight certain problems with the transformer health condition. For example, the

DGA can provide information about the fault type and possible transformer weakness points. OQA and FFA methods suffer from a similar limitation and thus cannot provide any completely informative feedback that can be used for an effective TAM plans. Thus, the problem with these individual CA methods is the fact that they can't give an overall understanding of the condition of the transformer unless those test parameters are integrated together in one index called the transformer health index (HI).

The transformer HI is a cumulative index value that represents a combination of the transformer operation condition outcomes produced by laboratory experimentations, code standards, site observations and expert judgments [26]. The HI method allows utility companies to input the outcomes of the standard CA methods into one model to give an index value that can be used as a reference for the condition of the transformer. The HI solves the problem of taking all the CA outcomes into consideration and allows for new inputs that are related to the transformer operation (loading and maintenance) to be used in the HI model. Moreover, the HI model design is based on a comprehensive set of international standards and expert experience that allows for generalized conclusions that are applicable in any region.

The main objective of using the HI is to enhance the understanding of the transformer's probability of failure, effective age and its remaining life. The HI provides a threshold based criterion which allows the utility to classify the transformer condition from being in a very poor to an excellent state of operation. This allows for a full understanding of the condition of the transformers in service, and therefore allows for prioritizing the transformers maintenance plans [26].

2.4.2. **Computation of the Health Index in industry.** Several computational techniques have been developed for calculating the HI of a transformer [26], [27]. These techniques have been developed by utility experts who work in asset management industry. Computation of the HI in industry is based on analyzing the condition of the individual transformer tests. These parameters can be either internal or external. Internal parameters are associated with the transformer oil-paper insulation systems, LTC, winding resistance and other condition parameters. Part of the external parameters includes the transformer's loading history and frequency of maintenance orders. Precisely, the computation of the HI is based on three major parts which are namely the insulation strength, service record and design.

The insulation strength is concerned with the mechanical and electrical strength of the transformer oil-paper insulation system. Analysis of the insulation strength can be done through the standard CA methods identified earlier. The strength of the oil insulation system can be evaluated through the OQA methods. The HI model requires the computation of the oil quality as a cumulative factor whose components are the oil's BDV, IFT, acidity, water content, color and DDF. A score is given to each of these components that represents the condition of the transformer oil. A weight value is then used along with the score to indicate the significance of that specific test in calculating the oil quality factor. The standards used in defining the scores and weights are based on the IEEE and IEC codes. In a similar manner, the DGA factor is computed with its components being the key gas elements. The score and weights are set after going through the IEC, IEEE and Bureau of Reclamation standards and Dorenburg's method. The strength of the paper insulation is assessed through measurement of the Furanic concentration and assessing its condition using the FFA factor.

The service record factor comprises of data related to the transformer age, fault history, loading and maintenance history. Such information is important to add a qualitative understanding of how the transformer was operated and maintained within its entire duration of service. Transient faults such as lightning and load switching reduce the life time expectancy of the transformer. For example, transformers operating in regions of a high likelihood of lightning (North America for example) should be dealt with differently than transformers in other regions. In addition, understanding the extent and duration of overloading can further enhance the knowledge of the life-expectancy of the transformer to understand the frequency of possible internal thermal faults. Finally, the maintenance history represents a record of the number of work orders that have been done for several parts of the transformers such as the connectors or bushing. The condition of the transformer is a function of the frequency of the work order and type. The service record factor allows to quantify all of these qualitative conclusions and observations to further support the HI model.

Finally, the design factor is an understanding of the manufacturer of the transformer and country of origin. This allows for a higher degree of freedom by allowing the HI model to use the manufacturer as an input. An expensive transformer made by well-established manufacturing companies has a higher life expectancy than cheaper transformers from low-profile manufacturers. Using the design of a transformer in the HI model can be omitted if all the concerned transformer assets in the utility are of the same manufacturer and origin.

The computation of the HI by different TAM based companies is similar with the exception of minor factors. As an example, the companies mentioned in [26] and [27] use a similar computational strategy in which the DGA, OQA and FFA are used as main components of the HI model. However, [26] takes into account the health condition of the other parameters that are not related to the insulation strength and thus is considered of a higher accuracy than the method used in [27]. Moreover, the weights and scores of the components vary based on different interpretations by utility experts.

Figure 3 illustrates the general computation of the HI in industry as a function of the transformer's insulation strength, service record and design.

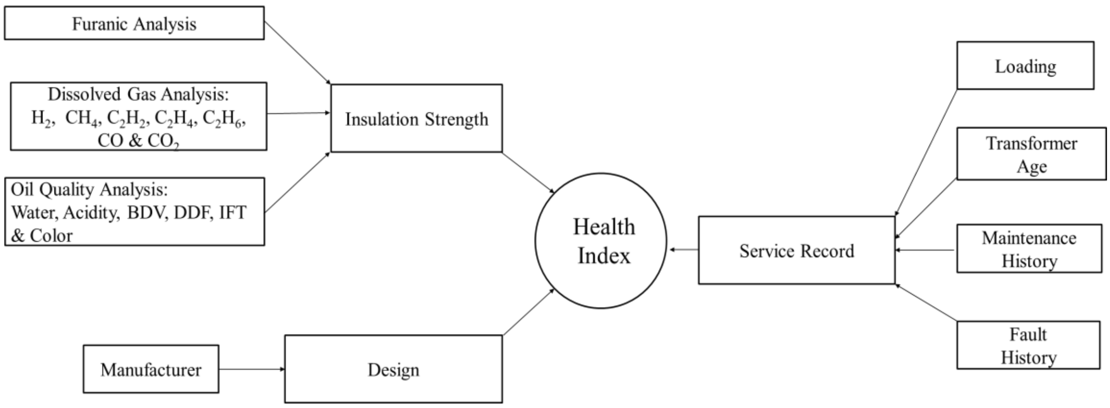

Figure 3: General computation of the Health Index using industry standards

2.4.3. **Computation of the Health Index using AI.** The drawback of using the industry based HI model as a CA method lies in the fact that the accuracy of the produced HI is a function of the number of given inputs. This adds a cost constraint of time and money to acquire these tests. Moreover, Furan and IFT are considered highly expensive test features [21, 25] that require the proper equipment, personnel and testing facility. These tests are considered highly important in the computation of the HI, and thus add an extra cost to the utility company.

In [28], a fuzzy based approach has been conducted to predict the HI value using the oil quality, dissolved gas and Furan content parameters as inputs. Six membership functions are designed to represent the condition of water content, acidity, BDV, DDF, dissolved combustible gas and Furan in the oil samples respectively. Each membership function was built based on a defined IEEE standard. The output membership function is the HI value of the oil sample. The fuzzy model ran on a set of rules that have been defined by utility experts. Each rule determines the consequence output (HI) based on the condition of the six membership functions. The final HI output is the result of defuzzification (using centroid method) of the truncated sums of the outputs of each rule. The set of 33 rules used, allows for a more general model that takes many scenarios into account. The success of the method was tested by comparing the predicted HI and condition of the transformer oil sample using the fuzzy model with that produced using a TAM-company's own algorithm. The reported classification success rate was 97% based on a three class condition classification. Though this result is good for a three class HI classifier problem, no success rate has been provided for the actual five class problem and no MAPE has been presented for predicting the HI values. Related work in [29] used IFT and oil acidity in a fuzzy logic model to predict the remnant percentage lifetime of the transformer based on a 40-year lifetime span. Assessing the validity of the model was done by comparing the predicted life-time results with the actual life-time expectation based on correlation measures with IFT and acidity.

In other related work [30], an ANN approach has been made to classify the condition of the transformer based on the predicted HI value. The input features used in this model are based on water content, acidity, BDV, Hydrogen, Methane, Ethane, Acetylene, Ethylene, Furan, DDF and color. The model was a feed-forward ANN with two hidden layers (four and two neurons respectively) that was trained on 67% of the available data. The pre-processing method of data normalization is used for all the input features. Based on the testing outcomes, 97% of the testing samples were correctly classified based on a three-class condition problem.

A novel method is presented in [31], where the authors apply a fuzzy-based SVM technique for classifying the condition of the transformer based on the HI of the oil samples (the five class problem). In this method, a generalized approach was followed in classifying the HI-condition of the training samples. This was done by combining a number of HI-condition model outcomes for a given oil quality, dissolved gas and Furan inputs. The individual HI models are based on industry standards, utility expert judgments, DGA interpretation and fuzzy membership functions (of each condition of the HI). A HI-condition for an oil sample was selected based on a majority vote of each of the individual HI models. A weight factor is introduced to eliminate the outlier samples (considered as noise) in the training database by assigning higher weights to the samples closer to the feature mean. This strengthens the model in the sense that it eliminates the possibility of faulty training of the SVM due to un-representative samples. Moreover, the pre-processing methods of oversampling, under sampling and SMOTE have been used to balance the training samples in each HI-condition class. The Fuzzy Support Vector Machine (FSVM) model was trained with 70% of the available samples. An average accuracy classification rate of 87.8% has been obtained based on a 10-trial training and testing procedure.

2.5. Objectives and Contributions of the Research

The main objective of TAM is to increase the reliability of the power system with efficient maintenance costs. To reach this objective, the utility company has to be fully aware of the overall operational condition of each transformer in the power system. The HI amongst other CA methods is considered the best practical tool that can provide a comprehensive meaningful quantity of the transformer condition and its remnant lifetime. However, the limitation associated with the standard HI computational methods (by industry standards) is the need of all input features for a high accuracy of HI prediction. From the literature review, some features such as the Furanic content and oil quality IFT are difficult to acquire. The difficulty lies in the expensive equipment and experienced personnel required to perform the laboratory CM testing procedure.

The objective of this thesis is to follow up with the AI-based approach that is developed by other researches in computing the HI. Precisely, this thesis presents an ANN Multi-Layer Perceptron approach that will predict the HI based on the oil-paper insulation characteristics. This objective will be accomplished with the objective of minimizing the input feature cost. Thus, a feature selection procedure will be followed based on the step-wise regression method to selectively remove the redundant features that have the least statistical significance on the HI outcome. This will reduce the cost of computing the HI and the overall TAM maintenance strategy. To complete the work and further reduce the TAM cost, this thesis aims at presenting a cost saving HI computational strategy where the utility can predict a costly input feature that can later be used in the main HI model.

In summary, three main outcomes and research contributions are presented in this work.

1. ANN MLP approach is used to compute the transformer HI based on the standard computational method indicated by [26]. Using step-wise regression, the required input oil-paper insulation features will be reduced and thus provide a cost-effective AI based HI model.
2. The designed HI model (using the data of one utility) will be tested on a new set of transformer oil data (of another utility) using only the reduced features to validate the generalization of the cost-effective HI model based on one industry computational method. Moreover, the reduced features will be generalized for building and testing the HI model using the data of any utility.
3. To further reduce the cost of the proposed HI calculation method, a cascaded ANN network will be used for predicting a particular feature in order to use it as an input in the final HI model. Precisely, one network will be used to predict the IFT feature using easily obtained oil-paper condition parameter inputs to pass it on the main HI ANN model.

Chapter 3. Materials and Methods

3.1. Transformer Oil Samples

The main required step towards building the HI predictor model lies in having a training database of sufficient number of transformer oil samples. A high number of available samples can insure a better performance of the predictor model. The desired input to be used in designing the model should include all the features indicated by [26]. However, obtaining all these features is difficult and impractical for most of the utilities. This is due to the fact that most of the utility companies are using the recognized standards of the CM and CA methods (DGA, OQA and FFA) in the TAM strategy. The concept of the HI has been recently developed and not common to a high percentage of the utility companies worldwide. In other words, the available data used in the transformer CA are the standard condition parameters of the oil-paper insulation system. This is attributed to the importance of the insulation system parameters in assessing the transformer condition. Nevertheless, the HI method has an advantage over the standard CA methods even with only having the oil-paper insulation parameters. This is again due to the fact that the HI method can comprehensively incorporate the outcomes of all the standard CA methods in one number representing the transformer health condition.

Hereinafter, all the features that will be used are those related to the dissolved key gases concentration, oil quality parameters and Furanic content. Two sets of transformer oil-paper insulation features have been acquired from two different utility companies that are named UTILA and UTILB. UTILA transformer oil samples are 66/11kV transformers of power ratings that range from 12.5 to 40MVA. While UTILB transformer oil samples are 33/11kV transformers of 15MVA power rating. The condition parameters (henceforth, feature inputs) acquired by one utility varies more than the other in terms of the conducted tests. UTILA conducted the complete CM tests to include: the seven key gases for the DGA, the six oil quality parameters for OQA, and the Furanic content for FFA. UTILB, on the other hand, have neither conducted the dissolved gases tests for Carbon Dioxide nor Carbon Monoxide. In addition, they have neither conducted the color nor the DDF tests. Since both the number of oil samples and input features are outnumbered by UTILA, the main HI model will be based on UTILA's database. Subsets of the samples acquired by UTILA and UTILB are shown in Table 4 and Table 5, respectively.

Table 4: Subset of UTILA data set

	UTILA Transformer Oil Samples						
	1	*2*	*3*	*4*	*5*	*6*	*7*
Hydrogen (ppm)	15.6	19.9	7.9	9.6	19.2	27.1	18.7
Methane (ppm)	14.4	15.6	0	2.8	2.4	236.8	2.5
Ethane (ppm)	2.5	3	0.7	7.1	2.3	202.1	41.4
Ethylene (ppm)	0.3	0.4	0.5	20.6	6.5	628.1	55
Acetylene (ppm)	0	0	0	0	34.8	27.7	58.2
Carbon Monoxide (ppm)	1,019.6	1,240.8	132.5	126.4	237.6	230.5	165.8
Carbon Dioxide (ppm)	2619.3	3,323.2	1,770.7	2,126.5	3,654.1	4,810.7	5,123.8
Water (ppm)	15.6	19.9	7.9	9.6	19.2	27.1	18.7
Acid (mgKOH/g)	14.4	15.6	0	2.8	2.4	236.8	2.5
BDV (kV)	2.5	3	0.7	7.1	2.3	202.1	41.4
DDF (25°C to 50Hz)	0.3	0.4	0.5	20.6	6.5	628.1	55
Color	0	0	0	0	34.8	27.7	58.2
IFT (mN/m)	1,019.6	1,240.8	132.5	126.4	237.6	230.5	165.8
Furan (ppm)	2,619.3	3,323.2	1,770.7	2,126.5	3,654.1	4,810.7	5,123.8

Table 5: Subset of UTILB data set

	UTILA Transformer Oil Samples						
	1	*2*	*3*	*4*	*5*	*6*	*7*
Hydrogen (ppm)	0	9	4	17	30	34	123
Methane (ppm)	6	8	4	1	13	60	148
Ethane (ppm)	0	1	2	0	4	66	138
Ethylene (ppm)	0	0	2	3	10	107	16
Acetylene (ppm)	0	0	0	0	25	1	0
Water (ppm)	4	8	10	18	27	17	41
BDV (kV)	64.4	55.3	39.4	78.8	26.2	36	23.5
Acid (mgKOH/g)	0.011	0.05	0.024	0.425	0.107	0.065	0.468
IFT (mN/m)	39.3	24.1	32.3	15.6	33.3	23.2	14.4
Furan (ppm)	0.009	0.29	0.62	6.09	1.93	0.76	22.6

The data set of Table 4 will be used as an example for the HI computation in the following section. Shown in Table 6 and Table 7, are the statistical parameters of the data sets of UTILA and UTILB respectively.

3.2. Computation of the HI (Industry Standards)

The aim of this work is to present a practical AI-based tool that can be used by utility companies to compute the HI of a transformer. The practiced exercise of computing the HI for a population of transformer is done through hiring specialized transformer CA companies. The hired company follows its own set of procedures for computing the HI, with reference to the standards that are set by the technical professional organizations such as the IEEE and CIGRE. A professional CA company would normally conduct the desired CM tests on the transformer's insulation system, and request the utility for additional information such as the loading conditions and the maintenance history. In this work, the HI for the available data will be computed through a practiced computational method which is exercised by one professional transformer CA company [26]. The computed HI will be based on the dissolved gas, oil quality and Furan content data only, which is similar to what has been done in other works such as in [31].

Computation of the HI will be based on calculating a set of three cumulative factors that represents the transformer condition based on DGA, OQA and FFA respectively. The HI value will be the cumulative calculation of each of the three factors with respect to their weights or significance towards the transformer's overall insulation health.

Table 6: Statistical parameters of UTILA data set

	Average	*Median*	*Minimum*	*Maximum*	*Variance*
Hydrogen (ppm)	27.24	19.7	1.9	605	1,020.63
Methane (ppm)	18.96	10.3	0	298.1	888.55
Ethane (ppm)	19.52	3.3	0	339.4	1,917.39
Ethylene (ppm)	6.47	1.9	0	628.1	945.67
Acetylene (ppm)	1.68	0	0	64.3	51.97
Carbon Monoxide (ppm)	501.83	440.6	22.2	1,621.9	92,073.6
Carbon Dioxide (ppm)	3,584.71	2,577.95	188.5	113,883	2.4e7
Water (ppm)	5.96	5	1	32	21.54
Acid (mgKOH/g)	0.02	0.005	0.005	0.261	1.18e-3
BDV (kV)	74.18	77	16	99.5	229.72
DDF (25°C to 50Hz)	8.77e-4	0	0	0.015	2.05e-6
Color	0.68	0	0	4	1.21
IFT (mN/m)	30.9	32	13	43	50.83
Furan (ppm)	0.49	0.01	0.001	11.03	1.43

Table 7: Statistical parameters of UTILB data set

	Average	*Median*	*Minimum*	*Maximum*	*Variance*
Hydrogen (ppm)	4.57	4	0	123	76.18
Methane (ppm)	16.03	8	0	702	1,849.16
Ethane (ppm)	25.30	1	0	1,096	7,959.04
Ethylene (ppm)	7.77	1	0	1,384	5,884.69
Acetylene (ppm)	0.48	0	0	25	6.54
Water (ppm)	6.85	5	2	41	23.25
BDV (kV)	62.55	67.8	12.5	95.3	281.97
Acid (mgKOH/g)	0.04	0.016	0.0099	0.471	0.005
IFT (mN/m)	34.09	35.2	9.2	75.1	72.20
Furan (ppm)	0.46	0.09	0.01	22.6	2.81

3.2.1. **Dissolved Gas Analysis Factor (DGAF).** The DGAF is a representative factor of the oil-paper insulation condition from the dissolved gases perspective. The DGAF is an integer score value from zero to four, which indicates the degree of danger or degradation of the oil-paper system based on DGA. A score of zero indicates a poor DGA condition, while a score of four indicates a healthy DGA condition. In order to calculate the DGAF, a scoring system for each of the seven key gas elements is used to indicate the extent of the respective dissolved gases in the oil sample. Table 8 indicates the score system which is used for the key gas elements. Based on the score value, the DGAF is calculated using [26]

$$DGAF = \frac{\sum_{i=1}^{7} S_i \times W_i}{\sum_{i=1}^{7} W_i} \quad (1)$$

where S_i is the score outcome of each of the seven key gas elements based on Table 8, and W_i is the key gas associated weight or significance factor. The DGAF will be a real positive number that will be converted to an integer value from zero to four based on the DGAF scoring system shown in Table 9. As an example, the DGAF for the key gas elements of the samples, shown in Table 4 (from UTILA), will be computed as per the explained procedure. Table 10 shows the obtained DGAF values for these samples.

Table 8: DGAF score and weight system [26]

Hydrogen		*Methane*		*Ethane*		*Ethylene*		*Acetylene*		*Carbon Monoxide*		*Carbon Dioxide*	
$W_1=2$		$W_2=3$		$W_3=3$		$W_4=3$		$W_5=5$		$W_6=1$		$W_7=1$	
ppm	S_1	*ppm*	S_2	*ppm*	S_3	*ppm*	S_4	*ppm*	S_5	*ppm*	S_6	*ppm*	S_7
<155	1	<103	1	<92.5	1	<75	1	<5	1	<500	1	<2,750	1
<225	2	<145	2	<95.5	2	<85	2	<15	2	<850	2	<3,500	2
<365	3	<240	3	<96.5	3	<95	3	<25	3	<1,050	3	<4,500	3
<585	4	<400	4	<97.5	4	<105	4	<35	4	<1,250	4	<6,000	4
<700	5	<600	5	<100	5	<130	5	<60	5	<1,400	5	<7,000	5
>700	6	>600	6	>100	6	>130	6	>60	6	>1,400	6	>7,000	6

Table 9: DGAF final scoring system [26]

DGAF Calculated	*DGAF Final*
< 1.2	4
< 1.5	3
< 2	2
< 3	1
≥ 3	0

Table 10: Computed DGAF for UTILA Data Subset

Sample	*Hydrogen*		*Methane*		*Ethane*		*Ethylene*		*Acetylene*		*Carbon Monoxide*		*Carbon Dioxide*		*DGAF Calc.*	*DGAF Final*
	ppm	S_1	*ppm*	S_2	*ppm*	S_3	*ppm*	S_4	*ppm*	S_5	*ppm*	S_6	*ppm*	S_7		
1	15.6	1	14.4	1	2.5	1	0.3	1	0	1	1019.6	3	2619.3	1	1.11	4
2	19.9	1	15.6	1	3	1	0.4	1	0	1	1240.8	4	3323.2	2	1.22	3
3	7.9	1	0	1	0.7	1	0.5	1	0	1	132.5	1	1770.7	1	1.00	4
4	9.6	1	2.8	1	7.1	1	20.6	1	0	1	126.4	1	2126.5	1	1.00	4
5	19.2	1	2.4	1	2.3	1	6.5	1	34.8	4	237.6	1	3654.1	3	1.94	2
6	27.1	1	236.8	3	202.1	6	628.1	6	27.7	4	230.5	1	4810.7	4	4.00	0
7	18.7	1	2.5	1	41.4	1	55	1	58.2	5	165.8	1	5123.8	4	2.28	1

3.2.2. **Oil Quality Factor (OQF).** The OQF represents the health condition of the transformer oil. The OQF value is also an integer value which ranges from zero to four. An OQF of four indicates an excellent level of oil quality, while an OQF of zero indicates a very poor level of oil quality. Computation of the OQF is based on the scores and weights of the six oil quality parameters respectively. The OQF calculation is done using [26]

$$OQF = \frac{\sum_{i=1}^{6} S_i \times W_i}{\sum_{i=1}^{6} W_i} \tag{2}$$

Table 11 indicates the weights (W_i) and score values (S_i) for the oil quality parameters. The OQF will be a real positive number that will be converted to an integer score from zero to four based on the OQF scoring system shown in Table 12. The computed OQF values for the UTILA data subset (of Table 4) is shown in Table 13.

Table 11: OQF score and weight system [26]

Water		*Acidity*		*BDV*		*DDF*		*Color*		*IFT*	
W_1=4		*W_2=1*		*W_3=3*		*W_4=3*		*W_5=2*		*W_6=2*	
ppm	*S_1*	*mgKOH/g*	*S_2*	*kV*	*S_3*	-	*S_4*	-	*S_5*	*mN/m*	*S_6*
≤30	1	≤0.05	1	≥45	1	≤0.1	1	≤1.5	1	≥25	1
≤35	2	≤0.1	2	>35	2	≤0.5	2	≤2	2	>20	2
<40	3	<0.2	3	>30	3	<1	3	<2.5	3	>15	3
≥40	4	≥0.2	4	≤30	4	≥1	4	≥2.5	4	≤15	4

Table 12: OQF final scoring system [26]

OQF Calculated	*OQF Final*
< 1.2	4
< 1.5	3
< 2	2
< 3	1
≥ 3	0

Table 13: Computed OQA for UTILA data subset

Sample	*Water*		*Acidity*		*BDV*		*DDF*		*Color*		*IFT*		*OQA*	*OQA*
	ppm	S_1	*mgKOH/g*	S_2	*kV*	S_3	*-*	S_4	*-*	S_5	*mN/m*	S_6	*Calc.*	*Final*
1	3	1	0.005	1	99	1	0	1	0	1	42	1	1.00	4
2	2	1	0.005	1	84	1	0	1	0	1	43	1	1.00	4
3	1	1	0.133	3	76	1	0.005	2	3	4	20	3	2.00	1
4	11	1	0.046	1	75	1	0.001	1	1	1	23	2	1.13	4
5	7	1	0.042	1	55	1	0.002	2	3	4	18	3	1.87	2
6	4	1	0.029	1	82	1	0.002	2	2	2	18	3	1.60	2
7	4	1	0.057	2	74	1	0.002	2	3	4	18	3	1.93	2

3.2.1. **Furan Factor (FFA).** One of the main numerical factors that indicates the extent of the paper insulation degradation is the FFA. Calculation of the FFA is based on the scoring system shown in Table 14. The FFA is an integer score value that ranges from zero to four. Similarly, a value of four indicates a healthy paper condition, while a value of zero indicated a very poor paper condition. The calculated FFA values for the UTILA data subset are shown in Table 15.

Table 14: FFA scoring system [26]

Furan - ppm	*FFA*
< 0.1	4
< 0.25	3
< 0.5	2
< 1	1
≥ 1	0

Table 15: FFA for UTILA data subset

Sample	*Furan*	*FFA Final*
	ppm	
1	0.01	4
2	0.01	4
3	0.28	2
4	4.45	0
5	0.54	1
6	2.18	0
7	2.73	0

3.2.2. **Final Health Index (HI) value.** Based on the final insulation parameter scores (DGAF, OQF and FFA), a cumulative calculation for the HI is done. Each factor is assigned with a weight value (W_{DGAF}, W_{OQF} and W_{FFA}) that indicates the significance of the factor in the overall health of the transformer.

The cumulative calculation for the transformer HI is done using the following formula [26]

$$HI\ (\%) = 100 \times \frac{(DGAF \times W_{DGAF}) + (OQF \times W_{OQF}) + (FFA \times W_{FFA})}{4(W_{DGAF} + W_{OQF} + W_{FFA})}\ \% \tag{3}$$

The final HI values for the UTILA data subsets are shown in Table 16. The overall HI computation is illustrated as a block diagram shown in Figure 4.

Table 16: Final HI for UTILA data subsets

Sample	*DGAF*	*OQF*	*FFA*	*HI (%)*
	$W_{DGAF} = 10$	$W_{OQF} = 6$	$W_{FFA} = 5$	
1	4	4	4	100.00
2	3	4	4	88.10
3	4	1	2	66.67
4	4	4	0	76.19
5	2	2	1	44.05
6	0	2	0	14.29
7	1	2	0	26.19

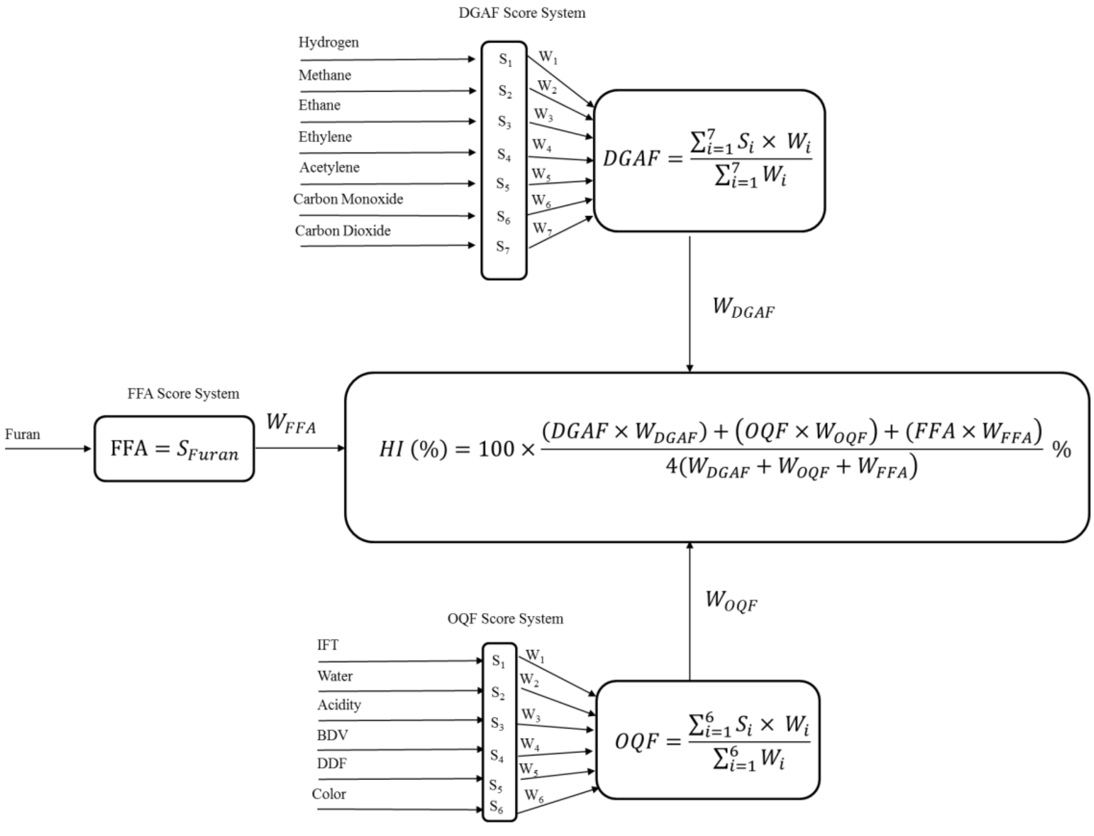

Figure 4: Overall HI computation using [26]

If an existing data set (such as that of UTILB) is missing certain insulation features, then the corresponding weight for these features is zero. Thus the features have no contribution in the relevant insulation factor and the final HI computation.

3.3. Artificial Neural Networks (ANN)

Pattern recognition methods have been used as AI techniques for identifying the response with respect to a given set of inputs. One of the well-known techniques used in pattern recognition is ANN. ANN systems are sophisticated responsive systems that mimic the neural behavior in the human body. Biologically, an environmental stimulus (input) triggers the transmission of electrical signals via neuron cells [32], [33]. The inter-connection links between the neuron cells are adjusted to produce a proper connection path. Triggered signals through the properly adjusted paths eventually result in the proper human response. Similarly, ANN consists of interconnected neuron units of adjusted weights to predict a response based on given input features. Modeling the ANN is similar to the human learning process. Based on a given input and a corresponding response output, the neuron units learn by continuous adjustment of the inter-connecting neural weights. The input to the ANN is a *d*-dimensional feature vector, in which each element represents a feature variable belonging to the input sample in question. Each feature element is fed to a corresponding input neuron, which is fed in the forward propagating direction towards the subsequent layer (hence a Feed-Forward mechanism). An ANN network can use the input features to predict one or more response outputs. Figure 5 shows a schematic diagram of the ANN system in terms of the neuron and layer components, where $X^{(i)}$ represents the i-th feature of the input feature vector [32]. W_{ij} and W_{ki} represent the link weights of the input-hidden and hidden-output neurons respectively. Z_k is the final response output that is produced from each neuron in the output layer. Each neuron in the hidden layer receives a net activation input from the *d* input neurons as [32]

$$net_j = \sum_{i=1}^{d} X_i W_{ji} + W_{j0} \tag{4}$$

where W_{j0} represents the link weight from the bias unit in the input layer [32]. Each neuron in the hidden layer has an activation function for the given net activation input. The output of the hidden neuron is

$$Y_j = f_H(net_j) \tag{5}$$

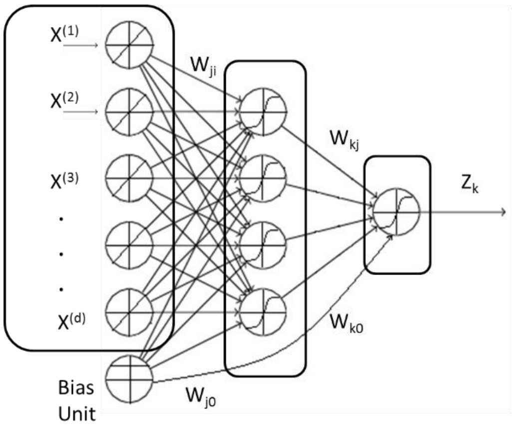

Figure 5: Schematic of a typical ANN network

The output Y_j is fed forward to the neurons of the output layer for computing the response variable [32]. Each neuron in the output layer receives a net activation input from the *L* hidden neurons bias W_{k0} as

$$net_k = \sum_{j=1}^{L} Y_i W_{kj} + W_{k0} \quad (6)$$

The response output is the computation of the activation function in the output neuron

$$Z_k = f_o(net_k) \quad (7)$$

Combining equations (4) to (7) will result in

$$Z_k = f_o(\sum_{j=1}^{L} f_H(\sum_{i=1}^{d} X_i W_{ji} + W_{j0})\, W_{kj} + W_{k0}) \quad (8)$$

For a given input feature vector and corresponding target output, a back-propagation learning process occurs where the weights of the neural links are continuously adjusted. The adjustment process occurs in order to minimize the training error constraint. For a given neural weights and corresponding response output Z_k [32], the training mean-square error function J for *M* samples is

$$J(W_{ji}, W_{kj}) = \frac{1}{2}\sum_{k=1}^{M}(\widehat{Z_k} - Z_k)^2 \quad (9)$$

where $\widehat{Z_k}$ is the desired response output for a given input. Thus, as it is apparent from equation (9), the training error in the ANN is a function of the neural weights. The method of gradient descent is used in order to solve the J error function as an optimization problem [32]. In any gradient descent problem, the gradient or the derivative of the optimized function is taken with respect to the changing variable. Since J is a function of the neural weights W_{ji} and W_{kj}, the partial derivative of the error function is

$$\frac{\partial J}{\partial W_{ji}} = -f_H'\left(net_j\right)X_i \sum_{k=1}^{L}(\widehat{Z_k} - Z_k)f_o'(net_k)W_{kj} \tag{10}$$

$$\frac{\partial J}{\partial W_{kj}} = -(\widehat{Z_k} - Z_k)f_o'(net_k)Y_i \tag{11}$$

The partial derivatives of equations (10) and (11) are used for the neural weight updates in the gradient descent problem [32].

$$W_{ji}{}^{x+1} = W_{ji}{}^{x} - \eta \frac{\partial J}{\partial W_{ji}} \tag{12}$$

$$W_{kj}{}^{x+1} = W_{kj}{}^{x} - \eta \frac{\partial J}{\partial W_{kj}} \tag{13}$$

where η is the combination coefficient that determines the step size of the gradient descent. The selection of the proper activation function depends on the relationship between the response variable and input features. For a non-linear complex relationship between the response and the input, a tan sigmoid function is used in the neurons of the hidden layer, while a linear transfer function is used in the neurons of the output layer as the activation functions.

3.4. Stepwise Regression

Stepwise regression is a feature selection tool that is commonly used to eliminate the redundant input features [34]. This is done by measuring the statistical significance of the input term in predicting the response variable. Assuming that the regression model is:

$$Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \epsilon \tag{14}$$

where β_k is the regression coefficient of the X_k input feature, $\beta_0$is the bias value and ϵ is a random error term. Measuring the statistical significance of an input feature is done by calculating the partial F-statistics [34].

The partial F-statistics of the jth term is given as:

$$F_j = \frac{SS_R(\beta_j|\beta_0,\beta_1,\ldots,\beta_{j-1},\beta_{j+1},\ldots,\beta_k)}{MS_E} \tag{15}$$

where $SS_R(\beta_j|\beta_0, \beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k)$ is the regression sum of squares due to the added term j, given that $\beta_0, \beta_1, \dots, \beta_{j-1}, \beta_{j+1}, \dots, \beta_k$ are in the model. MS_E indicates the mean square error of having $\beta_0, \beta_1, \dots, \beta_{j-1}, \beta_j, \beta_{j+1}, \dots, \beta_k$ in the regression model [34]. The computed p-value of the partial F-statistics is used in a comparative matter with respect to a threshold entrance and exit tolerance. The p-value represents the probability of observing a sample statistic as an extreme than the one observed underneath the assumption that the null hypothesis is true. A p-value below the entrance tolerance rejects the null hypothesis. On the other hand, a p-value above the exit tolerance confirms the null hypothesis.

Stepwise regression can be applied in a forward manner or the backward elimination manner. In the forward manner, the method starts with initiating a model with an initial single input feature. One can choose to start the procedure with a chosen initial term, or an initial term with the smallest p-value (largest partial F-statistics) added. Then, the subsequent input features are added or removed depending on the comparative study with the tolerance values. If the p-value of added term in question is less than the entrance tolerance, the null hypothesis is rejected and the term is added to the regression model. Else, if the p-value of the added term in question is greater than the exit tolerance then the null hypothesis is accepted and the term is removed. Stepwise regression stops when no term can be added or removed. Figure 6 illustrates the stepwise regression procedure in the forward manner [35].

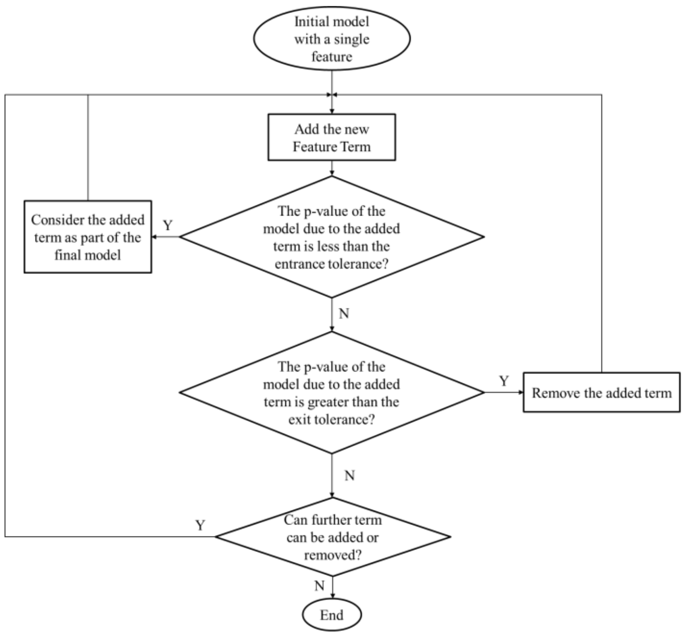

Figure 6: Stepwise regression procedure in the forward manner

The backward elimination procedure for stepwise regression works in a similar way as the forward manner procedure except that the initial model includes all the feature terms. With all the features being in the initial model, the effect of removing the feature in question is tested. If the p-value of model is reduced to the removal of the feature in question, the term is removed and the procedure proceeds with the following feature. The procedure stops when no further terms can be removed. Figure 7 illustrates the procedure followed in stepwise regression using in the backward elimination manner [35].

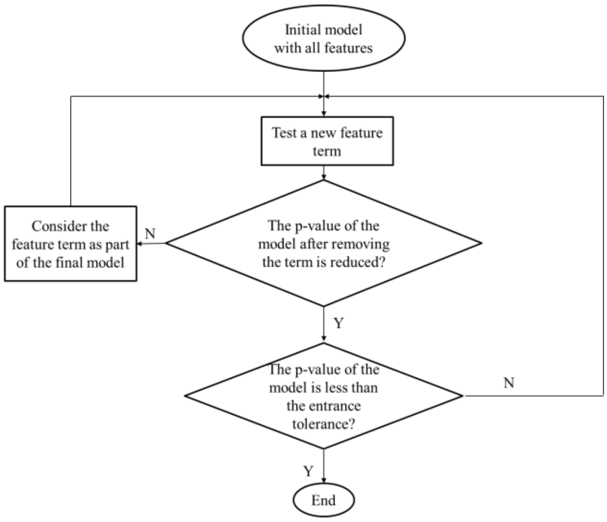

Figure 7: Stepwise regression procedure in the backward elimination manner

Selection of different initial terms can result in regression models of different selected features. Performance of the selected terms in the final regression model is done by means of the F-statistic and Adjusted-R^2 statistic [35]. The Adjusted-R^2 statistic ($\bar{R}^2$) is given as:

$$\bar{R}^2 = 1 - \frac{(1-R^2)(n-1)}{n-p-1} \tag{16}$$

where p is the number of final selected terms, n is the number of output samples and R^2 is the statistic error parameter that indicates the extent of variation between the predicted and desired response term [35]. R^2 is given as:

$$R^2 = 1 - \frac{\sum_{k=1}^{N}(\widehat{Z_k} - Z_k)}{\sum_{k=1}^{N}(\widehat{Z_k} - \overline{Z_k})} \tag{17}$$

where Z_k and $\widehat{Z_k}$ are the predicted and desired outputs respectively. $\overline{Z_k}$ is the mean value of the desired response outputs.

3.5. Research Methodology

The main objective of the research is to reduce the cost of using the HI parameter in TAM. Using the transformer HI value as a CA tool can only be useful if it accurately represents the transformer's operational condition. Based on the industry computational method used, the computed HI value can be only accurate if all the required inputs of that method are available. This, however, adds a maintenance cost burden to the utility for each existing transformer unit in the power system. Moreover, using a cost-effective feature reduction strategy would be considered very difficult using the existing industry computational method.

Thus, the aim is to develop an alternative model that can predict the HI value based on a given industry computational method. The literature review strongly supports the use of AI as an alternative to the standard transformer CA methods. AI was mainly used to either predict a CM test parameter [19]- [21] or the overall HI value [28], [30] and [31]. Therefore, the initial step is to develop the HI predictor model with the same input features using AI. The validity of the proposed model is tested by measuring the accuracy of re-producing the HI values. Once the validity of the proposed model is proved, a feature selection method is considered to reduce the required CM test features (thus the overall cost).

3.5.1. **HI prediction.** Accomplishing the main objective of this work requires the development of an AI based HI model. Based on the literature review, the use of multi-layer perceptron (MLP) ANN is highly recommended in predicting a desired response variable, which is theoretically correlated to the input features. This has been shown and proven in the works of [13], [16], [18], [19] and [30]. Thus, the first task is to design a HI predictor model using MLP ANN. The HI predictor model will produce the HI value of an oil sample based on the corresponding insulation CM parameter inputs. Selection of the insulation CM tests as the input features is due to the existing theoretical correlation between the transformer's HI, and the strength of the insulation system. Still, a non-linear and complex numerical relationship exists between the HI and the complete set of CM tests. Therefore, the use of multiple hidden layers (instead of one hidden layer) is required in order to improve the non-linear mapping of the predictor model.

Designing the HI predictor requires the availability of a large database, consisting of transformer oil samples and corresponding HI values. The aim of the HI predictor is to ideally reproduce the industry computed HI using the CM input features. The 730 transformer oil samples of UTILA will be used mainly in designing the HI AI predictor. Using the industry computational method of [26], the HI of the 730 transformer oil samples of UTILA will be computed. The HI values will be used as the target output values of the developed AI predictor model.

Once the HI is computed for all the transformer oil samples, training and testing data sets will be created to design the ANN. The training set will be mainly used for setting the non-linear HI mapping function of the input features. This will be done by setting the proper weights of the neural links. On the other hand, the testing set is used to evaluate the performance of the AI model in predicting the HI for untrained set of transformer oil samples. Figure 8 shows the feed-forward (FF) MLP ANN architecture which will be implemented in the HI predictor model.

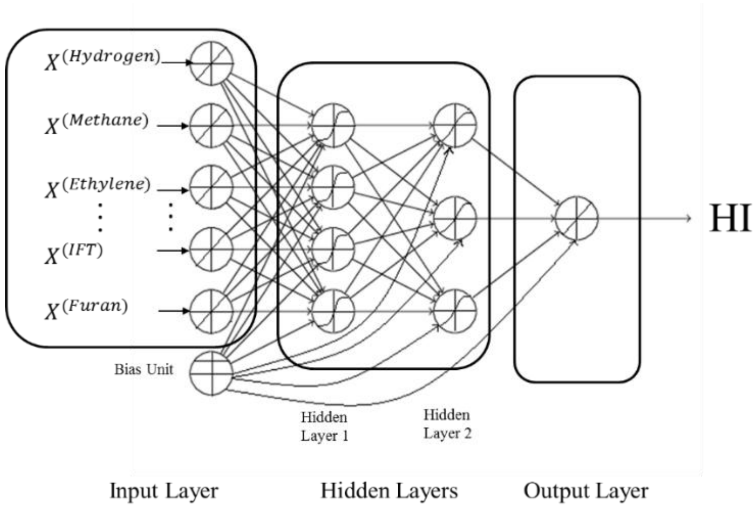

Figure 8: HI predictor with 14 CM input features

3.5.2. **Feature selection of the HI predictor.** Once the main HI predictor model is tested and verified, the following task is to reduce the required number input features (i.e. reduce the CM tests). Reduction of the input features is done through the use of feature-based exhaustive techniques and stepwise regression. In the exhaustive feature-based search technique, different sets of ANN HI models will be trained and tested based on single input features. Then, the ANN HI model will be designed with multiple input features of the highest feature-based performance in predicting the HI.

The performance of predicting the HI of these developed models will be compared against the stepwise regression based model.

Stepwise regression is used (forward and backward manner) to eliminate the redundant input features that have low statistical significance on the output response variable. The selected features with highest statistical significance on calculating the HI, will be only used in the main HI predictor model. The validity of the new feature-reduced model will be indicated through the HI prediction of the testing data set.

3.5.3. **Generalizing the HI predictor model.** One of the objectives of the presented work is to create an HI platform that can be generalized and used by different utilities. Based on one industry method, the HI computation should produce an accurate level of transformer CA regardless of the transformers' region of operation and utility owner. The same level of accuracy should be expected from the developed ANN-based HI predictor model. To support the claim of having a generalized HI predictor model, the original model should be tested with an unseen data from a different utility company. Thus, the developed HI predictor model (using the dataset of UTILA) will be tested against the new unseen dataset of UTILB. The testing will include the previously developed ANN models of the stepwise-based features. In addition, a generalized feature approach will be validated and tested in the presented work. The generalized feature approach indicates the general use of the selected features in the stepwise regression model, by different utility companies. Thus, the utility company can have a HI predictor model which is trained on its own data, with the preliminary knowledge of the required general features. Figure 9 is a schematic diagram of the HI generalizing procedure.

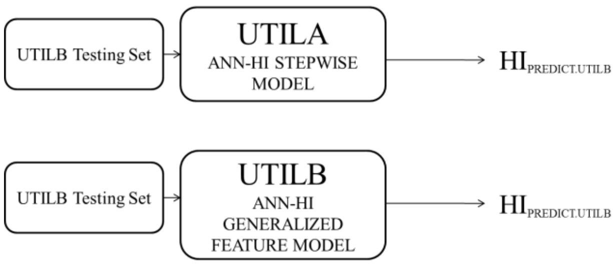

Figure 9: Generalizing the HI model

3.5.4. **Predicting HI using predicted feature.** Significant research in the literature review focused on predicting CM test features. This is attributed to the cost of acquiring these CM tests in terms of the availability of required equipment, proper testing environment and experienced personnel. Using AI techniques, researchers were able to predict some of the costly CM features like Furan or IFT. The aim of this work is to reduce the cost of the HI assessment through reducing the required features, and predicting the costly reduced features. That is, the cost objective will help the utility company in computing the HI with less CM test requirements, and improved CM test predictions. An FF-MLP ANN network will be designed to predict the IFT feature using other CM features, since IFT is considered as a costly test. Similar to predicting the HI, stepwise regression will be used to eliminate the redundant CM test features in predicting the IFT.

The main objective of the presented work is to predict the HI, with less selected features and predicted costly features. This is the main contribution of the presented work in terms of a much improved cost-efficiency in computing the HI for TAM. The developed stepwise regression based ANN model (of UTILA dataset) will use a predicted costly feature (along with the other reduced features) as an input to predict the HI. The performance of the final model will be assessed through predicting the HI of the testing dataset of UTILA. Figure 10 is a schematic diagram of the cost-effective HI model. Figure 11 briefly illustrates the methodology followed in the presented work

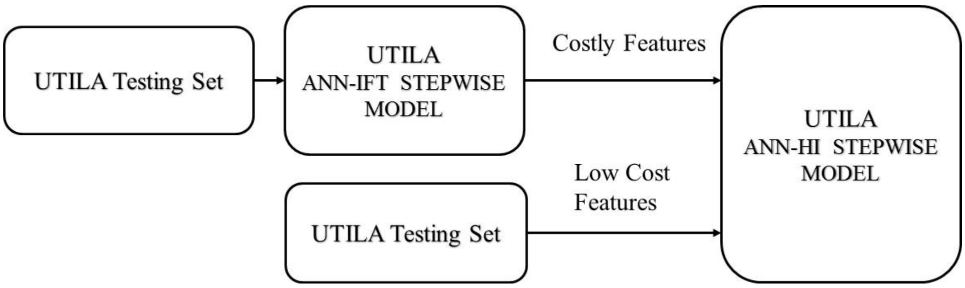

Figure 10: Schematic of a cost-effective HI model

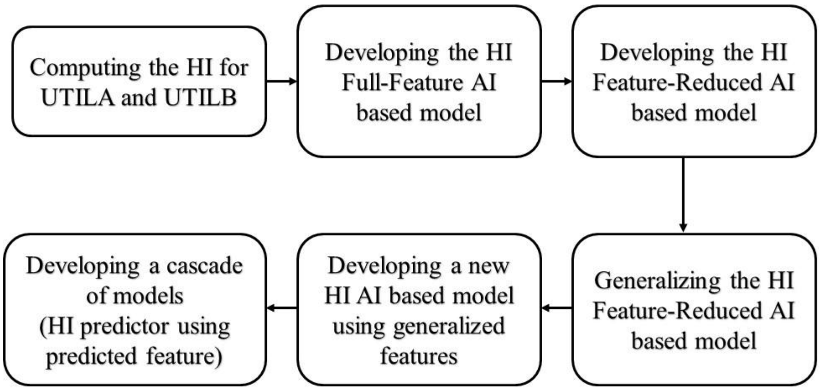

Figure 11: Research methodology procedure

3.6. Model Setting and Validation

Setting the ANN model for the HI predictor will be done through MATLAB [36]. In order to generalize the ANN used in the presented work, a wide range of neural combinations are applied to predict the HI. For a two hidden layer ANN, each layer can have 1 to 10 neurons. This allows for a 100 combination of neurons to be used in the ANN model. This was done in order to indicate the validity of a general neuron tuning scheme in predicting the HI. Table 17 shows a sample of the neuron table which will be used in this research for evaluating the performance of ANN models.

Table 17: Neural matrix combination for a two-hidden layer problem

		Number of Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Number of Neurons in First Hidden Layer	*1*	[1,1]	[1,2]	[1,3]	[1,4]	[1,5]	[1,6]	[1,7]	[1,8]	[1,9]	[1,10]
	2	[2,1]	[2,2]	[2,3]	[2,4]	[2,5]	[2,6]	[2,7]	[2,8]	[2,9]	[2,10]
	3	[3,1]	[3,2]	[3,3]	[3,4]	[3,5]	[3,6]	[3,7]	[3,8]	[3,9]	[3,10]
	4	[4,1]	[4,2]	[4,3]	[4,4]	[4,5]	[4,6]	[4,7]	[4,8]	[4,9]	[4,10]
	5	[5,1]	[5,2]	[5,3]	[5,4]	[5,5]	[5,6]	[5,7]	[5,8]	[5,9]	[5,10]
	6	[6,1]	[6,2]	[6,3]	[6,4]	[6,5]	[6,6]	[6,7]	[6,8]	[6,9]	[6,10]
	7	[7,1]	[7,2]	[7,3]	[7,4]	[7,5]	[7,6]	[7,7]	[7,8]	[7,9]	[7,10]
	8	[8,1]	[8,2]	[8,3]	[8,4]	[8,5]	[8,6]	[8,7]	[8,8]	[8,9]	[8,10]
	9	[9,1]	[9,2]	[9,3]	[9,4]	[9,5]	[9,6]	[9,7]	[9,8]	[9,9]	[9,10]
	10	[10,1]	[10,2]	[10,3]	[10,4]	[10,5]	[10,6]	[10,7]	[10,8]	[10,9]	[10,10]

Ideally, the ANN model for a given two-hidden layer neural network combination should predict the same value as the desired HI. In order to validate the performance of the ANN, a mean accuracy precision error (MAPE) index for *N* samples is given by

$$MAPE = \frac{\sum_{i=1}^{N}\left(\frac{|Y_{i.target} - Y_{i.predict}|}{Y_{i.target}}\right)}{N} \times 100\ \% \quad (18)$$

where $Y_{i.target}$ is the desired ith HI value, and $Y_{i.predict}$ is the predicted HI value. The followed procedure herein after is the following: For a given [n,n] neuron combination, the following will be computed:

- The prediction accuracy value for 10 trials where the prediction accuracy is

$$Prediction\ Accuracy = 100 - MAPE\ \% \quad (19)$$

- The average prediction accuracy value for the 10 trials.
- The variance of the prediction accuracy values for the 10 trials.

Chapter 4. Results and Discussion

The main outcomes of the proposed research are reducing the cost of calculating the HI, and generalizing the HI model. To achieve these outcomes, a step by step modeling approach of the HI prediction is followed. Such an approach ensures a proper transition for obtaining the HI using the industry computational method, to a cost-effective ANN based approach. This chapter starts with presenting the HI distribution of UTILA and UTILB using the industry computational method. Then, an ANN approach is followed using UTILA training samples to predict the HI using all oil-paper insulation features. Once the performance outcomes of the full-feature model are satisfied, feature selection methods are used to eliminate the relatively redundant test features. The exercised methods in this work include exhaustive single-feature modelling and stepwise regression. In single-feature modelling, ANN models will be trained and tested based on a single feature. Individual features of high modelling performance will be collectively used in a multi-feature model, whose HI prediction performance will be tested. Stepwise regression will be used to further explore the possibility of having higher prediction performance using a reduced-feature model. The selected features from both techniques will be assessed based on the HI predictor performance. Later, a generalized model is developed, in which the UTILA stepwise-based model is tested in predicting the HI of UTILB. Same selected features are used to develop and test a HI-ANN predictor model using only UTILB sample database. Then, stepwise regression is applied for predicting a costly feature (IFT) using other relatively low cost features. Hence, a cost-effective HI predictor model is developed using the feature-reduced HI model with the predicted costly feature as an input.

4.1. Predicting the HI Using all Test Features

The transformer HI value can be computed using several methods. The industry computational method used in [26] is used to compute the transformer HI of UTILA and UTILB. Table 18 and Table 19 show the transformer HI distribution of UTILA and UTILB respectively.

Table 18: HI distribution for 730 transformer samples of UTILA

Very Poor	*Poor*	*Fair*	*Good*	*Excellent*
$HI \leq 30$	*$30 < HI \leq 50$*	*$50 < HI \leq 70$*	*$70 < HI \leq 85$*	*$85 < HI \leq 100$*
11	66	111	60	482

Table 19: HI distribution for 327 transformer samples of UTILB

Poor	*Fair*	*Good*	*Excellent*
$30<HI\leq 50$	$50<HI\leq 70$	$70<HI\leq 85$	$85<HI\leq 100$
5	35	49	238

The database of UTILA will be mainly used in the developed model. The HI distribution is unbalanced due to the relatively higher number of samples in the excellent category of HI. In order to assure proper training of the developed models and avoid overtraining the model on the excellent samples, only a subset of 25-27% (130-140 samples) of the excellent HI samples will be used in training the ANN models. Along with the excellent samples, a set of 60%, 80%, 80% and 80% of training samples were used from the very poor, poor, fair and good HI categories respectively. Furthermore, 15% of the entire training sample group was used for model validation to avoid over fitting. After developing the HI ANN predictor model using all the input features, the remaining number of unused samples is used for testing. For 100 neural combinations in a two-hidden layer ANN model, the average and variance of prediction accuracy for 10 trials is produced in the full feature model. Table 20 and Table 21 indicate the average and variance results. Figure 12 shows one [4,5] model trial for the HI actual versus the predicted results for a given oil sample in the testing group.

Table 20: Average prediction accuracy result in full-feature HI predictor

		Average Prediction Accuracy of 10 Trials for a $N_1\times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	95.42	94.08	95.92	95.90	93.38	93.52	95.81	95.85	93.82	92.11
	2	76.57	93.27	95.44	95.63	95.99	94.28	88.14	95.80	95.04	95.30
	3	75.32	95.47	95.64	95.77	94.64	92.92	95.61	95.39	95.21	94.95
	4	88.16	95.17	95.69	90.95	95.63	95.21	96.14	95.44	95.68	95.43
	5	84.36	93.24	95.92	94.36	94.48	95.87	95.26	95.58	95.12	95.14
	6	96.25	88.29	94.81	90.90	95.05	92.04	94.91	95.30	93.00	95.40
	7	95.69	85.60	95.52	96.13	95.90	95.50	95.52	95.32	95.40	95.74
	8	94.71	92.05	92.37	95.87	95.27	85.49	93.15	93.68	95.36	93.34
	9	89.35	89.11	95.11	94.92	93.95	95.99	95.10	95.73	95.56	91.22
	10	81.72	88.55	91.81	91.30	95.71	95.48	94.80	95.88	94.59	95.38

Table 21: Variance of prediction accuracy result in full-feature HI predictor

		Variance of Prediction accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	0.04	32.61	0.10	0.30	49.29	56.69	0.13	1.88	15.83	28.13
	2	81.22	43.91	0.60	0.51	0.18	1.46	112.78	0.30	1.70	0.85
	3	21.74	0.28	0.56	0.75	4.45	38.71	0.47	1.81	4.39	1.67
	4	71.28	2.30	0.24	63.91	0.28	0.65	0.27	0.52	0.04	0.46
	5	59.28	17.86	0.09	3.21	3.73	0.24	0.43	0.49	1.75	0.76
	6	0.16	70.14	3.19	46.58	3.40	54.11	1.07	0.58	22.91	0.82
	7	1.07	34.01	0.10	0.04	0.12	0.58	0.35	0.56	0.32	0.12
	8	2.67	57.00	50.16	0.22	0.25	67.59	12.02	12.64	0.29	7.67
	9	113.08	32.78	1.54	3.84	10.02	0.16	2.79	0.10	0.45	53.19
	10	61.88	23.09	34.31	23.35	0.31	0.38	3.13	0.29	3.58	0.21

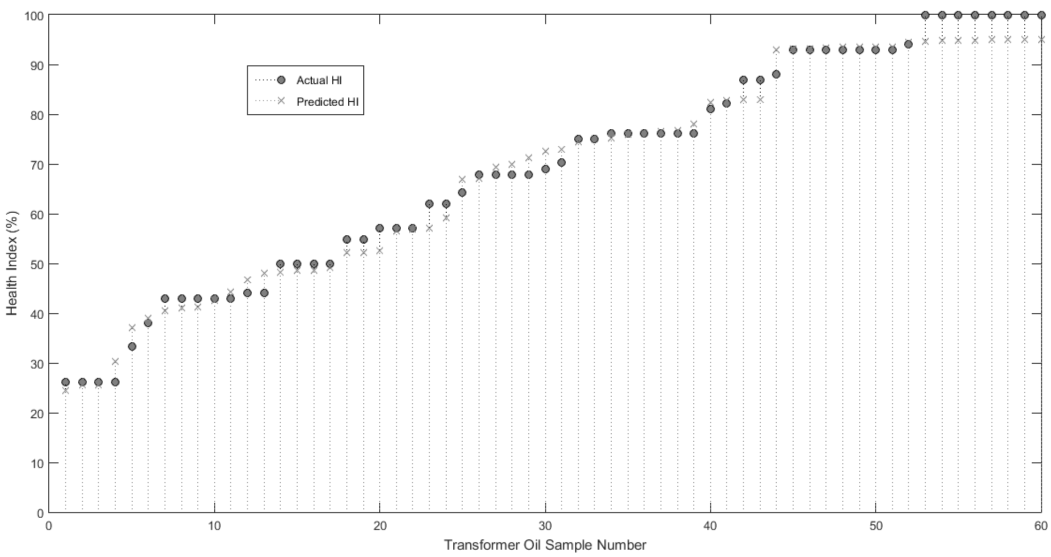

Figure 12: Actual vs. predicted HI for full-feature HI Predictor for selected transformers

The results indicate an excellent regression performance, with the average prediction accuracy being around 95% with a low variance value. This is due to the existing correlation between the HI value and the individual features of the insulation CM tests.

4.2. Exhaustive Single-Feature and Stepwise Regression

To explore the HI cost reduction using feature selection, ANN models were trained and tested based on single features. Table 22 indicates the obtained results using single feature ANN models. The indicated results are the 10-trial average prediction accuracy and its variance for the 100 two-hidden layer neural combination models.

Table 22: Single feature ANN model results

Feature	*Hydrogen*	*Methane*	*Ethane*	*Ethylene*	*Acetylene*	*Carb. Mon*	*Carb. Diox*	*Water*	*Acid*	*BDV*	*DDF*	*Color*	*IFT*	*Furan*
Average Prediction Accuracy Rate (%)	64	63	65	69	69	63	68	64	74	62	71	74	78	75
Variance of Average Prediction Accuracy	28	4	2	52	9	2	14	2	15	4	9	3	6	7

The produced results are relatively poor compared to the full-feature HI prediction results as depicted in Table 20. Generally, IFT and Furan produced the highest average prediction accuracy followed by color, acidity and the DDF. Moreover, acidity and DDF are excellent indicators of the oil-paper insulation condition. Based on the individual feature modeling results, a multi-feature model is built based on Furan, IFT, color, acidity and DDF test features. Table 23 and Table 24 show the obtained results for the 100 neuron combinations for the two-hidden layer multi-feature ANN problem. The average prediction accuracy is around 80% with a low variance value.

Table 23: Average prediction accuracy for multi-feature HI predictor

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	78.38	74.63	79.04	75.91	78.41	72.57	74.93	78.56	74.45	78.12
	2	77.97	80.32	79.16	77.10	81.00	80.93	70.97	79.48	80.00	75.37
	3	73.86	80.42	80.38	79.93	80.04	79.96	79.35	79.10	80.31	78.62
	4	68.79	79.44	80.55	80.18	75.33	79.98	79.25	79.53	78.65	80.00
	5	75.34	74.73	80.95	78.75	80.02	79.54	78.83	79.87	79.38	80.08
	6	80.86	80.01	80.48	80.17	80.17	80.03	79.90	80.59	80.45	78.71
	7	72.82	80.49	79.17	79.37	77.32	78.31	80.19	79.30	79.29	79.46
	8	76.18	77.14	79.39	80.24	79.89	80.35	79.51	79.95	81.03	80.98
	9	72.23	79.15	72.73	81.01	79.33	79.61	78.87	79.73	79.25	78.27
	10	69.65	80.32	77.14	80.00	78.69	80.63	79.24	80.90	79.50	79.47

Table 24: Variance of average prediction accuracy for multi-feature HI predictor model

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	0.10	29.76	0.91	36.02	0.58	67.19	7.62	0.22	20.03	9.47
	2	2.80	8.61	1.19	17.59	0.17	1.49	94.09	2.17	2.15	65.28
	3	57.46	1.63	3.19	1.54	0.51	1.12	1.89	1.51	1.62	5.67
	4	96.36	1.48	0.95	0.82	71.23	1.45	4.44	3.55	66.76	1.93
	5	59.21	55.69	0.99	0.63	3.95	2.42	2.90	2.48	1.23	2.29
	6	0.47	3.35	1.70	2.17	1.64	2.83	5.91	0.79	1.60	3.62
	7	74.25	0.81	2.00	4.09	11.40	3.58	1.40	0.63	3.39	0.40
	8	70.35	36.73	4.31	1.34	0.92	3.00	2.84	0.57	2.22	0.71
	9	51.06	3.89	50.17	1.59	2.17	2.15	1.77	2.63	2.79	5.98
	10	64.69	0.34	12.16	3.70	1.29	1.09	1.48	0.92	1.50	3.40

Stepwise regression is used to reduce the required oil-paper insulation CM parameters needed to predict the HI of a given oil sample. Forward stepwise regression is applied on the same training set used in the full-feature HI predictor model. The p-enter tolerance value was $1x10^{-5}$, while the p-exit tolerance was selected to be 0.05. The procedure was applied 14 times, with each feature chosen as the initial term in the model. Each term has a computed individual p-value which represents the regression effect of the term given that all other features are in the regression model. After the initial term is set in the model, the term with the lowest individual p-value is added to the model. In each step the overall p-value of the current model due to the added term is computed against the entrance and exit tolerance. Based on the final models of highest Adjusted-R^2 value and F-statistic, the reduced features of the corresponding models are chosen in the HI predictor. Table 25 shows an example of the stepwise process of the features in predicting the HI using Ethylene as the initial feature in the model. With Ethylene being the initial term, IFT is added since it has lowest individual p-value. The overall p-value of IFT as a part of the model containing Ethylene is tested against the tolerance values. The calculated overall p-value is less than the entrance tolerance and thus Ethane is accepted as a part of the model. The process continues until all the low individual p-value terms are added and tested. Then, the effect of removing the initial term from the final model is tested. If the overall p-value of having the term in the model is greater than the exit tolerance, the initial term is removed from the final model. The stepwise procedure stops until no term can be added or removed. Table 26 indicates the final selected features in the stepwise regression procedure.

Table 25: Example of forward stepwise regression for UTILA

	Action	*p-value of the current model*
Step 1	Added IFT	$1.79x10^{-74}$
Step 2	Added Ethane	$1.79x10^{-74}$
Step 3	Added Acetylene	$5.97x10^{-14}$
Step 4	Added Furan	$4.25x10^{-19}$
Step 5	Added Color	$5.87 x10^{-8}$
Step 6	Removed Ethylene	0.414

Table 26: Selected features for UTILA in forward stepwise regression

Initial Feature	*Hydrogen*	*Methane*	*Ethane*	*Ethylene*	*Acetylene*	*Carb. Mon*	*Carb. Diox*	*Water*	*Acid*	*BDV*	*DDF*	*Color*	*IFT*	*Furan*
Selected Features in the Final Model	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen
	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane
	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane
	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene
	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene
	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon
	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox
	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water
	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid
	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV
	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF
	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color
	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT
	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan
Adjusted R^2	0.832	0.837	0.824	0.824	0.824	0.831	0.824	0.831	0.824	0.830	0.828	0.824	0.824	0.824
F Statistic	248.99	220.46	281.85	281.85	281.85	246.86	281.85	246.43	281.85	244.64	241.67	281.85	281.85	281.85

The Adjusted-R^2 slightly varies over the 14 final models, thus the highest F-statistic was used to select the final features. According to Table 26 the highest F-statistic value is 281.85 with an Adjusted-R^2 value 0.824. The selected features accordingly are Furan, IFT, Color, Acetylene and Ethane. This is due to the high correlation that exists between the selected features and the HI. Particularly, Furan and IFT are directly related to the degradation of the paper insulation system and hence the overall value of the HI. Moreover, the color of the oil is a good visual CM parameter that can give a good estimate of the oil condition. Furthermore, Acetylene and Ethane are key dissolved gases that indicate the existence of thermal and electrical stresses experienced by the transformer. Using the selected features, the HI predictor is re-built and tested. Table 27 and Table 28 indicate the average prediction accuracy and variance results for the 100 neural combinations problem. An average prediction accuracy of 95% was obtained for the 100 neural combinations.

Applying stepwise regression in the backward elimination manner requires the use of all the 14 CM test features in the initial model. The exit tolerance is set to $5x10^{-5}$ while the entrance tolerance remains $1x10^{-5}$ to increase the feature selectivity of stepwise regression. Table 29 indicates the backward elimination procedure followed for UTILA CM test features in predicting the HI. The final selected features are Methane, Ethane, Acetylene, Carbon Monoxide, Color, IFT and Furan. Table 30 and Table 31 indicate the prediction accuracy results obtained using the backward elimination selected features.

Table 27: Average prediction accuracy for reduced-feature predictor (using forward stepwise regression)

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	93.72	93.52	92.02	94.14	93.48	82.48	91.14	94.11	90.8	93.7
	2	94.99	90.32	94.47	93.68	94.97	95.01	88.78	94.64	94.11	93.22
	3	89.32	92.48	94.92	94.5	94.03	95.57	93.45	94.51	94.67	93.05
	4	88.46	95.5	95.53	94.8	94.53	94.6	94.21	94.65	94.8	94.72
	5	82.3	92.46	95.19	95.1	95.23	94.8	95.29	95.29	95.08	95.08
	6	95.2	94.96	95.46	94.52	95.26	94.99	95.33	94.88	94.83	94.51
	7	83.45	95.19	94.59	94.51	93.36	95.46	95.07	94.12	94.69	94.43
	8	75.98	93.85	93.95	95.18	94.48	95.73	93.69	94.89	94.93	94.94
	9	92.33	87.67	85.59	94.98	95.12	94.6	95.21	95.42	94.94	94.7
	10	91.47	94.09	94.96	95.23	94.72	94.97	94.25	95.33	94.68	95.22

Table 28: Variance of prediction accuracy result in reduced-feature HI predictor (using forward stepwise regression)

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	0.11	0.83	40.31	0.30	1.35	97.65	29.71	0.26	7.23	2.56
	2	0.29	62.08	0.47	4.11	0.04	0.16	41.58	0.61	1.86	10.32
	3	35.49	34.87	0.79	0.77	2.01	0.07	3.13	1.64	0.42	7.36
	4	73.50	0.16	0.42	0.25	1.21	0.56	0.42	0.77	0.27	0.69
	5	45.04	17.14	0.42	0.22	0.26	0.54	0.18	0.42	0.32	0.30
	6	0.31	0.23	0.24	5.89	0.22	0.27	0.10	1.31	0.73	0.64
	7	93.15	2.00	0.82	1.17	10.04	0.22	0.31	0.19	0.69	2.06
	8	20.66	9.86	2.22	0.36	0.12	0.03	4.65	1.44	0.30	0.39
	9	7.93	91.13	85.87	1.15	0.59	1.61	0.22	0.13	0.40	0.67
	10	47.97	3.49	0.18	0.48	0.67	0.40	1.14	0.76	0.22	0.32

Table 29: Backward elemination stepwise regression on UTILA

	Action	*p-value of the current model*
Step 1	Remove Carbon Dioxide	0.34
Step 2	Remove Ethylene	0.21
Step 3	Remove BDV	0.11
Step 4	Remove Acid	0.02
Step 5	Remove Water	5×10^{-4}
Step 6	Remove Hydrogen	1×10^{-4}
Step 7	Remove DDF	5.8×10^{-4}
Final Features	Methane, Ethane, Acetylene, Carbon Monoxide, Color, IFT & Furan	
F-Statistic		274.42
Adjusted-R^2		0.85

Table 30: Average prediction accuracy for reduced-feature predictor (using backward elimination)

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	94.32	91.05	82.23	95.59	95.77	95.57	95.27	95.49	95.69	95.54
	2	93.77	96.23	95.84	96.20	95.98	95.95	92.67	94.98	95.58	90.24
	3	95.82	95.83	95.46	95.09	96.20	96.64	96.06	95.61	96.47	96.17
	4	90.67	96.03	96.13	95.82	96.51	95.74	96.17	96.42	96.24	96.10
	5	92.58	96.13	94.70	96.31	96.60	96.31	96.40	95.91	96.37	96.14
	6	96.02	87.04	96.54	95.23	91.64	95.81	95.95	96.45	95.98	95.71
	7	75.00	88.93	96.05	96.51	96.04	96.31	96.12	95.96	96.50	96.28
	8	88.03	90.53	96.05	92.32	94.62	95.36	96.31	96.29	95.95	96.03
	9	84.51	96.01	96.36	96.49	93.53	96.21	96.13	96.08	96.02	92.87
	10	94.78	96.52	95.70	93.81	96.27	96.48	95.65	96.27	96.15	96.22

Table 31: Variance of prediction accuracy result in reduced-feature HI predictor (using backward elimination)

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	8.44	73.66	105.06	0.72	0.24	0.20	3.47	0.14	0.56	0.51
	2	25.78	0.29	0.86	0.22	0.37	0.34	59.91	2.36	0.18	115.65
	3	0.64	0.49	1.12	9.46	0.37	0.04	0.50	0.76	0.03	0.40
	4	68.29	0.18	0.15	0.61	0.13	0.45	0.17	0.34	0.30	0.23
	5	44.99	0.22	4.50	0.49	0.20	0.05	0.19	0.21	0.05	0.28
	6	0.62	76.77	0.11	4.79	15.41	0.44	0.49	0.15	0.40	0.54
	7	551.84	42.06	0.28	0.08	0.30	0.30	0.53	3.40	0.40	0.42
	8	110.28	49.97	0.52	18.38	4.27	4.27	0.15	0.18	0.24	0.29
	9	107.98	0.35	0.07	0.16	19.32	1.10	0.35	0.18	0.40	14.81
	10	9.38	0.10	0.86	19.34	0.26	0.13	0.84	0.05	0.15	0.25

The average prediction accuracy results obtained using the backward elimination selected features were around 96% which is slightly more than what was obtained by the forward selection procedure. However, two additional test features of Methane and Carbon Monoxide are required. Since the prediction accuracy did not improve much and the number of required features are more using backward elimination, the forward stepwise features will be selected as the main reduced features in the presented work.

It is evident that the performance of the ANN model is enhanced using the forward stepwise features as compared to the exhaustive based multi-features. The average prediction accuracy was around 95%, which is close to the one obtained using the full-feature model. It is noticed that Acidity and DDF were not considered in the final stepwise model. This is due to the relative redundancy of these oil quality features in the presence of the more HI influential features like Furan and IFT. Figure 13 shows one [4,5] reduced model trial for the HI actual versus the predicted results for some transformer oil samples.

4.3. Generalizing the HI Model

A generalized HI model (for a given industry computational method) should be capable of producing high level of transformer HI accuracy for different utilities. UTILB transformer oil samples are used to assess the performance of the previously developed HI model (based on UTILA) as a generalized model. As mentioned earlier, the CM test features of UTILB are relatively less than those of UTILA. Particularly, UTILB database does not contain dissolved gas CM test data for neither Carbon Monoxide nor Carbon Dioxide. Moreover, oil quality test results for both DDF and color are not available. In the reduced-feature HI model of UTILA, color was considered as one of the five selected features. With color and other features being not available, a new stepwise regression procedure is developed for UTILA. The stepwise regression is only performed for the features that are available in UTILB.

Based on the same stepwise approach discussed earlier, Table 32 indicates the new forward stepwise regression results. The results indicate that Furan, IFT, Acetylene and Ethane are the features of highest statistical significance in predicting the HI value with the highest F-statistic with a value of 359. The same number of training samples are used for training the feature-reduced HI predictor model. The entire 327 data samples of UTILB are used for testing.

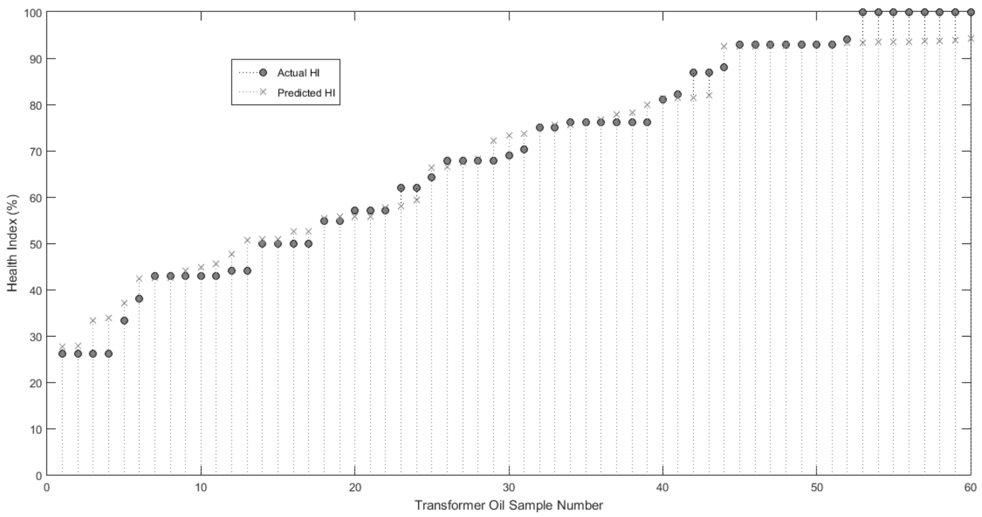

Figure 13: Actual vs. predicted HI for reduced-feature HI predictor

Table 32: Final selected features for reduced features of UTILA (using forward stepwise selection)

Initial	*Hydrogen*	*Methane*	*Ethane*	*Ethylene*	*Acetylene*	*Water*	*BDV*	*Acid*	*IFT*	*Furan*
Selected Features in the Final Model	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen
	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane
	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane
	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene
	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene
	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water
	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV
	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid
	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT	IFT
	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan
Adjusted R^2	358	359	359	359	359	358	358	359	359	359
F Statistic	0.85	0.84	0.84	0.84	0.84	0.84	0.85	0.84	0.84	0.84

Table 33 and Table 34 show the obtained results for the 100 neuron combinations for the two-hidden layer ANN problem. The average prediction accuracy is around 91% with a low variance value. The results indicate the high performance of the reduced-feature HI predictor (of UTILA) in predicting the HI value for UTILB data samples.

Table 33: Average prediction accuracy for reduced-feature HI generalized model

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	81.29	86.16	86.63	87.32	90.58	91.14	90.32	90.62	90.58	92.52
	2	84.63	91.89	88.20	91.35	92.13	88.03	92.50	92.32	91.33	91.55
	3	88.47	74.51	92.44	93.26	92.29	92.61	92.87	93.16	91.81	92.77
	4	91.45	91.89	92.68	93.64	93.82	92.21	93.88	92.87	91.38	92.29
	5	89.61	92.81	92.32	92.06	91.17	90.37	93.03	92.49	93.16	93.41
	6	86.27	92.30	43.63	88.41	93.00	92.68	92.17	92.43	93.57	94.05
	7	91.60	91.77	92.63	93.36	93.68	91.95	92.73	92.67	91.88	92.66
	8	83.08	91.11	88.98	91.70	93.61	92.82	93.09	93.34	91.38	91.60
	9	84.28	82.07	91.67	92.96	92.40	92.15	92.79	93.20	92.46	92.26
	10	92.84	90.61	93.02	91.01	92.19	92.56	92.67	91.79	90.12	91.59

Table 34: Variance of prediction accuracy for reduced-feature HI generalized model

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	21.43	44.48	33.22	65.66	2.31	0.67	8.49	3.67	0.94	0.43
	2	34.51	2.54	112.28	1.21	0.56	52.99	2.23	0.44	1.23	3.44
	3	13.78	119.30	0.68	0.69	1.76	0.31	1.81	0.22	2.85	0.60
	4	42.94	1.16	0.42	1.14	0.69	1.78	1.15	1.20	5.72	0.68
	5	40.08	2.11	7.36	19.13	22.39	28.87	1.56	2.62	0.97	0.71
	6	94.88	7.25	21845.65	15.66	0.77	0.67	0.39	2.89	0.32	0.25
	7	19.41	55.33	3.01	0.33	0.44	8.45	1.79	0.78	3.64	1.61
	8	30.63	7.34	17.73	0.83	0.23	0.79	2.62	0.91	1.66	1.09
	9	135.65	32.50	0.58	1.87	0.70	1.41	0.93	0.30	2.51	2.84
	10	3.77	54.11	0.47	22.70	1.60	1.28	2.47	10.11	8.24	0.89

Figure 14 shows one [4,5] reduced general model trial for the HI actual versus the predicted results, for a given oil sample in the testing group.

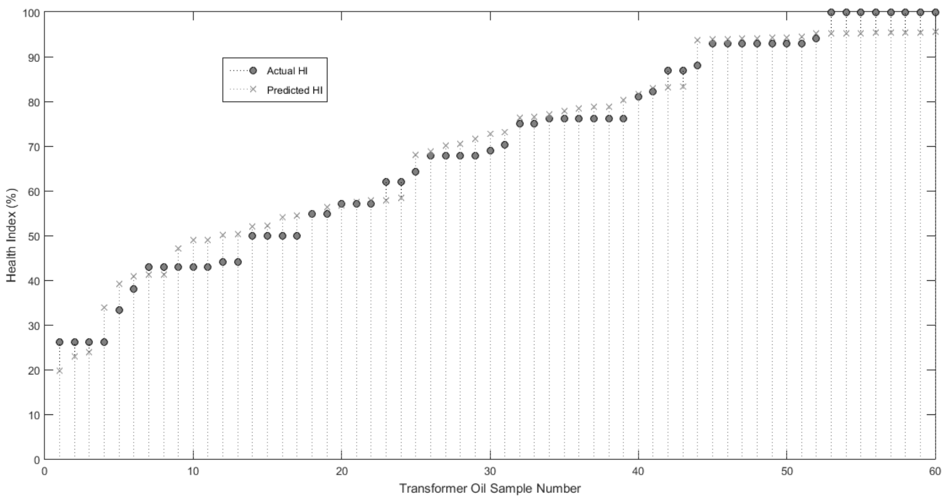

Figure 14: Actual vs. predicted HI for reduced-feature generalized HI predictor

Another scenario of generalization is that the HI prediction would be able to generalize the selected features instead of the model. That is, to use the selected features of Furan, IFT, Acetylene and Ethane to build the ANN model using the data of any utility company. In this part of the presented work, the selected features are used to build the ANN model using the training set of UTILB data only. The model will be tested against the HI prediction of the testing set of UTILB. 50%, 60%, 60% and 60% of UTILB data in the poor, fair, good and excellent categories respectively, which will be used as a training set. The remaining samples will be used for testing. Table 35 to Table 38 present the 100 neural combination results for the ANN UTILB models based on the full-features and selected general features.

Table 35: Average prediction accuracy for full-feature UTILB predictor model

		Average Prediction Accuracy of 10 Trials for a N_1×N_2 Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	84.25	89.07	76.14	85.18	91.66	79.83	89.02	91.11	88.60	80.74
	2	92.37	82.73	91.79	92.69	83.48	92.97	91.71	92.77	90.80	91.01
	3	89.02	91.79	91.50	94.00	93.11	91.39	90.72	93.03	91.95	92.03
	4	86.15	81.66	89.77	93.39	92.02	91.87	92.26	92.26	91.95	94.23
	5	94.00	85.38	93.57	93.01	93.36	93.77	88.69	91.08	91.94	90.50
	6	83.04	92.87	93.34	92.27	92.40	90.12	93.10	93.31	91.28	92.56
	7	93.58	88.84	93.65	92.32	91.74	90.95	93.08	92.37	92.19	92.15
	8	86.51	92.44	87.52	91.86	90.64	90.25	91.32	92.20	91.49	93.02
	9	93.64	89.76	86.01	92.69	92.24	91.04	91.58	91.01	91.18	90.83
	10	90.76	92.71	93.92	91.28	92.50	92.44	91.22	92.69	91.52	93.01

Table 36: Variance of prediction accuracy for full-feature UTILB predictor model

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	50.83	23.99	349.15	30.21	2.52	1349.26	45.23	4.66	3.3	83.24
	2	0.79	112.62	1.72	0.4	31.27	0.97	4.36	1.75	2.89	20.78
	3	42.48	9.77	5.94	0.94	3.9	7.84	3.5	1.29	4.13	1.33
	4	30.41	6	13.02	5.33	8.03	4.85	2.69	2.71	2.52	1.87
	5	0.58	60.12	3.54	2.35	1.96	2.11	12.31	9.81	6.14	13.33
	6	49.97	0.73	1.19	7.3	8.53	7.87	3.57	4.29	4.16	3.63
	7	1.02	11.59	1.03	3.04	3.93	12.35	0.62	6.25	4.95	2
	8	30.04	8.42	4.37	3.1	12.09	10.21	13.06	2.57	3.74	0.35
	9	0.92	11.58	21.62	0.82	3.77	5.47	10.87	1.13	5.4	4.9
	10	2.37	3.56	0.43	6.12	3.91	2.94	12.2	3.06	4.43	1.75

Table 37: Average prediction accuracy for the generalized-feature UTILB predictor model

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	89.42	91.44	91.64	84.37	89.16	92.38	91.16	90.51	87.09	87.61
	2	89.12	91.79	86.66	88.71	89.92	89.82	90.98	88.38	88.87	91.15
	3	81.25	83.91	91.25	88.21	89.09	90.88	90.32	91.74	86	91.08
	4	88.04	92.39	91.85	93.34	91.42	90.86	89.98	89.8	89.84	89.76
	5	90.91	91.37	89.76	87.53	89.21	89.12	91.8	92.57	90.62	90.62
	6	85.86	84.19	89.27	88.65	91.07	92.87	91.56	91.71	93.82	91.29
	7	89.58	91.44	90.4	91.73	91.59	91.93	90.97	91.09	90.2	91.37
	8	83.86	70.77	89.41	89.02	92.69	91.16	90.67	87.87	90.71	90.41
	9	84.64	80.25	87.84	93.05	91.31	93	89.68	90.73	89.97	90.06
	10	88.78	90.91	90.94	90.42	91.34	91.63	90.78	89.57	90.05	90.58

Table 38: Variance of prediction accuracy for generalized-feature UTILB predictor model

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	9.57	1.29	0.5	328.58	6.45	3.96	0.65	0.58	25.54	23.67
	2	25.27	0.11	91.2	8.1	13.92	20.3	0.44	57	0.91	21.34
	3	233.66	45.98	48.31	21.73	7.31	2.89	3.54	3.66	5.09	0.5
	4	9.19	0.66	0.53	1	0.48	3.32	2.35	4.13	2.78	16.78
	5	1.75	1.11	21.46	13.12	19.76	10.57	3.54	0.6	5.97	2.95
	6	23.06	182.99	5.81	13.96	2.91	1.44	2.61	1.34	0.46	1.53
	7	6.55	0.08	11.28	1.12	0.41	2.22	2.9	1.18	6.86	0.9
	8	32.41	655.64	13.52	47.03	1.94	3.16	12.47	20.22	4.1	4.23
	9	14.19	166.49	1.53	0.48	0.78	0.32	5.18	4.98	8.75	5.44
	10	14.32	0.2	1.12	12.72	1.99	0.33	5.47	11.29	9.01	0.73

The average prediction accuracy is around 91% with a low variance value for the full-feature and general-feature model. A high performance of HI prediction is indicated in both the UTILB full-feature and general-feature model. The produced results clearly indicate the efficiency of using the generalized features in terms of the lower CM tests required. Figure 15 shows one [4,5] reduced general model trial for the HI actual versus the predicted results, for a given oil sample in the testing group.

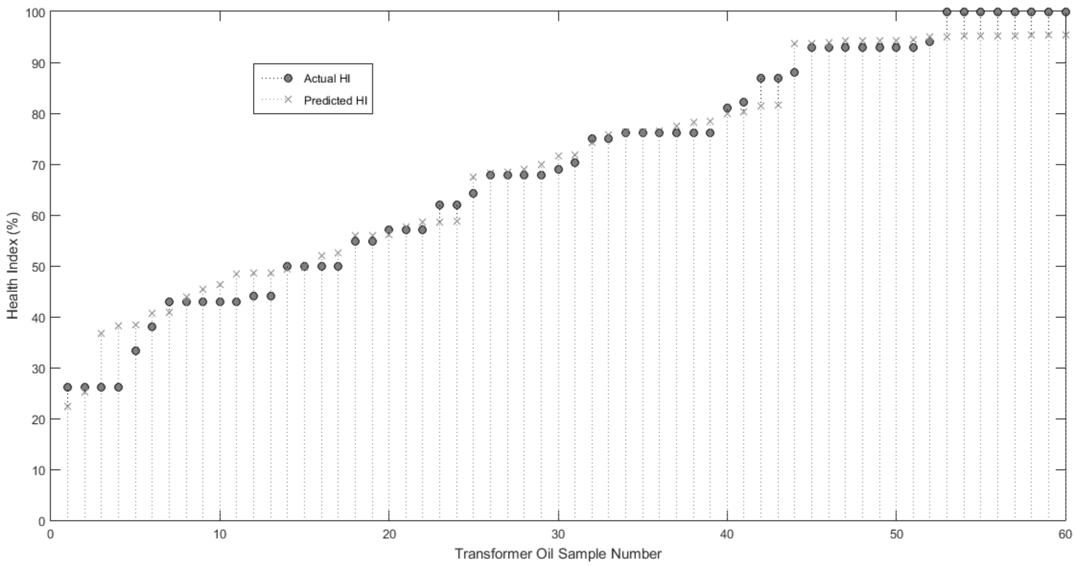

Figure 15: Actual vs. predicted HI for reduced-feature generalized feature HI predictor

4.4. HI Prediction Using Predicted IFT

Predicting Furan in an AI-based approach using other oil insulation parameters has been attempted in many works such as in the works of [23], [24] and [25]. Fewer attempts have been used in predicting the IFT value despite being a costly feature. The aim behind predicting IFT is to further reduce the cost of the transformer HI prediction. The utility will no longer require to conduct IFT CM tests. Instead, IFT is predicted using relatively cheaper oil parameter tests. Using the 13 insulation CM test features as the input features with the IFT being the target feature, stepwise regression is performed to obtain the optimum features required in IFT prediction. Table 39 indicates the stepwise regression results. The results indicate the selection of Ethane, acidity and color for predicting IFT for an F-statistic of 425.18. All of these factors are direct indicators of the paper degradation in the paper insulation system. Table 40 and Table 41 indicate the 100 neural combination results for predicting IFT.

Table 39: Selected features for IFT predictor

Initial Feature	*Hydrogen*	*Methane*	*Ethane*	*Ethylene*	*Acetylene*	*Carb. Mon*	*Carb. Diox*	*Water*	*Acid*	*BDV*	*DDF*	*Color*	*Furan*
Selected Features in the Final Model	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen	Hydrogen
	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane	Methane
	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane	Ethane
	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene	Ethylene
	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene	Acetylene
	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon	Carb. Mon
	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox	Carb. Diox
	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water	Water
	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid	Acid
	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV	BDV
	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF	DDF
	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color	Color
	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan	Furan
Adjusted R^2	0.81	0.79	0.79	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81	0.81
F- Statistic	348.11	422.40	425.18	283.26	283.29	348.11	348.11	348.11	348.11	348.11	286.92	348.11	348.11

Table 40: Average prediction accuracy for reduced-feature IFT predictor model

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	85.18	90.12	90.10	88.80	88.86	89.88	89.59	89.79	90.17	89.31
	2	90.00	90.19	90.31	88.98	89.84	89.83	90.24	89.94	89.25	90.20
	3	90.06	90.06	89.51	90.03	89.58	90.16	90.09	89.71	89.47	90.28
	4	90.13	90.15	88.87	90.15	90.12	90.16	89.97	88.02	90.05	89.35
	5	89.96	90.12	89.30	90.00	89.62	90.07	90.11	90.19	90.05	88.99
	6	89.09	89.69	89.56	89.55	90.34	90.30	90.31	89.81	90.31	89.67
	7	90.09	89.81	90.10	90.17	90.25	89.89	90.10	89.88	89.61	89.89
	8	89.82	88.23	88.05	90.15	89.96	90.16	90.06	89.71	90.09	89.70
	9	88.70	86.54	90.33	90.17	89.65	90.36	90.17	90.12	89.88	89.68
	10	89.71	89.87	89.79	89.79	89.85	90.05	89.82	90.02	89.81	90.18

Table 41: Variance of prediction accuracy for feature-reduced IFT predictor model

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	66.25	0.26	1.20	0.19	0.27	49.33	0.25	6.61	4.34	0.21
	2	4.02	3.78	1.85	4.72	0.13	0.07	1.27	0.51	0.54	0.27
	3	26.01	0.28	0.19	0.31	0.27	0.83	0.81	3.08	0.15	0.26
	4	52.08	2.32	34.80	6.78	0.11	0.78	2.22	0.23	0.55	0.29
	5	0.08	9.94	0.04	0.19	0.24	0.37	1.14	0.14	0.31	0.63
	6	1.04	0.25	10.12	0.19	0.23	0.07	1.75	0.12	0.54	0.13
	7	30.22	8.68	17.41	0.26	0.28	0.27	0.13	0.21	0.32	0.87
	8	0.10	20.93	0.33	0.29	1.04	0.16	7.19	1.30	0.11	0.37
	9	46.42	0.22	0.07	15.32	0.22	0.12	0.03	0.35	0.35	0.23
	10	22.84	5.85	0.12	0.39	0.49	0.54	2.23	6.00	0.58	0.23

From the results, the stepwise-based features can be used to predict the IFT, with a relative lower average prediction accuracy of 89% (as compared to 95% HI prediction). The selected features are relatively cheap features that can be acquired through a proper experimental procedure. Figure 16 shows one [4,5] reduced model trial for the actual IFT versus the predicted results, for a given oil sample in the testing group.

Having a validated feature-reduced HI and IFT predictor models, the objective now is to design the overall cost-effective HI predictor. Using UTILA data samples, the same training and testing groups are created (as was done in HI prediction). The training group will be used in training the HI predictor and the IFT predictor. For IFT prediction, the IFT stepwise-based features are used to train the predictor model. On the other hand, the HI stepwise based features (Furan, actual IFT, color, Ethane and Acetylene) will be used to train the HI predictor model. The testing procedure is used to validate the performance of the overall model. For a given testing oil sample, the IFT predictor model is used to predict IFT using the stepwise-based features of the sample. The predicted IFT along with the other HI selected features (of the oil sample) are fed to the HI model, for the final HI value. Figure 17 shows a schematic block diagram of the testing and training procedure used to obtain the results.

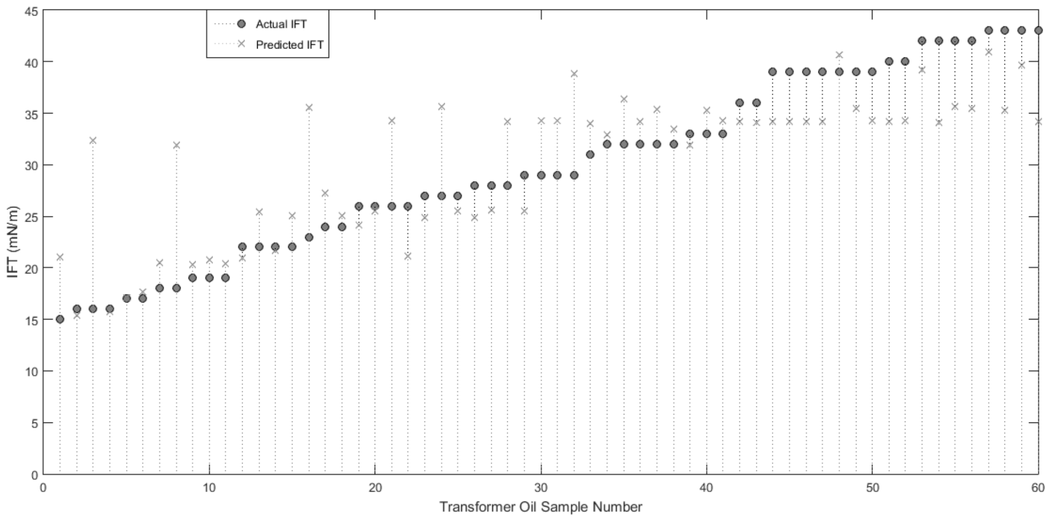

Figure 16: Actual vs. predicted IFT for transformer oil samples

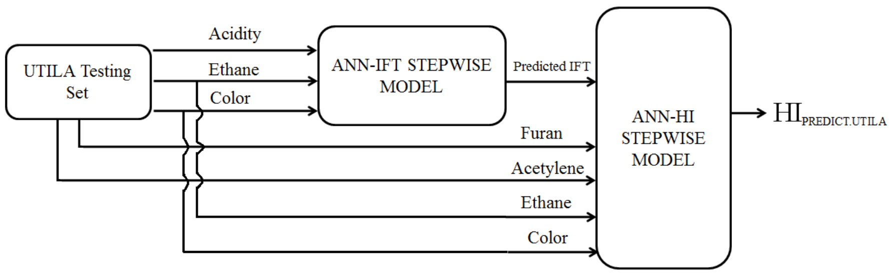

Figure 17: Training and testing procedure for the cost-effective HI predictor

Table 42 and Table 43 show the 100 neural combination results of predicting the HI value. With an average prediction accuracy of 95%, the final results clearly indicate the high performance of the overall cost effective HI predictor model. Figure 18 shows one [4,5] reduced model trial for the actual HI versus the predicted results, for a given oil sample in the testing group.

Table 42: Average prediction accuracy results for overall cost-effective HI predictor

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	94.48	83.36	90.21	94.56	94.55	92.77	93.63	93.38	92.53	92.49
	2	94.61	93.29	95.13	92.55	95.11	95.12	94.96	94.96	95.08	94.06
	3	90.23	93.25	95.33	95.46	94.56	94.99	95.42	95.23	95.15	95.24
	4	80.11	95.51	95.65	94.85	95.48	95.33	94.92	95.32	94.88	94.92
	5	88.66	88.94	95.62	95.19	95.22	95.66	94.65	95.21	95.45	95.19
	6	95.58	95.54	94.61	95.08	95.72	95.78	95.65	95.54	95.55	94.93
	7	78.73	95.38	95.44	95.77	94.73	95.76	95.42	95.27	95.52	95.44
	8	80.15	93.65	95.52	95.61	95.46	95.75	94.95	95.2	95.77	95.62
	9	95.62	94.39	90.98	95.68	95.76	94.29	95.54	95.31	94.82	95.04
	10	81.11	95.23	93.25	95.43	95.27	95.33	95.4	95.48	95.29	95.08

Table 43: Variance of prediction accuracy results for overall cost-effective HI predictor

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	0.02	86.22	83.82	0.27	0.06	3.44	3.89	3.47	2.73	2.18
	2	0.38	14.25	1.6	35.97	0.12	0.23	0.16	3.69	0.2	3.58
	3	49.54	40.4	0.32	0.11	4.24	1.47	0.37	0.19	0.2	0.62
	4	102.66	0.07	0.23	0.69	0.15	0.24	1.76	0.73	0.68	0.47
	5	72.19	80.02	0.04	0.25	0.26	0.08	1.99	0.2	0.22	0.41
	6	0.04	0.15	0.72	0.85	0.15	0.03	0.07	0.16	0.19	0.53
	7	70.52	1.32	0.32	0.15	1.12	0.07	0.26	0.29	0.18	0.19
	8	62.33	2.43	0.14	0.23	0.14	0.09	0.2	0.65	0.26	0.24
	9	0.07	3.12	39.63	0.12	0.1	4	0.18	0.36	1.03	1.13
	10	26.21	0.45	21.61	0.48	0.77	0.09	0.67	0.66	0.29	0.56

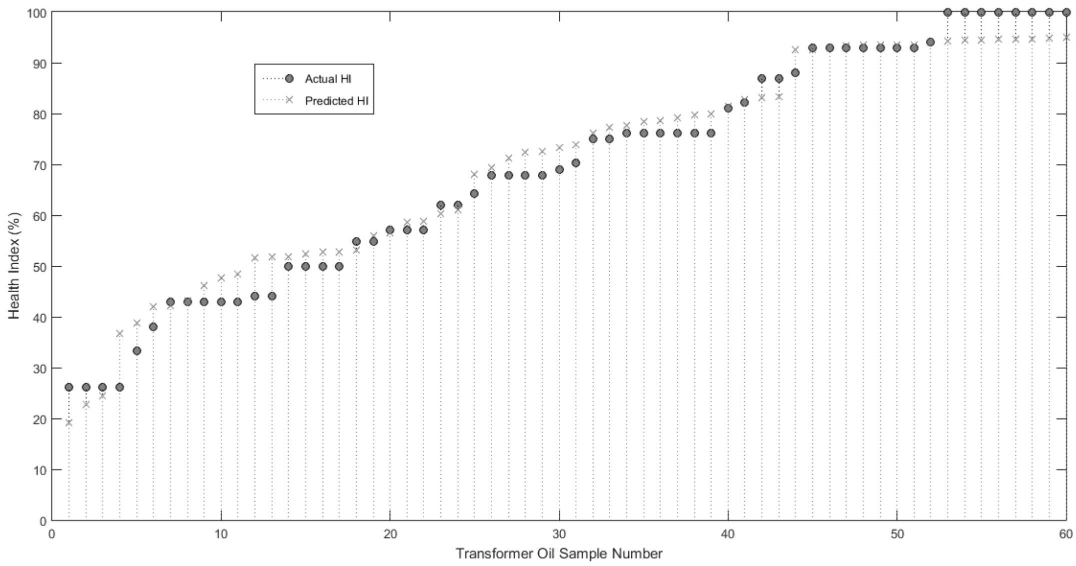

Figure 18: Actual vs. predicted HI for overall predictor model

Figure 17 indicates that Ethane and color are used in both the IFT and HI predictor model. Acidity however, is the additional feature which is only used in predicting the IFT value and not the HI. The forward stepwise procedure used earlier in this chapter indicated the elimination of acidity due to its relative redundancy in the presence of IFT. Thus an alternative cost-effective method would suggest to explore the use of acidity as an input to HI model rather than IFT. Accordingly the cost-effective model is modified as shown in Figure 19.

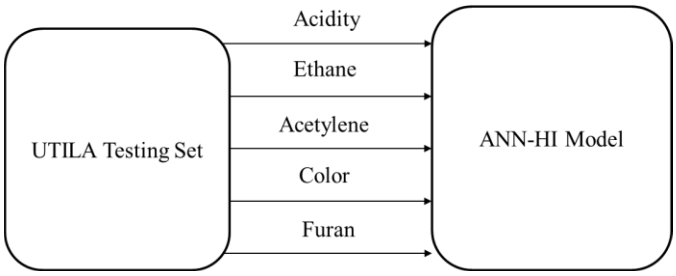

Figure 19: Alternative cost-effective HI predictor using acidity

Table 44 and Table 45 indicate the prediction accuracy results obtained for the alternative cost-effective model. The average prediction accuracy obtained is around 93% which is slightly less than the 95% accuracy obtained using the original cost-effective model with IFT as one of the reduced inputs. The obtained results validate the used of acidity as an alternative reduced feature to IFT in the cost-effective HI predictor model.

Table 44: Average prediction accuracy results for the modified cost-effective HI predictor

		Average Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	92.46	90.80	92.62	92.90	89.24	91.28	93.40	92.42	89.13	92.08
	2	92.80	93.04	92.84	90.44	93.14	92.69	83.11	93.09	94.12	92.71
	3	81.65	92.08	91.13	92.87	92.22	93.02	93.35	93.67	93.43	92.79
	4	89.37	93.04	93.69	93.77	93.06	93.78	93.51	93.60	93.95	93.57
	5	88.76	91.34	93.59	93.48	93.49	93.49	93.46	93.83	93.70	93.62
	6	92.16	93.74	93.66	93.76	93.50	93.46	93.71	93.47	92.73	93.33
	7	87.54	94.11	93.75	92.89	93.24	93.46	93.02	93.49	93.78	93.15
	8	93.31	91.28	93.27	91.93	93.83	93.80	93.12	93.54	94.29	93.59
	9	90.13	82.80	85.71	94.20	93.85	93.27	93.96	93.51	92.81	91.31
	10	91.96	93.67	93.82	93.90	93.84	93.34	93.66	93.80	93.40	93.29

Table 45: Variance of prediction accuracy results for the modified cost-effective HI predictor

		Variance of Prediction Accuracy of 10 Trials for a $N_1 \times N_2$ Neural Combination									
		N_2 Neurons in Second Hidden Layer									
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
N_1 Neurons in First Hidden Layer	*1*	0.18	38.70	3.34	0.94	72.44	26.55	0.22	1.17	8.16	5.82
	2	0.29	0.46	0.22	7.49	0.93	0.72	99.13	0.43	0.12	1.42
	3	72.43	3.46	38.44	0.61	8.42	1.18	0.39	0.26	0.15	0.43
	4	65.31	0.52	0.13	0.24	1.65	0.18	0.18	0.47	0.51	0.49
	5	46.14	37.92	0.22	0.79	0.30	0.48	3.81	0.23	0.46	0.14
	6	25.69	0.08	0.38	0.45	0.27	0.10	0.46	0.27	1.99	1.34
	7	80.62	0.27	0.35	3.61	1.43	0.59	0.88	0.12	0.31	0.97
	8	0.73	13.48	0.58	8.14	0.35	0.37	0.49	0.85	0.28	0.30
	9	16.08	257.03	54.09	0.07	0.20	0.73	0.39	0.36	0.58	7.19
	10	40.85	0.15	0.54	0.17	0.13	0.17	1.14	0.31	1.05	0.54

Chapter 5. Conclusion and Recommendation

Transformer Asset Management (TAM) is a planning method which involves monitoring and maintaining the energized transformer assets in the power system. The standard Condition Monitoring (CM) methods are used to acquire the condition parameters of the transformers in question. The Condition Assessment (CA) methods on the other hand, are used to process and analyze the CM outcomes to give a proper estimate of the transformers' expected lifetime. However, the individual CA outcomes are not correlated in the sense that they cannot collectively represent the transformer's overall operational conditions. Thus, the Health Index (HI) is used to provide a comprehensive understanding of the transformer condition based on a complete diagnostics of the individual components of the transformer. The computation of the HI is done through industry computational methods that are conducted by specialized TAM companies. However, the drawback of these methods lies in the need for all the CM test features for a high accuracy of HI value outcome.

5.1. Outcomes of the Thesis Work

The presented work provides a CM test-reduced HI predictor alternative that can effectively reduce the cost of TAM. The data used was for a number of transformer oil samples acquired from a utility company. The presented work started with applying Artificial Neural Networks (ANN) for predicting the HI (based on one industry computational method), using the CM test input features. Feature selection techniques that involve feature-based search and stepwise regression, were later used to eliminate the CM test features of least statistical significance in predicting the HI. The selected features of stepwise regression proved a validated success in predicting the HI. Another outcome of the presented work was the validation of the use of the built model (from one utility) for unseen data samples of a different utility company. Moreover, the selected features of stepwise regression can be considered as generalized features that can be used for modeling the HI predictor using the data of any utility company. The major outcome presented in this work is the idea of having a cost-effective HI predictor model, which uses a predicted costly feature. Such an outcome allows the utility companies to reduce the TAM cost of conducted tests, and to predict costly tests features using relatively cheaper ones.

5.2. Recommendations for Future Work

The presented work provides a cost-effective HI predictor model that can be used in the area of TAM. The model is based on one industry computational method that is practiced by the company mentioned in [26]. A continuation of the presented work in this thesis would be to follow all the conducted simulations for other industry computational methods. This should allow for a more generalized HI outcome, in which any utility company can choose the industry method that best fits its standards and requirements. Applying this work for further methods is considered as a major contribution in the area of TAM.

References

[1] X. Zhang and E. Gockenbach, "Asset Management of Transformers Based on Condition Monitoring and Standard Diagnostics," *IEEE Insulation Magazine,* vol. 24, no. 4, pp. 26-40, July 2008.

[2] M. Wang, A. J. Vandermaar and K. D. Srivastava, "Review of Condition Assessment of Power Transformers in Service," *IEEE Insulation Magazine,* vol. 18, no. 6, pp. 12-25, 2002.

[3] A. Abu-Elanien and M. Salama, "Asset Management Techniques for Transformers," *Electric Power Systems Research,* vol. 80, no. 4, pp. 456-464, 2010.

[4] AB Nynas Napthenics, Transformer Oil Handbook, Stockholm, 2004.

[5] A. d. Pablo, "Furfural and Ageing: How are they Related," in *IEE Colloquium on Insulating Liquids (Ref. No. 1999/119),* Leatherhead, 1999.

[6] S. Singh and M. N. Bandyopadhyay, " Dissolved Gas Analysis Technique for Incipient Fault Diagnosis in Power Transformers: A Bibliographic Survey," *IEEE Electrical Insulation Magazine,* vol. 26, no. 6, pp. 41 - 46, 2010.

[7] "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers," IEEE Std. C57.104-2008, 2009.

[8] "IEEE Guide for Acceptance and Maintenance of Insulating Oil in Equipment," IEEE Std. C57.106-2006, 2006.

[9] IEC 60422, "Mineral Insulating Oils in Electrical Equipment-Supervision and Maintenance Guidance," Edition 4, 2012.

[10] J. Singh and Y. R. Sood, "The Influence of Service Aging on Transformer Insulating Oil Parameters," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 19, no. 2, pp. 421 - 426, 2012.

[11] A. J. Kachler and I. Hohlein, "Aging of Cellulose at Transformer Service Temperatures. Part 1: Influence of Type of Oil and Air on the Degree of Polymerization of Pressboard, Dissolved Gases, and Furanic Compounds in Oil," *IEEE Electrical Insulation Magazine,* vol. 21, no. 2, pp. 15 - 21, 2005.

[12] M. Mirzaie, A. Gholami and H. Tayyebi, "Thermal Degradation of Cellulose Paper Insulation in Power Transformers," in *IEEE International Conference on Solid Dielectrics,* Winchester, UK, 2007.

[13] F. R. Barbosa, O. M. Almeida, A. P. S. Braga and C. M. Tavares , "Artificial Neural Network Application in Estimation of Dissolved Gases in Insulating Mineral Oil from Physical-Chemical Datas for Incipient Fault Diagnosis," in *15th International Conference on Intelligent System Applications to Power Systems, ISAP '09.*, Curitiba, 2009.

[14] A. Abu-Siada , S. Hmood and S. Islam, "A New Fuzzy Logic Approach for Consistent Interpretation of Dissolved Gas-In-Oil Analysis," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 20, no. 6, pp. 2343 - 2349, 2013.

[15] I. K. Annapoorani and D. B. Umamaheswari, "Fault Prediction Based on Dissolved Gas Concentration from Insulating Oil in Power Transformer Using Neural Network," in *IEEE 10th International Conference on the Properties and Applications of Dielectric Materials (ICPADM),* Bangalore, 2012.

[16] F. Zakaria, D. Johari and I. Musirin, "Artificial Neural Network (ANN) Application in Dissolved Gas Analysis (DGA) Methods for the Detection of Incipient Faults in Oil-Filled Power Transformer," in *IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Penang, 2012.

[17] A. K. Mehta, R. N. Sharma, S. Chauhan and S. Saho, "Transformer Diagnostics Under Dissolved Gas Analysis Using Support Vector Machine," in *International Conference on Power, Energy and Control (ICPEC),*, Sri Rangalatchum Dindigul, 2013.

[18] K. Shaban , A. El-Hag and A. Matveev, "A Cascade of Artificial Neural Networks to Predict Transformers Oil Parameters," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 16, no. 2, pp. 516 - 523, 2009.

[19] R. A. Ghunem, A. H. El-Hag, K. Assaleh and F. Al Dhaheri, "Estimating Transformer Oil Parameters Using Artificial Neural Networks," in *International Conference on Electric Power and Energy Conversion Systems,. EPECS '09.*, Sharjah, 2009.

[20] K. Assaleh and A. El-Hag, "Estimating Transformer Oil Parameters Using Polynomial Networks," in *International Conference on Condition Monitoring and Diagnosis, CMD 2008.,* Beijing, 2008.

[21] N. Abu Bakar and A. Abu-Siada, "A Novel Method of Measuring Transformer Oil Interfacial Tension Using UV-Vis Spectroscopy," *IEEE Electrical Insulation Magazine,* vol. 32, no. 1, pp. 7 - 13, 2016.

[22] M. Arshad, S. M. Islam and A. Khaliq, "Fuzzy Logic Approach in Power Transformers Management and Decision Making," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 21, no. 5, pp. 2343 - 2354, 2014.

[23] R. A. Ghunem, K. Assaleh and A. H. El-hag, "Artificial Neural Networks with Stepwise Regression for Predicting Transformer Oil Furan Content," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 19, no. 2, pp. 414 - 420, 2012.

[24] K. Benhmed , K. B. Shaban and A. El-Hag, "Cost Effective Assessment of Transformers Using Machine Learning Approach," in *IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA),* Kuala Lumpur, 2014.

[25] A. Abu-Siada, S. P. Lai and S. M. Islam, "A Novel Fuzzy-Logic Approach for Furan Estimation in Transformer Oil," *IEEE Transactions on Power Delivery,* vol. 27, no. 2, pp. 469 - 474, 2012.

[26] A. Jahromi, R. Piercy, S. Cress, J. Service and W. Fan, "An Approach to Power Transformer Asset Management Using Health Index," *IEEE Electrical Insulation Magazine,* vol. 25, no. 2, pp. 20 - 34, 2009.

[27] L. En-Wen and S. Bin, "Transformer Health Status Evaluation Model Based on Multi-Feature Factors," in *International Conference on Power System Technology (POWERCON),* Chengdu, 2014.

[28] A. E. B. Abu-Elanien, M. M. A. Salama and M. Ibrahim, "Calculation of a Health Index for Oil-Immersed Transformers Rated Under 69 kV Using Fuzzy Logic," *IEEE Transactions on Power Delivery,* vol. 27, no. 4, pp. 2029 - 2036, 2012.

[29] S. Forouhari and A. Abu-Siada, "Remnant Life Estimation of Power Transformer Based on IFT and Acidity Number of Transformer Oil," in *IEEE 11th International Conference on the Properties and Applications of Dielectric Materials (ICPADM),* Sydney, NSW, 2015.

[30] A. E. B. Abu-Elanien, M. M. A. Salama and M. Ibrahim, "Determination of Transformer Health Condition Using Artificial Neural Networks," in *International Symposium on Innovations in Intelligent Systems and Applications (INISTA),* Istanbul, 2011.

[31] A. D. Ashkezari, H. Ma, T. K. Saha and C. Ekanayake, "Application of Fuzzy Support Vector Machine for Determining the Health Index of the Insulation System of In-Service Power Transformers," *IEEE Transactions on Dielectrics and Electrical Insulation,* vol. 20, no. 3, pp. 965 - 973, 2013.

[32] E. Alpydin, *Introduction to Machine Learning,* MIT Press, 2004.

[33] Stanford University, *UFLDL Tutorial:Multi-Layer Neural Network* [Online]. Available: http://ufldl.stanford.edu/tutorial/supervised/MultiLayerNerualNetworks.

[34] T. Shanableh and K. Assaleh, "Feature Modelling Using Polynomial Classifiers and Stepwise Regression," *Neurocomputing,* vol. 73, no. 10-12, p. 1752–1759, 2010.

[35] D. Montgomery and G. Runger, *Applied Statistics and Probability for Engineers*, Wiley, 1994.

[36] "Mathworks, MATLAB User Guide," The MathWorks, Inc., 2016.

Vita

Alhaytham Alqudsi was born on December 9, 1989, in Sharjah, in the United Arab Emirates. He was educated in a private school and was of the top ranking students in the best three AS British curriculum subjects across the country in 2007. He was awarded the Dean's list recognition for six semesters in the American University of Sharjah, from which he graduated with cum laude, in 2011. His degree was a Bachelor of Science in Electrical Engineering.

Mr. Alqudsi worked as a site electrical engineer for three years at several contracting companies. In 2011, Mr. Alqudsi began a Master's program in Electrical Engineering at the American University of Sharjah.

Mr. Alqudsi is a student member of the IEEE organization and the Emirati Engineers Association.