Health Estimation of Electrical Transformer

Project report submitted to

Visvesvaraya National Institute of Technology, Nagpur in partial fulfillment of the requirements for the award of the degree

Bachelor of Technology
In
Electrical and Electronics Engineering

by

Ganesh Rohit (BT18EEE012)

Nishit Khandelwal (BT18EEE015)

Vinit Kavalekar (BT18EEE061)

Aadhithya Iyer (BT18EEE063)

Arihant Gaur (BT18EEE066)

Vyankatesh Muley (BT18EEE068)

under the guidance of

Dr. Ashwin Dhabale

Department of Electrical Engineering
Visvesvaraya National Institute of Technology
Nagpur 440 010 (India)

2022

Health Estimation of Electrical Transformer

Project report submitted to

Visvesvaraya National Institute of Technology, Nagpur in partial fulfillment of the requirements for the award of the degree

Bachelor of Technology
In
Electrical and Electronics Engineering

by

Ganesh Rohit (BT18EEE012)	**Nishit Khandelwal (BT18EEE015)**
Vinit Kavalekar (BT18EEE061)	**Aadhithya Iyer (BT18EEE063)**
Arihant Gaur (BT18EEE066)	**Vyankatesh Muley (BT18EEE068)**

under the guidance of
Dr. Ashwin Dhabale

Department of Electrical Engineering
Visvesvaraya National Institute of Technology
Nagpur 440 010 (India)

2022

Department of Electrical Engineering
Visvesvaraya National Institute of Technology, Nagpur

Declaration

We, **Ganesh Rohit, Nishit Khandelwal, Vinit Kavalekar, Aadhithya Iyer, Arihant Gaur, Vyankatesh Muley**, hereby declare that this project work titled **Health Estimation of Electrical Transformer** is carried out by us in the Department of Electrical Engineering of Visvesvaraya National Institute of Technology, Nagpur. The work is original and has not been submitted earlier whole or in part for the award of any degree/diploma at this or any other Institution / University.

Date: 05/05/2022

Sr. No.	Enrollment No.	Names	Signature
1	BT18EEE012	D Sri Ganesh Rohit	
2	BT18EEE015	Nishit Khandelwal	
3	BT18EEE061	Vinit Kavalekar	
4	BT18EEE063	Aadhithya Iyer	
5	BT18EEE066	Arihant Gaur	
6	BT18EEE068	Vyankatesh Muley	

Department of Electrical Engineering
Visvesvaraya National Institute of Technology, Nagpur

Certificate

This is to certify that the project titled **Health Estimation of Electrical Transformer**, submitted by **Ganesh Rohit, Nishit Khandelwal, Vinit Kavalekar, Aadhithya Iyer, Arihant Gaur, Vyankatesh Muley** in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology in Electrical and Electronics Engineering**, VNIT Nagpur. The work is comprehensive, complete and fit for final evaluation.

Dr. V. B. Borghate
Professor and Head,
Department of Electrical Engineering
VNIT, Nagpur
Date: 05/05/2022

Dr. Ashwin Dhabale
Assistant Professor
Department of Electrical Engineering
VNIT, Nagpur

ACKNOWLEDGEMENT

First and foremost, we express our sincere gratitude to our guide, Dr. Ashwin Dhabale, Assistant Professor, Department of Electrical Engineering, for his constant support and guidance during the course of the entire project. We were greatful for the suggestions received during our regular meetings and seminars. We learnt a great deal from his guidance.

We would like to thank Dr. Bhimrao Umre, Professor, Department of Electrical Engineering for his invaluable feedbacks during the progress seminars.

We would also like to thank the esteemed faculty members at the Department of Electrical Engineering, VNIT Nagpur for their colossal teaching throughout our graduation. We feel privileged to have interacted with these professors.

Lastly, we would like to sincerely thank our parents, for their constant support and help. A major portion of this thesis was written during the COVID - 19 lockdown. We are greatful to them for not only supporting us during our stay at home, but also for their constant encouragement.

ABSTRACT

Transformer is an essential component in the distribution power system. Therefore, any damage to the transformers will hamper the distribution of electricity towards the consumers. For example, high temperature in transformers can result in the degradation of transformer insulation, in turn causing its failure. This will also cause a disturbance in electrical power system and result in a major economic loss. To circumvent the disaster, transformers need to undergo preventive and reactive maintenance, until it is no longer efficient. Many evaluation methodologies have been investigated and developed to evaluate transformer health. These technologies will assist transformer operators in forecasting the distribution transformer's state and successfully responding.

This work reviews advances in estimating the health of a three phase distribution transformer. It also attempts to evaluate recent methods in health index estimation and monitoring, primarily related to machine learning. Finally, a dataset generation paradigm is discussed, which is also used for testing and comparing different approaches. This dataset is built upon a set of prescribed standards by various governing bodies, in the field of Electrical Engineering.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Energy has become a basic necessity today. We can't imagine a day without electricity. Be it a person working in an IT company or a farmer, everyone requires electricity. It is a necessary commodity for modern economies to survive. As a consequence, a healthy power system is critical for ensuring a steady supply of electricity.

Transformer plays an important role in transmission of electricity. It links the energy transmission from generation station to all the way to end users like us. But during the event, transformers may fail due to various reasons which results in utilities experiencing major loss such as loss of revenue and market backlash. These problems aren't limited to utilities only, it affects consumer as well. The consumer could face an electrical shortage or fluctuations in supply can damage the equipments installed. This could lead to shutdown of industries, hampering production and leading to unemployment. As a result Transformer Asset Management is of prime importance to prevent suddenly occurring failures of transformers.

A proper asset management will allow quality assessment of conditions and to develop future management strategies of transformers. But for that we first need to understand and identify the root cause of failures in transformers. Authors in [1] presented the statistical data of component failures from 350 transformers to establish a three-level model of failure mechanism, failure linkages, and failure modes. It was found out that the most critical to power transformer health is insulation with an incident rate of about 41%; then, components showing high failure rates are windings, 14%, bushings, 13%, and on-load tap changers at about 10%. Other components such as the cooling system,

Table 1.1: Number of power transformer failure per voltage population during 2009–2013

Years	400–230kV	230–110kV	110–66kV	110–33kV	110–22kV	110–11kV	66–11kV	33–11kV	Total Failures
2009	0	0	0	0	3	7	0	11	21
2010	0	1	1	9	9	6	0	15	41
2011	0	1	0	10	8	11	0	14	44
2012	0	0	0	10	7	8	0	16	41
2013	0	2	1	13	8	9	1	15	49
Total	0	4	2	42	35	41	1	71	196

core, and operational errors do not have a significant impact. This is shown in Figure 1.1.

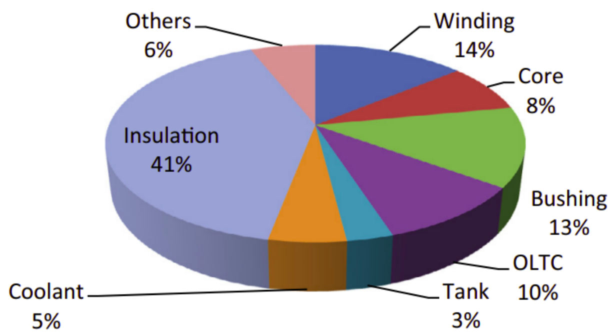

Figure 1.1: Failure statistics of power transformer component based failures [1]

Investigating data from transformers supervisors in utilities, component failures of transformers, failure rate and operation impact level is shown in Table 1.1

1.1 Condition Monitoring

Condition Monitoring (CM) is a field in electrical engineering, dealing with monitoring of parameters of a machine, to identify significant changes, that might be indicative of a developing fault. It is a major component of Predictive Maintenance (PdM). The monitoring of the electrical system can either be done offline or online.

1.1.1 Offline Condition Monitoring

Offline Condition Monitoring is the monitoring of a machine or production process under the outage condition i.e. whenever the equipment is in off condition. A shutdown is necessary for off line Condition Monitoring. The Offline Condition Monitoring analysis system will generate and collect data during start-up and shutdown of the installation. There are different kind of test that are carried out in offline condition monitoring which has been listed as follows:

1. **Routine Tests**: Routine tests are performed during manufacturing on all equipment after the active part assembly completed.

2. **Type Tests**: Type tests are tests which are made on equipment representative of other equipment to demonstrate that they comply with specified requirements not covered by routine tests.

3. **Special Tests**: Special Tests are test, other than Routine or Type tests, agreed between manufacturer and purchaser.

Beside these tests there are some other test which do not fall under any of the above mentioned categories and which are done at site:

1. Pre - commissioning Tests

2. Periodic/Condition Monitoring Tests

3. Emergency Tests

Offline test can be performed when the machine is not in working condition. It requires the electrical equipment to be in off state in order to be able to perform the test. Moreover, a number of test has to be performed, which requires more number of sensors, in order to get a clear picture of the health of electrical equipment. This increases the cost of condition monitoring which is not desired. So we go for online condition monitoring techniques.

1.1.2 Online Condition Monitoring

Online Condition Monitoring is the monitoring of a machine or production process under the starting as well as running condition i.e. whenever the equipment is in on condition. The Online Condition Monitoring analysis system will generate and collect data during starting and running or performing condition. Major components of online Condition Monitoring are:

1. Sensors for measurement of various parameters
2. Data Collection and Processing System for Decision making
3. Display/Execution System for Implementation of Decision

Our thesis deals primarily with online condition monitoring, using latest proposed approaches with machine learning. However, to shed some light on the most commonly use offline techniques, we move on to explaining some of them.

1.1.3 Sweep Frequency Response Analysis (SFRA)

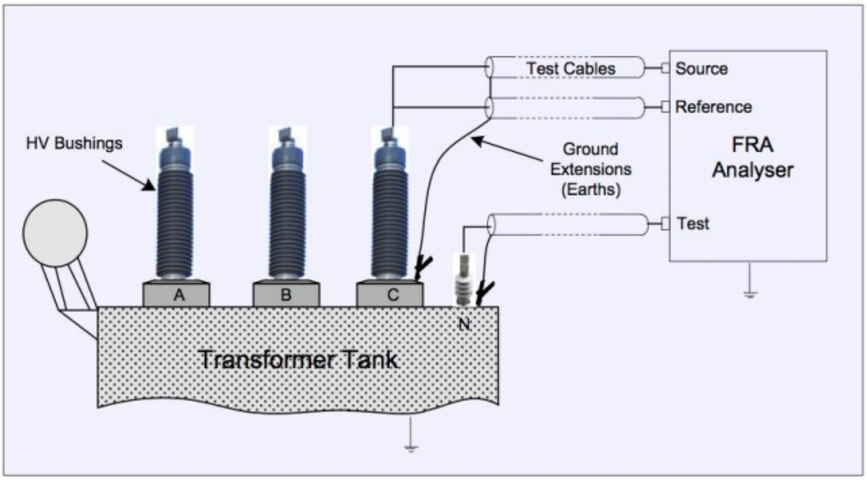

Figure 1.2: SFRA Analysis in Transformer

The windings of transformer or alternator rotors maybe exposed to mechanical stresses during unsafe transportation, lightning impulse, heavy short circuit faults, transient switching impulses, and DC component. SFRA Test is used to find out physical condition of transformer windings and Alternator Rotor's windings.

The SFRA idea is very simple and straightforward. As is well known, two parallel electrodes work as a capacitor even when the dielectric medium is air. However, in this case, the windings are insulated with paper and mica or varnish. Insulation will function as a dielectric medium. The winding itself causes capacitance between them, resulting in the formation of a distributed network. Furthermore, all electrical devices have certain resistance, inductance, and capacitance values, thus each of them can be thought of as a complex RLC circuit (Figure 1.3). Resistance, inductance, and capacitance should all

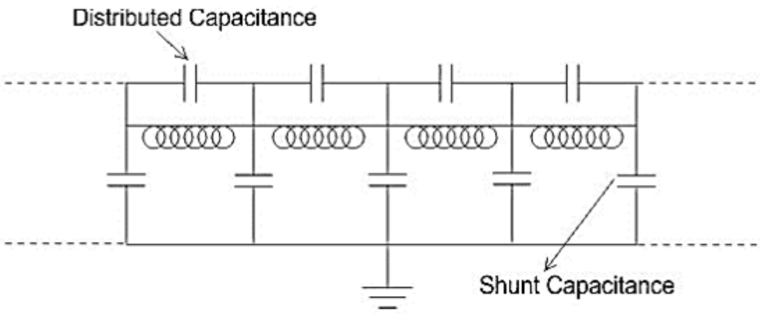

Figure 1.3: Equivalent Circuit as Interpreted by SFRA Test

be equal to zero. In reality, however, an ideal machine cannot be built. An equipment's resistance, inductance, and capacitance are not equal to zero.

The transformer is seen as a complex impedance circuit. Open (magnetization impedance) and Short (short-circuit impedance) responses are measured over a wide frequency range and the results are presented as magnitude response (transfer function) in dB. Changes in the impedance/transfer function can be detected and compared over time, between test objects or within test objects. The method is unique in its ability to detect a variety of winding faults, core issues and other electromechanical faults in one test.

Experience has shown that different sub-bands are dominated by different internal components of the transformer and are subsequently more sensitive to different types of failures. Table 1.2 shows some band frequencies and their interpretations.

A sample of the frequency spectrum obtained is shown in Figure 1.4.

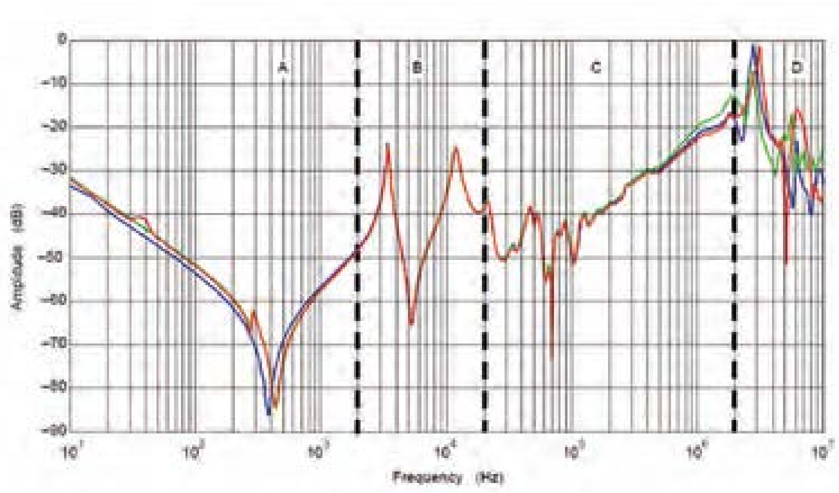

Figure 1.4: Equivalent Circuit as Interpreted by SFRA Test

More information regarding the test is available on IEC 60076-5: 'Power Transformer-

Table 1.2: Frequency Sub - band Sensitivity

Region	Frequency Sub - band	Component	Failure Sensitivity
1	<2 kHz	Main core bulk and winding inductance	Core deformation, open circuits, shorted turns and residual magnetism
2	2 kHz - 20 kHz	Bulk component and shunt impedances	Bulk winding movement between windings and clamping structure
3	20 kHz - 400 kHz	Main windings	Deformation within the main or top windings
4	400 kHz - 1 MHz	Main windings, top windings and internal leads	Movement of the main and top winding, ground impedance variations

Part 5: Ability to withstand Short Circuit' [2].

1.1.4 Dissolved Gas Analysis (DGA)

DGA is an examination of electrical transformer oil contaminants. Insulating materials (oil, paper, cotton tapes, press boards, wood, etc.) within electrical equipment liberate gases as they slowly break down over time.

Generally, the gases found in the oil in service are hydrogen (H2), methane (CH_4), Ethane (C_2H_6), ethylene (C_2H_4), acetylene (C_2H_2), carbon monoxide (CO), carbon dioxide (CO_2), nitrogen (N_2) and oxygen(O_2).

The composition and distribution of these dissolved gases are indicators of the effects of deterioration, such as pyrolysis or partial discharge, and the rate of gas generation indicates the severity.

By analyzing the volume, types, proportions, and rate of production of dissolved

Table 1.3: Key Gas Concentrations as per Different Kinds of Faults

Gas	Normal Limit (ppm)	Action Limit (ppm)	Potential Fault Type
Hydrogen (H_2)	150	1,000	Corona, Arcing
Methane (CH_4)	25	80	Sparking
Acetylene (C_2H_2)	15	70	Arcing
Ethylene (C_2H_4)	20	150	Severe Overheating
Ethane (C_2H_6)	10	35	Local Overheating
Carbon Monoxide (CO)	500	1,000	Severe Overheating
Carbon Dioxide (CO_2)	10,000	15,000	Severe Overheating
Total Combustibles (TDCG)	720	4,630	

gases, much diagnostic information can be gathered. Since these gases can reveal the faults of a transformer, they are known as "fault gases". Gases are produced by oxidation, vaporization, insulation decomposition, oil breakdown and electrolytic action. Table 1.3 shows the type of fault, associated with different gas concentrations.

As the value exceeds the 'Normal Limit', sample frequency should be increased with consideration given to planned outage in near term for further evaluation. If the value goes beyond the 'Action Limit', removal of transformer from the surface must be considered.

While DGA is categorized under offline condition monitoring, it is possible now to even perform this test in online condition.

More information regarding the test is available on IEC 60599 [3], IEEE C57.104 [4] and CIGRE 296 [5].

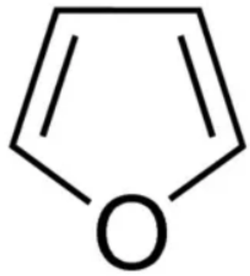

Figure 1.5: Chemical Structure of Furan C_4H_4O

Table 1.4: Furan Compounds against Different Kinds of Faults

Compound Name	Abbreviation	Potential Fault Type
5-Hydroxymethyl-2-furaldehyde	5H2F	Oxidation
Furfuryl Alcohol	2FOL	High Moisture
2-Furaldehyde	2FAL	Overheating, old faults
2-Furyl methyl Ketone	2ACF	Rare, lightening
5-Methyl-2-furaldehyde	5M2F	Local, severe overheating

1.1.5 Furan Analysis and Degree of Polymerization (DP)

Furan is a heterocyclic chemical molecule made up of a five-membered aromatic ring with four carbon atoms and one oxygen atom (Figure 1.5).

This chemical substance is generated when a cellulose molecule depolymerizes (breaks into smaller lengths or ring configurations). The overall DP of paper insulation can be deduced with a high degree of certainty by examining the quantity and kinds of furans present in a transformer oil sample. Furan types and concentrations in oil samples can also indicate abnormal stress in a transformer, whether it's intense, short-term overheating or long-term, general overheating. The type of defect caused by the presence of distinct chemical structures of the furan compound is shown in Table 1.4.

Degree of Polymerization (DP) test is a measure of length of cellulose chains. Paper which is used as an insulation in transformer. Its insulation degradation is accelerated by temperature, water and oxygen and byproducts formed are Water, Carbon Oxides, Furans, Alcohols, Acids and Aromatic and Aliphatic Hydrocarbons. A DP value of $800 - 1200$ is an indication of paper insulation of healthy transformer whereas DP value less than 200

indicates end of life criteria for paper insulation.

1.1.6 Thermography

The term "Thermography" refers to the capturing of thermal patterns and data emitted by an object with the use of an infrared imaging and measurement camera. An image is then produced with the camera that can give you data that is otherwise unattainable.

Infrared radiation is that portion of the electromagnetic spectrum that extends from the long wavelength, or red, end of the visible-light range to the microwave range (Figure 1.6).

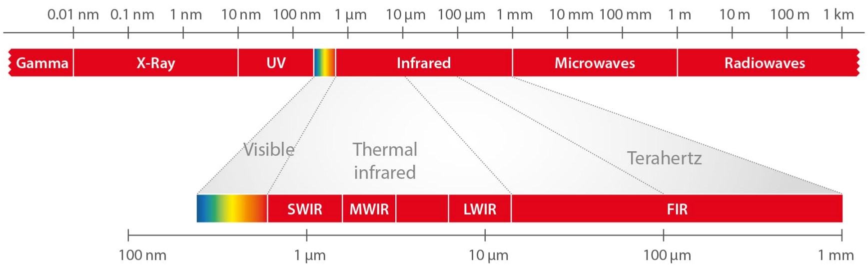

Figure 1.6: Electromagnetic Spectrum

Principle: Every object whose surface temperature is above absolute zero (0 K) radiates energy at a wavelength corresponding to its surface temperature. Utilizing our highly sensitive infrared cameras, it is possible to convert this radiated energy into a thermal image of the object being surveyed.

The primary goal of infrared thermography is to confirm machinery is running normally and to detect abnormal heat patterns within a machine, indicating inefficiency and defects. An increase in electrical resistance will cause local heating. The heat will be conducted away from the local resistance creating a thermal gradient, generally we can trace the hottest area and locate the anomaly. Typical hotspot is shown in Figure 1.7.

Normally, infrared photographs are colourized so that objects that emit more heat radiation than others seem brighter (yellow, red, and white). Cooler objects have darker blue, purple, or green hues. Although thermal imaging is typically used to detect only surface temperatures, infrared signals frequently show temperatures within structures.

We have seen different offline approaches for condition monitoring. These approaches

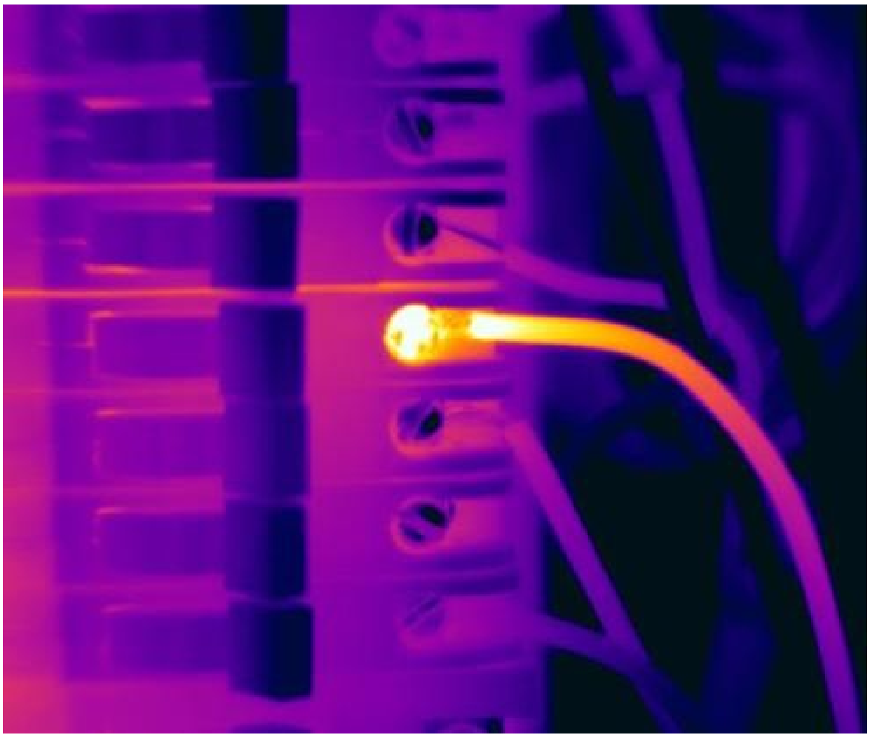

Figure 1.7: Hotspot Generation

have, in general, one thing in common. They require additional sensor stack. Thermography, for example, would require a specialized infrared camera, that is able to capture wavelengths emitted from the hotspot, and subsequently, register its temperature. Similarly, DGA would require spectroscopy equipment to determine the concentration for each gas component. There are two main disadvantages:

1. Offline condition monitoring requires the machine to be kept off the grid. This would mean loss of revenue for the end consumers, usual maintenance cycle would always require interruption of the power supply.

2. The above methods require specialized equipments, that are often expensive and come under additional costs.

With this view, our focus went towards online condition monitoring system (OCMS). At the same time, we planned to study whether it is possible to monitor the health of the electrical transformer, just using the basic parameters that can be obtained using common measurement devices.

The question still remains, as to how can the instantaneous condition of the transformer be quantified. This will be shown in the following section.

1.2 Health Index

There are many approaches in transformer asset management to tackle such issues and plan and prioritize the predictive maintenance of transformers. One of the approaches is a useful calculation technique known as Health Index (HI) calculation. This method not only allow us to plan maintenance strategies for transformers but also help us to identify risk and opportunities. Generally there are three parts of HI formulation which are Input, Computational Algorithm and Output or the interpretation of the calculation part. This has been summarized in Figure 1.8 [6].

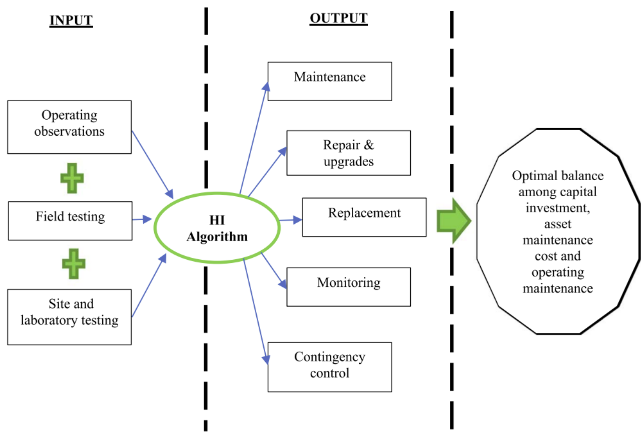

Figure 1.8: Number of power transformer failure per voltage population during 2009–2013

Different test have been conducted on transformers and on basis of that, inputs are being taken. Some have been discussed in the previous section. The input for HI is usually obtained from the operating observations, field, site and laboratory testing.

A lot of fomulations have been defined for Health Index. Ballal et. al. [7] represented health index as,

$$\mathrm{HI} = \frac{\sum_{c_i=1}^{n} S_{P_i} W_{P_i}}{S_{max} \sum_{c_i=1}^{n} W_{P_i}} \tag{1.2.1}$$

Where HI is the health index metric, S_{P_i} is the score of each assessment condition that is defined based on measured data, S_{max} is the maximum score of assessment condition, W_{P_i} is its corresponding weight and n are the number of such conditions.

Different parameters have different weightage according to the degree of importance given to any particular parameter that affects the condition of transformer [8, 9, 10, 11, 12]. Parameters like Load History and Power Factors have higher weightage as compared

Table 1.5: Weighing Factor of Various Transformer Parameters

Parameters	Weighting Factor
Dissolved Gas Analysis	10, 3
Load History	10
Power Factor	10
Global loss factor	10
Thermo Scan	10
Infrared Thermography	10, 8
Conductivity factor	10
Polarisation Index	10
Furanic Compound content	9, 6, 5
Oil Quality	8, 6
Overall transformer condition	8, 6
Leakage reactance	8
Winding resistance	8, 6, 2
Bushing Condition	7,5
Frequency Response Analysis	6
DGA of tap changer oil	6
Turns ratio	5,2
Tap changer contact condition	5
Overall LTC condition	5, 2
Age	4
Paper Insulation factor	4
Internal faults history	4
Dielectric Breakdown test	4
Water content test	3
Surge arrester	3
Cooling equipment condition	3
Tap changer oil quality	3, 2
Location	3
Main tank corrosion	2
Insulation Resistance test	2, 1
Core to ground connection	2
Oil leaks	2

to weightage given to Age or Location. Different researchers have considered different weightage for same factor. The number of parameters used in the calculation of HI is also different among researchers. This has been shown in Table 1.5.

Table 1.6: Output of Health Index as shown in [12]

HI	Condition	Action
85-100	Very good	Normal maintenance.
70-85	Good	Normal maintenance.
50-70	Fair	Increase the number of diagnostic tests, corrective maintenance or need of replacement, depending on the criticity.
30-50	Poor	Start planning the replacement process or repair, taking in account the risk.
0-30	Very Poor	Immediate risk assessment, replacement or repair, depending on the case.

Several mathematical equations or techniques for the formulation of HI have been proposed in prior works. Despite the fact that they use the same basic equation, certain adjustments have been made to make the equations more dependable and scientifically proven.

The final HI output will be subjected to a specified range, and appropriate preventive action will be performed. Researchers give varied ranges, and there is no standard for identifying the range and the preventive measure done. Two specific examples [12, 13] have been shown in Table 1.7 and 1.6.

The HI method have some limitations as well:

1. The accuracy depends heavily on weighted parameters

2. The condition monitoring may be costly and the results only reflect the preferences of the human-expert

3. Low accuracy for the systems and devices are controlled linguistically, or have a contradictory condition

To alleviate some of the issues, we attempt a thorough literature review of recent advances in health index estimation of transformer.

Table 1.7: Expected Lifetime based on Health Index as shown in [13]

Health Index	Condition	Description	Approximate Expected Lifetime
85-100	Very good	Some aging or minor deterioration of a limited number of components.	More than 15 years
70-85	Good	Significant deterioration of some components.	More than 10 years
50-70	Fair	Widespread significant deterioration or serious deterioration of specific components.	Up to 10 years
30-50	Poor	Widespread serious deterioration.	Less than 3 years
0-30	Very poor	Extensive serious deterioration.	At end of life

Chapter 2

Related Work

The monitoring system must make physical measurements and analyse the data in the context of specific environmental conditions in order to offer information about the transformer's state of health and detect incipient faults. They are crucial for asset management because they assist in identifying, prioritising, and scheduling essential capital and maintenance spending.

Effective ways for monitoring the state and health of distribution transformers could help utilities avoid failures and degradation before they happen. This will increase dependability and lower the cost of power service in the long run. It's especially crucial now since distributed solar, electric vehicles, and other energy resources are rapidly changing the grid's operation and putting additional strain on service transformers. Since then, a variety of methodologies have been utilised to estimate the distribution transformer's health. We present a list of approaches for estimating the transformer's health that have been published in the literature.

2.1 Health Index Calculation

Health index (HI) calculation is a valuable methodology; it is the most basic method used to develop transformer maintenance strategies, according to [14]. This method converts the representative indexes of the transformer's operation and statement into a quantitative index and evaluates the transformer's overall condition. A health index calculation method is used to completely examine the distribution transformer conditions. The transformer's statement is graded from "perfect health" to "very poor condition." In Section 1.2, a generic formulation of HI is offered.

2.2 Thermal Model and Loss Life Calculation

Aging or deterioration of the transformer(insulation) is closely related to temperature, humidity level and the amount of oxygen in the air. The factor which greatly affects the life of the transformer is temperature. Temperature distribution in transformer is not uniform, so to find the point whose temperature is the most influencing, a hot spot or the hottest point on a transformer is used. The aging of the transformer can be evaluated using HST (Hot spot temperature). The TOT(Top oil temperature) rise in a transformer depends on the AT(Ambient temperature). HST(Hot spot temperature) cannot occur at a fix point or location or by any fixed empirical formula. It can be estimated by multiplying the temperature gradient with hot spot location distance from top oil point and adding to top oil temperature gives HST. The increase in the TOT value, which is also an increase in HST, has an effect that can reduce the insulation life of the transformer. Abnormal conditions, such as overloading, supplying non-sinusoidal loads, and the influence of high ambient temperatures, can accelerate the aging of the transformer. So it can be concluded that the increase in TOT and HST can shorten the life of the transformer. When the transformer is energized and loaded with ambient temperature (AT), the dissipation caused by core loss, winding loss, and stray losses in the tank, harmonics present as well as metal support structure, as sources of heat, will cause oil temperature and winding temperature to increase. To take the harmonic distortion into account, the power loss is modified with the current factor. The core density, core dimensions, frequency and voltage is taken into account to calculate the flux density. Top oil temperature is calculated by the method of heat balance and heat stored method and according to these two factors, Top oil temperature and thereby the Hot spot temperature is obtained. Based on the TOT, the aging factor FAA is calculated as,

$$FAA = \exp^{\frac{15000}{383} - \frac{15000}{\theta_h + 273}} \tag{2.2.1}$$

The loss of life factor is incorporated over a period of time to evaluate the insulation heating effect. Where FAA has a value larger than 1 for winding hottest-spot temperatures greater than $110°C$ and less than 1 for temperatures less than $110°C$. Since insulation aging is a cumulative effect, the percent loss of life per day is the summation of the percent loss of life [15].

2.3 Fuzzy Logic

Fuzzy logic has been offered as a feasible way to address the constraints of the health index calculating method. Fuzzy logic is intended to be used for representing vague notions and unclear information, particularly when standard logic techniques are ineffective. A comprehensive fuzzy control system has three steps: fuzzification, inference, and defuzzification. Fuzzification calculates fuzzy values from exact values at the input in the first stage. To determine the fuzzy value for the output, the fuzzy inference employs all available fuzzy rules. The defuzzification process extracts the exact output value from the fuzzy result generated in the fuzzy inference phase [16].

2.4 Machine Learning

2.4.1 Neural Networks

An artificial neural network is a collection of neurons, connected to other neurons in a layerwise fashion. Each connection is assigned a weight, which is determined by training the neural network, using training data and a validation set, to check for overfitting (that is, the network should be generalizable to other scenarios, rather than memorizing the training set). In this particular case, the inputs to the neural network can be various transformer parameters. The paper [17] proposes the use of ANNs , with four parameters as input (voltage, load current, oil level and oil temperature). There are two hidden layers (with logsig and purelin as the nonlinearities respectively), with the output layer denoting the health status of transformer. The method provided an accuracy of 97.2% on data provided by MSEDCL on a 15 kVA, $400/400$V three phase distribution transformer. Their experiment setup is shown in Figure 2.1 and their network architecture in 2.2. As with online condition monitoring systems, the setup expects an energized transformer. The input voltage and load current are obtained from the trivector energy meter. Temperature and oil level are obtained from their respective sensing devices. This is fed to the ANN OCMS system, that converts the continuous signals into digital form.

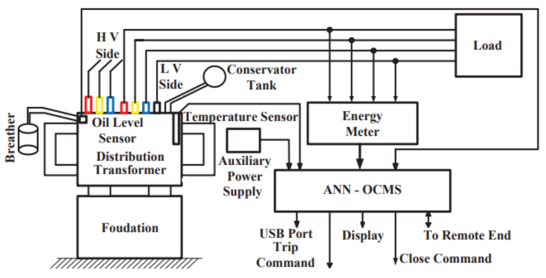

Figure 2.1: FBD of ANN based OCMS

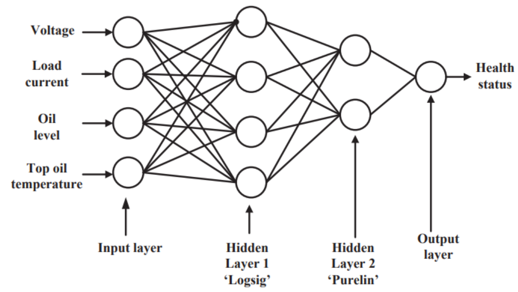

Figure 2.2: ANN Architecture

Another work estimates the health of transformer using two sets of ANNs [18], with input parameters to the first network being DGA for gases, Furan, insulation power factor, O & M (Oil and Maintenance), and age. The output from first network is fed to the second network, along with a few other parameters like turns ratio test, short circuit impedance, DC winding resistance, FRA and degree of polymerization. Test results showed an accuracy of a max of $92.4\ \%$, on a transformer with primary voltage as 150 kV, under MSEDCL Indonesia. Its architecture is shown in Figure 2.3.

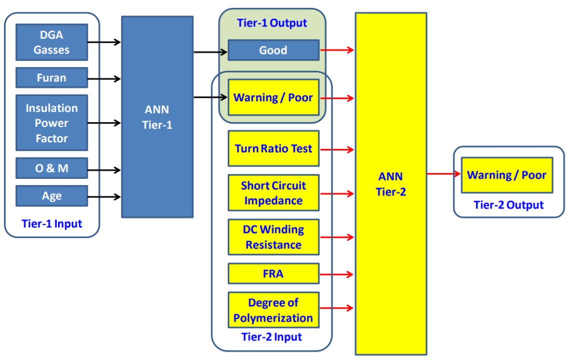

Figure 2.3: 2 - Tier MLP Network

2.4.2 kNN (k-Nearest Neighbour)

The kNN algorithm is a supervised learning technique that has been used in many ML applications. It classifies objects based on the closest training examples in the feature space. The idea behind kNN is to find a predefined number of training samples closest in distance to a given query instance and predict the label of the query instance from them.

An example is been shown in Figure 2.4. There are three clusters shown, ω_1, ω_2 and ω_3. The class assigned to each data point is the cluster to which it belongs. This is made by measuring the distances of one data point from other data points. If that distance falls below a set threshold, then it can be considered to be a part of the same class as the anchor class. So, if an unknown class is given to us, it is possible to determine its class by measuring the distances to this point from the centroids of the clusters. It would be assigned the class from which its distance is the least.

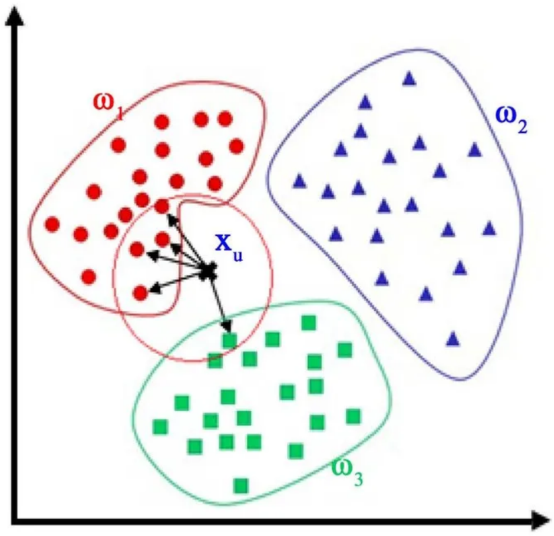

Figure 2.4: An example of kNN [19]

2.4.3 Support Vector Machines (SVMs)

The objective of the support vector machine [20] is to find a hyperplane in the N-dimensional(N is the number of features) space that distinctly classifies the data points. It is generally used for binary classification problems. For Multi class classification problems the SVM is used by breaking the multi class problem into individual binary classification problems.

In the context of health index estimation, the problem can be considered as a regression problem (SVR). A small illustration has been shown in Figure 2.5. Note that these points are in a higher dimensional space, just shown in 2D for visualization. The goal is to determine the 'decision boundary', at some set distance from the regression plane, such that points within the boundary lines can be considered to lie on this plane. This can be considered as a margin of tolerance. Lower its value, better will be the fit of the hyperplane.

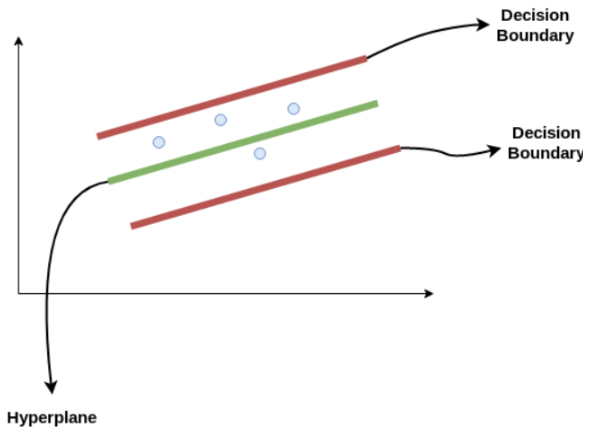

Figure 2.5: Support Vector Regression

2.4.4 Decision Trees

Decision Tree is a Supervised learning technique that can be used for both classification and Regression. A decision tree is a tree-like structure composed of internal nodes, branches, and leaf nodes. where internal nodes contain dataset properties, branches represent decision rules, and each leaf node represents the result. Each internal node represents a test on an attribute, each branch represents the result of a test, and each leaf node represents a class label. The root node is the tree's highest node. A decision tree is a class discriminator that separates the training set recursively until each partition is wholly or primarily composed of samples from one class. Several papers [21, 22] presented this technique for health estimation of transformers and compared it with techniques like random forest, support vector machines and kNN.

2.4.5 Random Forest

Random forest [23] is a classification and regression ensemble learning approach that involves merging numerous classifiers to tackle a complex problem and enhance the model's performance. Random Forest is a classifier that combines a number of decision trees on different subsets of a dataset and averages the results to increase the dataset's predicted accuracy. Instead of depending on a single decision tree, the random forest collects the forecasts from each tree and predicts the final output based on the majority votes of predictions. The bigger the number of trees in the forest, the more accurate it is and the problem of overfitting is avoided. This method has been used to estimate the health of transformer in [21]. This particular paper aims to gauge the health of a service transformer (in the form of four categories - good, fair, poor and very poor) from loading profile, vibration and oil temperature. The output is obtained from a machine learning algorithms, particularly support vector machine, decision trees, random forest and k - Nearest Neighbors. The inference is that Random Forest Algorithm performed the best, along with kNN.

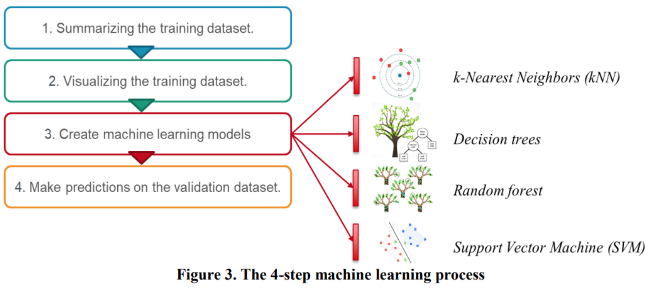

Figure 2.6: Algorithms used in [21]

2.5 Linear Regression

In linear regression, the dependent variable is a linear combination of the parameters. It has been widely used for predictive modeling. The equation of linear regression can be written as follows:

$$y = \theta_1 x_1 + \theta_2 x_2 + ... + \theta_n x_n \tag{2.5.1}$$

Where y is the dependent variable, x are the independent variables and $\theta_i, i = 1, ..., n$ are the coefficients. These are the weights assigned to to independent variables.

The overall goal of linear regression is to find the line of best fit (Figure 2.7) to the data points. For a given hypothesis line, it is possible to find the residuals (distance of the data points from the line) and minimize it. There are three ways of quantifying these residuals:

1. Sum of Residuals $\sum(Y - h(X))$. It allows positive and negative errors to be cancelled.

2. Sum of absolute value of residuals $\sum|(Y - h(X))|$. The use of absolute value prevents the positive and negative error to be cancelled.

3. Sum of square of residuals $\sum(Y - h(X))^2$. This method ensures that higher values of error are penalized more. This is the most used residual.

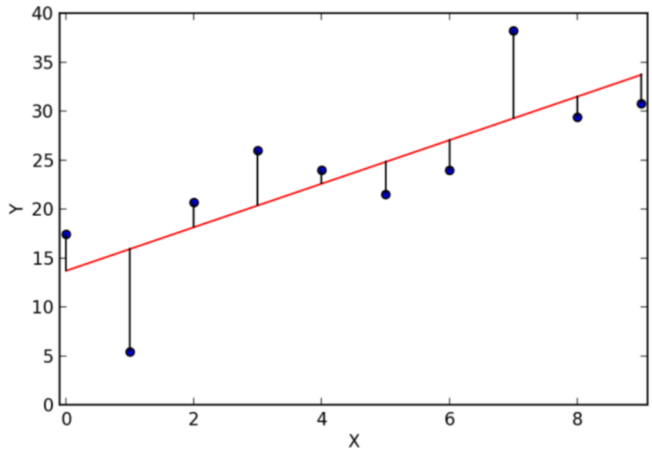

Figure 2.7: Line of Best Fit for 2D Data [24]

For each point, a corresponding residual value can be obtained. Adding it for all the points in the data, we get required cost function.

$$J(\theta_1, ..., \theta_n) = J(\Theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})^2 \tag{2.5.2}$$

Where m is the number of data points, h is the hypothesis linear function to be determined, parameterized by $\theta_1, \theta_2, ..., \theta_n$. Notice that the equation is the same is sum of squared residuals, except the addition of $\frac{1}{2m}$ to ease the mathematics.

The cost function can be minimized using a technique known as 'Gradient Descent'. It takes into account the instantaneous slope at a point, and then updates the weights, along the direction of the gradient, in proportion to its magnitude. It can be represented as,

$$\theta_i := \theta_i - \eta\frac{\partial J}{\partial \theta_i} \tag{2.5.3}$$

Where J is the cost function, η is known as learning rate. This formula updates the parameter θ_i. Similar thing will be done for all parameters.

Care must be taken to choose an appropriate value of learning rate. Keeping a high learning rate would mean that gradient descent would take larger steps, and might miss the minima, resulting in further divergence (overshoot). Having a lower learning rate would mean that step size would be lower, therefore more number of iterations. Therefore, an optimal value of η is to be required.

2.6 Hidden Markov Model

Hidden Markov Models (HMM) can be used to transform various data collected from substation equipment into failure probabilities. A quantitative decision tool based on these failure probability might be developed and utilised in system-level simulation and testing. One of the prediction approaches that may be used to determine the future states of transformers based on HI is the Markov Model (MM). The Markov decision process is a memoryless process that uses a probabilistic estimate to anticipate the future condition of equipment. This paper [25] discusses the use of hidden markov model for prediction of health of transformer.

Chapter 3

Dataset

3.1 Overview

For the field test of transformer, many organizations such as Maharashtra State Electricity Distribution Company (MSEDCL) [26], National Electrical Manufacturers Association of USA (NEMA) [27], Asset Management and Health Assessment Consulting Company (AMHA) [28], etc., have compiled a lot of data from many transformers (both power and distribution). However, much of their data is restricted to individuals affiliated with the respective organizations. To alleviate the issue of lack of data, we propose a novel dataset for distribution transformer, along with the health index for each instant.

We acquire the electrical specifications from Maharashtra State Electricity Distribution Company Limited (MSEDCL), for various distribution transformers, classified on the basis of their rating. We use three sets of distribution transformers. Their nameplates are shown in Figure 3.1, 3.2 and 3.3.

Additionally, we consider BEE (Bureau of Energy Efficiency) star ratings of 3 and 5. Therefore, we get a total of six scenarios, for different power and BEE star ratings. This would impact the constant power loss occurring in the transformer. Higher the BEE rating, lesser will be the loss. Relevant standards have been set by Maharashtra DISCOM, pertaining to power loss corresponding to a particular 'kVA' range and BEE rating can be seen in [26].

We propose the transformer setting in the city of Nagpur, for a period of ten years (2011 - 2020), where readings will be taken every three hours. This means the number of

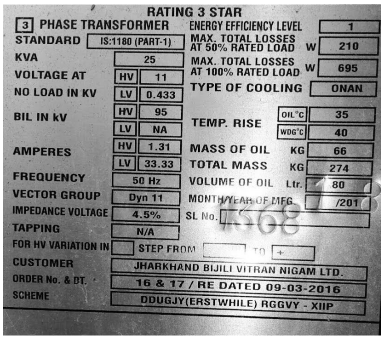

Figure 3.1: Transformer 1

Figure 3.2: Transformer 2

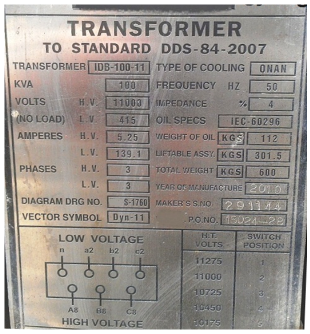

Figure 3.3: Transformer 3

readings for this setting will be $29,224$ per transformer.

3.2 Ambient Temperature

We obtain the ambient temperature, from the data open - sourced by the Indian Meteorological Department (IMD) [29]. The data is presented as a binary file, consisting of grids in the map of India, as shown in Figure 3.4.

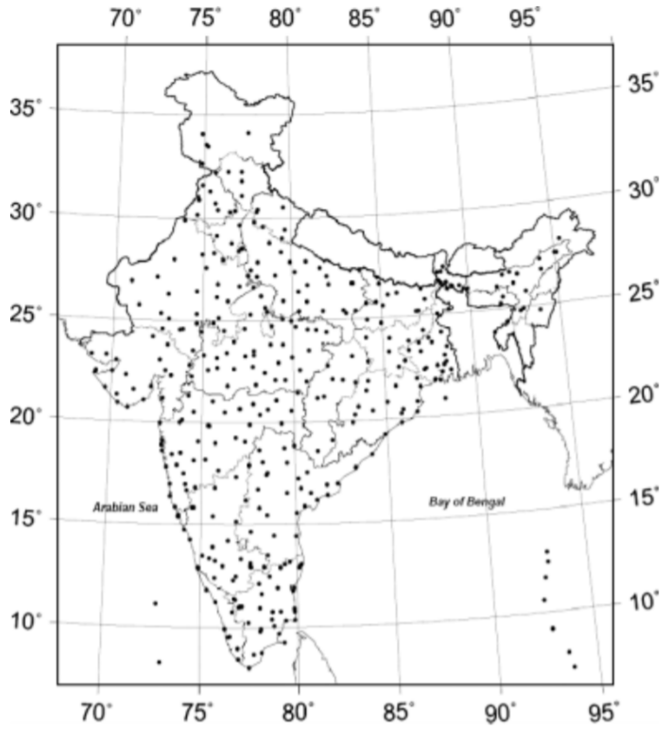

Figure 3.4: Map of India divided in grids for Temperature Measurement

We convert the gridded binary file to a text file, having maximum and minimum temperatures for each area. We consider the area in and around Nagpur (we take Lat. 21.50°N, Long. 79.50°E) and consider temperatures for the month of April.

The data provided by IMD consists of the maximum and minimum temperature for each day. Therefore, we linearly interpolate to ensure we have a total of 8 readings per day. Given two points (x_1, y_1) and (x_2, y_2), we can obtain the reading for the point between them as,

$$y = y_1 + (x - x_1)\frac{y_2 - y_1}{x_2 - x_1} \tag{3.2.1}$$

Intuitively, this represents traversing a line segment between two points (x_1, y_1) and (x_2, y_2). Therefore for a day, we have two temperature readings, maximum and minimum temperature.

3.3 Oil Temperature Rise Calculation

We assume the transformers has a cold start (the initial temperature rise is zero, that is the temperature of transformer is the same as ambient temperature). As per IEEE Standard C57.91-1995, [30, 31], we can obtain the thermal capacity of transformer as,

$$C_{DT} = 0.0272W_c + 0.0272W_t + 7.305\Theta_{oil} \tag{3.3.1}$$

Where W_c represents weight of core and coil assembly (in kilograms), W_t is the weight of tank and fittings (in kilograms) and Θ_{oil} represents the volume of oil (in litres).

The general expression of top oil time constant is,

$$\tau_{TO} = \tau_{TO,R}\frac{\left(\frac{\Delta\Theta_{TO,U}}{\Delta\Theta_{TO,R}}\right) - \left(\frac{\Delta\Theta_{TO,i}}{\Delta\Theta_{TO,R}}\right)}{\left(\frac{\Delta\Theta_{TO,U}}{\Delta\Theta_{TO,R}}\right)^{\frac{1}{n}} - \left(\frac{\Delta\Theta_{TO,i}}{\Delta\Theta_{TO,R}}\right)^{\frac{1}{n}}}, \tau_{TO,R} = C\frac{\Delta\Theta_{TO,R}}{P_{T,R}} \tag{3.3.2}$$

Where $\tau_{TO,R}$ is the top oil time constant at rated load, $\Delta\Theta_{TO_U}$ represents ultimate top oil rise over ambient temperature for load L, $\Delta\Theta_{TO,i}$ represents initial top oil rise over ambient temperature, $\Delta\Theta_{TO,R}$ is the ultimate top oil rise over ambient temperature at rated load, C is thermal capacity and $P_{T,R}$ is the power loss at rated load. Here, n represents the exponent of heat loss q in the expression $\Delta\Theta_{TO} = kq^n$. If the assumption of direct proportionality holds for heat loss, $n = 1.0$,

$$\tau_{TO} = \tau_{TO,R} \tag{3.3.3}$$

The top-oil temperature rise at a time after a step load change is given by the following exponential expression,

$$\Delta\Theta_{TO} = (\Delta\Theta_{TO,U} - \Delta\Theta_{TO,i})\left(1 - \exp\left(\frac{-t}{\tau_{TO}}\right)\right) + \Delta\Theta_{TO,i} \tag{3.3.4}$$

For each time step, the initial and ultimate temperature rise can be obtained as,

$$\Delta\Theta_{TO,i} = \Delta\Theta_{TO,R}\left[\frac{K_i^2 R + 1}{R+1}\right]^n, \Delta\Theta_{TO,U} = \Delta\Theta_{TO,R}\left[\frac{K_U^2 R + 1}{R+1}\right]^n \tag{3.3.5}$$

Where R is the ratio of load loss at rated load to no load loss, K is the ratio of specified load to rated load. From Section 3.1, we have the transformer loss at different loadings (50% and 100%). The ratio R can be determined as,

$$R = \frac{K-1}{0.25-K} + 1; K = \frac{P_{T,R/2}}{P_{T,R}} \tag{3.3.6}$$

Where $P_{T,R/2}$ is the loss at half the rated load. Similarly, K can be obtained as,

$$K = \frac{I}{I_r} \tag{3.3.7}$$

Where I_r is the rated current and I is the instantaneous current per phase. For the three phases, the value of K can be averaged across them.

3.4 Loading

The most variable parameter for any transformer is the load. The scenario we have taken load variation for all three transformers. Load profiles for three different scenarios have been shown in Figure 3.5, 3.6 and 3.7.

The load variation for the residential load has been designed taking into account that during the evening hours, the power consumption is relatively high than afternoon or the

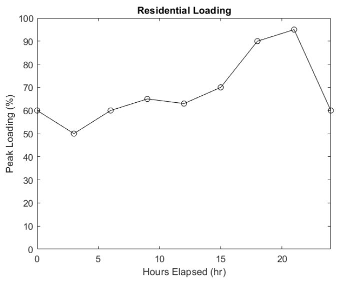

Figure 3.5: Residential Load %

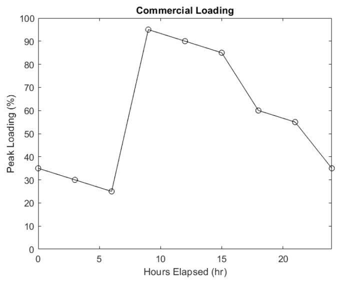

Figure 3.6: Commercial Load %

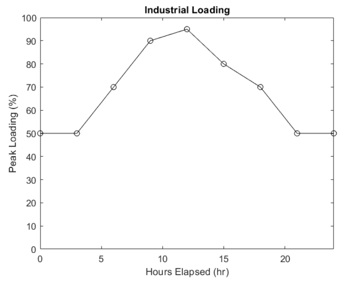

Figure 3.7: Industrial Load %

morning hours. However, during the early morning hours, the load is quiet low as the daily chores are almost over are yet to start and only a few constantly running appliances are consuming the power.

For the commercial load, profile has been designed taking into account the average market rush trends. Therefore, during morning working hours, the load is high. During the late evening hours, there is a maximum demand of power for lighting as well for the increased consumer count.

In the industrial load, taking into account the maintainable and shutdown work during the midnight, and the main production during the day hours, the load profile has been made.

These load variations have directly been applied in the current profile for all three phases. Along with this the random Gaussian noise has been added in this characteristics. The noise has been added in only one phase and other two phases have been designed in such a way that the total deviation from the mean/ rated current value remains zero.

To take into account the error in the measurement of the transformer parameters, the we have added random values in between the rated current with deviation as mentioned above. For these we have chosen around 1% of the total data points at a random have added abruptly high and/or low values at this points as depicted below. At the same time, a total of 10 fault locations are added at random locations, where the load collapses to zero.

3.5 Oil Leakage Factor

It is always observed that during the working tenure of any machine, there is leakage in its component liquid constituents. May it be insulation oil or be it fuel. In case of our three transformers, to consider the oil leakage, we have added a function that updates itself with each iteration and hence degrades oil volume and eventually thermal constant over the period of time.

Now the loading and ambient temperature data being ready with us we update the top oil temperature rise with every iteration and accordingly update the transformer temptation sire and the transformer temperature as per the previously explained equations.

3.6 Per Phase Current and Voltage Calculation

We have the rated voltage per phase as $0.433/\sqrt{3}$kV on the LV side. To incorporate variation in loading, we add Gaussian Noise to the resultant current from the given power rating and LV side voltage. The random sequence is generated by sampling from a normal distribution, and added to voltage for each phase. Following [7], we also calculated the unbalanced voltage V_U. As per NEMA, the unbalanced voltage is defined as,

$$V_U = \frac{\sigma_{\max}(V_{ab}, V_{bc}, V_{ca})}{\mu(V_{ab}, V_{bc}, V_{ca})} \tag{3.6.1}$$

Where $\sigma_{\max}$ is the maximum deviation and μ is the mean of line voltages. The unbalance in harmonics and voltages results in the unbalance and harmonics in currents. This can increase the core, copper and eddy current losses.

To incorporate degradation with time, we increase the variance of the sampling distribution, so that the randomness increases with time.

We also calculate current and voltage variation for three phase system [14].

$$\Delta I_{DT} = \sqrt{\frac{1}{3}(|I_a - I_r|^2 + |I_b - I_r|^2 + |I_c - I_r|^2)} \tag{3.6.2}$$

$$\Delta V_{DT} = \frac{1}{3}\left(\left|\frac{U_r - U_a}{U_r}\right| + \left|\frac{U_r - U_b}{U_r}\right| + \left|\frac{U_r - U_c}{U_r}\right|\right) \tag{3.6.3}$$

3.7 Health Index Estimation

Since dataset will require a ground truth, we propose to determine a ground truth health index for each reading. This will help in the supervised training for learning based methods.

For the calculation of health index, the following parameters are utilized: Phase voltages and currents, top oil rise temperature with ambient temperature, transformer loading and 3 phase current variations. This accounts for 11 different characteristics.

Define the data matrix X of size 29224×11.

We used the health index values on a yearly basis for the transformer from [25]. We then linearly interpolate the values. Additionally, we penalize the places where the transformer is abnormally overloaded (1% of the total data points). Also, the HI value is slashed to zero for 10 potential fault points. A more refined statistical approach to estimate ground truth health index using data points could be a future scope of this work.

3.8 Data Visualization

All the experiments have been performed using MATLAB, except ambient temperature acquisition, which was done using Python. A subset of the readings have been shown here.

3.8.1 Phase Voltages

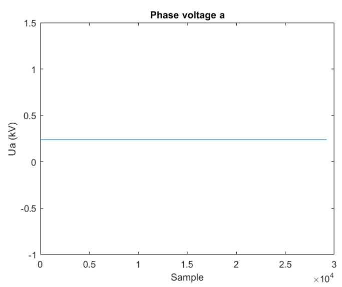

Figure 3.8: Phase Voltage a

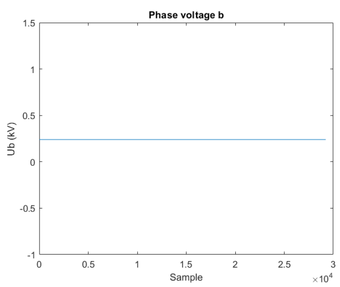

Figure 3.9: Phase Voltage b

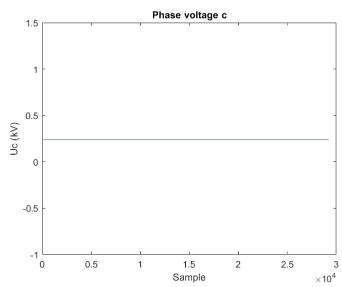

Figure 3.10: Phase Voltage c

3.8.2 Phase Currents

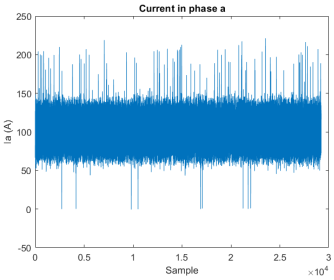

Figure 3.11: Phase Current a

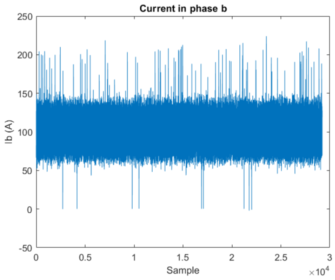

Figure 3.12: Phase Current b

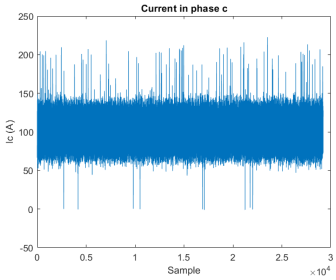

Figure 3.13: Phase Current c

3.8.3 Transformer Oil Temperature

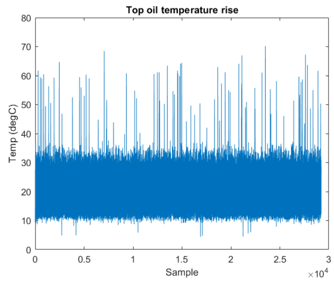

Figure 3.14: Top Oil Rise

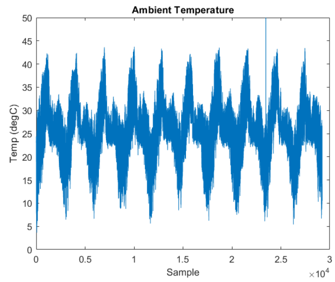

Figure 3.15: Ambient Temp.

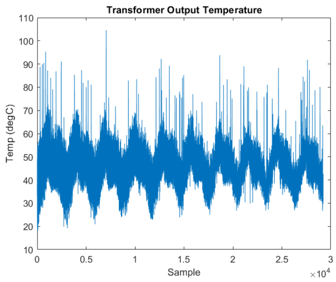

Figure 3.16: Oil Temperature

3.8.4 Miscellaneous

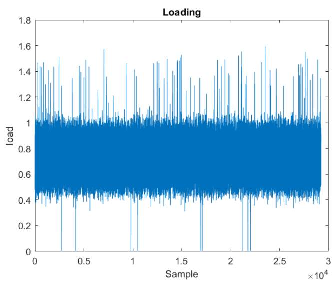

Figure 3.17: Loading

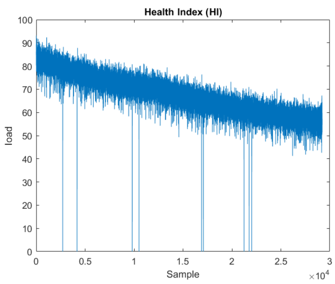

Figure 3.18: Health Index

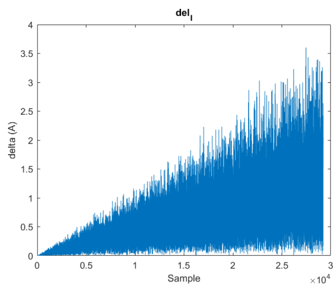

Figure 3.19: Current Deviation

Chapter 4

Results and Evaluation

With the obtained dataset, it is possible to evaluate different approaches under Chapter 2. The dataset for all six cases of a single distribution transformer is considered overall as a single set. The set is divided into train and test set, in the ratio $7 : 3$ (70% allocated for training and the rest for testing).

4.1 Benchmark Details

In order to compare different methods, we use Coefficient of Determination. It provides a measure for goodness of a fit of model. In other words, it tells us how well the regression line fits the actual data. Mathematically, it can be represented as,

$$R^2 = 1 - \frac{\text{SSR}}{\text{SST}} = 1 - \frac{(y_i - \hat{y}_i)^2}{(y_i - \overline{y})^2} \tag{4.1.1}$$

Where SSR is the sum squared regression (sum of residuals) and SST is total sum of squares. The range of R^2 is between 0 and 1.

It is possible to obtain negative values of R^2, that is, trend of the data will be converse of the trend observed by the regression model. Negative values imply that the data fitting is extremely poor. An example is shown in Figure 4.1. From this, we can infer that having a value close to 1 is required. It would mean that the regression line is fitting the dataset well.

In real - life, however, getting R^2 close to 1 is not possible, due to the presence of outliers in the data. As shown in the previous chapter, the proposed synthetic dataset

Table 4.1: Coefficient of Determination R^2 using Neural Network

Method*	25kVA	63kVA	100kVA
NN + Linear	-0.2538	0.2542	0.2450
NN + Logsig	0.5642	0.7452	0.7056
DTR	0.6029	0.5740	0.6234
RFR	0.7511	0.7228	0.8066
LiR	0.7946	0.7966	0.7993
RR	0.7946	0.7966	0.7993
LaR	0.7943	0.7946	0.7974

**NN: Neural Network, DTR: Decision Trees Regression, RFR: Random Forest Regression, LiR: Linear Regression, RR: Ridge Regression, LaR: Lasso Regression*

incorporates fault and overloading conditions as well. Therefore, such points can never be fitted to the regression line. Nevertheless, we hope to get R^2 as high as possible. Subsequent sections would elaborate on the implementation and results obtained from different methods.

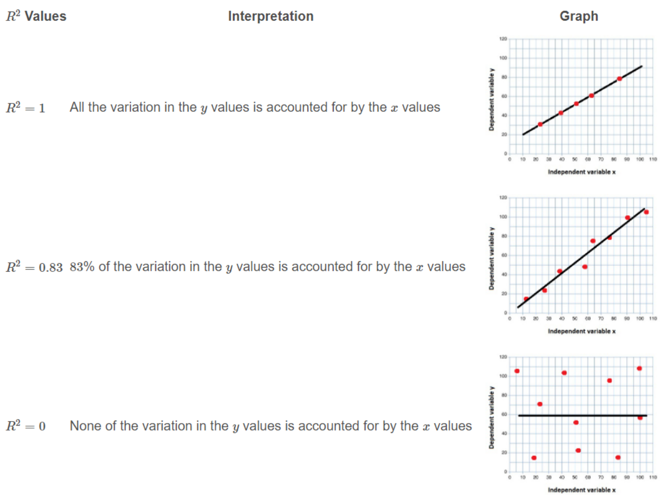

Figure 4.1: Coefficient of Determination Cases

A summary of the results is shown in Table 4.1.

4.2 Neural Network

We have considered a two layered neural network. The input is the 11 - dimensional input vector, representing the current state of the transformer. The output is a single value, representing the health index. There is an intermediate layer having 50 neurons. The model architecture is shown in Figure 4.2.

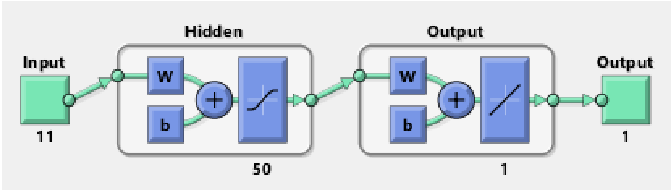

Figure 4.2: Neural Network Architecture

The training data consists of $1,22,740$ data points. A total of $5,000$ epochs are set and training is done using Gradient Descent. Total training time is close to 5 minutes, per transformer. The results are shown in Figure 4.3, 4.4 and 4.5.

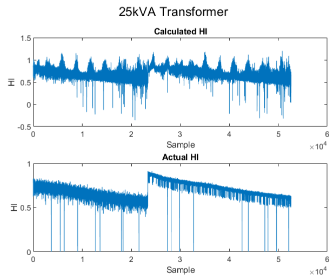

Figure 4.3: NN: 25kVA

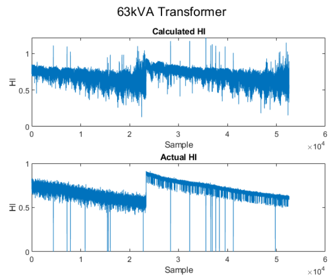

Figure 4.4: NN: 63kVA

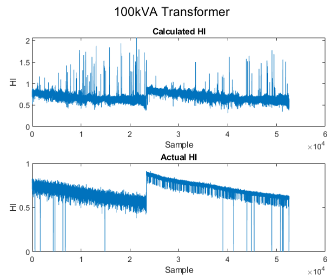

Figure 4.5: NN: 100kVA

The actual health index, shown on the bottom half of the graph, represents two regions of the transformer. The first half represents the declining health of transformer, during its final stage. The second half represents the health of a newly commissioned transformer. The upper half represents the health index calculated by the proposed method.

It can be seen that neural network is not able to model faults properly. Additionally, in Figure 4.5, health index values go way past 50% of the maximum possible value.

To alleviate this, we used a sigmoid function. Given a value x, its sigmoid can be represented as,

$$\text{sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{4.2.1}$$

MATLAB represents this function using 'logsig'. The overall network can now be shown as in Figure 4.6.

Training parameters are the same as before. The results are shown in 4.3, 4.4 and 4.5.

The issues are similar to what is observed without 'logsig'. The faults aren't being detected. At the same time, while the methods was able to learn the overall trend of the graph, it cannot detect the appropriate range of health index values.

The coefficient of determination for the two cases are shown in Table 4.1. It can be

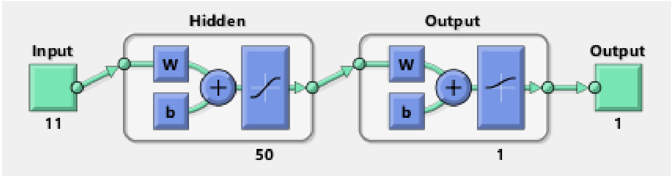

Figure 4.6: Neural Network Architecture with Sigmoid Function. Notice the Second Layer

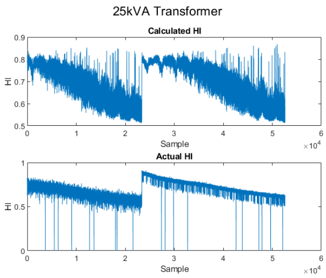

Figure 4.7: NN: 25kVA

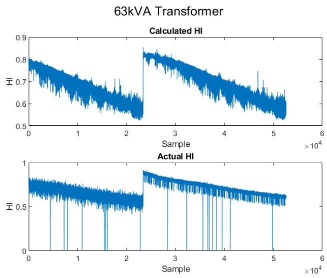

Figure 4.8: NN: 63kVA

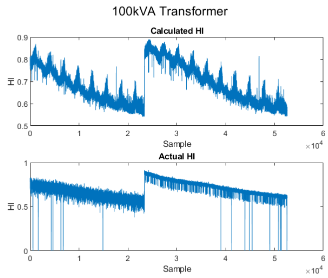

Figure 4.9: NN: 100kVA

seen that using sigmoid ensures that the health index values remain between 0 and 1. This improves the overall R^2 value drastically.

4.3 Support Vector Regression

We have implemented SVR with python. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is two, then the hyperplane is just a line and if the number of input features is three, then the hyperplane becomes a 2-D plane. A library, known as sklearn, has been used for this purpose. We used the Radial Basis Function (RBF) kernel. The coefficient of kernel has been determined as,

$$\gamma = \frac{1}{n * \text{var}(X)} \tag{4.3.1}$$

Where n represents number of features and var(X) represents variance of the input data. To avoid overfitting (discussed in Linear Regression), we used L2 regularization, with a coefficient of 1. The results for the same are shown in Figure 4.10, 4.11 and 4.12.

Notice that the overall health index obtained doesn't show stochastic nature. This is due to the fact that SVR fits the plane in a higher dimensional space. Therefore, only the points in and around the plane are registered. Therefore, all the points are obtained

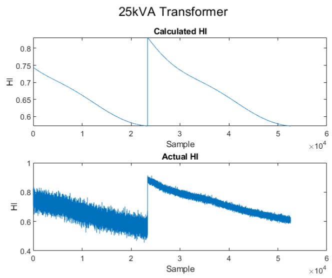

Figure 4.10: SVR: 25kVA

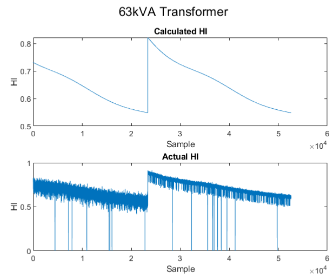

Figure 4.11: SVR: 63kVA

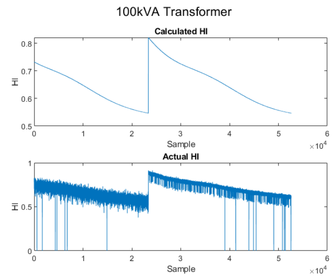

Figure 4.12: SVR: 100kVA

Table 4.2: Resultant Decision Tree for Different Transformers

Characteristic	25kVA	63kVA	100kVA
No. of Leaves	$1,19,308$	$1,19,562$	$1,19,530$
Depth	70	74	78

against the fitted line.

4.4 Decision Trees

Training is done using python, with sklearn library. After training, the number of leaves, along with the depth of the tree are registered, shown in Table 4.2. Results are shown in Figure 4.13, 4.14 and 4.15.

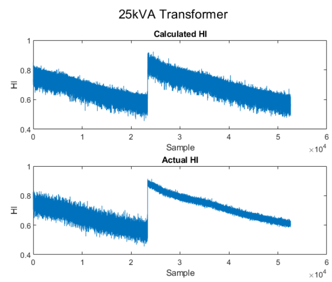

Figure 4.13: DT: 25kVA

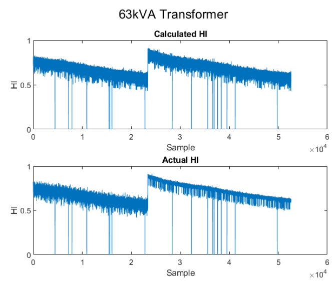

Figure 4.14: DT: 63kVA

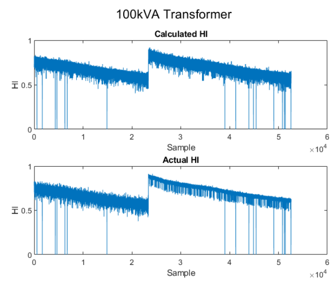

Figure 4.15: DT: 100kVA

While the R^2 score isn't as high as neural network with sigmoid, it is able to determine faults perfectly. Surprisingly, the fault detection accuracy of decision trees is 100%. So while the method is more accurate, the reason for lesser R^2 is lesser precision. The output is comparatively more noisy than the ground truth.

4.5 Random Forest

Similar setting for random forest is used as for decision trees. The number of trees in the forest are set to 100. Results are shown in Figure 4.16, 4.17 and 4.18.

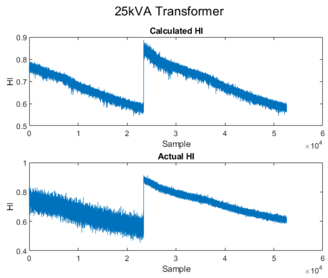

Figure 4.16: RFR: 25kVA

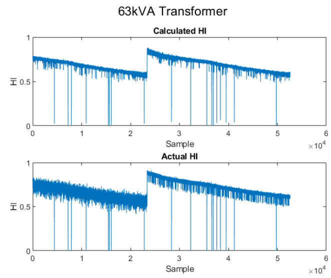

Figure 4.17: RFR: 63kVA

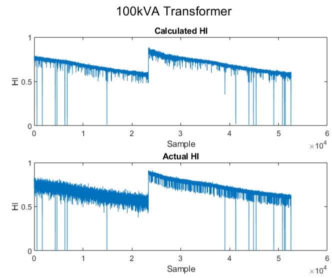

Figure 4.18: RFR: 100kVA

The results are similar to decision trees as well, however the noise is much lower. Additionally, the R^2 value is much higher (Table 4.1), even more then neural network based approach. The fault detection accuracy is again 100%. It will later be inferred that random forest is the overall best method in terms of coefficient of determination.

4.6 Linear Regression

As discussed earlier, Linear Regression will be used to determine 11 parameters, each representing linear combination of the input feature. The training will be done using gradient descent. Implementation is done using MATLAB. The results are shown in Figure 4.19, 4.20 and 4.21.

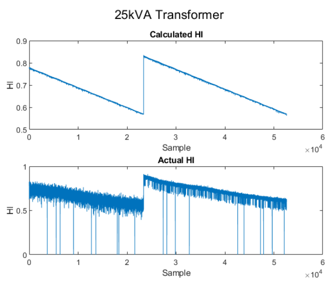

Figure 4.19: LiR: 25kVA

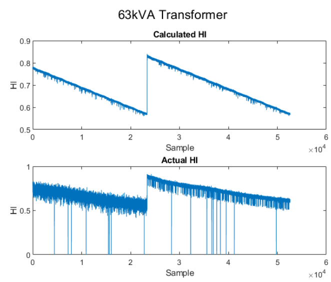

Figure 4.20: LiR: 63kVA

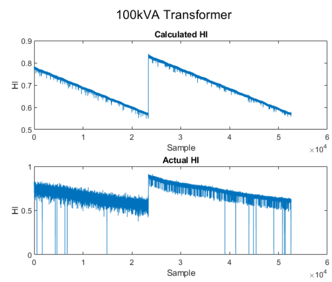

Figure 4.21: LiR: 100kVA

Additionally, we try two different variations of linear regression, namely ridge and lasso regression.

4.6.1 Ridge Regression

Before understanding ridge regression, it is important to understand regularization. Often there is a condition, where the algorithm has learnt the data too much. However, the problem is that the method would not generalize well to new set of data. A model is supposed to learn the general trend of the data, rather than memorizing the entire dataset and fitting each and every one of the points. In machine learning, such a condition is known as overfitting. It is not easy to exactly pinpoint where exactly the model starts to overfit, as the training progresses.

To alleviate this, there is a technique known as regularization, where additional loss term is added to the cost function. Mathematically, it can be represented as,

$$J(\theta_1, ..., \theta_n) = J(\Theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})^2 + \lambda||\Theta||_2 \tag{4.6.1}$$

Where λ is the penalty factor, indicating the contribution of the regularization factor. It should be important to set the λ parameter right. Having the value too high would imply that the model would underfit, which is the converse of overfitting (model doesn't fit the data properly).

Ridge regression, as shown in Equation 4.6.1, uses L2 regularization technique. We use $\lambda = 1.0$. Results for the same are shown in Figure 4.22, 4.23 and 4.24.

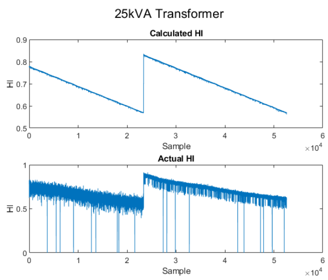

Figure 4.22: RR: 25kVA

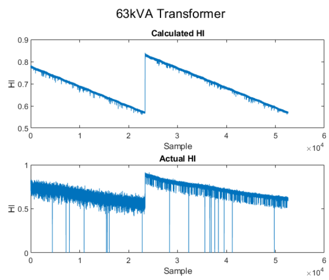

Figure 4.23: RR: 63kVA

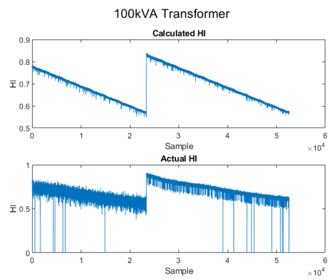

Figure 4.24: RR: 100kVA

4.6.2 Lasso Regression

Lasso (or LASSO) stands for Least Absolute Shrinkage Selector Operator. It is quite similar to ridge regression, but it uses L1 regularization technique.

$$J(\theta_1, ..., \theta_n) = J(\Theta) = \frac{1}{2m}\sum_{i=1}^{m}(h_\Theta(x^{(i)}) - y^{(i)})^2 + \lambda||\Theta||_1 \tag{4.6.2}$$

Similar to ridge regression, we use $\lambda = 1.0$. Its results are shown in Figure 4.25, 4.26 and 4.27.

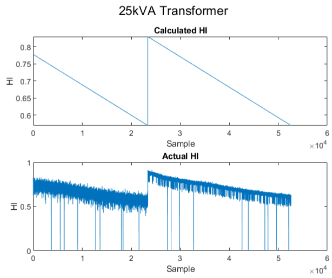

Figure 4.25: LaR: 25kVA

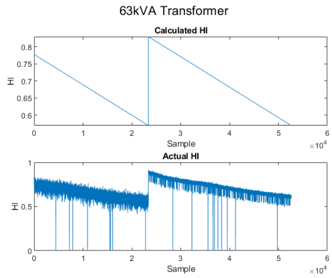

Figure 4.26: LaR: 63kVA

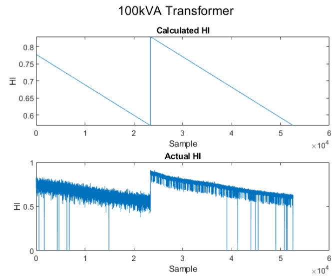

Figure 4.27: LaR: 100kVA

As it turns out, using regularization would not significantly change R^2. On the other hand, the use of Lasso regression improves the health index graph and makes it smoother.

From the above results, we can imply that while linear, ridge and lasso regression boast higher coefficient of determination, it is random forest that is able to detect all of the faults, while being the best correlation amongst random forest and decision trees.

Chapter 5

Conclusion and Future Work

From the previous chapter, we draw the following conclusions:

1. Neural Network gives good R^2 value, and is able to give a general overview about the health index of transformer. However, it isn't precise.

2. Using 'logsig' transfer function instead of the linear output improves the results to a great extent. However, the problem mentioned in first point still arises.

3. Support Vector Regression and Lasso Regression give the smoothest health index curve with time.

4. Decision Trees and Random Forest are able to model 100% of the faults. No other method was able to do this.

5. Linear Regression and its variations obtain the best fit to the model. However, their design restricts them from detecting outliers, specifically faults.

6. Overall, we can conclude that Random Forest is the optimal approach for detecting faults, and estimating an optimal range of health index.

In this work, we propose a new synthetically created dataset using the transformer parameters, based on standards. We also manage to test our dataset on a series of machine learning based approaches. Despite our best efforts to fabricate a dataset and modeling it closer to real transformer, it is not immune to certain flaws. First is the artificial addition of noise in the data. While this will give a certain randomness to the data, it is certainly not the most accurate representation of the actual working of transformer. Second is the health index calculation scheme, which can certainly have potential areas of improvement, particularly the use of refined statistical approaches. There can be further investigation that can be done on this particular aspect, since current approaches mostly rely on the opinion of experts, which incorporates subjectivity in the grading of transformer health.

Due to the pandemic, a majority portion of our time was spent in fabrication of the dataset and testing on different models. We feel that it is possible to extend this work to a real time transformer, in the form of Online Condition Monitoring System (OCMS). Additionally, this approach can be extended further for time series analysis, anomaly detection and forecasting.

Bibliography

[1] Raji Murugan and Raju Ramasamy. "Failure analysis of power transformer for effective maintenance planning in electric utilities". In: *Engineering Failure Analysis* 55 (2015), pp. 182–192.

[2] *IS 2026-5 (2011): Power Transformers, Part 5: Ability to Withstand Short Circuit [ETD 16: Transformers]*. URL: https://law.resource.org/pub/in/bis/S05/is.2026.5.2011.pdf.

[3] *Mineral oil-impregnated electrical equipment in service – Guide to the interpretation of dissolved and free gases analysis (IEC 60599-2015)*. URL: https://webstore.iec.ch/publication/23323.

[4] *IEEE Guide for the Interpretation of Gases Generated in Mineral Oil-Immersed Transformers)*. URL: https://standards.ieee.org/ieee/C57.104/7476/.

[5] *Recent developments on the interpretation of dissolved gas analysis in transformers (CIGRE 296)*. URL: https://e-cigre.org/publication/296-recent-developments-on-the-interpretation-of-dissolved-gas-analysis-in-transformers.

[6] A Azmi et al. "Evolution of transformer health index in the form of mathematical equation". In: *Renewable and Sustainable Energy Reviews* 76 (2017), pp. 687–700.

[7] Makarand S Ballal et al. "Online condition monitoring system for substation and service transformers". In: *IET Electric Power Applications* 11.7 (2017), pp. 1187–1195.

[8] A Naderian et al. "An approach to determine the health index of power transformers". In: *Conference Record of the 2008 IEEE International Symposium on Electrical Insulation*. IEEE. 2008, pp. 192–196.

[9] Muhammad Arshad, Syed M Islam, and Abdul Khaliq. "Power transformer asset management". In: *2004 International Conference on Power System Technology, 2004. PowerCon 2004.* Vol. 2. IEEE. 2004, pp. 1395–1398.

[10] J Haema and R Phadungthin. "Development of condition evaluation for power transformer maintenance". In: *4th International Conference on Power Engineering, Energy and Electrical Drives.* IEEE. 2013, pp. 620–623.

[11] Bogdan Gorgan et al. "IS 2026-5 (2011): Power Transformers, Part 5: Ability to Withstand Short Circuit". In: *Annals of the University of Craiova, Electrical Engineering Series* 34 (2010), pp. 13–18.

[12] M Augusta Martins. "Condition and risk assessment of power transformers: a general approach to calculate a Health Index". In: *Ciência & Tecnologia dos Materiais* 26.1 (2014), pp. 9–16.

[13] IG N Satriyadi Hernanda et al. "Application of health index method for transformer condition assessment". In: *TENCON 2014-2014 IEEE Region 10 Conference.* IEEE. 2014, pp. 1–6.

[14] Gajanan C Jaiswal et al. "Genetic algorithm–based health index determination of distribution transformer". In: *International Transactions on Electrical Energy Systems* 28.5 (2018), e2529.

[15] Sarah Afifah et al. "Prediction of Power Transformers Lifetime Using Thermal Modeling Analysis". In: *2019 IEEE International Conference on Innovative Research and Development (ICIRD).* IEEE. 2019, pp. 1–6.

[16] Abdel Ghani Aissaoui and Ahmed Tahour. "Application of fuzzy logic in control of electrical machines". In: *Fuzzy Logic-Controls, Concepts, Theories and Applications* 28 (2012), p. 428.

[17] Gajanan C Jaiswal, Makarand Sudhakar Ballal, and DR Tutakne. "ANN based methodology for determination of distribution transformer health status". In: *2017 7th International Conference on Power Systems (ICPS).* IEEE. 2017, pp. 133–138.

[18] Prasetiyono Hari Mukti, Feby Agung Pamuji, and Buyung Sofiarto Munir. "Implementation of artificial neural networks for determining power transfomer condition". In: *5th International Symposium on Advanced Control of Industrial Processes (ADCONIP2014), Hiroshima, Japan*. Citeseer. 2014, pp. 473–477.

[19] Wenhao Zhang et al. "Machine learning approaches to predicting company bankruptcy". In: *Journal of Financial Risk Management* 6.04 (2017), p. 364.

[20] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.

[21] Quynh T Tran, Kevin Davies, and Leon Roose. "Machine learning for assessing the service transformer health using an energy monitor device". In: *IOSR J. Electr. Electron. Eng* 15 (2020), pp. 1–6.

[22] Alhaytham Alqudsi and Ayman El-Hag. "Application of machine learning in transformer health index prediction". In: *Energies* 12.14 (2019), p. 2694.

[23] Leo Breiman. "Random forests". In: *Machine learning* 45.1 (2001), pp. 5–32.

[24] *A comprehensive beginners guide for Linear, Ridge and Lasso Regression in Python and R*. URL: https://www.analyticsvidhya.com/blog/2017/06/a-comprehensive-guide-for-linear-ridge-and-lasso-regression/.

[25] Muhammad Sharil Yahaya et al. "Estimation of transformers health index based on the Markov chain". In: *Energies* 10.11 (2017), p. 1824.

[26] *Maharashtra State Electricity Distribution Co. Ltd.* URL: https://www.mahadiscom.in/en/home/.

[27] *National Electrical Manufacturers Association of USA*. URL: https://www.nema.org/.

[28] Ahmed EB Abu-Elanien, MMA Salama, and M Ibrahim. "Calculation of a health index for oil-immersed transformers rated under 69 kV using fuzzy logic". In: *IEEE Transactions on Power Delivery* 27.4 (2012), pp. 2029–2036.

[29] AK Srivastava, M Rajeevan, and SR Kshirsagar. "Development of a high resolution daily gridded temperature data set (1969–2005) for the Indian region". In: *Atmospheric Science Letters* 10.4 (2009), pp. 249–254.

[30] Bernard C Lesieutre, Wayne H Hagman, and JL Kirtley. “An improved transformer top oil temperature model for use in an on-line monitoring and diagnostic system”. In: *IEEE Transactions on Power Delivery* 12.1 (1997), pp. 249–256.

[31] Dejan Susa, Matti Lehtonen, and Hasse Nordman. “Dynamic thermal modeling of distribution transformers”. In: *IEEE transactions on power delivery* 20.3 (2005), pp. 1919–1929.