Conditional Based Health Monitoring of Transformer

By

Muneeb Islam

NUST00000170779MSMEE2016

Supervisor

Dr. Sajjad Haider Zaidi

Department of Electronics and Power Engineering

Pakistan Navy Engineering College(PNEC)

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

April 2020

Conditional Based Health Monitoring of Transformer

By

Muneeb Islam

NUST00000170779MSMEE2016

Supervisor

Dr. Sajjad Haider Zaidi

G.E.C Committee

Dr. Bilal Muhammad Khan

Dr. Lubna Moin

A thesis submitted in partial fulfillment of requirements for the the degree of

Master of Science (MS) in **Electrical Engineering**

With specialization in Control

Department of Electronics and Power Engineering

Pakistan Navy Engineering College(PNEC), Karachi

National University of Sciences and Technology (NUST), Islamabad

April 2020

Declaration

I, *Muneeb Islam* declare that thesis titled âĂIJConditional Based Health Monitoring of Transformer" also, the work introduced are my own particular and has been created by me accordingly my own unique research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a Master of Science degree at NUST
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at NUST or any other institution, this has been clearly stated
3. Where I have consulted the published work of others, this is always clearly attributed
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work
5. I have acknowledged all main sources of help

Muneeb Islam,

NUST00000170779MSMEE2016

Copyright Notice

This thesis is dedicated to *my beloved parents and teachers*

Abstract

Conventional Grid in current scenario has no self-diagnosing capacity. Transformer installed in this system for the purpose of transmission and distribution is a worthy asset of any power network. The average life of a transformer running under normal loads is usually around 20-25 years. Immense research has been done from earliest times till date to maximize the use of these assets and enhance reliability and sustainability of the electrical system at large. The main goal of any power system is to provide uninterrupted power along with more cost savings at the end of the utility. These objectives can only be achieved if we are more conscious and aware of what is running in the system and how well is it executed.

Health based monitoring of power transformers is currently installed and executed successfully on all units. However due to financial constraints, this has not been updated on the Distribution transformers yet. The presence of a large no of transformers in the distribution system spread over a wide area in the power system, proves to be quite a credible area for further research work and regard this as an important and crucial asset installed in the power system. A cheaper system is thus designed which precisely evaluates the health index of the particular unit in real time. The suggested scheme will help in diagnosing and self-monitoring of the transformer with respect to non-intrusive and later intrusive faults by accurate data classification.

Keywords: *Condition Monitoring, Wavelet, Debaucies, Reliability, Health Index, classification, Intrusive Faults*

Acknowledgments

I am highly thankful to Almighty Allah for providing me such a great opportunity to pursue my research work in a renowned institute. I would also like to show my gratitude to my supervisor Dr. Sajjad Haider Zaidi for his patience, motivation, enthusiasm and continuous support in completing my masterâĂŹs research work. Furthermore my gratitude extends to my GEC thesis committee members, Dr. Bilal Khan and Dr. Lubna Moin for granting their valuable time and support for my thesis work.

I am also indebted to all my colleagues and friends whom I work with and get a chance to exchange knowledge and learn new skills. My honest acknowledgements to my family members for their support in the execution and conclusion of my thesis.

Finally, I am truly grateful to National University of Science and Technology (NUST) for providing such a competitive and healthy environment and facilities to conduct the research work in Power Research Lab developed at PNEC.

Contents

List of Figures

List of Abbreviations

Abbreviations

EMF	Electro Motive Force
ONAN	Oil Natural Air Natural
ONAF	Oil Natural Air Forced
OFAF	Oil Forced Air Forced
FFT	Fast Fourier Transform
CWT	Continuous Wavelet Transform
DWT	Fast Fourier Transform
PCA	Principle Component Analysis
CBM	Condition Based Monitoring
DGA	Dissolved Gas Analysis
RTD	Resistance Temperature Detector

Chapter 1

INTRODUCTION

1.1 Introduction

Distribution Transformer is an electrical equipment used in the network for voltage step down in order to lower down the voltage magnitude to some safe limits. Once the safe limits are achieved at the secondary side, electricity is distributed to residential consumer as per the load demand. However in order to ensure un-interruptible power supply is supplied to our consumer, the unitâĂŹs health is a major concern for all. Any disruption caused due to the breakdown of the unit supply can result in major setback and leave the consumer dis-satisfied.

Currently the systems installed only gauge the electrical parameters of the unit, while neglecting the intrusive parameters like temperature, vibrations and internal pressure of the unit. However in this scenario, vibration sensor shall be placed on the respective unit such that all the acoustic parameters of the unit shall be recorded in the normal state as well as in different fault states that we tend to incorporate in our test library for the monitoring system. Furthermore the same methodology shall be used for some other sensors to prepare a diverse library with other variants like temperature, oil level and voltage monitoring as well. The performance of these parameters for the health of the system decisions shall be taken to either keep the network asset in service or to execute the procedure for preventative maintenance.

Once developed this system shall be made mandatory with the consent of the local utility company to be incorporated as an essential element in every new purchase of distribution transformer unit. Once done, this shall bring design changes in all OEMâĂŹs, while replacing all old designs. As this shall be incorporated by the manufacturer themselves rather than the utility company, it shall serve as a cost effective solution to the production line as well. This will help us in the long run towards more cost savings and maximum asset utilization. Thus effective and regular maintenance is most essential for continual operation when the system demands for it at the right time for better service life.

Signal analysis shall be used to compute the resulting signal values for all faults and then a complete set of data shall be finally used to analyze the result for the need of preventive maintenance or complete asset replacement. A great cost saving opportunity for the utility company this can save more than a mere replacement cost that is done when neglected.

1.1.1 Maintenance Strategies

Health monitoring systems are usually too expensive to be implemented and hence there is a dire need to bring some cost effective solution to avoid any kinds of breakdown and focus more on preventive maintenance. We shall discuss at a later stage the feasibility of the techniques that might be used in order to cater a corrective maintenance scenario and convert the subject situation towards preventive maintenance program that might be predictable, planned, scheduled and more cost effective resulting in less sudden shutdowns. These ill planned breakdowns convert into huge business losses resulting into a cascaded breakdown. During the time period the transformer is energized, several oil test are also conducted to evaluate the approx. health of the unit and how often and how intense PM activity to be conducted.

The health of any system is linked to the relative faults that it experiences throughout a period of time. Faults in electrical systems or to be exact, transformer units are any changes or irregularities that are found in the normal parametric patterns observed previously. The subject abnormality experienced shall be identified and mitigated in the least possible time to lessen the MTBF (Mean Time Before Failure) for any transformer unit.

1.1.2 Impacting the Bigger Picture

In the current scenario of our country where there remains a huge demand supply gap of electrical energy, we can make small improvements to get benefits in the bigger picture at large. Timely predictive maintenance and proper planned outages may help in reducing shutdown costs with effective planning and shall also lead us to uplift our vision of supplying un interruptible power supply to our consumer.

Apart from this, timely preventive maintenance of electrical assets will not only save it from total breakdown but shall also help the authorized personnel updated in identifying the weak links in the circuit causing the losses and thus shall the replacement be carried out respectively of actual equipment rather than any huge investments.

1.2 Background Study

Atefeh Dehghani Ashkezari in his paper Investigation of feature selection techniques for improving efficiency of power transformer condition assessment [1] discusses conditional based health monitoring system by evaluating the deterioration of the insulation of the transformer by conducting several tests on oil of the unit and fuzzy set theory was applied to determine the health index of the unit.

Md Mominul Islam in his paper: Calculating a Health Index for Power Transformers Using a Subsystem-Based GRNN Approach, [2] discusses the classification of data extracted from 345 transformer to determine the same health index of the unit.

Kaixing Hing in his paper: A vibration measurement system for health monitoring of power transformers, [3] uses the methodology of diagnostics information retrieval for the systems. Three in-service transformers were selected and four different vibration analysis techniques were applied to obtain a health metrics based on variant perspectives of the same transformer.

Pavlovsky in his patent: Hydrogen Sensor for Oil Transformer Health Monitoring, [4] uses the principle of hydrogen gas detection using palladium nanoparticles. These particles are proved to be quite sensitive to the presence of hydrogen in minimal concentrations. Presence of hydrogen indicates possible corona discharge in the transformer unit within the windings. This proves to be more feasible and more economical gas detection technique as compared to the DGA test set usually engaged by the field teams on site.

T. Bhavani Shanker in his paper [5] discusses the results of DGA tests conducted on two in service transformers currently live with different types of faults detected within the units installed in thermal power stations in INDIA. Analysis method used for DGA is key gas method and later verified by gas ratio methods (Rogers ratio and Doernenburg ratio) as well.

Chapter 2

THEORY OF OPERATION

A transformer is a static electrical device designed to transfer electrical energy from one circuit to another. The transfer may be at different or same voltage levels, depending upon the application required. The process of this transfer is executed through the process of electromagnetic induction.

2.1 Governing Law of Induction

Faradays law of electromagnetic induction is the basic law which governs the working principle of a transformer. The basic principle involved is that electromotive force (EMF) shall be produced within the circuit once the magnetic field shall link with the electric field. A current will be induced within the conductor that is exposed to any alternating magnetic field. This law is also similarly applicable for motors, generators and inductors.

In case of a transformer, mutual induction between two or more windings is executed which allows electrical energy to be transferred between both circuits. The alternating current between the windings results in a changing and alternating flux to be produced. The resultant flux linkage is observed in the second winding or coil where an EMF will be induced and a current will flow through it, thus transferring the electrical energy from one winding to another. This process is executed in complete electrical isolation between both windings.

One of the main reasons that alternating current is available at our homes for use is that it can easily be generated and transmitted at a convenient and safe voltage levels. At high voltage levels, the magnitude of the current decreases in a given power case. This implies that our I2*R losses are cut down by a significant amount.

2.2 Types by Design

The placement of windings with respect to the transformer construction can be divided into two major categories.

2.2.1 Core Type Transformer

Cylindrical type of windings are wounded in this construction type of the unit. The core will be a rectangular shaped around which the winding shall be formed. The windings are such formed such as to fit over a cruciform core section. This enables the transformer to be constructed with overall a good mechanical strength.

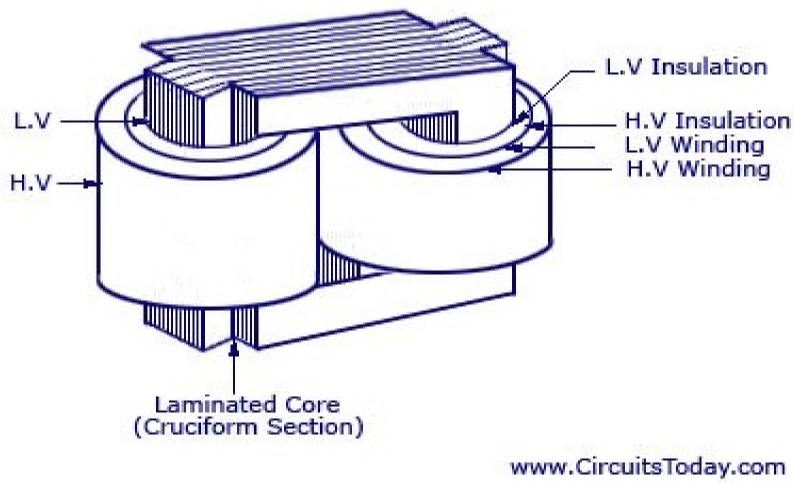

Figure 2.1: Core Type Transformer Cruciform Section

2.2.2 Shell Type Transformer

Shell type of windings as seen in below picture 2.2 are constructed in such a manner that the core surrounds a significant portion of the windings. The coils are wound in a multi layered fashion, while paper is used to insulate the different layers of the coil. The insulated spaces serve the purpose of insulating ducts used for horizontal cooling.[6]

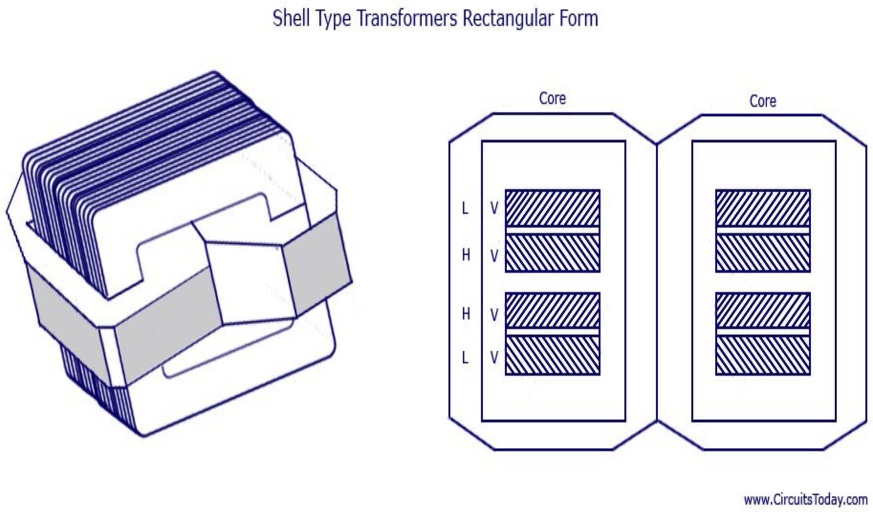

Figure 2.2: Shell Type Transformer Rectangular Form

2.3 Types by Cooling Methods

Transformers when operative tend to heat up and need to cooled down effectively and efficient to avoid burn out. Various cooling methods are employed in this regards and vary as per application and cost comparison methods.

2.3.1 ONAN [Oil Natural Air Natural]

Transformer operative on ONAN principle house the active part of the transformer in an oil tight steel tank filled with insulating oil. The oil is used to transfer the heat generated by the active part when the unit is live. This process is speeded up by radiator fins welded to the steel tank and experiencing a greater surface area to combat the heat with the surrounding environment. Heat is then radiated from the fins with the passage of time. This type of cooling is usually used in smaller distribution transformers.

2.3.2 ONAF [Oil Natural Air Forced]

Transformer operative on ONAF principle houses the active part similarly in an oil filled steel tank as mentioned earlier. The main difference here is that external fans are mounted on the radiator fins / tubes to accelerate the cooling process. Natural air flow is supported by forced air cooling to expedite the heat to be radiated out of the transformer unit from the active part to the surrounding environment in less time. This type of cooling is usually used in Grid Power Transformer / Auto Transformers.

2.3.3 OFAF [Oil Forced Air Forced]

Transformer operative on OFAF principle also houses the active part as same as in an oil filled steel tank. The difference in the steel tank is that additional pumps are installed within to force the oil circulation or movement within the transformer unit through the radiators and the tank. This immensely speeds up the heat radiation process as the oil that reaches the exteriors of the fins / tubes is again able to radiate heat even further due to the forced cooling caused by the external fans installed exterior to the transformer body. This type of cooling is used in huge special type of transformers with dedicated application and sensitive / critical usage.

2.4 Types by Application

The windings in both coils of the transformer (primary / secondary) placed in input and output sides of the transformer are such designed to change voltage levels. A step up transformer is as such in which the primary coil shall have less no of turns as compared to the secondary coil thus inducing more flux on the secondary side and thereby increasing or stepping up the voltage levels. The same is opposite in step down transformer where the secondary coil has less turns as compared to primary coil.[7]

Apart from differences there exists another type of transformer termed as Isolation transformer or Impendence transformer which has equal no of turns on both sides of the windings. The main application of the transformer is impedance matching or the isolation of adjoining circuits.

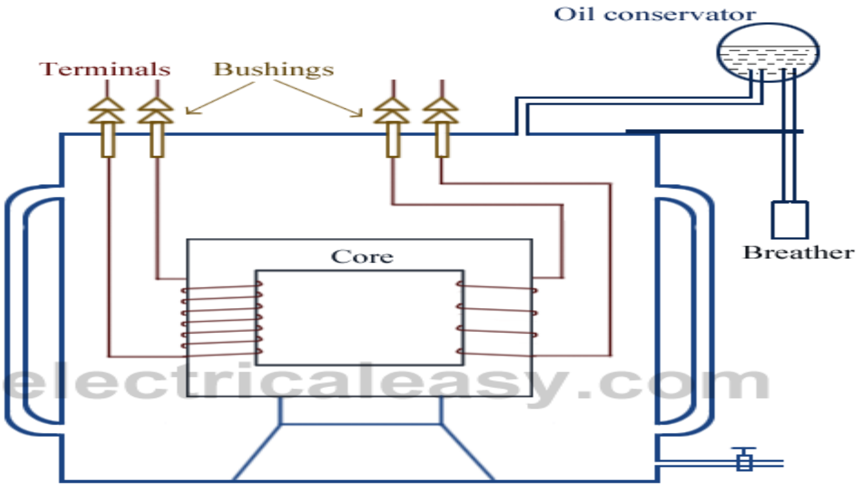

Figure 2.3: Transformer Model

2.5 Transformer Turns Ratio

$$TTR = n = \frac{Vp}{Vs} \tag{2.5.1}$$

where,

Np = primary coil turns

Ns = secondary coil turns

EMF generated = Turns x Rate of change

$$E = n = \frac{d\phi}{dt} \tag{2.5.2}$$

$$E = n\omega\phi max\,(\cos(\omega * t)) \tag{2.5.3}$$

$$E_{\max} = n * \omega * \phi_{\max} \tag{2.5.4}$$

$$E_{rms} = \frac{n\omega}{\sqrt{2} * \phi_{max}} \tag{2.5.5}$$

$$E_{rms} = 4.44 * f * n * \phi_{max} \tag{2.5.6}$$

The above equation 2.5.6 derived is known as the transformer EMF equation where,

f = flux frequency

n = no of coil windings

Îę = flux in webers

From the above equation 2.5.6 we can analyze that if a transformer is connected to DC supply having zero frequency, the inductive reactance shall be zero and the effective impedance will be very low as well. This impedance shall be equal to the copper used in the windings drawing a high magnitude of current from the DC supply causing the unit to overheat and eventually burn out.

2.6 Transformer Power

The power rating of the transformer remains same on both primary and secondary sides and this is the reason that increasing or decreasing one of the parameters voltage or current implies a counter effect on the other parameter. Power is measured in VI and is obtained thus my multiplying both parameters. Ideally or theoretically this power loss is none and the power transmitted at the primary is equal to the power received at the secondary side.

2.6.1 Copper Losses

Practically this power is not the same at both ends and power at the secondary side is always less than primary side. This is due to two main type of losses involved in the operation of the transformer unit. Copper losses or also known as I2R losses are the contributing factor for power loss due to heat production within the transformer. This is mainly caused by the circulating currents of the transformer windings. Copper losses account for the major part of the losses in a transformer. These losses are thus variable with the load current.

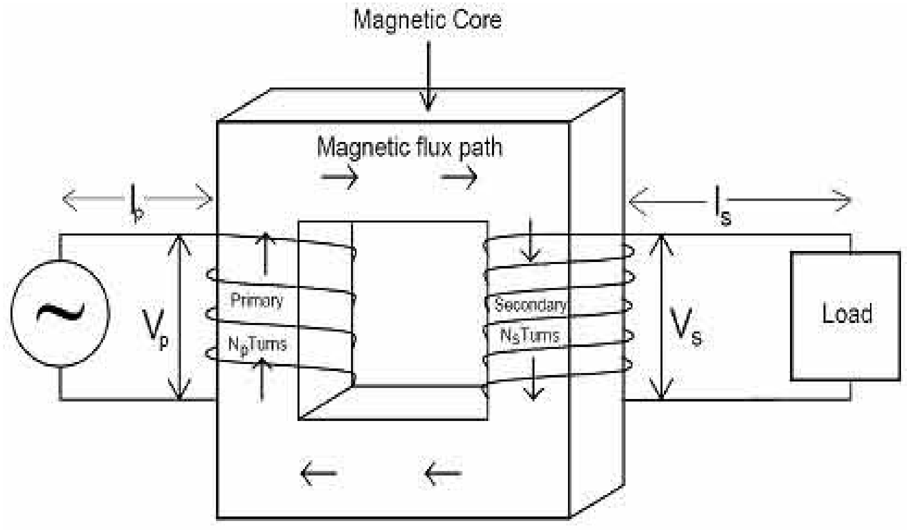

Figure 2.4: Transformer Magnetic Circuit

2.6.2 Iron Losses

Iron losses secondly is a less yet significant contributing factor towards the overall transformer losses. These losses are also known as the hysteresis loss which occur within the core of the transformer unit as a response to the alternating magnetic flux. With the alternating flux, magnetic molecules require power to reverse or alternate their poles. This reversal process is based on the attainment of appropriate amount of flux. The conversion shall produce heat due to the process of friction involved which shall lead to IRON LOSSES. These losses are hence fixed for a given frequency and particular material of core. The ferromagnetic material is basically the reason for the losses incurred on the unit and results in iron losses to be categorized as follows.

2.6.2.1 Hysteresis Loss

This type of loss is present in the system due to the process of reversal of magnetization in the transformer core. In other words, it is the continuous magnetization and demagnetization of the core that causes hysteresis loss within the unit. The transformer unit operates on alternating current, which flows in alternate directions for every half cycle. This continuous shift in the polarity of the supply results in the core to take shifts in the magnetization process at a frequency of 50 / 60 Hz. This phenomenon shall create a huge lag between magnetic flux and the magnetic field, causing a residual magnetic flux to be left behind in the core. Despite the fact that the magnetic field is cut off, additional energy is still required to remove this residual flux from the core. This additional energy required is then termed as Hysteresis Loss as seen in the picture 2.5. Depending on the quality or grade of the core material this type of loss may vary along with the frequency of the magnetic reversals.

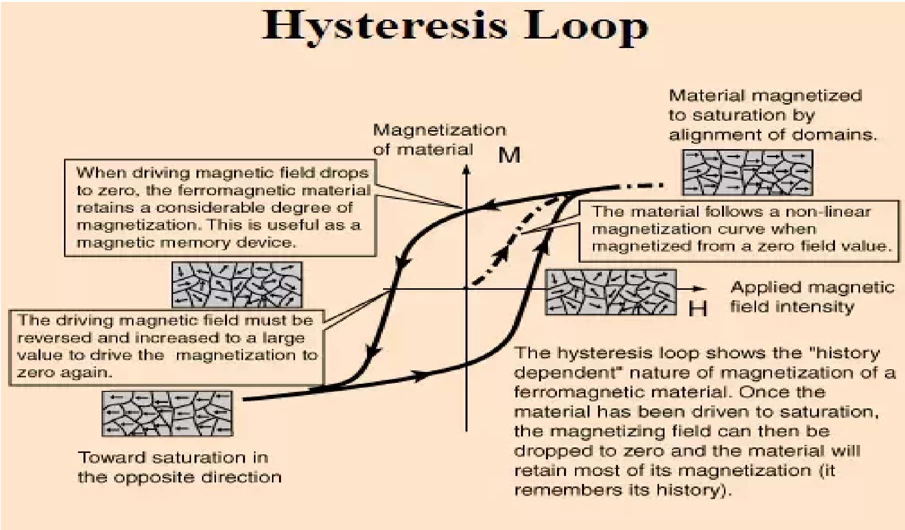

Figure 2.5: Hysteresis Loss

$$W_h = \eta * \beta_{max} * 1.6 * f * V \tag{2.6.1}$$

where,

n = Steinmetz hysteresis constant

V = volume of the core in m3

2.6.2.2 Eddy Current Loss

Another type of iron loss to be considered is the eddy current loss. This type of loss is produced due to the circulating current that mobilize within the core material. The eddy current loss shall me maximum if the core material is used as a single metallic block over which winding shall be wound. However technically we ensure the usage of laminated core to be used for the subject purpose. This reduced the magnitude of the eddy current losses as circulating currents are unable to cover a major portion of the circuit and thereby reducing the intensity of the losses induced within them.

Some of the flux that is induced in the secondary winding gets induced with the body of the transformer causing a current to be circulating. This current produces losses within the core material.

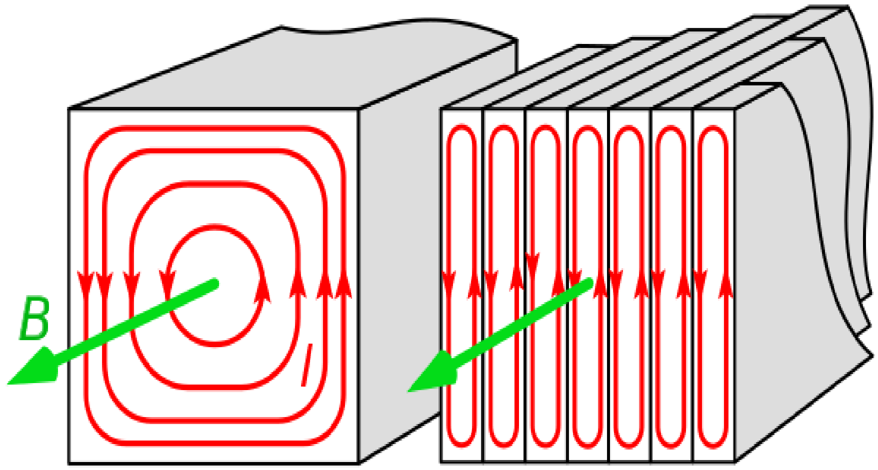

Figure 2.6: Eddy Current Loss

These losses emerge from the fact that the core itself is composed of a material that conducts and hence the voltage that is being induced in it by varying flux results in circulating currents to be produced within the material. This type of loss mainly depends upon the rate of change of flux as well as the path resistance involved in the circuit.

These losses then ultimately reduce the output of the transformer as mentioned earlier and has a quite negative impact on calculating the efficiency of the transformer

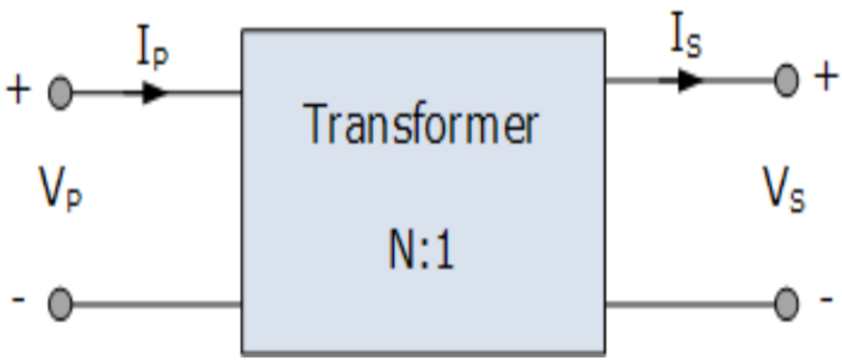

Figure 2.7: Ideal Transformer

The efficiency of a transformer unit can be defined similarly to any machine as output / input. Most transformers have a full load efficiency between 95 to 99 percent This parameter is maximum when the above mentioned copper losses are equal to the iron losses.

2.7 Transformer Configuration

2.7.1 Wye - STAR CONFIGURATION

A Y configuration is such that all neutral/ground points are shorted together at a single point. This point is later grounded properly. Most of the transmission and distribution transformers are connected in Y configuration in their load side. Here the line current is equal to the phase current while on the other hand the phase voltage is the line voltage divided by root 3. There is a separated fourth neutral conductor which may be grounded or floating ground to conduct any imbalance current from the phases. With regard to application, this configuration provides low starting current. It works for long distanced electrical network.

2.7.2 Delta - MESH CONFIGURATION

A Delta Configuration is such that it has no neutral point. The return path of one phase is shorted with the phase of the other terminal. Here the line voltage is equal to the phase voltage while on the other hand the phase current is the line current divided by root 3. The three phases represent a triangle shaped circuit. All three coils in a three phase system are connected in series to form a closed circuit. This type of configuration is usually used in distribution network system. With regard to application, this configuration provides high starting torque. It works for short distanced electrical network. All the single phase load is concentrated on one phase as all connections are shorted out.

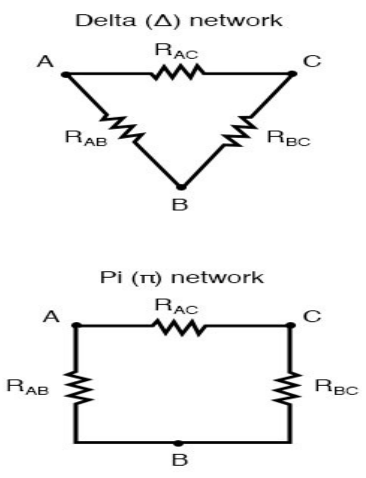

Figure 2.8: Transformer Configurations

2.8 TYPES OF FAULTS INDUCED

2.8.1 Line to Line Faults

A LL or a line to line fault occurs in the system when two phases of a live system are shorted and a short circuit condition is created. A three phase system with lines a, b, c are considered in which if b and c are shorted out, the LL fault imitates a symmetrical fault with respect to the reference phase a. In that case the zero sequence component of current of un-faulted phase a shall be zero. [8]

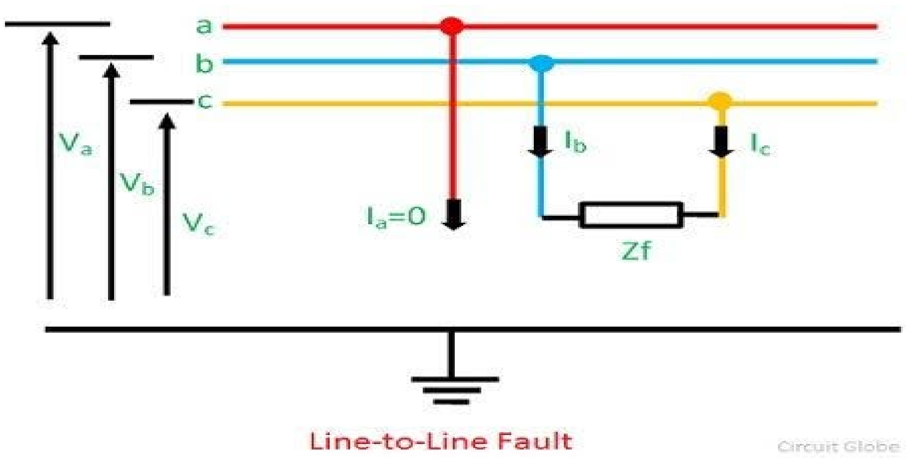

Figure 2.9: Line to Line Fault

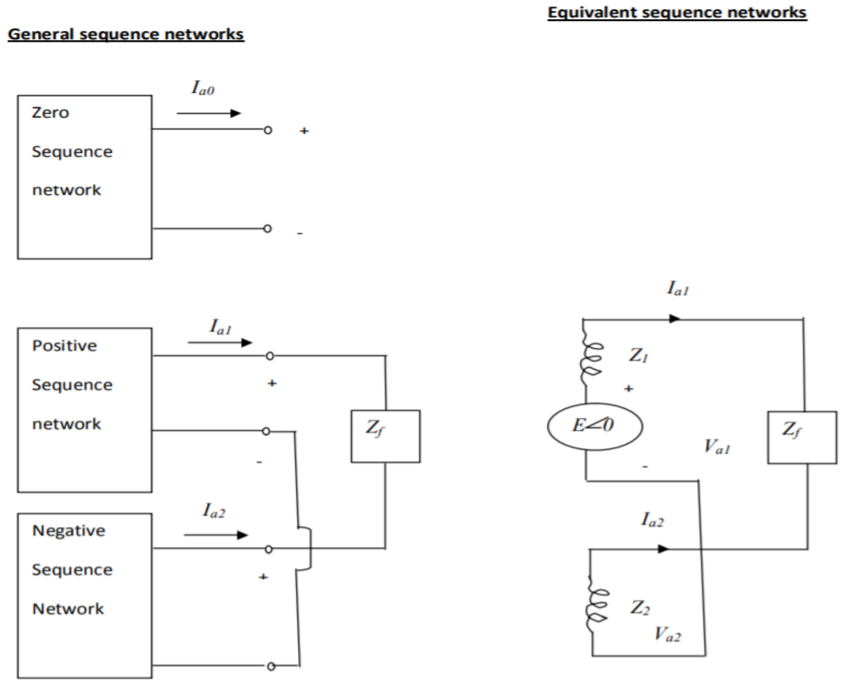

Figure 2.10: Line to Line Fault - Sequence Network

2.8.2 Line to Ground Faults

A LG or a line to ground fault occurs in the system when one phase of a live system falls and shorts with the ground or earth. Same scenario shall be observed when the line comes in contact with a tree or and other object. The system is reported to have higher resistance values as compared to LL faults. It can also be the other way round when the tree or object falls on the live line thus creating a fault and disrupting the system. This is the most common type of fault occurring in an electrical network.

The single line to ground fault may exceed the 3 phase short circuit fault when executed near a D-Y transformer unit. The reason behind is that the delta winding effectively blocks the zero sequence impedance tracked from the source. In event of the source impedance to be zero, the major contributing factor is the transformer impedance. The zero sequence component shall rapidly increase until the value of the Line to Ground fault current becomes less than three phase fault current. [9]

If the fault occurs in delta wye transformer having wye neutral grounded, then the zero sequence model is the transformer impedance to the zero sequence neutral. This shall in in return make the zero sequence impedance smaller as compared to positive and negative sequence impedances.

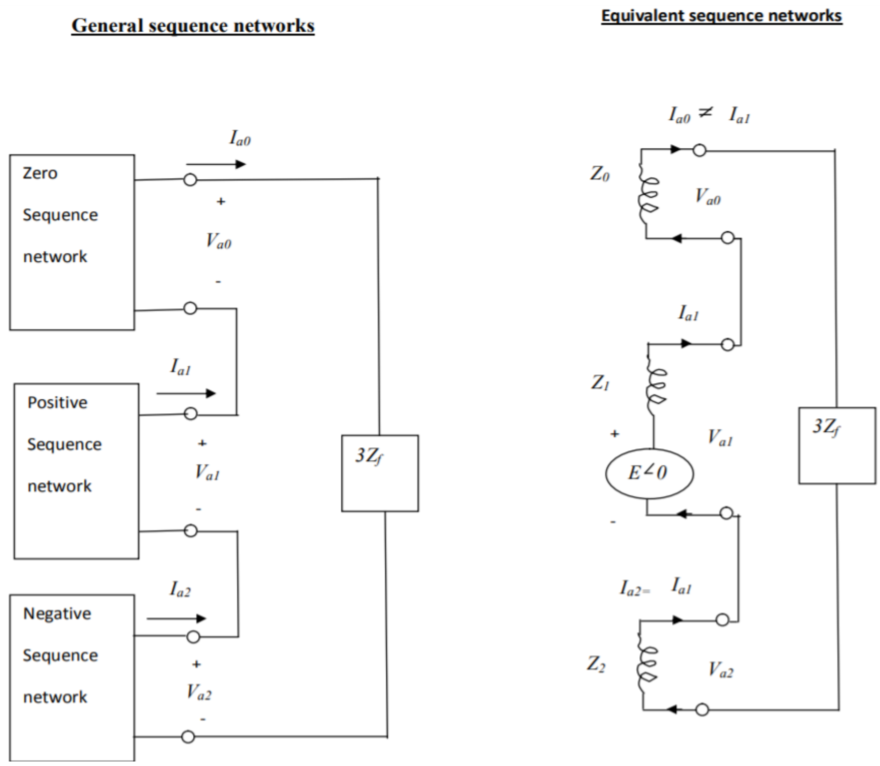

Figure 2.11: Line to Ground Fault - Sequence Network

A three phase system with lines a, b, c are considered in which if a is grounded, the LG fault imitates a symmetrical fault with respect to the reference phase a. In that case the zero sequence component of current of un-faulted phase a shall be non-zero, while other phase components are zero.

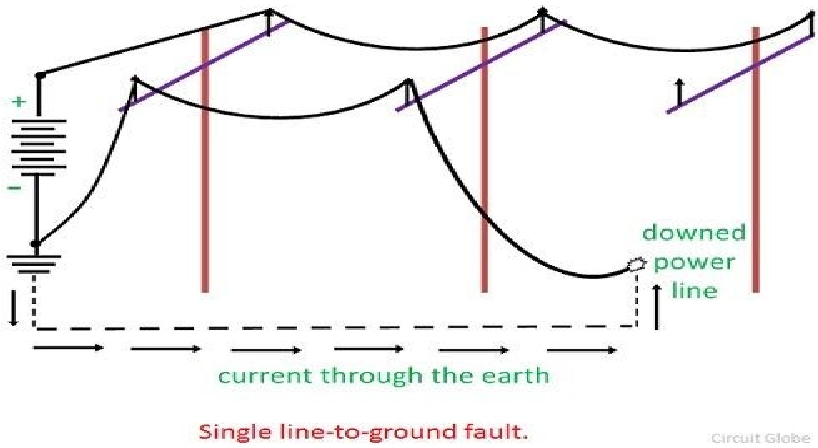

Figure 2.12: Line to Ground Fault

Earth fault relays along with differential protection actually are the main front protection devices in the event of such failures.

2.8.3 Double Line to Ground Faults

A LLG or a double line to ground fault occurs in the system when two phases of a live system are shorted and grounded. In real life situations, such scenarios are thus created when LL fault is prolonged and aggravated causing the third phase to break and be grounded.

A three phase system with lines a, b, c are considered in which if b and c are shorted out, the LL fault imitates a symmetrical fault with respect to the reference phase a. The third phase a then breaks away and is grounded.

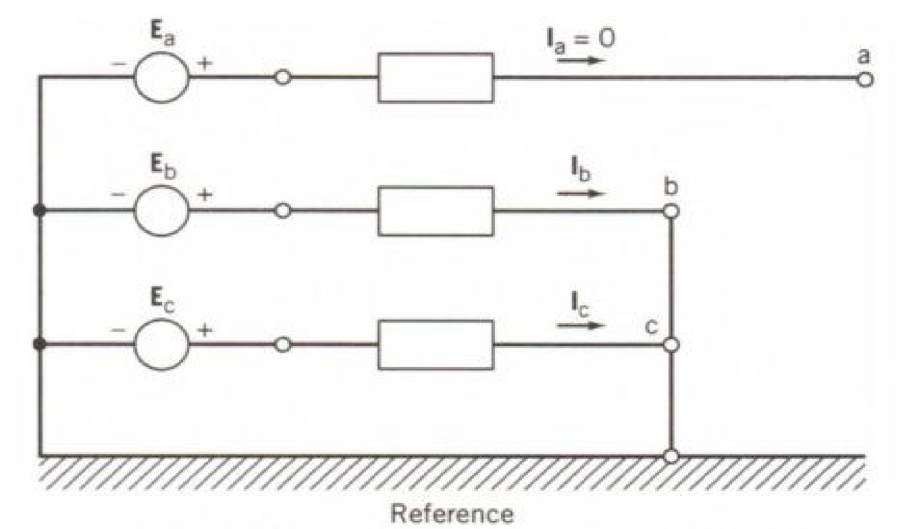

Figure 2.13: Double Line to Ground Fault

Same scenario shall be observed when the line comes in contact with any nearby tree or and other object. It can also be the alternatively if the tree or object happens to fall on the live line thus creating a fault and disrupting the system. Such fault scenarios do have another aspect when two lines comes into contact with each other but to break or fall to the ground creating ungrounded line to line faults.

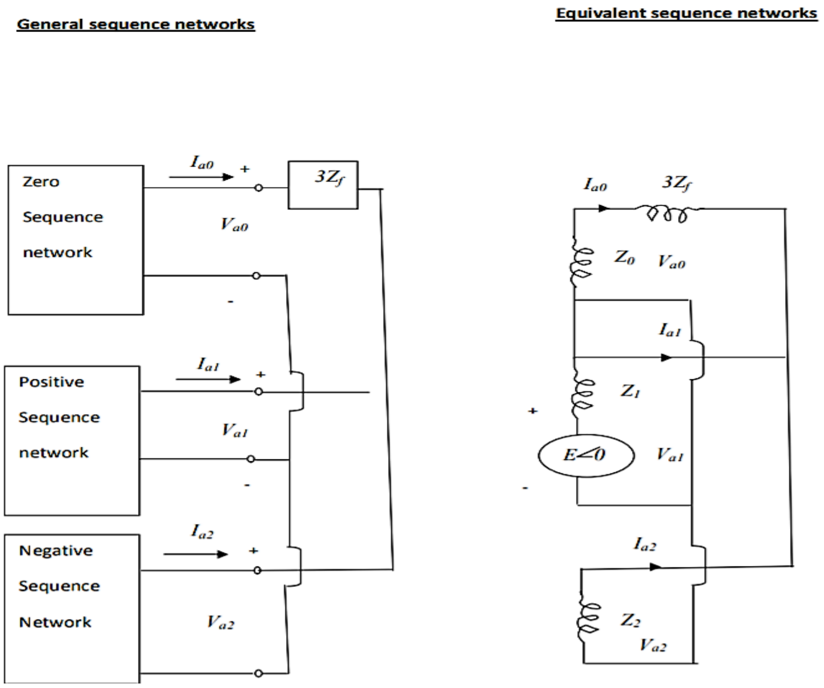

Figure 2.14: Double Line to Ground Fault - Sequence Network

2.8.4 Triple Line to Line Faults

A LLL or a triple line to line fault or 3ph fault occurs in the system when all three phases of a live system are shorted and an extreme short circuit condition is created. In real life situations, such scenarios are rarely created due to the intense nature of the fault.

A three phase system with lines a, b, c are considered in which if all a, b and c are shorted out, the LLL fault imitates a symmetrical fault with respect to the reference phase neutral.

2.8.5 Inter Turn Winding Faults

This type of fault occurs when by winding flashovers produced from surges or short circuit. High currents tend to flow within the unit in case of any fault scenario.

Chapter 3

SIGNAL ANALYSIS TECHNIQUES

3.1 Wavelet Transform

Faults can be rectified if they are detected and analyzed at the proper time and well before any probable breakdowns. Detection of fault signature shall be discussed later in the report using relevant sensors> however these signatures need to be identified as fault signatures as well and be identified different from healthy signatures. This analysis involves many techniques used worldwide, but we shall use the technique of Wavelet Transform with Debauchees levels for subject analysis in this report.

Wavelet theory or wavelet analysis as it is commonly known has been quite an effective tool lately in the field of signal processing. It has proved to be quite an active research domain in related fields. The wavelet tool is quite used with signals application involving FOURIER TRANSFORM. It is then utilized to obtain accurate frequency information. Such nature of information is usually obtained where varying transient signals are being involved within the system. Compared to Fourier Transform, wavelet analysis transforms the related functions into smaller wavelets rather than trigonometric polynomials. These wavelets do basically originate from the main function which is known as the mother wavelet. The wavelets thus obtained are localized in time and frequency domain ensuring enhanced numerical stability in reconstruction.

The fundamental objective of wavelet is to determine the basic wavelet functions and later deduce efficient computational methods. It can also be shown that wavelets are also used to produce localized frequency information from any application using Fast Fourier Transform (FFT). These wavelets are thus useful for processing non-stationary signals encountered by any system. A signal is termed as non-stationary if the finite duration signal, in actual terms known as transient signal is non-stationary. Here transient signal is such a signal carrying a length shorter to the observer signal.

In recent times, mathematicians in relevant communities have been facing some inevitable difficulties that are found in Fourier analysis. In this context the discovery of smooth mother wavelets forms a basis for discrete translations and dilations that have bounded energy denoted by the function below

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt > \infty \tag{3.1.1}$$

The wavelet technique traces its origins back to early nineties when it was realized that most of the signal occurring in nature were time localized transients and hence couldn't be adequately identified via use of stationary signal analysis techniques like Fourier transform.

The wavelet, via use of the filter banks, decomposes a signal into a series of frequency bands containing some part of the original signal; hence enabling to identify which band of frequencies were present in the signal at a particular instance of time. This enables us to locate the time localized transients via monitoring of various bands which is not possible via the Fourier transform.

The time localization nature and detection can be better understood by the waveforms shown below in figure 3.1. The original wave has a transient signal in middle of the main signal and as such it leads to transients which exist only for a specific amount of time. If we take the FFT of this signal, it would not be able to distinguish between transients and the main signal, however, the wavelet enables us to monitor the time - frequency domain activity of the signals by showing the frequency presence in each individual bands.

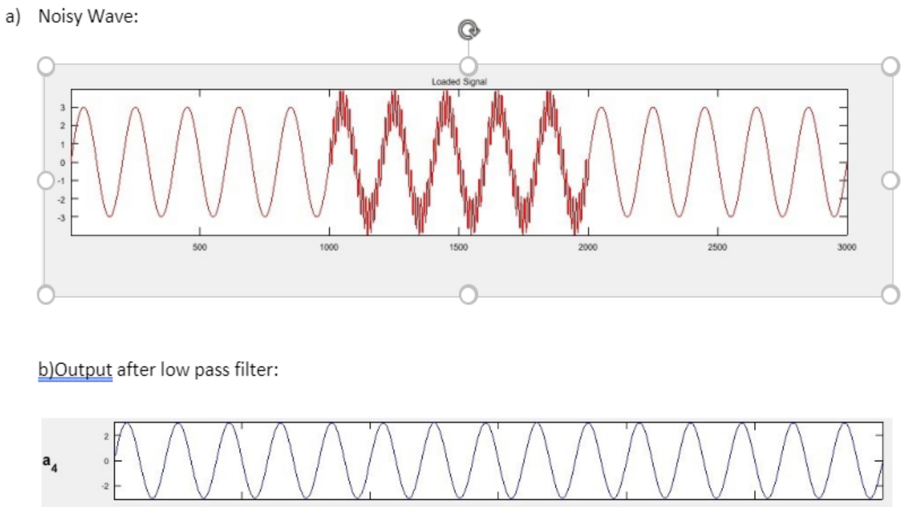

Figure 3.1: Time Localization

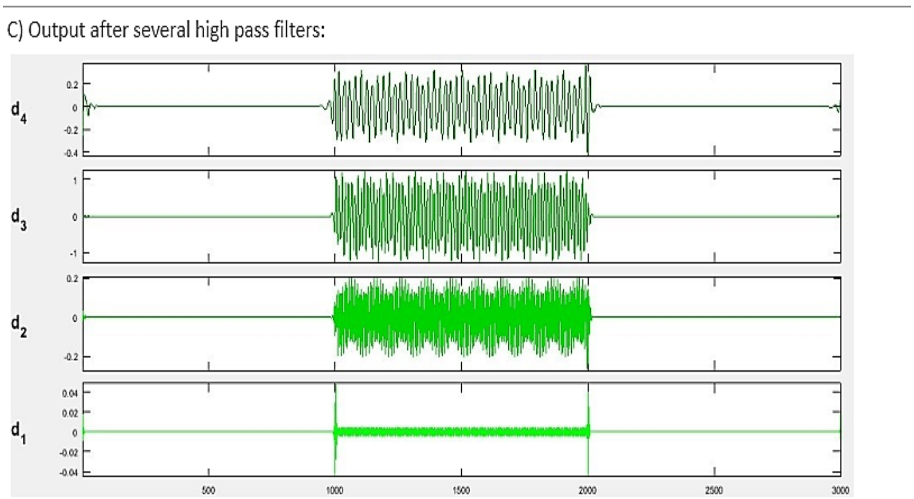

Figure 3.2: HPF OUTPUT

FOURIER TRANSFORM

$$X(F) = \int_{-\infty}^{\infty} |f(t)|\, exp^{-j2\pi * ft} dt \tag{3.1.2}$$

WAVELET ANALYSIS

$$X(a,b) = \int_{-\infty}^{\infty} |f(t)|\, \psi_{a,b}(t) dt \tag{3.1.3}$$

Where x is the real signal, ÏĹ is an arbitrary mother wavelet, a is the scale and b is the translation (X is the processed signal of course). [10]

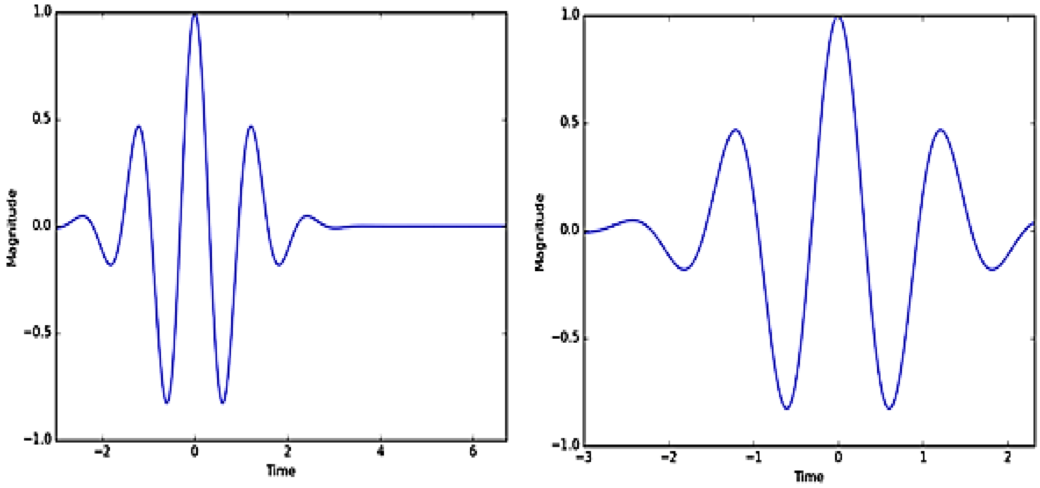

Figure 3.3: Illustration of how the mother wavelet expand when the scale gets bigger and bigger

Fourier Transform decomposes the signal into sines and cosines whereas Wavelet transform functions are localized in both real and Fourier spaces. It is an infinite set of various transforms. In this regards, wavelets are divided into categories based on its orthogonal properties. Orthogonal wavelets are used for discrete wavelet transform while non-orthogonal wavelets are used for continuous wavelet transform. Fourier transform is unable to represent true data efficiently due to its property of the components not being localized in time and space. Wavelets on the other hand are found much localized in time and frequency. Wavelet Transform can thus be categorized as below:

3.1.1 Continuous Wavelet Transform (CWT)

Continuous Wavelet Transform returns the output data array further one dimension more than the input data array. Data is highly co-related due to its property of non-orthogonal set of wavelets. The principle of operation in this transform type is the basic computation of convolution of the signal with the wavelet. This can also be achieved by convolution by means of multiplication in Fourier space.

3.1.2 Discrete Wavelet Transform (DWT)

Discrete Wavelet Transform differentiates with Continuous Wavelet Transform in the process of discretizing the scale of parameters. DWT restricts the value of this discretizing in the power of 2 (e.g, 1,2,4,8,âĂę) unlike some arbitrary value to be selected in CWT. [11] The application of DWT enhance the accuracy level of the subject signal quite better for decomposition and analysis. Amongst many different forms of transitions available for DWT, we have used Daubechies wavelet for our project. The DWT decomposes in a set of wavelets that are orthogonal to its translation and scaling. The advantage of reduction in the redundancy of information is obtained here making it quite reliable for signal processing and compression. It is considered as a suitable tool for electrical noise elimination that in return replaces the attenuation procedures involving low pass filters of the systems. [12]

3.1.3 Wavelet Types

A wavelet exists for a finite duration and has zero mean. There are several types of wavelets used as per the application. [13] The list is as follows.

- Morlet [CWT]
- Daubechies [DWT] (USED IN THESIS))
- Coiflets
- Biorthogonal
- Mexican Hat [CWT]
- Symlets

Daubechies wavelet has a distinctive scaling restriction with an enhancement of its fundamental Father Wavelet. The function of this wavelet is to decompose the real signal while removing the actual noise part and remove it. [14]

The concept of regularity provides a measure of smoothness for all wavelet or scaling functions. The simplest of the transforms are the Haar Wavelets. Daubechies wavelet extends the Haar wavelets with the help of longer filters. This results in the production of smoother scaling functions and wavelets. As we increase the k factor (vanishing moment index) allows to better compress regular parts of the signal but alos increasing the size of the support for wavelets. [15]

3.2 Principle Component Analysis

Principle Component Analysis (PCA) as it is commonly known, is a procedure of orthogonal transformation used to convert correlated variables into a set of linearly un-correlated variables. These variables are then known as principle components. This tool was invented first in 1901 by Karl Pearson. PCA is usually used in data analysis in predictive modelling designs. In simpler terms it aids in visualizing relatedness or connections between set of different data populations.

The basic methodology of PCA is eigenvalue decomposition of a covariant or correlated data set. Its operation can be related as revealing the variance of a data set in the best structure possible. Mathematically speaking, PCA is defined as orthogonal linear transformation. Here the data is transformed such that the coordinates with the greatest variance lies on the first coordinate. This first coordinate si termed as First Principle Component.

The transformation is usually defined by a set of p-dimensional vectors of weights or coefficients.

$$w(k) = (w_1, \cdots, w_p) \tag{3.2.1}$$

In simpler terms Principle Component Analysis is a method of feature extraction. In a large set of data this methodology aids us in retaining useful and important information. The new variables obtained after PCA are independent variables, which helps in the assumption of any linear model. It pools our predictors thus allowing us to drop the non-significant Eigen vectors. The PCA method, employs the covariance matrix to measure the association of every variable with one another.

Using eigen vectors would be easier to deduce the distribution of the data that is involved in our test set. On the other hand, the eigen values obtained indicates the relative importance of the information extracted by eigen vectors earlier. This technique has overall found many applications in the field of face recognition, patters identifications and image compression. Thus PCA is regarded as a powerful tool for data analytics. [16]

3.3 Data Classifiers

Data Classifiers are used in areas dealing with large data sets to cater the need for an algorithm to map the data into defined categories. Some of the commonly used classifiers are mentioned below.

- Perceptron
- Decision Tree
- Naive Bayes
- Ensemble
- K-Nearest Neighbor (KNN)
- Logistic Regression
- Artificial Neural Network / Deep Learning
- Support Vector Machine

Linear models produce analogous outputs for a given set of data. However machine learning methods give us different results irrespective of data classification. The reason behind is the methodology of supervised learning using random simulation. This supports the idea of getting the desired output as the training of the machine is continued. Hence the error ration significantly drops with the passage of time. Some of the important classifiers are discussed briefly below.

3.3.1 Support Vector Machine

3.3.2 Ensemble

The ensemble classifier works on the principle of obtaining weighted prediction votes from the data that is to be classified. It includes error correction in their coding that is to be determined for coding. [17] Overall ensemble classifiers perform better than any other single classifier where data does not over fit rapidly.

3.3.3 Support Vector Machine

The SVM classifier works on finding an optimal point to increase the differences between two classes. In other words, it focuses on more enhancing the boundaries between classes.

3.3.4 K-Nearest Neighhbor (KNN)

The KNN classifier also known as the lazy learner, uses the concept of unsupervised learning technique. The Euclidean distance between various classes is further classified to better identify the data category properly.

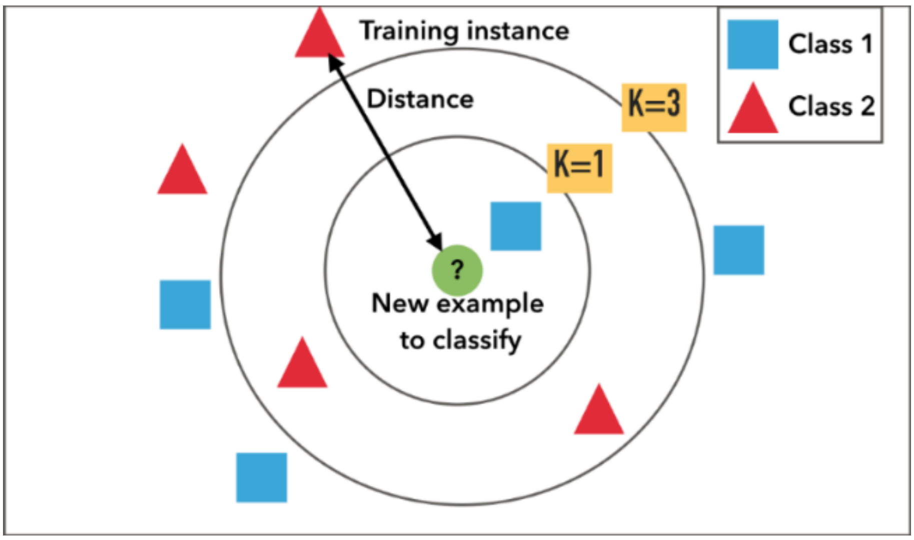

Figure 3.4: K-Nearest Neighbor

3.3.5 K-Fold Cross Validation

The data that is to be considered is to be divided into at least two parts. These observations form the machine are then used to train the machine learning model where an input is then provided to obtain the desired output as per the input. Initially the error will be observed high resulting in high variation in the desired output. However with supervised learning the deviation minimizes for the test data provided. This concept of dividing the data into âĂŸKâĂŹ no of folds or types is an important aspect of machine learning and is known as K fold cross validation concept. Here the initial fold is training set while the rest are testing data sets.

3.4 Confusion Matrix

To evaluate our data classifier, we shall judge the performance of the obtained output given a known input data set. An error matrix known as confusion matrix is used where co-variables of the data sets are compared to evaluate or categorize our output into the following categories.

- a. True Positive
- b. True Negative
- c. False Positive
- d. False Negative

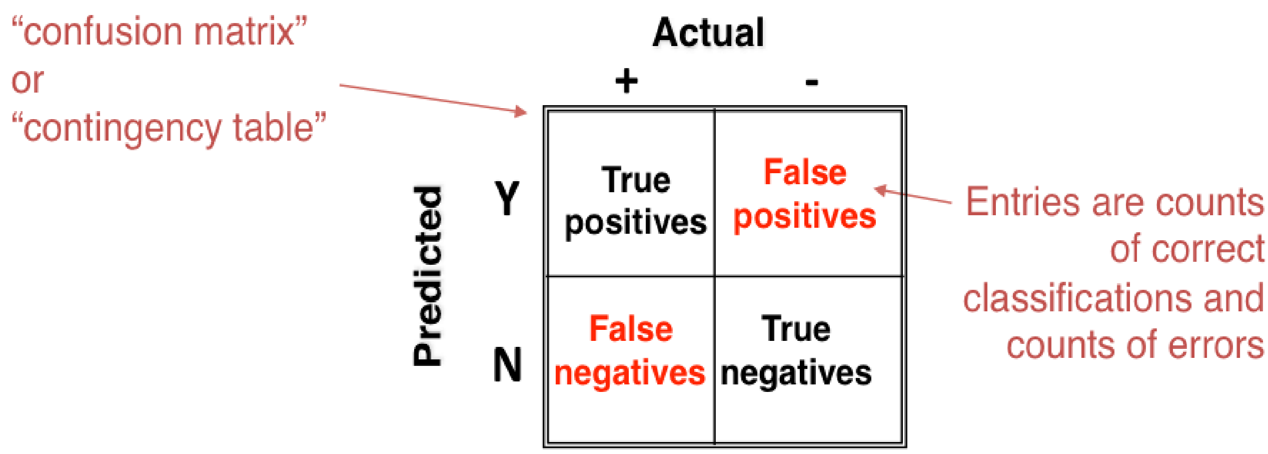

Figure 3.5: Confusion Matrix

For example in a data set of 100 observation or samples, the desired output should have been 50 in Category A while 50 in Category B. However, the obtained output is such that Category A was classified as 65 while the rest as Category B.

In such cases 50 output of Category A shall be termed as True Positive, while the rest 15 of Category A shall be classified as False Positive. On the contrary, the 35 data of Category B shall be termed as True Negative which were not detected or classified in Category A in output. In simple terms this matrix helps us to identify the correct and miss-classifications accordingly. An ideal model shall have all the output data in the True Region, which is ultimately achieved by increasing the data set and selecting the right classifier as per the requirements. Furthermore proper training of the machine leering model shall boost up the idea to achieve subject goal.

Chapter 4

CONDITIONAL MONITORING THEORY

4.1 Condition Based Monitoring

The basic theme of the project revolves around the concept of Condition Monitoring Principle. Before proceeding we shall dive in to the subject concept and glance at the theories in this regards more closely. Condition Monitoring is fundamentally the process of monitoring or observing the parameters of an object / device on continuous run time basis. The basic objective to be met is to check for any type of abnormality or irregularity experienced by any system which needs attention. Rectification or Troubleshooting decision can then be further taken by the authorized personnel associated with the linked system. It is often the observed that such decisive actions when not taken at the right time and the right place lead to destructive or catastrophic outcomes. [18] Condition Monitoring is the next step of Preventive Maintenance of any asset. Here in CBM we observe and analyze the running trends of healthy and fault data of any equipment, which leads us to the evolved concept of Predictive Maintenance. As seen in the figure 4.1 in Predictive Maintenance the authoritative can predict or estimate the life of the equipment or asset in service which shall support him/her in scheduling the maintenance activities henceforth accordingly.

In CBM, the system is linked to several sensors deployed to acquire parameter monitoring over a period of time. The data obtained can then be used to determine and establish signature trends for a particular equipment.

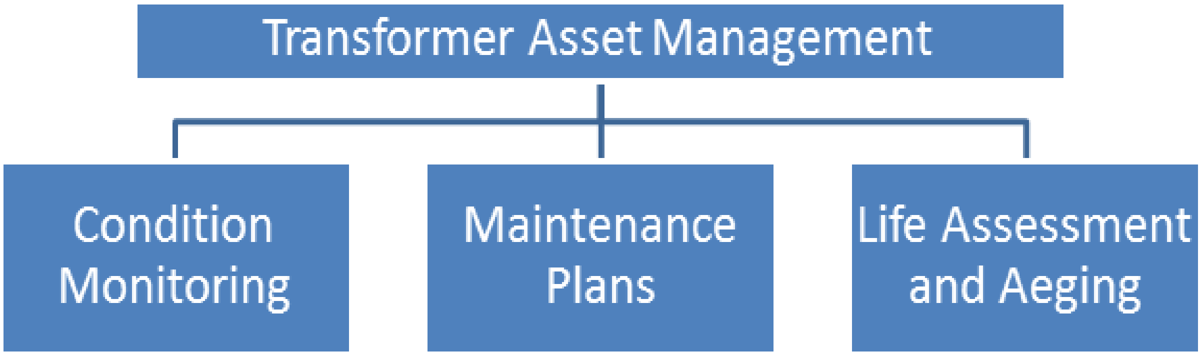

Figure 4.1: Types of Asset Management

4.1.1 Project Application Theory

The same concept of condition based monitoring was engaged in our project of transformers. Globally much work has been done on the monitoring, measurement and protection of power transformers. Due to the huge cost of transformers in that case, the cost of the protection system does not highlight and systems are in place. However the small transformers used in electrical distribution network are not engaged with any of such systems are the cost of the system does indicate a considerable investment for the OEM. We aim to develop a cost effective solution for this problem for the distribution network where the concept of condition monitoring of transformers is carried out. Some work has been done in this regards internationally as mentioned earlier but still needs some improvements and enhancements.

This thesis combines the previous works of different experiments conducted earlier along with improvements in the procedures and algorithms used. Condition Monitoring of transformers shall be carried out with two broad categories.

- NON-INTRUSIVE (PHASE-I)
- INTRUSIVE (PHASE-II)

An experimental setup was established with 10 x transformers (10 KVA) and 05 x transformers (05 KVA) to conduct condition based monitoring and developing a library of faults of several signatures for the 15 units. The phase 1 of the thesis project has been already completed. This phase was subject to non-intrusive monitoring current parameter monitoring of the transformers. Current signatures were obtained for the transformers in health condition. Next faults were induced of different nature on the same units. Initially temporary faults were induced followed by permanent damages on the unit. Signatures were established and stored in the memory for future. This helps us in predictive maintenance in case repetition of same nature faults are observed again, the systems employed can be used for self-diagnosing capability, resulting in minimization of faults and maximization of asset reliability in the long run. Any types of catastrophic failures leading to asset becoming non-functional can thus be avoided reducing maintenance and outage cost and time.

The 2nd phase of the setup is then initiated where we employ several sensors on the experimental setup with regard to intrusive monitoring. Intrusive testing or monitoring is any procedure executed on the unit within the surface contact. In this regards, all the relevant sensors are then mounted or physically connected with the live unit. This helps us to obtain real time parameters data from all the sensors.

Our monitoring system as explained later in the thesis report is based on the following parameter detections systems.

- Thermal analysis
- Voltage signatures
- Winding / Core vibration analysis
- Tank pressure check

The actual aim of executing such a monitoring system ensures the availability of a cost effective solution to achieve better asset saving and management capabilities. This shall in turn lead to reduction in un-informed or un-planned shutdowns of the network. The goal of any utility of supplying un-interruptible power supply to its consumer can thus be fulfilled effectively and efficiently.

The interpretation and understanding of the data obtained through this system are used in junction with International Test Standards to analyze and ascertain the actual condition of the transformer unit. These techniques are majorally categorized into two broad groups as conventional and non-conventional methods. Some of the methods used currently are Dissolved Gas analysis, Furan Analysis, Frequency Response Analysis, etc. However as discussed in the beginning these methods are not quite cost effective for distribution transformers.

Chapter 5

CURRENT MONITORING TECHNIQUES

5.1 Dissolved Gas Analysis

DGA [Dissolved Gas Analysis] is an oil testing method used to detect and quantify the gases dissolved in the transformer oil. With the passage of time and multiple operations/faults experienced by the transformer unit, several gases are developed inside the unit as the oil is decomposed or reacts with different materials having variant compositions. These gases consist of harmful and harmless nature of effects for the transformer itself. Presence of these gases helps condition monitoring engineers to determine the nature of faults that might have occurred recently.

These gases if undergone repeated faults or experience under extreme thermal and electrical stresses, may create enhanced pressure thus actuating the Bucholz Relay. This relay is placed on the top of the transformer unit externally and serves as a protection device for extreme pressure cases. However if the magnitude of the same gases are not sufficient enough to activate this protection, the unit remains in service and continues to supply uninterrupted power. These minimal concentrations of gases do penetrate and dissolve within the oil. The dissolved gases thus serves as a silent destroyer of the electrical equipment. Timely detection of these gases needs to be done with periodic tests conducted for all in service transformers.

The DGA test is thus carried out for online fault diagnosis with routine oil sampling from transformers. Possible mechanisms of gas generation include intense heavy arcing, low energy sparking, discharges from corona, overheating of insulation resulting from severe overloading, and failure of built in cooling systems. [19]This test is hence carried out which detects the minimal concentrations in ppm of these nine critical gases mentioned below

- 1. H2 (Hydrogen)
- 2. CH4 (Methane)
- 3. CO (Carbon monoxide)
- 4. CO2 (Carbon dioxide)
- 5. C2H4 (Ethylene)
- 6. C2H6 (Ethane)
- 7. C2H2 (Acetylene)
- 8. O2 (Oxygen)
- 9. N2 (Nitrogen)

5.2 Analysis Methods

The next step to analyze the fault from the data of test results obtained earlier is based empirical assumptions and practical knowledge gathered over a large period of time. In this context some of the approved and verified methods used worldwide are mentioned

- Key Gas Method,
- Dornenburg Ratio Method,
- Rogers Ratio Method,
- Nomograph Method,

- IEC Ratio Method,
- Duval Triangle Method
- CIGRE Method.

5.2.1 Key Gas Method

The basic DGA analysis method measures individual quantities of gases that are being emitted after a fault within the transformer unit. KEY GASES are the significant and proportion of each gases measured. This method is mostly frequently used compared to the others mentioned above. The gases that are produced within the unit are grouped into three main categories namely Hydrogen and Hydrocarbons, Carbon Oxides and non-fault gases as mention in below figure 5.1. The presence and hence the percentage of each gas is thus measured and determined using this method. The major faults detected are overheating of oil, overheating of insulation, partial discharge and arcing.

Key Gas Method (IEEE Std. C57.104-2008)		
Key Gas	Fault Type	Typical Proportions of Generated Combustible Gases
C_2H_4	Thermal oil	Mainly C_2H_4; Smaller proportions of C_2H_6, CH_4, and H_2; Traces of C_2H_2 at very high fault temperatures
CO	Thermal oil and cellulose	Mainly CO; Much smaller quantities of hydrocarbon; Gases in same proportions as thermal faults in oil alone
H_2	Electrical Low Energy Partial Discharge	Mainly H_2; Small quantities of CH_4; Traces of C_2H_4 and C_2H_6
H_2 & C_2H_2	Electrical High Energy (arcing)	Mainly H_2 and C_2H_2; Minor traces of CH_4, C_2H_4, and C_2H_6; Also CO if cellulose is involved

Figure 5.1: KEY GAS METHOD

5.2.2 Doernenburg Ratio Method

This is the 2nd most frequent method used for DGA testing and analysis. Ratio of gas concentrations are measured and analyzed for identifying thermal faults, corona discharge and arcing. Ratio values for CH4/H2, C2H2/CH4, C2H4/C2H6 and C2H2/C2H4 are determined and each ratio value if exceeding a particular limit value is termed to specific type of fault. If all four ratios lie within the defined limits,5.2 the diagnosis is confirmed. The drawback of this method however is that many results obtained in practical terms relate to NO INTERPRETATION at the end of the diagnosis activity. [20] Thermal Decomposition, partial discharge and arcing are some of the faults that are easily detected using this DGA testing method.

Fault Diagnosis	CH4/H2	C2H2/C2H4	C2H6/C2H2	C2H2/CH4
Thermal Decomposition	>1	<0.75	<0.3	>0.4
Corona (Low Intensity PD)	<0.1		<0.3	>0.4
Arcing (High Intensity PD)	>0.1, <1	>0.75	>0.3	<0.4

Figure 5.2: DOERNENBURG RATIO METHOD

5.2.3 Rogers Ratio Method

Another important ration method which diagnosis thermal faults better than Dornenburg Ratio Method. Th four ratios measured in this method are CH4/H2, C2H6/CH4, C2H4/C2H6 and C2H2/C2H. The faults detected here are ageing, partial discharge, electrical and thermal faults of variable severity. [20] The Rogers ration method is an effective tool for DGA diagnostics as it deals with gas analysis of each case. The major constraint visible is the non-consideration of the dissolved gases below concentration values which leads to misinterpretation of data in our analysis method. However a major constraint is that the ratio results do not fall into any of the suggested fault types.

Ratios for Key Gases – Rogers Ratios Method				
Case	Ratio 2 (R2) C_2H_2/C_2H_4	Ratio 1 (R1) CH_4/H_2	Ratio 3 (R3) C_2H_4/C_2H_6	Suggested Fault Type
0	<0.01	<0.1	<1.0	Normal
1	≥1.0	≥0.1, <0.5	≥1.0	Discharge of low energy
2	≥0.6, <3.0	≥0.1, <1.0	≥2.0	Discharge of high energy
3	<0.01	≥1.0	<1.0	Thermal fault, low temp <300 °C
4	<0.1	≥1.0	≥1.0, <4.0	Thermal fault, <700 °C
5	<0.2	≥1.0	≥4.0	Thermal fault, >700 °C

Figure 5.3: ROGERS RATIO METHOD

5.2.4 Duval Triangle Method

Another important method developed by Michael Duval in 1974, serves as a power full tool for data analysis from DGA tests. This method s also discussed in IEC standard. Here 03 hydro carbon gases are under review with utmost focus. Methane, Ethylene and Acetylene levels correspond to the increasing energy levels with the gas formation process. The most powerful strength of this method is that there are no undefined or indeterminate cases using the triangle method.[20]

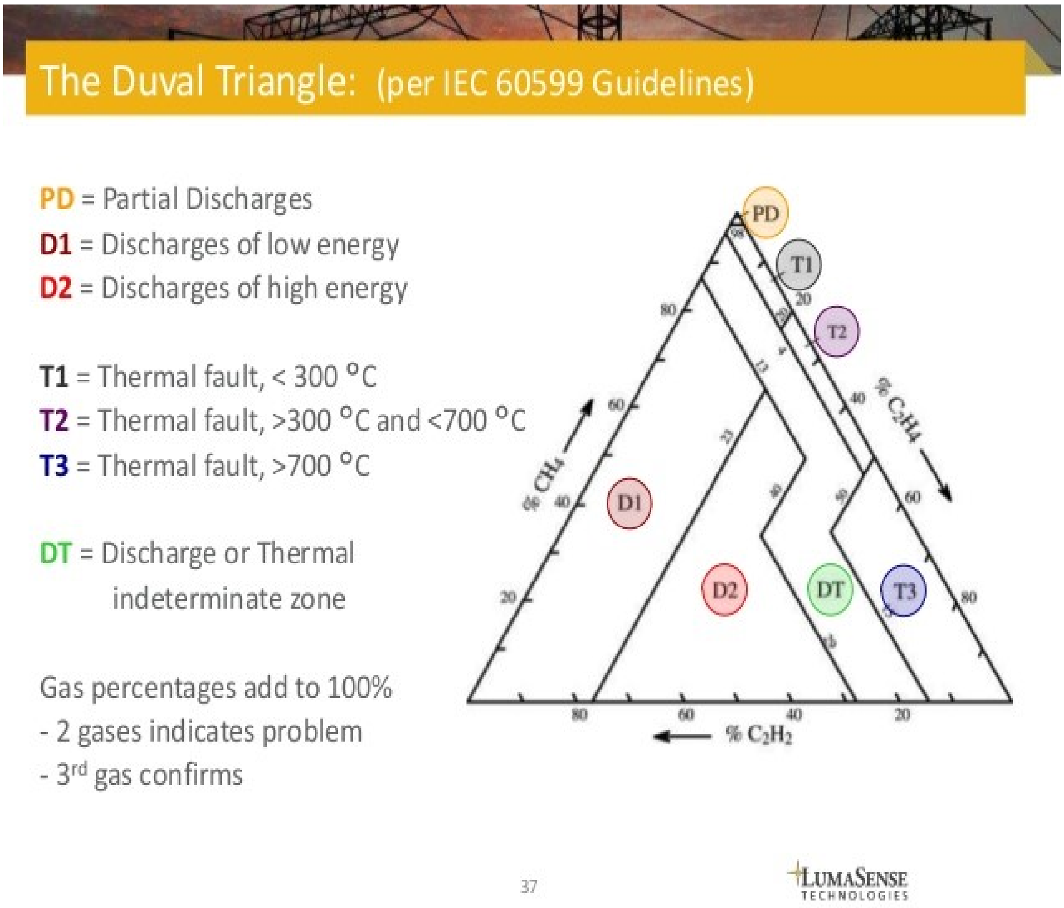

Figure 5.4: DUVAL TRIANGLE METHOD

There are some other methods as well in DGA testing but not commonly used and hence out of our point of discussion at the moment. Frequently used methods have thus been discussed briefly.

Chapter 6

HARDWARE IMPLEMENTATION

6.1 SENSORS / COMPONENTS DEPLOYED

6.2 RTD sensor (Resistance Temperature Detector)

Based on resistance measurement. The temperature coefficient of platinum resistance is positive which means that it increases with the increment in temperature. The change in the resistance is executed as a function of temperature, 0.39/1°C. The sensor proves to be stable for long terms as compared to most of the other methods of temperature measurement, precisely better than 0.2/0°C in one year.

With an effective measuring range of -58 to 572 ° F, RTDs usually have greater accuracy and enhanced repeatability as compared to thermocouples, with the drawback of slower response times. The working of the sensor is such that it tends to slightly warm up the sensing element of the device with the passage of the measurement current through the RTD sensor. This phenomenon is called self-heating. The temperature of the sensing element is thus dependent on the time it is being exposed to the relative heat, and hence is directly proportional to the length of time along with the magnitude of the measurement current involved in the subject process. The process of self-heating largely depends on the sensor's structure along with its thermal resistance with respect to its surroundings as well. It is also to be noted for an obvious fact that a small measurement error be observed from the self-heating process within the sensor.

Figure 6.1: RTD SENSOR

The current measured here is typically max of 1 mA using PT100 sensor for relevant measurements, but it can be as low as 100 ÂţA or even lower. According to standards (such as IEC 60751), the process of self-heating must be limited to maximum of 25% of the specifications mentioned as sensors tolerance. With respect to construction RTD elements have a fine coiled wire which is wrapped around a core made of ceramic or glass. Due to its fragile composition, its placement is usually done within a sheathed probe for better protection and efficient process execution.

6.3 PSA-CO1 (Automatic Pressure Switch)

Pressure Switch with the advantage of high accuracy and digital control was used for subject application here. The unit can be used to measure negative and compound pressure as well along with the standard pressure measurements requirements. With an analogue output of 1-5VDC the PSA series sensor can fit with the transformer conservator vent quite conveniently. The sensor used in our project was specifically designed to measure compound pressure of up to 100kPa such as the demand of the application. Once the sensor is mounted and tightly packed with pressure developing region like conservator top or transformer top plate, the pressure starts building up. This pressure is then sensed and translated on the high luminance screen. With an IP40 rating the sensor is protected from low pressure water jets experienced from any direction.

Figure 6.2: PSA SENSOR

6.4 ADXL335 (Accelerometer Vibration Measurement)

The ADXL335 is a small, thin, low power, complete 3-axis accelerometer with voltage outputs in the form of conditioned signals. The sensor measurement of acceleration is done with a minimum full-scale range of Âś3 g. [21] It can measure the static acceleration of gravity in applications where tilt-sensing is required, as well as dynamic acceleration caused by motion, vibration or any other type of shock.

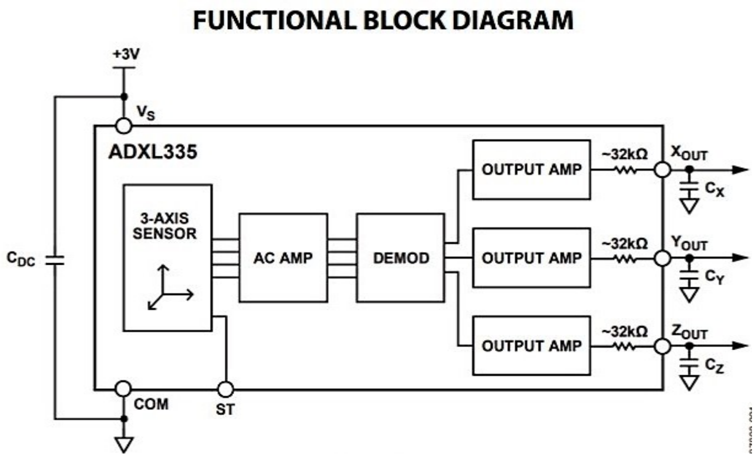

Figure 6.3: Functional Block Diagram

The main objective of this sensor is to detect any changes in the physical orientation of the unit if experienced with any actual fault. In case of any jerk or shock the physical orientation of the sensor placement shall be changed in one of the three axis which shall serve as a signature for the fault being induced the next time on the same unit. Accelerometer measure orientations in case of stress undergone by any object in terms of g, which is translated in all three axis as sensitivity in units of mV/g. The absolute maximum ratings that may cause permanent damage to the sensor device is 10000g and thus may affect the reliability of the output.

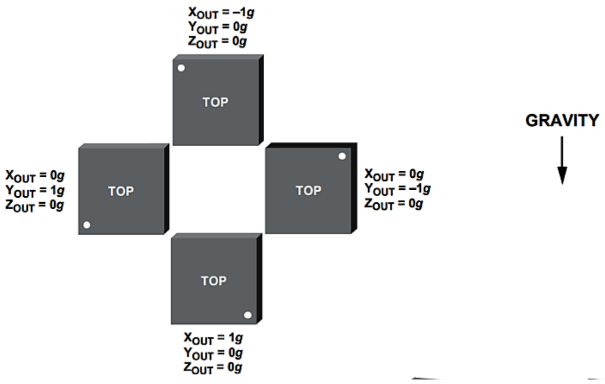

Figure 6.4: Orientation Chart

6.5 Potential Transformer (Voltage Measurement)

Potential transformer (PT) or Voltage Transformer (VT) are types of instrument transformer that are designed to get an accurate voltage measurement at the load side and obtaining secondary metering. The usage in this project relates to the real time voltage measurement to detect any abnormal faults that may affect the voltage trends of the unit being monitored. These transformers are thus used to step down the line voltages from 220 V to safe monitoring levels of 0-5V.

The Arduino can only measure DC voltages between 0 to 5V. Since we are working on a 3 phase AC System, voltage levels have to be brought into acceptable levels for the microcontroller. To measure the AC RMS Voltage of each phase potential transformers are used for stepping down voltages to 12V AC RMS. This step down AC voltage is then converted into DC voltage using a Full Bridge Rectifier and filtered using a capacitor. This gives a DC voltage of 16.97V. A voltage divider circuit is used after the Full Bridge Rectifier which gives a maximum peak voltage of 5V DC for corresponding maximum peak of single phase AC Voltage i.e. 311V. Arduino reads this DC voltage in the range of 0 to 1023 which corresponds to 0 to 5V respectively.

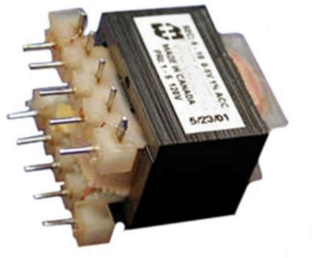

Figure 6.5: Potential Transformer

6.6 Current Transformer (Current Measurement)

Despite the fact that this current sensor is part of phase I of the project we have continued the data acquisition from this sensor as well for accurate measurements and enhancement of the monitoring system. Like Potential Transformer, Current Transformers are also types of instrument transformer that are manufactured to obtain an accurate measurement of the current values from the system installed at the load side in order to achieve an enhanced protection system to operate and isolate the unit from the fault effectively. The usage in this project relates to the real time measurement of current values to detect or identify any abnormal faults that may alter the current trends of the unit being monitored in normal or healthy conditions. The safe current values normalized by the CT are sent to the circuit via CT sensor module for accurate measurements and protect the circuit itself from any short circuit currents. ACS712T-MOD [22] has the ability of providing economical and accurate solutions for detecting and identifying AC/DC current in commercial, industrial and communications systems..

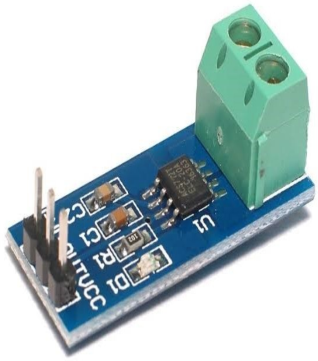

Figure 6.6: Current Transformer

6.7 BASIC ARCHITECTURE

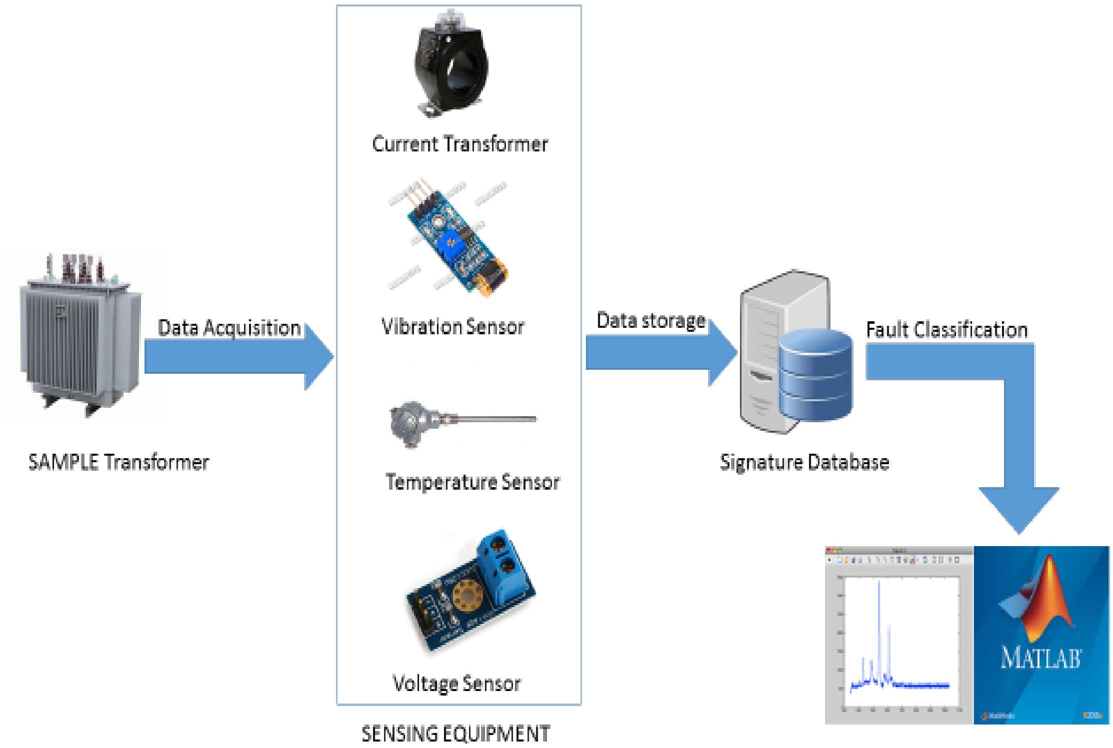

Figure 6.7: BASIC ARCHITECTURE

6.8

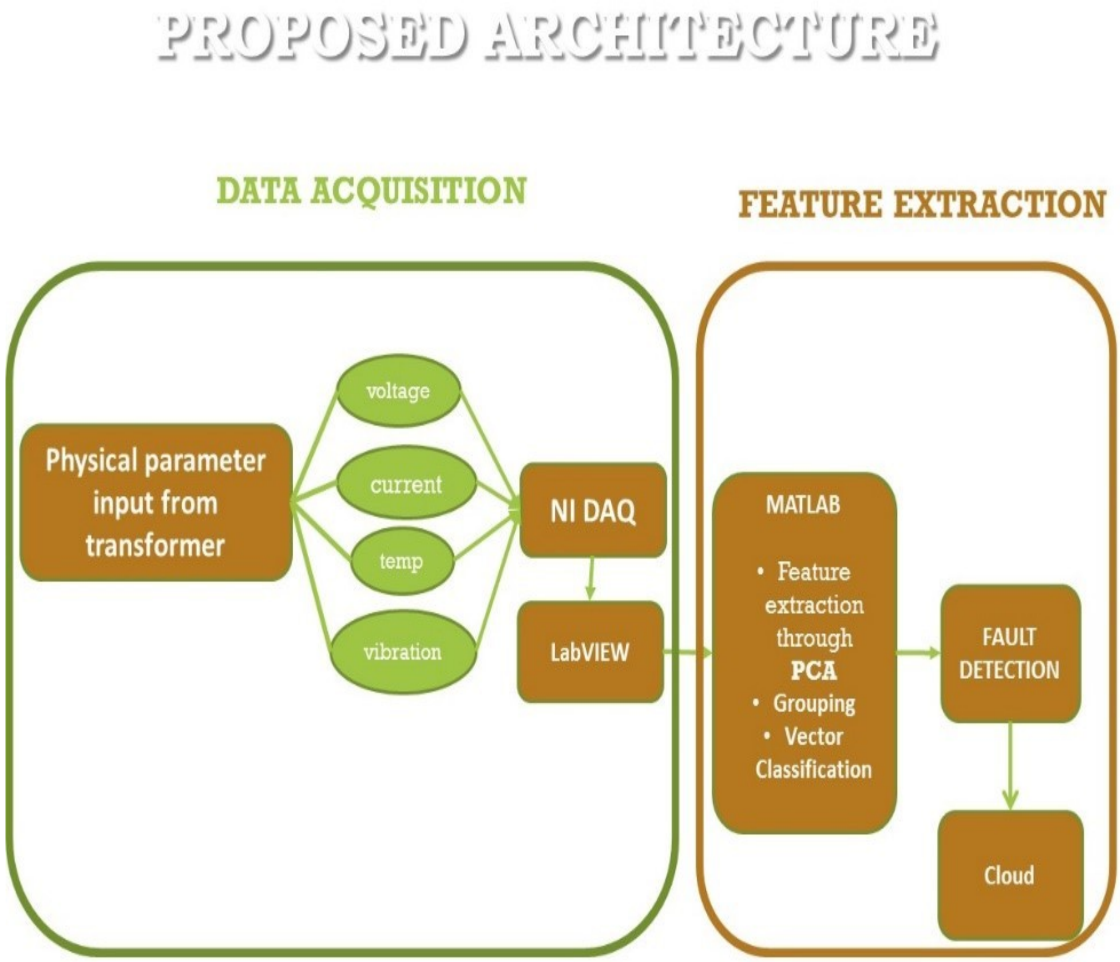

Figure 6.8: METHODOLOGY

Once the data from the sensor is obtained from the transformers in healthy condition, it is then stored in server libraries and systems. Later on faults are induced in to the system that may occur in real time imitating actual scenarios of the electrical network. Similar procedure is then repeated with the same sensors installed in place but in created fault conditions. Here we are expecting a noticeable and distant separation or identification in the parameters signatures obtained. All the signatures captured and then sent to the DAQ card. NI DAQ card 6009 is used for this purpose in this project. The card has the ability to connect 14 different channels with all the sensors at a single time without any disconnections. The data sampling is carried out a rate of 48kS/sec. DAQ card is thus used for signal conditioning and parameters value mapping as it exploits the processing power, display, productivity and connectivity capabilities with the live subject system leading to an enhanced and more powerful, cost effective and flexible solution for measurement. Concluding the DAQ card serves the purpose of an interface between the computer system and the signals received from the outside world.

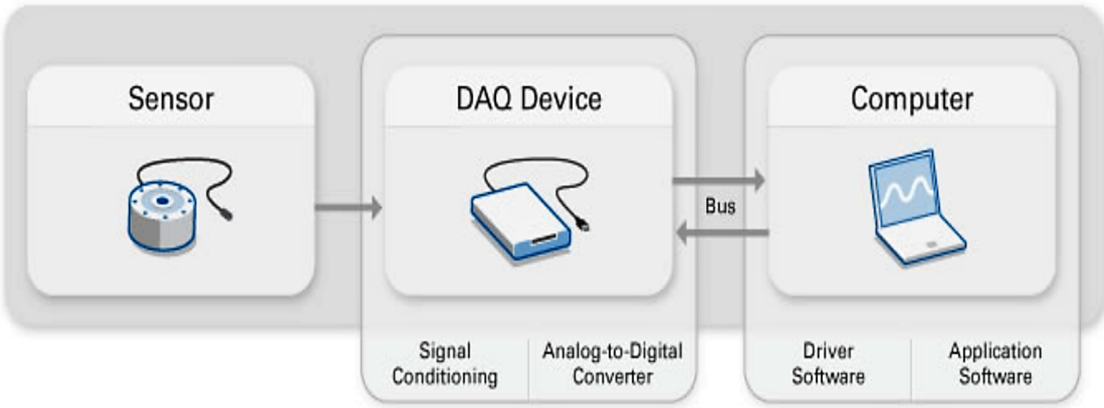

Figure 6.9: DATA FLOW

6.8 EMBEDDED SYSTEM MODULE

In order to industrialize the product we embedded all the monitoring system in an embedded system to be mounted with the transformer unit to ease of use and portability to ensure user friendly of the final product. In this context the parameter monitoring prototype device was designed and fabricated.

Parameter readings are sensed and transmitted via GPRS using a postpaid SIM card used with the SIM800C module. The GPRS signals are used to transmit the data on a cloud-based location. This location is directly connected to the server which displays the same data of the last interval on the web page (THINGSPEAK). Same data can also be obtained on a compatible mobile application that connects with the main server providing real time data with very little time delay.

Trends that are observed in the software graphs prevailing for the 3-4 readings taken over a specific period of time may guide us in taking efficient decision for our electrical assets. The device is also made from locally available sensors and proves to be a quite inexpensive designed monitoring module. This module helps the operator to take preventive measures in case of any fault or abnormal situation in a timely manner, having the risk of asset damage or loss to be minimized to a huge extent. Fault data and logs can be stored and later retrieved from the memory card to take preventive measures for avoiding any transient repetitive faults recorded throughout a span of time. Some of the main components used notable are discussed below.

6.8.1 ARDUINO MEGA

The Arduino MEGA 2560 is an open source microcontroller board based on the Microchip ATmega8U2 instead of the FTDI chip. The board is equipped with sets of digital and analog input/output (I/O) pins that may be linked to different expansion boards (shields) and other circuits. This results in faster transfer rates added twice the flash memory. The board has 54 digital I/O pins (fourteen capable of PWM output), 16 analog I/O pins, 16 MHz crystal oscillator, 4 UARTs (hardware serial ports), a power jack, an In Circuit Serial Programming header, a USB connection, and a button to reset the program. It holds everything required to support the microcontroller unit, and shall be activated by simply connecting it to a computer or power it up with an AC-to-DC adapter / battery. A type B-USB cable can thus be utilized to program the Mega with the Arduino IDE (Integrated Development Environment).

The power requirement for the MEGA can be sufficed by a USB cable or an external 9-volt battery, although the tolerance limit in voltages are between 7 and 20 volts. There is a huge variety of designs in Arduino board using microprocessors and controllers. The boards are interfaced with sets of analogue and digital input/output (I/O) pins that may be interfaced to various expansion boards (shields) or breadboards (for prototyping) and other circuits designing.

6.8.2 THINGSPEAK (IoT Platform)

ThingSpeak™ is an IoT platform service for analytics that allows to combine, envisage and evaluate live data streams in the cloud. ThingSpeak provides instant evaluations of data transmitted by devices to ThingSpeak platform. Online analysis and data processing can be easily executed with processing ability of MATLAB® code in ThingSpeak as the data gets imported in. [23] ThingSpeak is usually utilized for the objective of prototyping and proofing of the concept for IoT systems that require analytics.

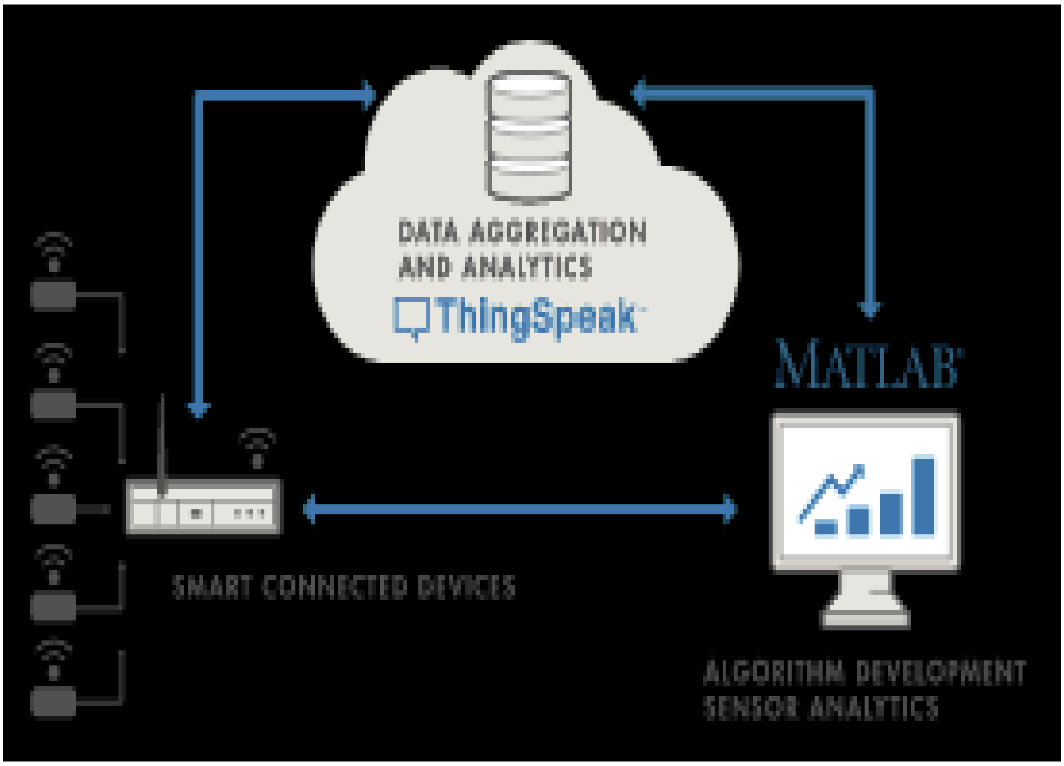

Figure 6.10: THINGSPEAK

The devices mounted at the left of the module 6.12 are the sensors that are being used in our embedded system These are used to collect the data after being mounted on the actual unit Next comes the cloud part where the data that is being gathered is aggregated and analyzed in real time by a special platform 6.11 designed for this purpose. Later the aspect of algorithm development gets into the picture where an insight is gained using the help of historical data gathered. This in turn relates to the linked engineer or the operator to act as a trigger to normalize any abnormal situation or take preventive timely action.

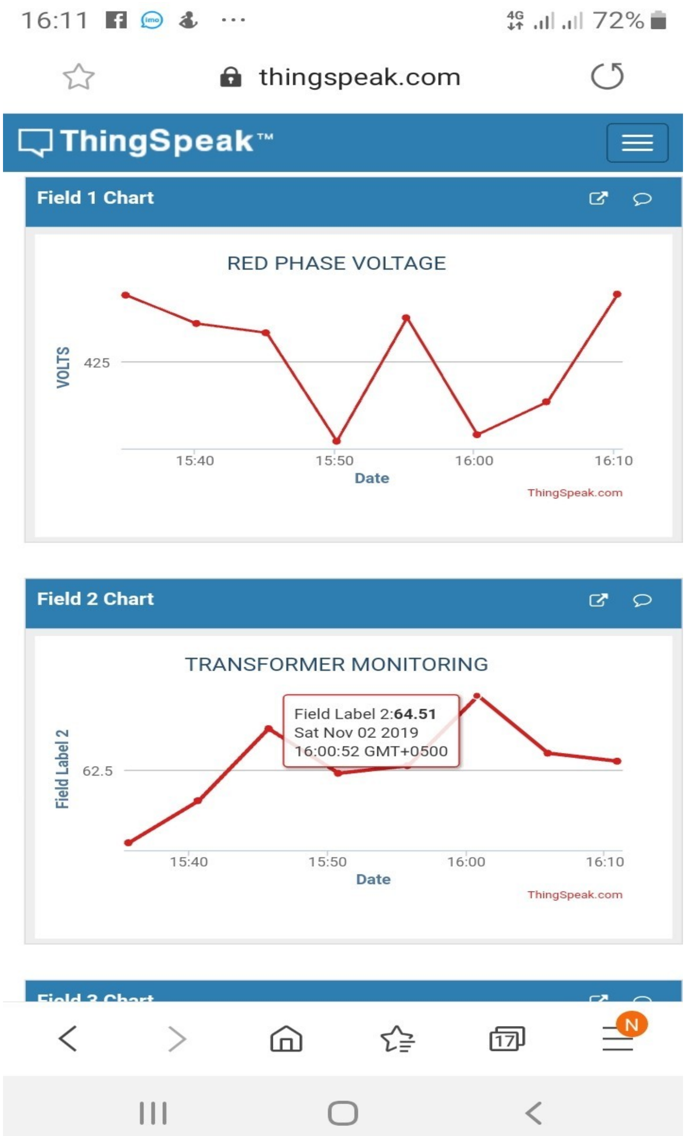

Figure 6.11: SOFTWARE OUTPUT

Figure 6.12: EMBEDDED MODULE

Chapter 7

TESTING, SIMULATION INTERFACING and RESULTS

7.1 LABVIEW INTERFACE

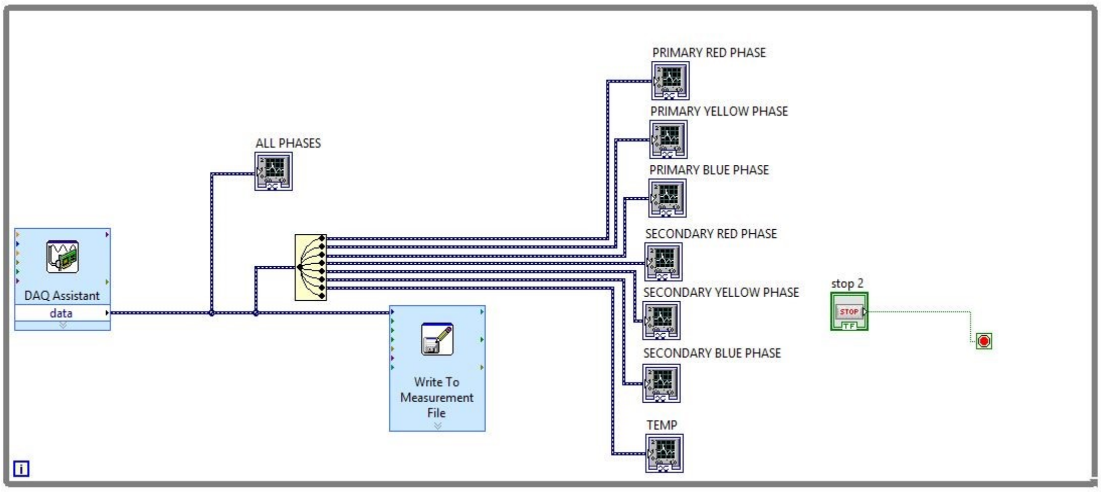

Figure 7.1: LABVIEW INTERFACE DESIGNED

The above figure indicates the lab view interface deigned as per the project needs. As seen this framework was the inter connection of the hardware (transformer units) and the software (MATLAB). Labview is a engineering software with a destined acronym Laboratory Virtual Instrument Engineering Workbench. It is designed for system applications that do require testing, measurement and controlling instrumentation with readily access and approachability to the systems hardware and its data insights.

7.2 MATLAB feature extraction

BFeature extraction from the data set has been done through PCA. Principal component analysis (PCA) is a statistical procedure that engages an orthogonal conversion to transform a set of values of variable that could be correlated (entities which portray similarity on various numerical values based on data patterns) into a set of linearly non-correlated values, also termed as principal components.

Principal Component Analysis is a methodology used to illuminate the variance-covariance structure of a data set obtained from variables through linear combinations. It is often used as a technique for dimensionality-reduction of the set of values.

The main idea of PCA is to decrease the data set dimensionality containing several variables interconnected with each other, either densely or lightly, while retaining the variations present in the dataset, up to the maximum possible extent. As a layman, it is a method of summarizing data. In simple terms PCA can be modularized into the following basics steps.

- 1. Accumulate co-related data
- 2. Evaluate / Calculate the mean data or value from the data set.
- 3. Compute Co-variance matrix
- 4. Obtain the eigen values and eigen vectors
- 5. Pick the highest eigen values.
- 6. Project the data points to the subject eigen vectors
- 7. Segregate the un-correlated data.

7.3 Data Classification

Data decomposition was then achieved after complete sampling of 25 testing intervals for all cases individually. The process was achieved using Daubechies-4 with 6 level decomposition. The matrix data was then dimensionally reduced using the tool of PCA. Later feature selection process was executed using the following classifiers:

- a. Tree
- b. Support Vector Machine (SVM).
- c. K- Nearest neighbor
- d. Ensemble

Accuracy classes were then determined and evaluated to classify the data accordingly as per respective algorithms individually. Optimal Classifier was then identified after repeated testing and training of the data set using types mentioned with multiple data variants.

A total of 20,001 entries were recorded in every data instance file recorded making a total data entries of while measuring several parameters like voltage, current, vibration and temperature making a total 25,601,280 data entries to be recorded, analyzed, segregated and later classified to develop an effective and efficient training data set for library of a transformer health monitoring system

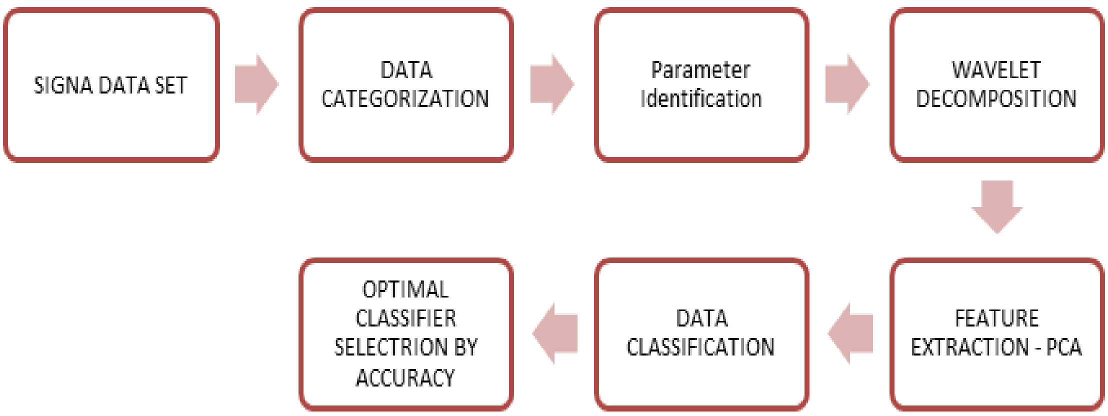

Figure 7.2: FLOW OF DATA SET FROM RAW FORM TO CLASSIFIED INFORMATION FOR TRAINING LIBRARY

7.4 TESTING RESULTS

As per our programming code where Wavelet and PCA tools were applied, we gathered data from 15 transformer units while obtaining 128 data sets accumulating around 384 sample data for 3 phases. The cumulative data sets were combined and used to classify it as per the methodology of unique feature selection. Data was classified in to five major categories to identify healthy data from different types of faulty data.

- 1. Healthy
- 2. Line to Ground
- 3. Line to Line (Grounded)
- 4. Line to Line (Non-Grounded)
- 5. Triple Line

As per the results different classifiers were applied on the results set after application of necessary tools and ENSEMBLE classifier produced most effective results with 72.7% overall accuracy achieved. On the contrary 27.3% error was observed where the Line to Line (Non-Grounded) faults exhibited the most erroneous data. This was mainly because prominent features were not obtained in subject scenario to categorize it from other faults or other data effectively as other four types of data classes defined.

Within the sampled data transformer parameters like Voltage, Current, Vibration, and Temperature were monitored to detect and analyze unique features emerging in each type of scenario for better classification. The signals were sampled at a rate of 20Ks/sec for undefined time length. Several repetitions for testing were executed to cater best results be included in the classification and thereby avoiding any non-important data that could not be imported in our program.

Our approved classifier was Ensemble classifier with the algorithm type of Bagged Trees. Bagged Tree is a tool used to reduce the variance of a decision tree. [24] Like other machine learning techniques, Bagged tree is also a supervised machine learning technique to classify the data as effectively as possible while ensuring the best predictive performance in the process. Ensemble uses the basic concept of data classification in such means as many weak learners combine to form a strong learner.

The idea behind the Bagged tree algorithm is the creation of several subsets of data from the randomly chosen training sample with replacement. Each subset data is then later used to train their decision trees. Hence we can conclude with an ensemble of different models having the average from all the predictions of different trees to be used emerging as a more robust model than as compared to a single decision tree. An algorithm that has high variance are actually decision trees like classification and regression trees.

Bagging is the application of bootstrap procedure to a machine learning algorithm having high variance. In the classification of bagging with decision trees the system or the relevant user is less troubled about individual trees over fitting the data set used for training. [25] Hence the individual trees are grown deep and are not cropped. The resulting trees shall have high variance and low bias. The consideration factor here is the no of samples and therefore the no of trees to include. The trade off point must then be decided after running successful cross validation test to determine the point when increasing data sets no longer increases the accuracy of the training system.

As mentioned below the data set confusion matrix for each classifier is discussed below with its accuracy and error ratio clearly defining the true positive identities where the prediction class is similar to the predicted class and vice versa for false identification identities.

7.4.1 ENSEMBLLE CLASSIFIER - Bagged Trees

Algorithm - Bagged Trees As seen in the classification learner screen shot below, the overall accuracy of 72.7% makes this as an optimal classifier for our training data set. Here in the algorithm the healthy data set denoted by *"H"* has been able to classify 84% of the data as true positive identities where the data of healthy transformer has been classified and predicted correctly while 16% of the data has been classified in correct where the predictive model was unable to classify the data as Healthy based on its features to be selected. Furthermore the confusion matrix shown below is summarized as per the true classes identified. On the contrary, the data set of Line to Line (Un-Grounded) has been the most poorly identified data set with only 66.2% accuracy class relating to true positive identities. It is also to be noted that here the classifier had an error ratio of 33.8% making this group of data set the least accurately classified. There would have several reasons behind this but to sum it all, this particular data set failed to produce unique features for the classifier to be identified.

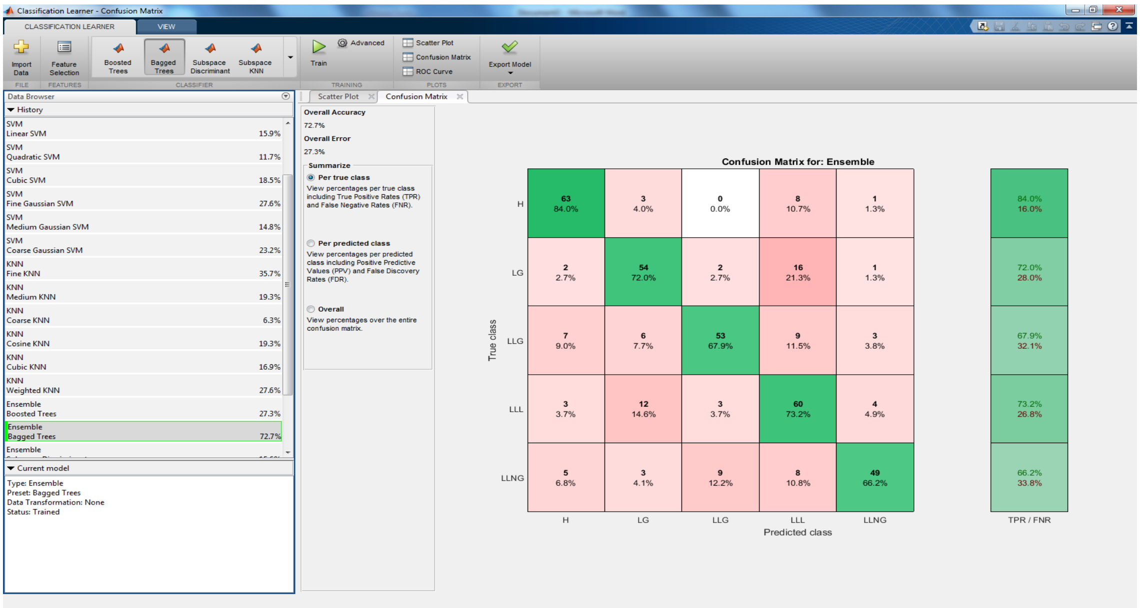

Figure 7.3: Classification Learner - Per True Class (Optimal Classifier)

Total Identifications = Total Positive Identifications + Total False Identifications

The figure below shows the same classifier performance with regard to predicted cases indicating positive predicted rates and false discovery rates of all classified groups.

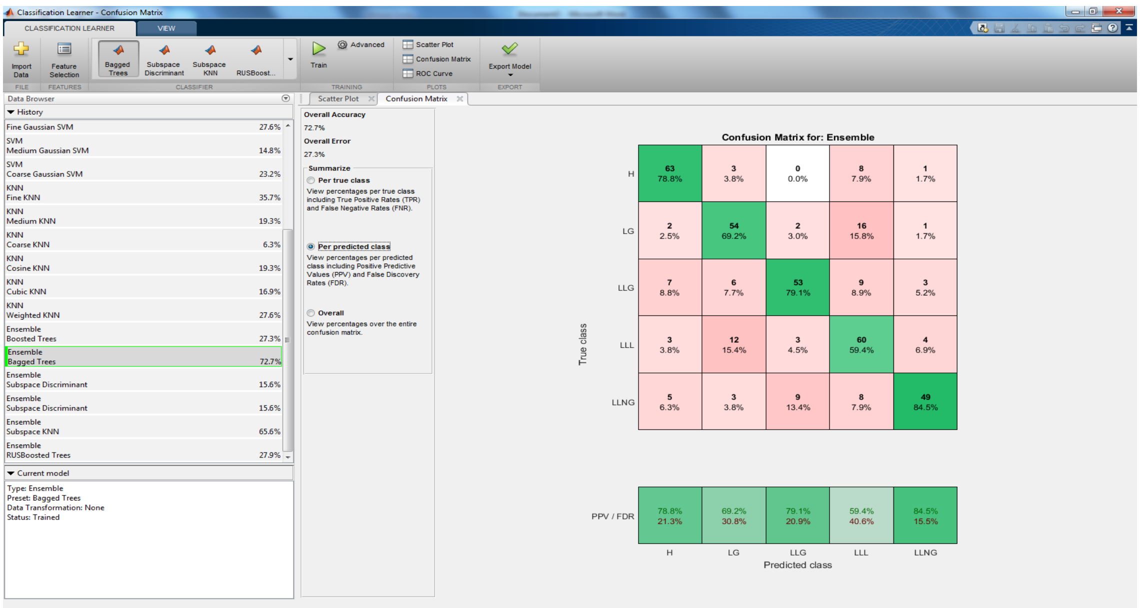

Figure 7.4: Classification Learner - Per Predicted Class (Optimal Classifier)

An overall summary of the classification can be viewed in the picture below.

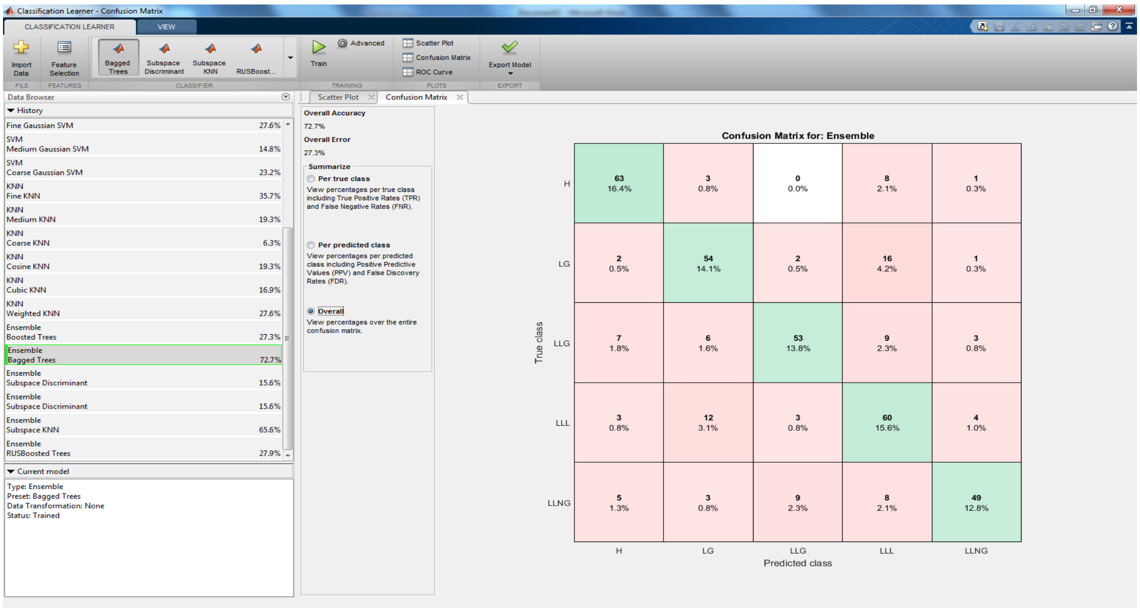

Figure 7.5: Classification Learner - OVERALL (Optimal Classifier)

7.4.2 ENSEMBLLE CLASSIFIER - Boosted Trees

Algorithm - Boosted Trees (ACC - 27.3%) Boosted Trees correctly classified 105 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 7 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

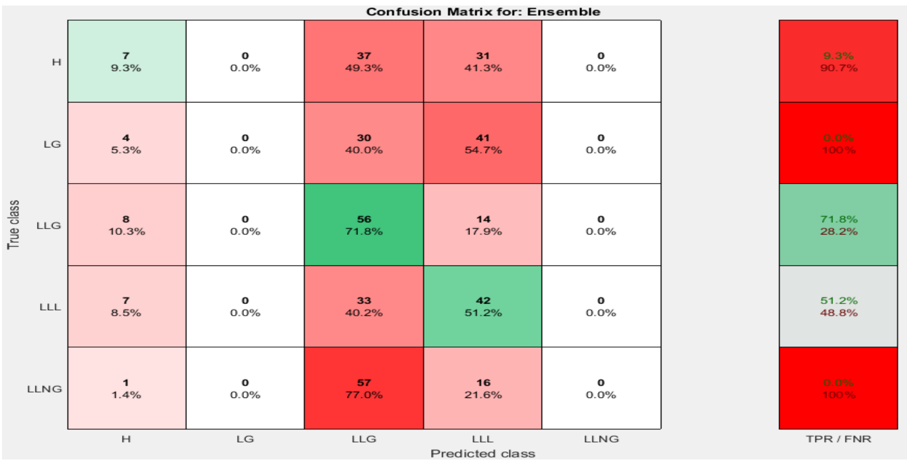

Figure 7.6: Classification Learner - Per True Class (Boosted Classifier)

7.4.3 ENSEMBLLE CLASSIFIER - Subspace Discriminant

Algorithm - Subspace Discriminant (ACC - 15.6%) Subspace Discriminant correctly classified 60 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 15 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

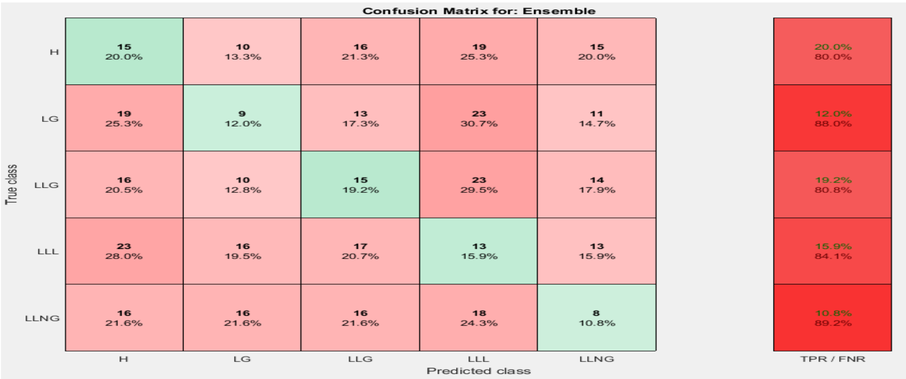

Figure 7.7: Classification Learner - Per True Class (Subspace Discriminant)

7.4.4 ENSEMBLLE CLASSIFIER - Subspace KNN

Algorithm - Subspace KNN (ACC - 65.6Subspace KNN correctly classified 252 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 58 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

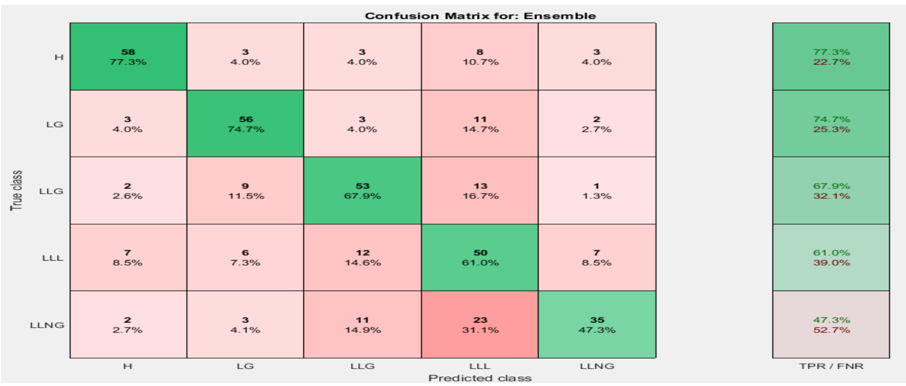

Figure 7.8: Classification Learner - Per True Class (Subspace KNN)

7.4.5 ENSEMBLLE CLASSIFIER - RUS Boosted Trees

Algorithm - RUS Boosted Trees (ACC - 27.9%) RUS Boosted Trees correctly classified 107 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 5 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

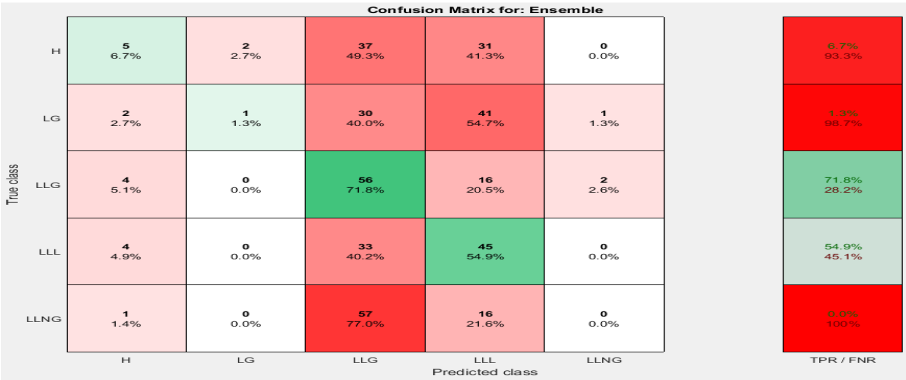

Figure 7.9: Classification Learner - Per True Class (RUS Boosted Trees)

7.4.6 TREE CLASSIFIER - Complex Tree

Algorithm - Complex Tree (ACC - 55.2%) Complex Tree correctly classified 212 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 37 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

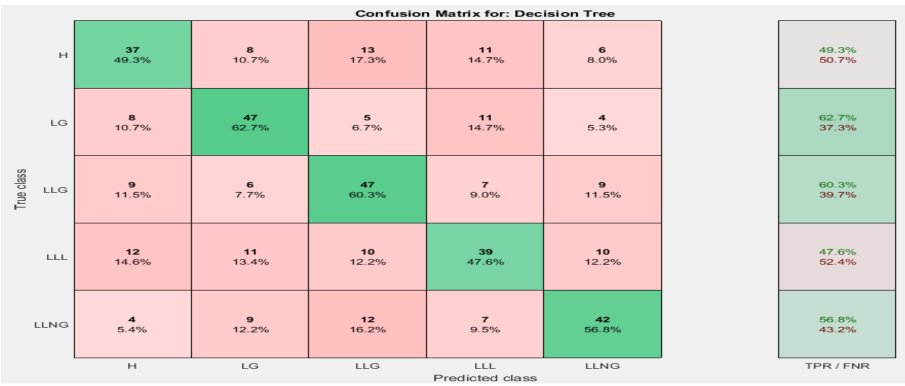

Figure 7.10: Classification Learner - Per True Class (Complex Tree)

7.4.7 TREE CLASSIFIER - Medium Tree

Algorithm - Medium Tree (ACC - 44.0%) Medium Tree correctly classified 169 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 39 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

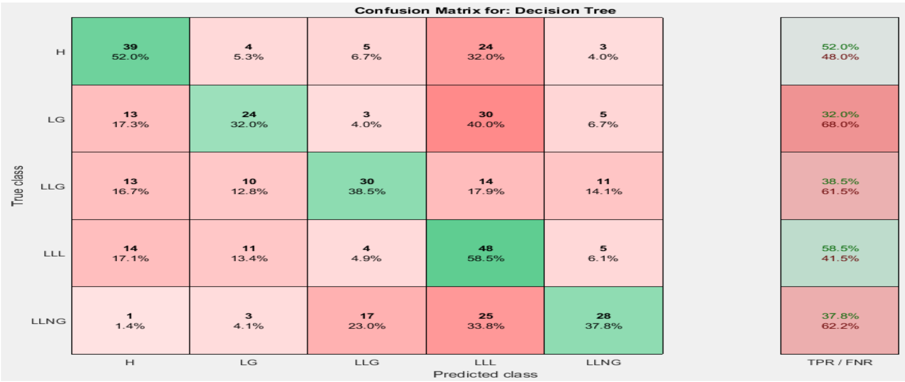

Figure 7.11: Classification Learner - Per True Class (Medium Tree)

7.4.8 TREE CLASSIFIER - Simple Tree

Algorithm - Simple Tree (ACC - 32.0%) Simple Tree correctly classified 123 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 29 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

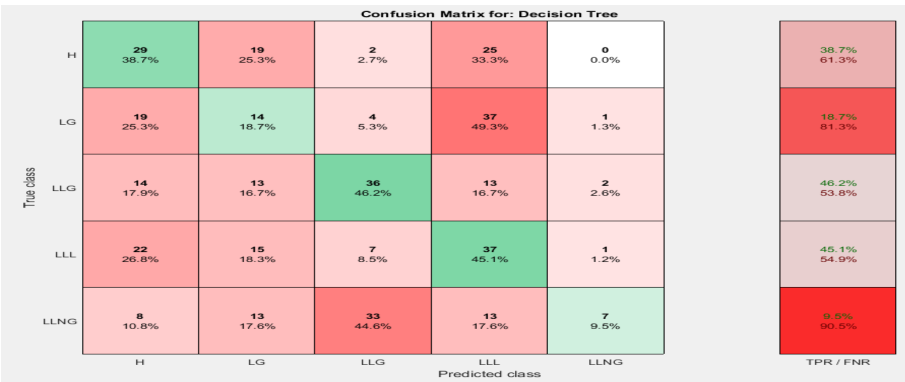

Figure 7.12: Classification Learner - Per True Class (Simple Tree)

7.4.9 SUPPORT VECTOR MACHINE - Linear SVM

Algorithm - Linear SVM (ACC - 15.9%) Linear SVM correctly classified 61 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 13 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

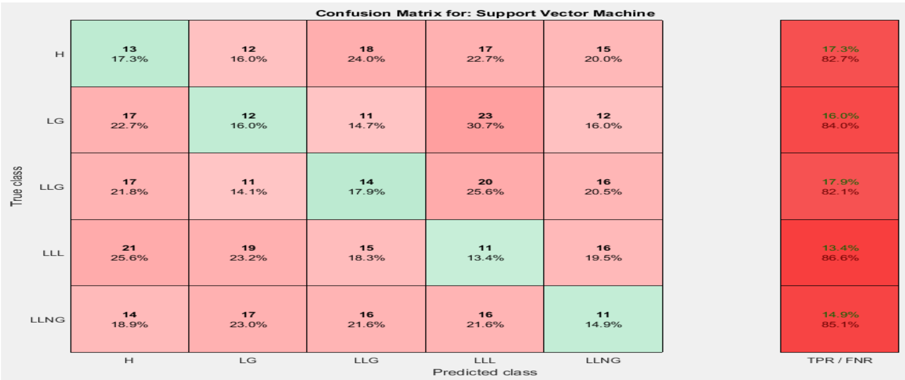

Figure 7.13: Classification Learner - Per True Class (Linear SVM)

7.4.10 SUPPORT VECTOR MACHINE - Quadratic SVM

Algorithm - Quadratic SVM (ACC - 11.7%) Quadratic SVM correctly classified 45 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 7 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

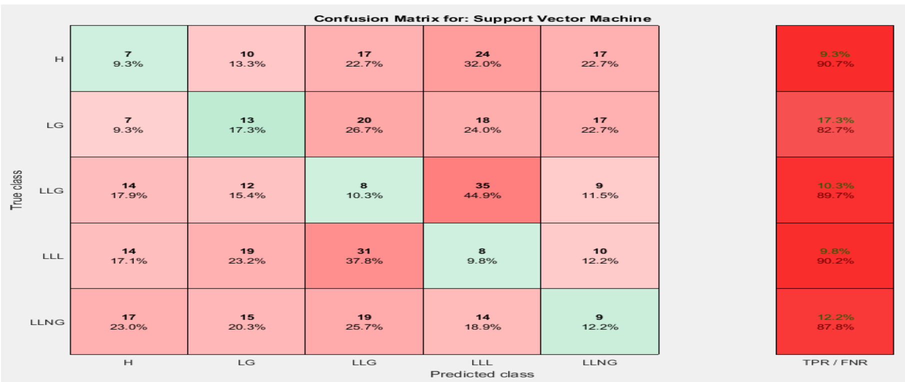

Figure 7.14: Classification Learner - Per True Class (Quadratic SVM)

7.4.11 SUPPORT VECTOR MACHINE - Cubic SVM

Algorithm - Cubic SVM (ACC - 18.5%) Cubic SVM correctly classified 71 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 16 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of faulty and healthy signal data at wavelet level 6.

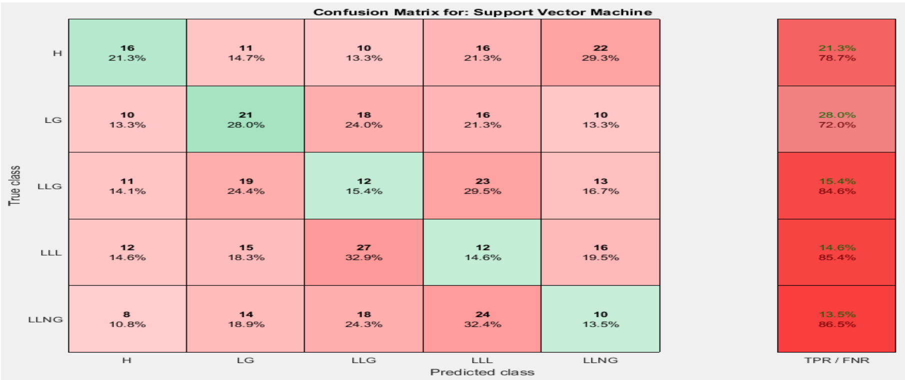

Figure 7.15: Classification Learner - Per True Class (Cubic SVM)

7.4.12 SUPPORT VECTOR MACHINE - Fine Gaussian SVM

Algorithm - Fine Gaussian SVM (ACC - 27.6Fine Gaussian SVM correctly classified 106 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 02 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

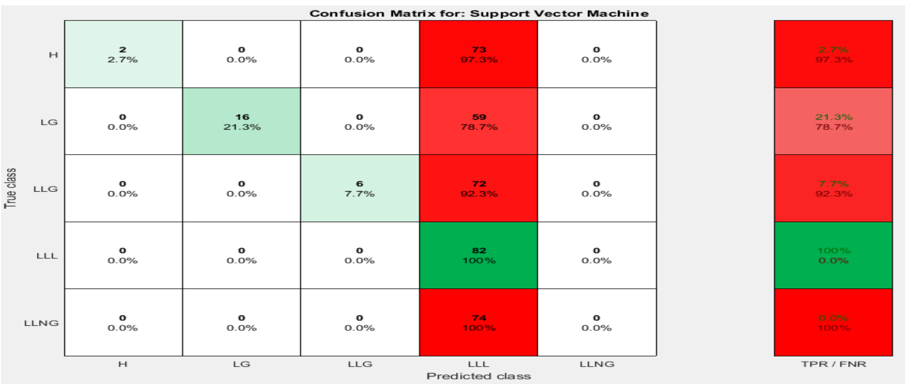

Figure 7.16: Classification Learner - Per True Class (Fine Gaussian SVM)

7.4.13 SUPPORT VECTOR MACHINE - Medium Gaussian SVM

Algorithm - Medium Gaussian SVM (ACC - 14.8Medium Gaussian SVM correctly classified 57 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 11 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

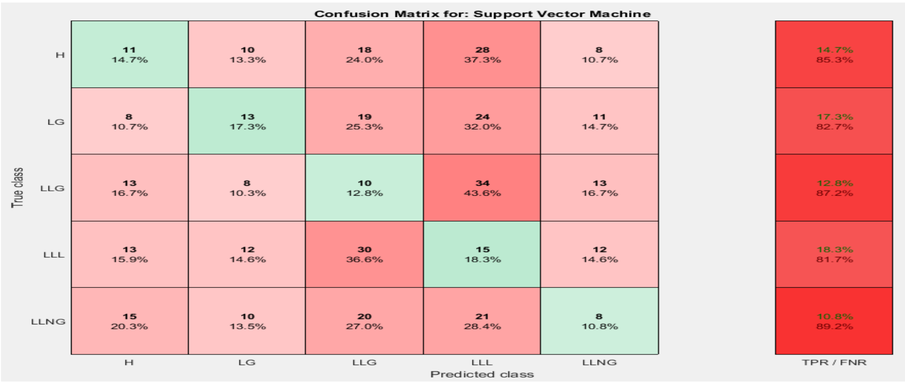

Figure 7.17: Classification Learner - Per True Class (Medium Gaussian SVM)

7.4.14 SUPPORT VECTOR MACHINE - Coarse Gaussian SVM

Algorithm - Coarse Gaussian SVM (ACC - 23.2Coarse Gaussian SVM correctly classified 89 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 02 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

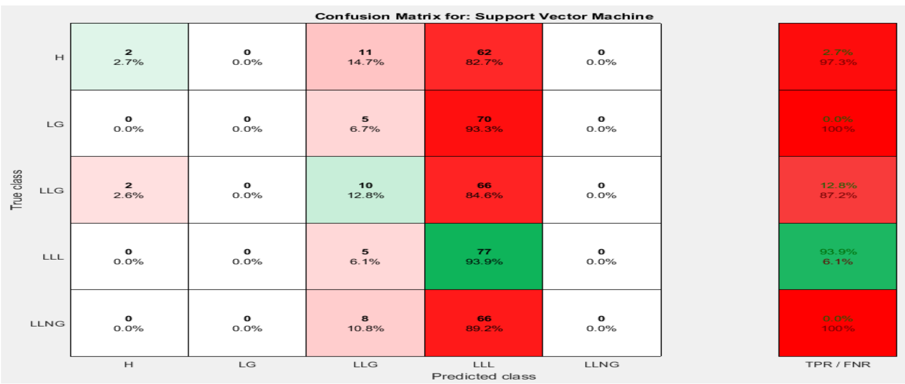

Figure 7.18: Classification Learner - Per True Class (Coarse Gaussian SVM)

7.4.15 K-NEAREST NEIGHBOR - Fine KNN

Algorithm - Fine KNN (ACC - 35.7Fine KNN correctly classified 167 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 41 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

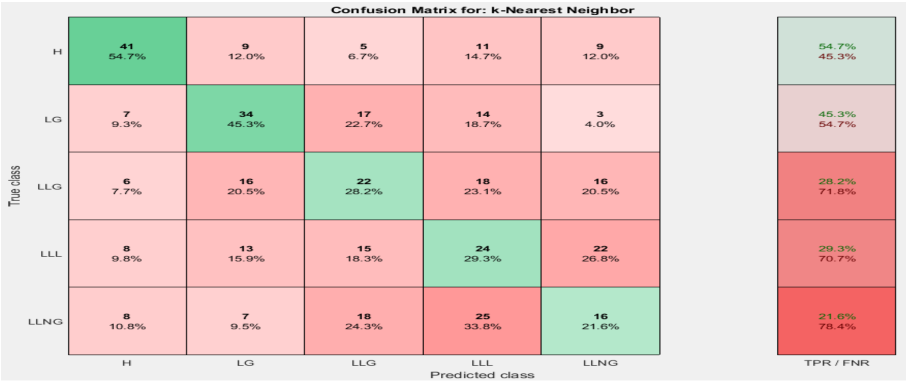

Figure 7.19: Classification Learner - Per True Class (Fine KNN)

7.4.16 K-NEAREST NEIGHBOR - Medium KNN

Algorithm - Medium KNN (ACC - 19.3Medium KNN correctly classified 74 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 25 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

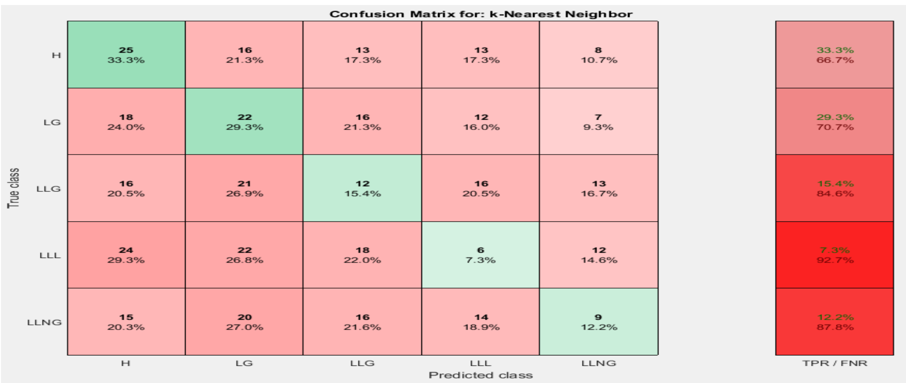

Figure 7.20: Classification Learner - Per True Class (Medium KNN)

7.4.17 K-NEAREST NEIGHBOR - Coarse KNN

Algorithm - Coarse KNN (ACC - 6.3Coarse KNN correctly classified 24 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 05 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

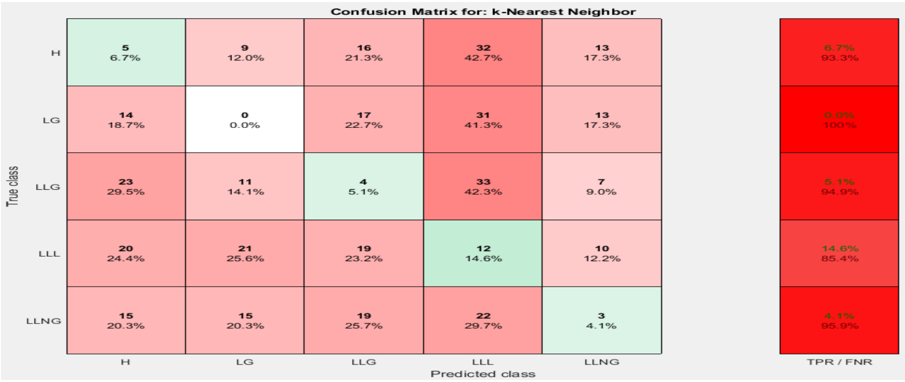

Figure 7.21: Classification Learner - Per True Class (Coarse KNN)

7.4.18 K-NEAREST NEIGHBOR - Cosine KNN

Algorithm - Cosine KNN (ACC - 19.3Cosine KNN correctly classified 74 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 27 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

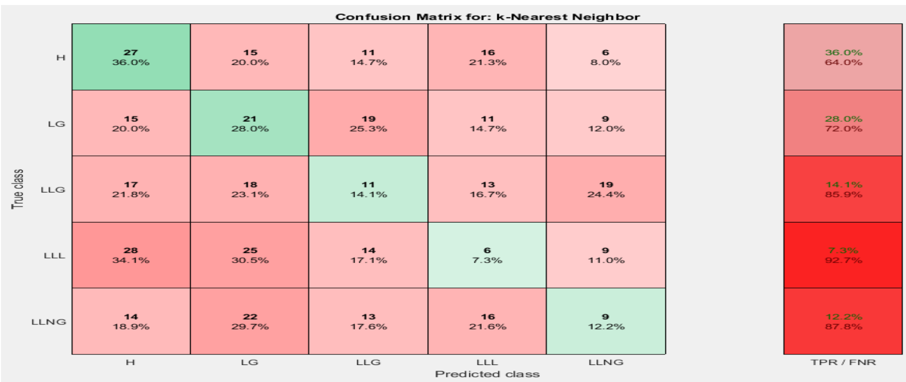

Figure 7.22: Classification Learner - Per True Class (Cosine KNN)

7.4.19 K-NEAREST NEIGHBOR - Cubic KNN

Algorithm - Cubic KNN (ACC - 16.9Cubic KNN correctly classified 65 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 26 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

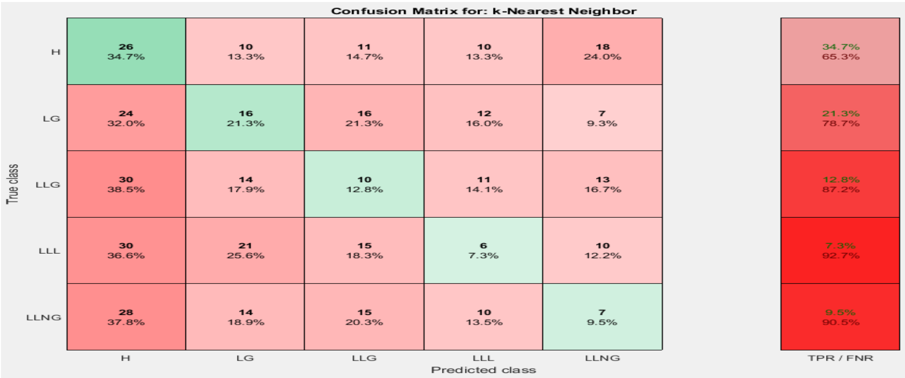

Figure 7.23: Classification Learner - Per True Class (Cubic KNN)

7.4.20 K-NEAREST NEIGHBOR - Weighted KNN

Algorithm - Weighted KNN (ACC - 27.6Weighted KNN correctly classified 106 data sets healthy + faulty events out of a total of 384 events. Correct classification of healthy events was 35 events out of a total of 75 events. This classification of events was performed on wavelet decomposition of fault and healthy signal data at wavelet level 6.

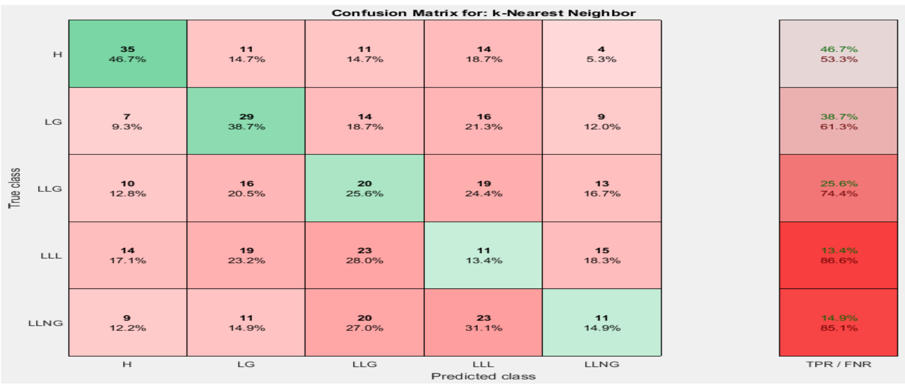

Figure 7.24: Classification Learner - Per True Class (Weighted KNN)

7.5 A Closer Look at the Optimal Classifier

AS discussed earlier we have deduced that our optimal classifier emerged from the subject data set was ENSEMBLE with Bagged Tree algorithm implemented.

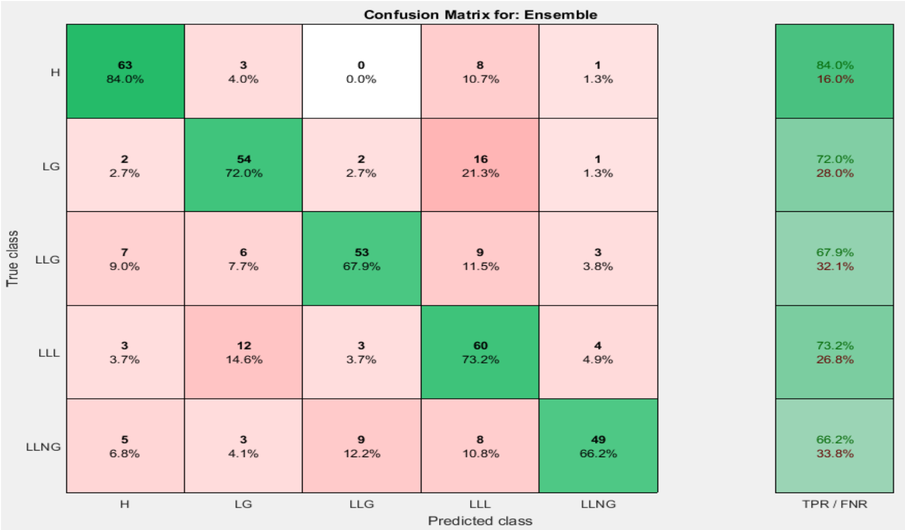

Figure 7.25: Classification Learner - Per True Class (Bagged-PTC)

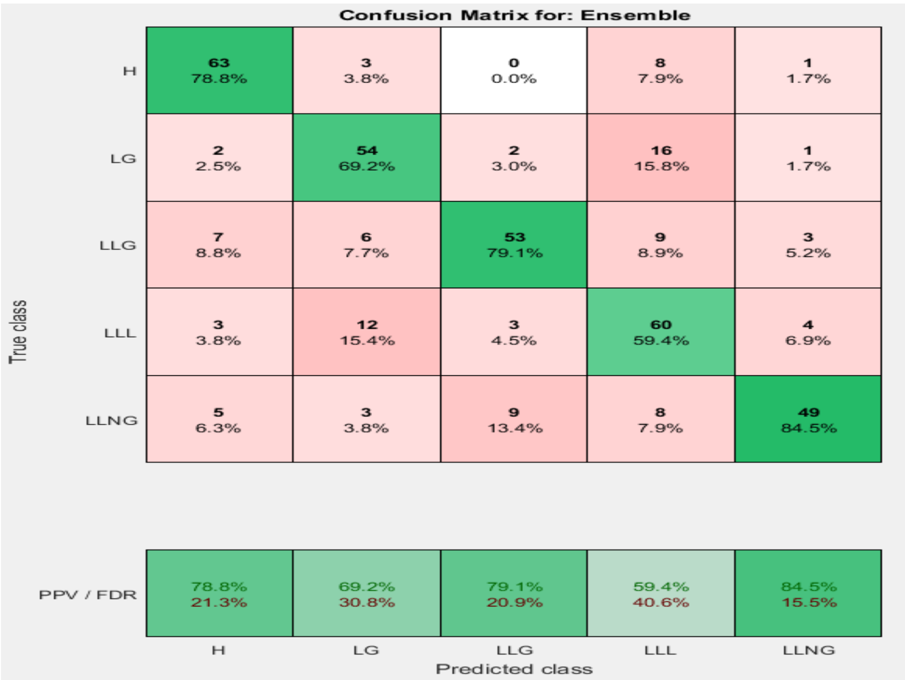

Figure 7.26: Classification Learner - Per Predicted Class (Bagged-PPC)

A closer look into the results indicates the following observations. (PER TRUE CLASS)

For healthy data, the Bagged tree algorithm detected 84% true positive identifications while declaring 63 data sets. On the other hand 10.7% results were false identifications where healthy data was predicted as triple line to line fault.

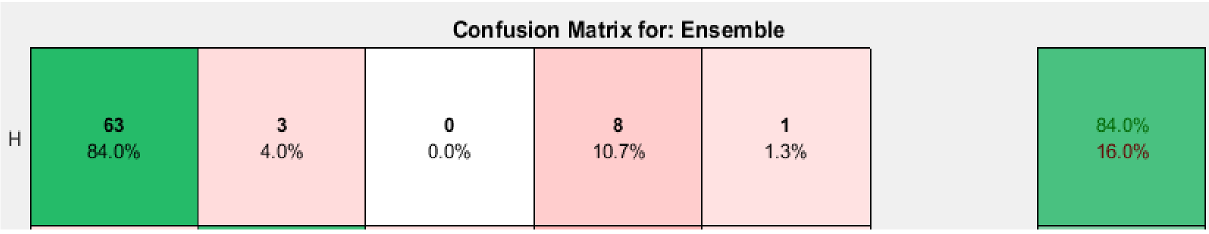

Figure 7.27: Healthy Signals Comparison (Bagged-PPC)

For faulty line to ground data, the Bagged tree algorithm detected 72% true positive identifications while declaring 54 data sets. On the other hand 21.3% results were false identifications where line to ground fault data was predicted as triple line to line fault.

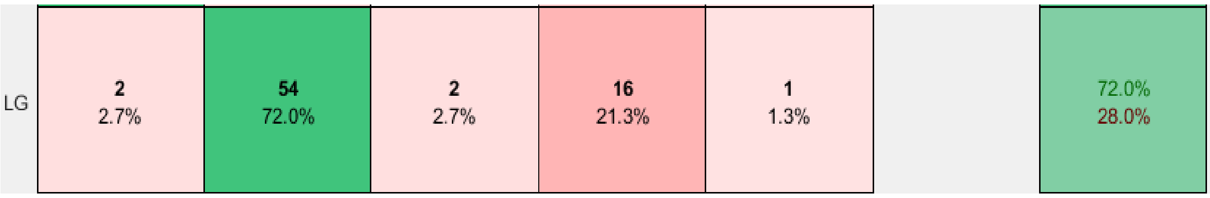

Figure 7.28: Line to Ground Fault (Bagged-PPC)

For faulty double line to line (ground) data, the Bagged tree algorithm detected 67.9% true positive identifications while declaring 53 data sets. On the other hand 11.5% results were false identifications where double line to line (ground) fault data was predicted as triple line to line fault.

Figure 7.29: Double Line to Ground Fault (Bagged-PPC)

For triple line to line data, the Bagged tree algorithm detected 73.2% true positive identifications while declaring 60 data sets. On the other hand 14.6% results were false identifications where triple line to line fault data was predicted as line to ground fault.

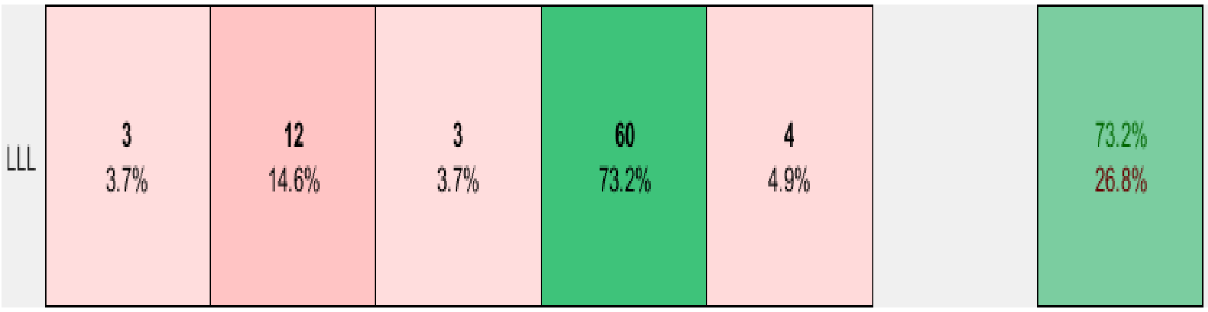

Figure 7.30: Triple Line to Line Fault (Bagged-PPC)

For faulty double line to line (un-ground) data, the Bagged tree algorithm detected 66.2% true positive identifications while declaring 49 data sets. On the other hand 12.2% results were false identifications where double line to line (un-ground) fault data was predicted as double line to line (ground) fault.

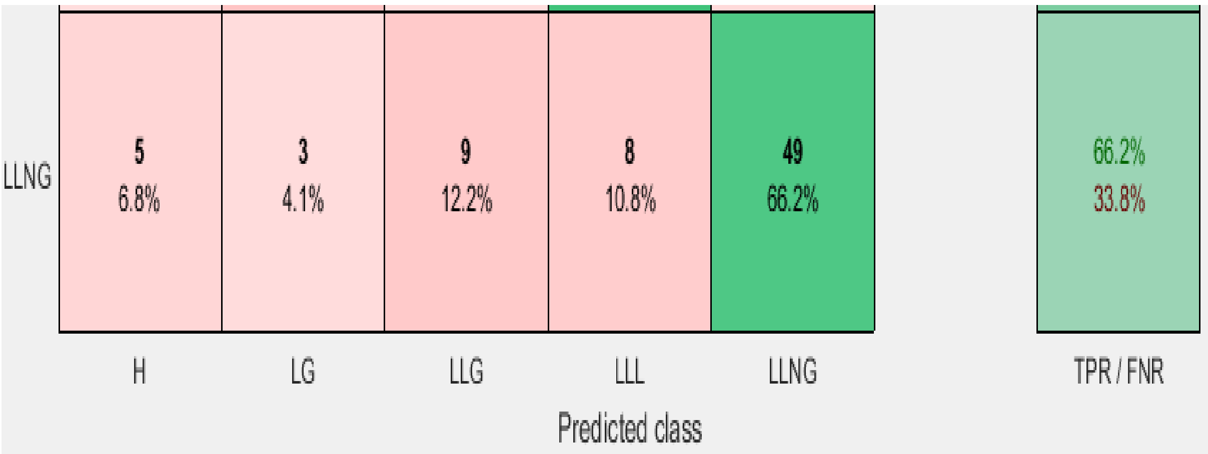

Figure 7.31: Double Line to Line Fault (Bagged-PPC)

Best (Highest) True Positive Rates The algorithm detected 84% true positive identifications of healthy while declaring 63 data sets.

Worst (Highest) False Negative Rates The algorithm exhibited worst 21.3% results as false identifications where line to ground fault data was predicted as triple line to line fault.

References

[1] Tapan Kumar Saha Yi Cui Atefeh Dehghani Ashkezari, Hui Ma. Investigation of feature selection techniques for improving efficiency of power transformer condition assessment.

[2] Sujeewa Nilendra Hettiwatte Kerry Williams Md Mominul Islam, Gareth Lee. Calculating a health index for power transformers using a subsystem-based grnn approach. *IEEE Transactions on Power Delivery*, 33(4):1903–1912, 2017.

[3] Yaqiong Fu Jianping Zhou Kaixing Hong, Hai Huang. A vibration measurement system for health monitoring of power transformers. *Journal of the International Measurement Confederation (IMEKO)*, 33(4):135–147, 2017.

[4] Igor Pavlovsky. Hydrogen sensor for oil transformer health monitoring. pages 211–213, 2008.

[5] T. Bhavani Shanker ; H. N. Nagamani ; Deepthi Antony ; Gururaj S Punekar. Case studies on transformer fault diagnosis using dissolved gas analysis. pages 01–03. IEEE, 2017.

[6] John. A complete guide to transformers. *Circuits Today*, 2018.

[7] Kiran Daware. Electrical transformer - basic construction, working and types.

[8] CircuitGlobe. Ll. *CIRCUIT GLOBE*, 2017.

[9] Single line-to-ground fault and 3-phase short circuit fault. *GoHz*, 2016.

[10] Muhammad Ryan. What is wavelet and how we use it for data science. *Towards Data Science*, 2019.

[11] Christopher Anderson Petr Klapetek, David NeÄDas. Chapter 4. data processing and analysis.

[12] 1 Roumen Zlatev 1 Michael Schorr Wiener 1 Rogelio Ramos, 1 Benjamin Valdez-Salas and Jose MarÃŋa Bastidas Rull2. The discrete wavelet transform and its application for noise removal in localized corrosion measurements.

[13] IEEE Bin Yu Senior Member IEEE S. Grace Chang, Student Member and IEEE Martin Vetterli, Fellow. Adaptive wavelet thresholding for image denoising and compression.

[14] MATLAB. Continuous and discrete wavelet transforms.

[15] Gabriel Peyre. 2-d daubechies wavelets.

[16] Matt Brems. A one-stop shop for principal component analysis.

[17] Thomas G Dietterich. Ensemble methods in machine learning. Technical report.

[18] V. Sokolov ; Z. Berler ; V. Rashkes. Effective methods of assessment of insulation system conditions in power transformers: a view based on practical experience. page 10.1109. IEEE, 2002.

[19] * Alberto Arroyo Pablo Castro Alberto Laso Sergio Bustamante, Mario Manana and Raquel Martinez. Dissolved gas analysis equipment for online monitoring of transformer oil: A review. page 4057, 2019.

[20] Jeff Golarz. Understanding dissolved gas analysis (dga) techniques and interpretations. Magazine, Electric Energy Online, 2015.

[21] ANALOG DEVICES. Small, low power, 3-axis Âś3 g accelerometer. Datasheet, ANALOG DEVICES, 2010.

[22] Allegro. Fully integrated, hall effect-based linear current sensor. Datasheet, Allegro MicroSystems, Inc, 2007.

[23] MathWorks. Learn more about thingspeak. Article, THINGSPEAK.

[24] Anuja Nagpal. Decision tree ensembles- bagging and boosting. Article, Towards Data Science, 2017.

[25] Jason Brownlee. Bagging and random forest ensemble algorithms for machine learning. Article, Machine Learning Mastery, 2016.