An-Najah National University

Faculty of Graduate Studies

Development of an Expert System for Power Transformers Fault Diagnosis Using Random Forest Technique

By

Ghazi Arar

Supervisor

Dr. Tamer Khatib

This Thesis is Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Electrical Power Engineering, Faculty of Graduate Studies, An-Najah National University, Nablus, Palestine.

2020

Development of an Expert System for Power Transformers Fault Diagnosis Using Random Forest Technique

By

Ghazi Arar

This Thesis was Defended Successfully on 28/7/2020 and approved by:

Defense Committee Members	**Signature**
– **Dr. Tamer Khatib / Supervisor**	**………..………….**
– **Dr. Jaser Sa’ed / External Examiner**	**……………………**
– **Dr. Anas Toma / Internal Examiner**	**……………………**

Dedication

To whom I have always been blessed by God for being part of my life,

My parents Riyad and Radiya

My brothers Hamza, Basil and Khaled

My life partner and wife Ola

My children Bailasan, Riyad, Maryam, and future ones

My dear friends and colleagues

For the souls of my grandfather and my grandmother, may God have mercy on them

And to whom I have always been blessed by God for their positive thinking

Acknowledgment

Praise be to God for guiding my way and lighting my sight through this journey and throughout my whole life.

I would like to express my deepest appreciation to Dr. Tamer Khatib for his valuable guidance, insightful suggestions, and relentless support. As well as, for his profound belief in my work and abilities.

I would also like to extend my gratitude to Electrical Power Engineering department lecturers who by their unparalleled knowledge contributed in this work comes to light.

الإقرار

أنا الموقع أدناه مقدم الرسالة التي تحمل عنوان:

Development of an Expert System for Power Transformers Fault Diagnosis Using Random Forest Technique

أقر بأن ما شملت عليه الرسالة إنما هو نتاج جهدي الخاص، باستثناء ما تمت الإشارة إليه حيثما ورد، وأن هذه الرسالة ككل أو أي جزء منها لم يقدم من قبل لنيل أي درجة أو لقب علمي أو بحثي لدى أي مؤسسة علمية أو بحثيّة أخرى.

Declaration

The work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degrees or qualifications.

Student's name:	**اسم الطالب: غازي عرار**
Signature:	**التوقيع:**
Date:	**التاريخ: 2020/7/28**

Table of Contents

List of Tables

List of Figures

List of Acronyms and Abbreviations

CT	Current Transformer
OOB	Out of Bag
CV	Cross Validation
CART	Classification and Regression Tree
BIL	Basic Insulation Level
ANN	Artificial Neural Network
FFNN	Feed Forward Neural Network
MLFFNN	Multi-Layered Feed Forward Neural Network
BPNN	Back Propagation Neural Network
FFBPNN	Feed Forward Back Propagation Neural Network
CFBPNN	Cascaded Forward Back Propagation Neural Network
RBFNN	Radial Basis Function Neural Network
PNN	Probabilistic Neural Network
PSO	Particle Swarm Optimization
BC	Bayesian Classifier
GSA	Gravitational Search Optimization
IGSA	Improved Gravitational Search Optimization
FD	Frechet Distance
CFD	Continuous Frechet Distance
IMF	Intrinsic Mode Function
EMD	Empirical Mode Decomposition
MHSI	Modified Hyperbolic S-Transform
SVM	Support Vector Machine
RBF	Radial Basis Function
STA	S-Transform Amplitude
MAC	Maximum Amplitude Curve
SDC	Standard Deviation Curve
HMM	Hidden Markov Model
LP	Likelihood Probability
MODWT	Maximal Overlap Discrete Wavelet Transform
Cf	Correlation Coefficient
DT	Decision Tree
MM	Mathematical Morphology
SE	Structuring Element

RMS	Root Mean Square
MDL	Minimum Description Length
AC	Alternating Current
DC	Direct Current
TNEB	Tamilnadu Electricity Board
Adaboost	Adaptive Boosting
SAMME	Stage-wise Additive Modelling Using a Multi-Class Exponential Loss Function
MDA	Mean Decrease Accuracy
MDC	Mean Decrease Gini
RF	Random Forest

Development of an Expert System for Power Transformers Fault Diagnosis Using Random Forest Technique

By
Ghazi Arar
Supervisor
Dr. Tamer Khatib

Abstract

This research will contribute to the field of power system protection. As traditional protection has failed to overcome its limitations to classify and discriminate different statuses of transformer, a need has risen to find new techniques to solve the problem. In this thesis, ensemble techniques are used to solve this issue. Hence, the differential protection constructed by using ensemble techniques to provide protection element via a trip and no trip actions. And further, conditional monitoring functions are used to distinguish five different statuses of transformer including normal, inrush, over-excitation, current transformer-saturation and internal fault.

By capturing practical transformer rating models for 20 different transformers with 5 different operating cases, 100 examples were provided as a data set to train and test models with 1600 observations. The 100 original and raw data were used to train random forest, then it has been validated with internal measures including out-of-bag error, margin, confusion matrix, and outliers. Afterward, an updated and weighted data set was generated to be used in training and testing random forest. OOB error and margin were captured for weighted examples to be compared with raw examples.

Different train to test, which are 80-to-20 and 60-to-40, have been used to validate system strength and reliability. Moreover, a faster version of random

forest models constructed with different sizes of data window included ¾, ½, and ¼ cycles, resulted in an accurate protection and high accurate conditional monitor. Besides, two different versions of random forest in terms of individual trees depth have been tested concerning the greedy and limited size.

Boosting technique was also applied to both, original data set and weighted data set with different train to test ratio including 80-to-20 and 60-to-40 to validate the model. And yet, the model has been tested conditioned with optimum number of trees by using out-of-bag error and cross validation folds. Due to that, the variable importance was achieved by using the optimum number of trees.

It is worth noting that the variable importance was captured by using ensemble techniques, and therefore conclusion for signal importance at different instances investigated.

In conclusion, random forest and boosting have shown promising results and showed the ability to classify the suggested problem. Thus, it provides accurate, fast, and reliable results.

Chapter One

Introduction

1.1 Background

The increasing demands of electrical energy by different consumers lead to increase the dependability on these sources, which reveals the great importance of protection systems. The protection systems are subjected to the toughest conditions along with introducing new loads and generation. Moreover, they are required to provide rapid, simple, selective, sensitive and reliable “meaning dependable and secure” response to assure continuity requirements. Also, the modern vast electrical system with a variety of equipment interacting with one another has subjected traditional protection to a bad compromise as a lot of data containing mixed pieces of information need to be handled, the traditional protection has failed to extract and analyse these pieces of information efficiently.

The electrical system infrastructure contains multiple elements which are: generator, transmission line, reactor, capacitor bank, circuit breaker, etc. However, the transformer has extreme importance amongst the electrical system elements since it links the whole system. And as each element has some degree of singularity to deal with, different protection regimes were suggested overtime to follow equipment singularity and development. Yet to add that the transformer exhibits different oscillatory power flow features including fault, likewise fault and normal features. The features differences are assigned mainly due to the nonlinearity of the magnetic core, which attains difficulties to diagnose and discriminate these different features. Therefore, mal-operation occurs.

Different combinations of protection are being used, depending on the transformer size. For instance, low size transformer may be used only with over-current fusing elements and arcing horns, rather than larger size transformers, which have more complicated protection elements with surge arrestor. However, the transformer has many protections with differential and over current as electrical protection. Hence, the mechanical protection includes Buchholz, pressure relief, oil, and winding thermometer.

The electrical transformer protection, specifically differential protection, has been under study for a long time, accordingly, different topologies were suggested for differential current discrimination including artificial intelligence, transitory feature detection, hybrid systems, and many others. Whereas traditional differential protection considers current transformer (CT) transformation errors, CT mismatch and taps variation. Where it follows an increasing curve with a positive specific slope for the relation between differential and restrain current. And whenever point of operation above this curve fault is presumed and tripping is imposed on the circuit breaker. While over-current used as another back-up electrical protection, it operates whenever current value increases above some threshold following a specific curve, which may be time delayed or instantaneous. The Buchholz protection is being used as a main mechanical protection, and its operation is considered very crucial and dangerous to the transformer. In cases of fault, gases will be formed inside the transformer. Due to Buchholz's design and position upside transformer, the Buchholz trap gases inside which accumulates leads to decrease oil inside and float contacts to drop in two states, these gases can be tested too. Moreover, Buchholz protection has another aspect that provides protection against the oil flow rate through using a flat piece of metal across the oil path located between the main tank and reservoir at at 0° to 5° slope. The importance of Buchholz protection is also revealed in its use as a protection for tap-changer. Same as the transformer, the Buchholz protection is located between the tap-charger tank and its perspective reservoir.

The Pressure relief valve is another mechanical protection that uses oil pressure inside the transformer instead of gases or oil flow. This protection behaves like a valve that opens when the pressure inside the transformer increased; creating a path for oil, from the main tank to reach the under-ground reservoir, and by that, it prevents destructive action and reliefs transformer.

1.2 Problem Statement

Today, transformer traditional protection, namely percentage differential protection, is a common practice for differential protection (Oliveira, Bretas & Ferreira 2014; Gethanjali, Raja Slochanal & Bhavani 2008). The discrimination between transformer statuses: Internal fault, normal, magnetizing inrush current, over-excitation, and CT saturation is not an easy job due to their similarity (Yazadani-Asrami et al. 2015). For a long time, harmonic restrain was used to prevent mal-operation by using a ratio of second and fifth harmonic with fundamental protection (Oliveira, Bretas & Ferreira 2014; Gethanjali, Raja Slochanal & Bhavani 2008). However, internal faults generate a high level of second harmonics (Yazadani-Asrami et al. 2015). So, tripping operation is dependable on the level of this harmonic which may result in mal-tripping (Oliveira, Bretas & Ferreira 2014; Yazadani-Asrami et al. 2015).

Many research works have been done in the scope of differential protection. Hence, the adaptive differential protection based on transient signal analysis using wavelet transform showed limited capabilities (Oliveira, Bretas & Ferreira 2014). Furthermore, Žarkovic and Stojkovic (2017) used a multi-input fuzzy based algorithm to predict the probability of fault and urgency of intervention. As well as the optimized ANN scheme using a swarm-based algorithm shows promising results (Gethanjali, Raja Slochanal & Bhavani 2008; Yazadani-Asrami et al. 2015). However, an artificial technique, namely random forest, provides more accurate and fast responses (Ibrahim & Khatib 2017). The random forest technique will be employed to discriminate these different features. Thus, several trees and leaves (nodes) will be specified as well as input data with classes. Random forest classifier is an ensemble of decision trees (CART, ID3). These trees start with root node having random data subset and by splitting this node and grow the trees down for full depth (Breiman 2001). The best split is done by measuring impurity, which has multiple measures and no pruning. Therefore, the Random forest has three parameters which are the number of decision trees located within the forest, the number of leaves within this tree generated from split action and depth. Moreover, some data classification is done based on the majority vote of decision (CART) trees.

1.3 Objectives

1. To construct ensemble techniques (random forest and boosting) of differential protection to classify different transformer statuses, which are no trip (normal), inrush, over-excitation, current transformer (CT) saturation, and trip (internal fault) based on differential current samples as input.
2. To construct condition monitoring technique of transformer, which are: normal, inrush, over excitation, (CT) saturation, and internal fault.

1.4 Methodology

WP.1 Data Collection	1. T1. To collect 20 different transformer model
	2. T2. Each transformer model will generate 5 different cases of differential current: normal, inrush, over excitation, CT-saturation and internal fault.
	3. T3. To have 100 set with 16 samples
WP.2 Literature Review	1. T1. Traditional differential protection review: percentage differential protection
	2. T2. Artificial neural network review: feed forward, back propagation, probabilistic neural network
	3. T3. Random forest algorithm review
WP.3 Development of Model	1. T1. Data arrangement to form dataset that produce minimum error
	2. T2. Model training using random forest and boosting with different train to test ratios
	3. T3. Model validation using out-of-bag error and different train to test ratios
WP.4 Analysis of Suggested Model	1. Using Out-of-bag error, cross validation, Variable importance, outlier, margin and confusion matrix.
WP.5 Thesis Writing	

1.5 Significance of the Work

Random Forest technique is important as it solves the modern systems protection dilemma. It will provide new scope differential protection with conditional monitoring. Furthermore, it does not need extensive mathematical analysis nor requires high computation burden to operate, rather simple processing. Thus, it will be able to discriminate and classify transformer status: normal, inrush, over-excitation, CT-saturation and internal fault.

This study will contribute to the improvement of differential protection not only to the power system manufacturer but also to all parties in the electrical field. Including working personnel that will highly benefit from this solution. They will capture the cause of disturbance faster, thus helping them to fast and accurate decision making.

1.6 Thesis Organization

The thesis consists of five chapters, as follows:

Chapter 1. This chapter includes the introduction, background, problem statement, objectives, methodology and significance of the work.

Chapter 2. This chapter is a review of literatures used in this study including: journals, articles for transformer principle and statuses, state of art. It also includes research gap.

Chapter 3. This chapter is composed of research approach – transformer model, data extraction, and sampling. Also, methods models utilized in this study. It also mentions the sources of data.

Chapter 4. This chapter explains this research analysis methodology, progressing steps to reach results, the final model testing and different features have been studied. It also includes comparison of proposed technique with other techniques.

Chapter 5. This chapter summarize the research conclusion as well as future work.

Chapter Two

Literature Review

2.1 Transformers

2.1.1 Introduction

Transformers are equipment used to move energy from one side to another by changing voltage and current. In addition, it forms a basic element in our modern power system, as well as, primary equipment in some bulk power systems like substations. Transformers are used to linkage power generation with the transmission, distribution, protection, and control systems. Therefore, transformers have multiple sizes and types to match with the system it is used in, and are classified under the following categories:

- Power transformer.
- Distribution transformer.
- Auto transformer to connect close voltage levels.
- Instrument transformer, i. e: voltage transformer, current transformer.
- Grounding transformer zigzag.

The transformer has a different core of topologies, which are: single-core, triplex core, five-legged wound core, three-legged stacked core, five-legged stacked core, and shell core. Also, the winding has different topologies, which are: concentric, interleaved, and pancake design (Martinez et al. 2005).

2.1.2 Transformer Operation Principle

Transformer main components are: primary coil, secondary coil made from conducting isolated cupper turns, and ferromagnetic core consist of isolated ferromagnetic material lamination as indicated in figure 2.1:

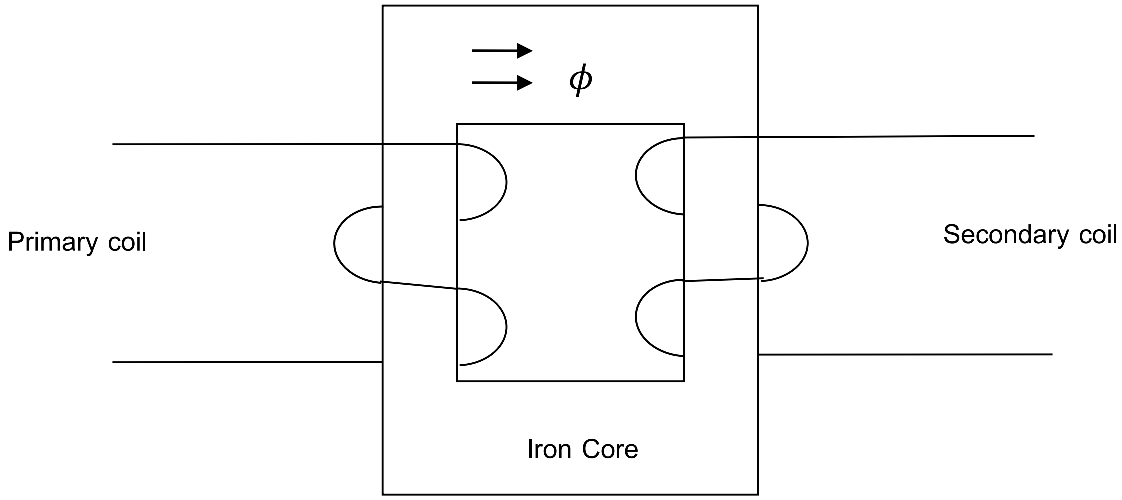


Figure 2.1: Transformer construction

The variable current flows through a primary coil to produce varying flux by the magnetic induction principle. The flux flows through the iron core to reach secondary winding, and then the electromotive force is produced by the act of the varying field. To investigate these phenomena's, starting with low amperes of a magnetic field, to calculate the magnetic field produced by current flowing through the coil. By means of partitioning the coil into several wires, each element will contribute with two magnetic fields: self-flux linkage for each element and mutual flux linkage that links all elements. In contrast, field produced propagate through secondary winding and produce electromotive force following Faraday and Lenz laws:

$$E = -N\frac{d\phi}{dt} \tag{2.1}$$

Faraday describes the magnitude and shape, whereas Lenz describes direction of generated electromotive force giving by minus sign in the equation.

2.1.3 Working Principle

Transformer parameters described in Figure 2.2, as follows:

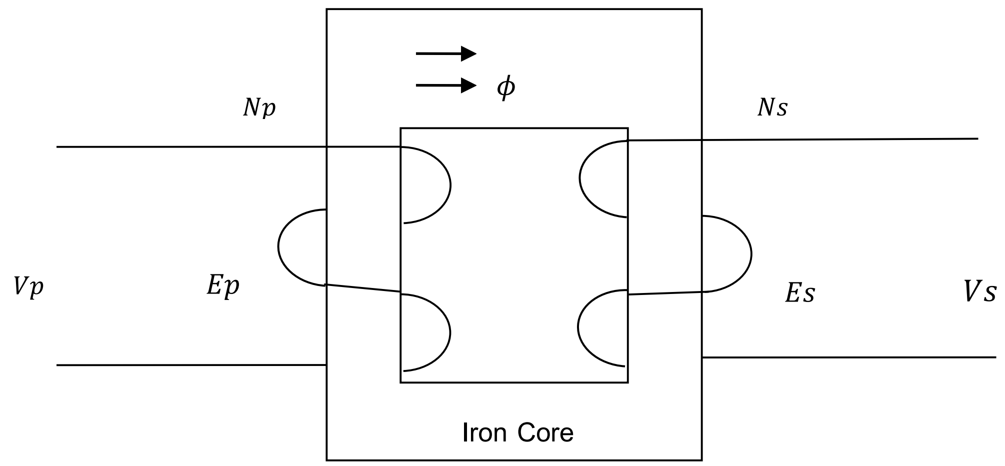


Figure 2.2: Transformer parameters

Where:

p: primary side notation

s: secondary side notation

V: applied voltage

E: electromotive force

N: number of turns.

Let us take the transformer working under no load to find Transformer basic modelling equation. Through feeding transformer by sinusoidal signal at the primary side, high current limited by wire resistance will generate high flux value. Then, the limiting value of electromotive force will start to generate limiting current and hence flux. Meanwhile, secondary electromotive force generated where its value following the changes applied to flux. For now, let us neglect resistance and assume ideal transformer then this assumption can hold:

$$Vp = Ep \tag{2.2}$$

$$Vs = Es \tag{2.3}$$

Following Faraday's law and removing negative sign since it indicates those induced voltages oppose the varying field that produce it:

$$Ep = \ Np\frac{d\phi_p}{dt} \tag{2.4}$$

$$Es = \ Ns\frac{d\phi_s}{dt} \tag{2.5}$$

$$\phi_p = \ \phi_s = \phi \tag{2.6}$$

The following conclusion can be reached by dividing both equations:

$$\frac{Ep}{Es} = \ \frac{Vp}{Vs} = \ \frac{Np}{Ns} = a \tag{2.7}$$

where:

a: Turns ratio and is constant for the same tap

Another formula can be written for ideal transformer:

$$Sp = Ss \tag{2.8}$$

Sp, Ss indicate apparent power for primary and secondary sides.

2.1.4 Transformer Steady-State Model

Steady-state equivalent circuit of transformer, constructed by taking into account heating and magnetic effect as in figure 2.3:

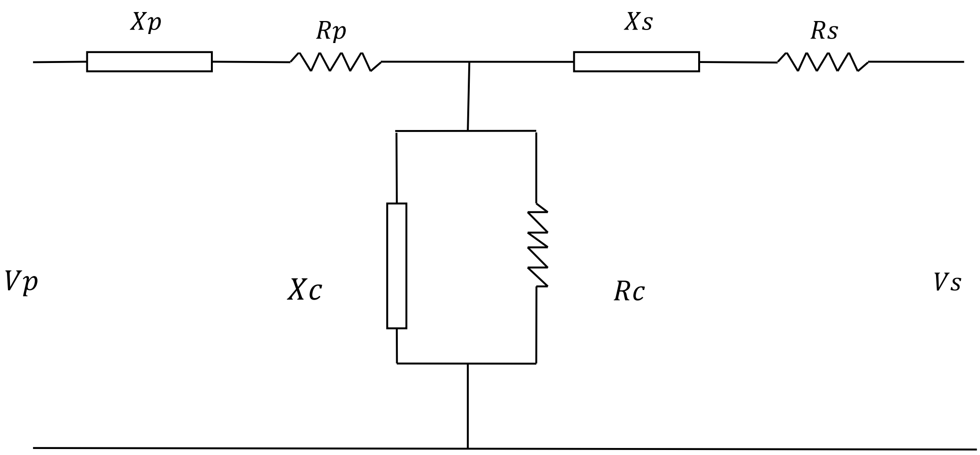


Figure 2.3: Transformer equivalent circuit

Same as before, p indicate primary notation and s is a secondary notation, the series branches indicate conductor losses and magnetic effect, while parallel branch refers to core losses and magnetic effect. The previous circuit needs to refer one side to another as in the following formulas where:

$$\frac{V_p}{V_s} = a \tag{2.9}$$

$$\frac{I_s}{I_p} = a \tag{2.10}$$

$$\frac{Z_s}{Z_p} = a^2 \tag{2.11}$$

To estimate equivalent circuit parameters two tests can be performed clarified in figure 2.4 and figure 2.5, namely:

- Open circuit test: This test achieved by feeding nominal voltage and open-circuited the other side. The excitation of current flow includes core loss and magnetizing component. Core losses include eddy and hysteresis effects, which are not dependable on load.
- Short circuit test: This test achieved by short circuit one side and applies a low voltage to the other and increasing the voltage until the nominal current flow through the short circuit. The short circuit losses are cupper losses which are linearly changed with the load. The skin and proximity effect can be ignored in a steady-state analysis as skin, proximity, eddy, and hysteresis effect are non-linear and frequency-dependent effects.
-

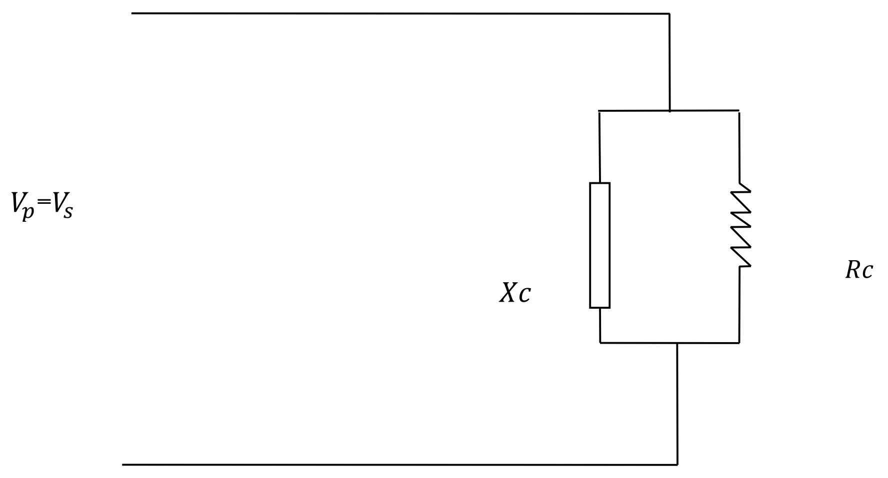


Figure 2.4: Open circuit test

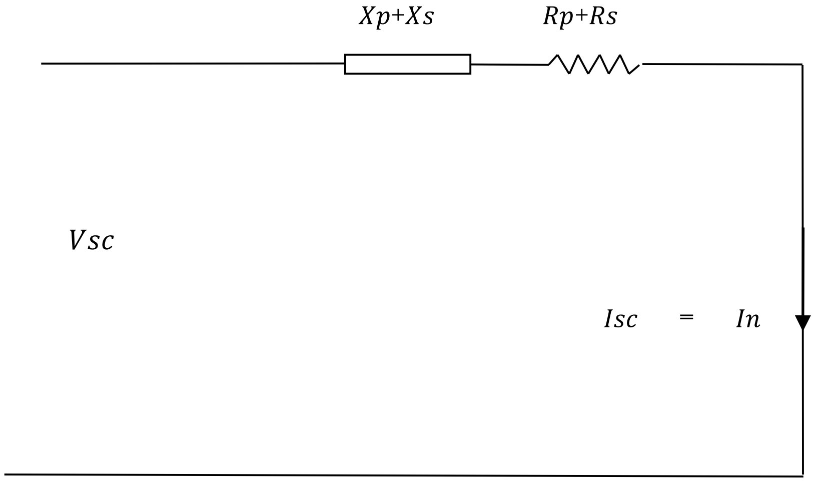


Figure 2.5: Short circuit test

To further investigate hysteresis losses, start with the virgin core, where magnetic current density and magnetic flux density are zeros. Increasing magnetic current density hence magnetic field density increases in an approximately linear manner until reaching the point of saturation, where a large step increase in current result small increase in field density. On the other hand, decreasing the current will decrease the field density until the current is zero. On the controversy, the core will still have remnant flux. When the current has some negative value, the field density reaches zero. This relation follows curve can be determined where the loss is the area entrapped. To calculate the energy and power of the hysteresis loop this equation can be used:

$$\int dE = \int HdB = \int \frac{B}{\mu} dB = \frac{1}{2} \times B \times H \quad , per\ unit\ of\ volume \tag{2.12}$$

$$P = \frac{1}{2} \times B \times H \times V_{core} \times f \tag{2.13}$$

The above equation coincides with the previous discussion that hysteresis losses are non-linear and frequency-dependent. As permeability increases, less value of current is needed to produce the same field; smaller area entrapped, and power losses.

Eddy current losses have reduced by the use of lamination to cut the current paths and have high core material resistance as possible.

Impedance achieved from short circuit test voltage. The short circuit test voltage is an important element of the transformer and used for different calculations. i.e.: short circuit analysis and parallel operation. Voltage impedance can be given as follows:

$$Z_{k_{pu}} = \frac{V_{sc}}{V_n} = \frac{\frac{V_{sc}}{I_n}}{\frac{V_n}{I_n}} = \frac{Z_k}{Z_n} \qquad , in\ per\ unit \qquad (2.14)$$

where V_{sc} indicates voltage measured by short circuit test, V_n and I_n are the nominal voltage and current respectively and Z refers to impedance.The above equation shows division by base value on the transformer rating, which proves that it is per unit value. To find the real impedance voltage multiplies it by the nominal voltage:

$$Z_k = \frac{Z_{k_{pu}} \times V_n}{100} \qquad (2.15)$$

2.1.5 Three-Phase Power Transformer

Three-phase power transformers handled in similar manner to the previous discussion taking into account the combination of the three-phases. Primary and secondary connections can be star or delta, were these connections alter phase-line current and voltage transformation by inserting $\sqrt{3}$ depending on the connection, and 3 for power calculation. Another important aspect is the vector group which arises from three parts:

- Two of the three parts indicate Type of coils connection can be D, d: delta, Y, y: star, z: zigzag.
- The number indicates the phase difference between phase voltages in the primary and the secondary winding.

The three-phase power transformer can have these specifications:

• Manufacturer	• Tape information
• Power	• Frequency
• Serial number	• Weights
• Date of manufacturing	• Transformer connection including taps
• Voltage nominal levels	• BIL
• Cooling type	• Vector group
• Impedance voltage	• No-load losses

2.2 Types of Current in Transformer

The transformer nonlinearity nature gives many difficulties to the system since it is in up normal status, and the parameter has no more steady-state nature. The transformer normal operation mode can easily handle in most common steady-state methods were linear system available, hence superposition and multiplication are applicable. Therefore, the relation is well known and predictable. However, there are other behaviours where nonlinearity and frequency dependability introduced, including internal fault, inrush magnetizing, over-excitation, and CT saturation.

2.2.1 Internal and External Faults

The main internal faults are due to insulation deterioration and breakdown. Faults in the winding, core, tap changer, cooling, bushing, and casing are a form of internal faults. While external fault current cause high current pass through transformer sides, a through fault of 10 times the rated current (with a tap changer at end position) can cause a differential current of 1-2 times the power transformer rated current (ABB BA THS / BU Transmission Systems and Substations, pp. 163-174). Moreover, high through current may cause internal faults due to overheating of insulation.

2.2.2 Over-Excitation Current

Over-excitation means the increase of flux flowing through core above some design limit, transformer sides have two currents differ by the value of magnetizing current. Magnetizing current characteristics follow core characteristics distorting the current signal. The current from source to load has two components since:

$$I_{Total} = I_{Magnetizing} + I_{Load} \tag{2.16}$$

where $\mathrm{I_{Magnetizing}}$ is the core magnetizing current and $\mathrm{I_{Load}}$ represents load current.

Transformer Over-excitation in transmission and distribution networks is caused by over-voltages in the network (ABB BA THS / BU Transmission Systems and Substations, pp. 163-174). To demonstrate this principle in figure 2.6 transformer feeding load from a big system normally has a constant voltage at the transformer source side, while current needed by the load and core magnetizing current drawn from the system. Whenever overvoltage appears across the transformer due to some source, i. e: surge, switching. The driving system will apply increased voltage to the transformer, hence increasing magnetizing current; this will increase the flux overexciting core. This current has a high percentage of fifth harmonic.

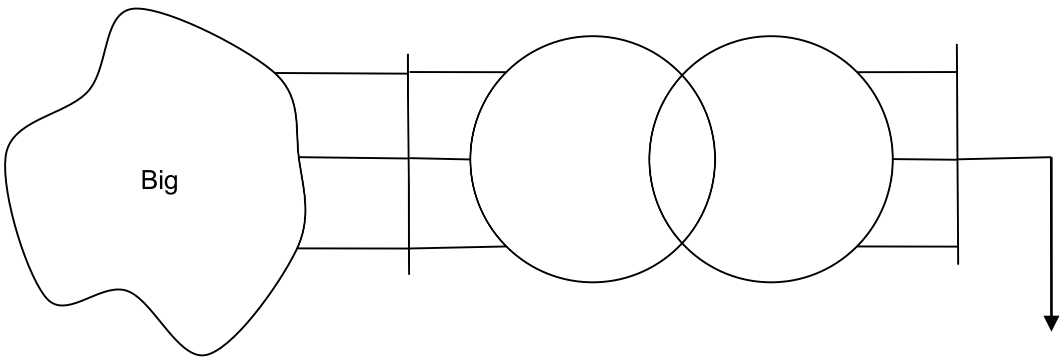


Figure 2.6: Transformer feeding load via big system

2.2.3 Inrush Current

Inrush current has an over-excited core with a special case of saturation during the initial excitation. Steady-state open-circuited power transformer energized with the sinusoidal signal at the primary side will generate flux through the core which lags the voltage approximately by $90°$:

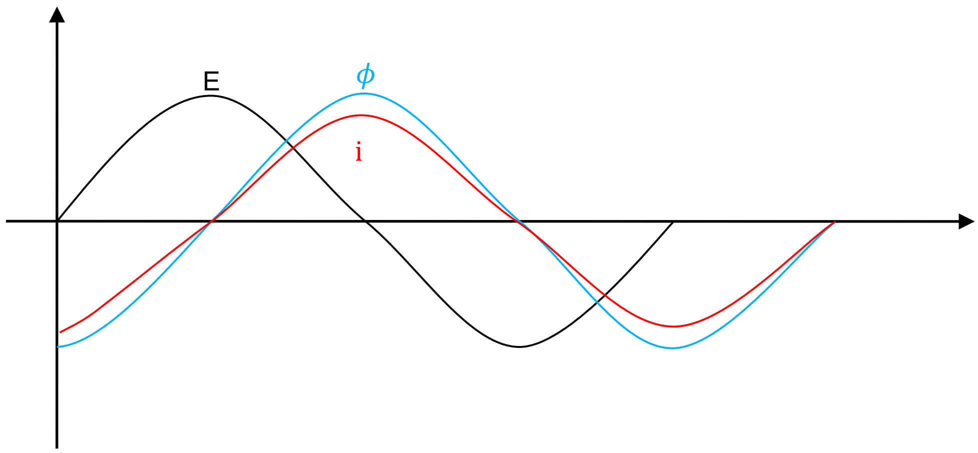


Figure 2.7: Voltage, current and flux waveform

Where:

E: induced voltage and

ϕ: magnetic flux

i: current

Every switch on operation has different inrush current values. The most severe situation when voltage begins at zero, where flux will be maximal. To explore the effect of inrush current, at time of switching the electromotive force will equal the applied voltage:

$$E = Vp \sin wt = N \frac{d\phi}{dt} \qquad (2.17)$$

$$\phi = \int \frac{Vp \sin wt}{N} dt \qquad (2.18)$$

$$\phi = \int \frac{Vp \sin u}{N} \times \frac{du}{w} \qquad (2.19)$$

$$\phi = \frac{Vp}{N \times w} \int \sin u \, du \qquad (2.20)$$

Following this formula, the value of flux at point of interest is the area under-voltage curve from the point of switching to the point of interest scaled by the value of ($\frac{1}{N\times w}$). Common sense suggests two integral limits that can be put into integration that will give a maximum depth of sine function, which is 2 that double the value of maximum flux (ϕ_{max}). As in next:

$$\phi = \frac{Vp}{N\times w}\int_0^{\pi} \sin u \ du \qquad (2.21)$$

$$\phi = \phi_{max}\,(\cos 0 - \cos \pi) = 2\,\phi_{max} \qquad (2.22)$$

Where:

$$\phi_{max} = \frac{Vp}{N\times w} : \text{maximum flux} \qquad (2.23)$$

By flux density versus current density, the curve can be noticed that flux may pass a saturation point, which results in an un-uniform current signal with high value. This high amplitude distorted signal has high a percentage of the second harmonic with a bad impact on the overall system include integrity and power continuity.

2.2.4 Deep Saturation

It is similar to external fault removal near the transformer, and to energizing transformer with the load. The transformer is driven more into a saturation dead angle disappear, and the differential current becomes more and more sinusoidal, so the ratio of second harmonic decrease as well as mal-operation may occur.

2.3 Classification Importance

Power transformer protection is a crucial element in the power system since it assures the goal of power system operators: the safety of personnel, integrity of equipment, and continued power supply. These criteria could be achieved by speed of protection, simplicity, selectivity, sensitivity, dependability and security. Meanwhile, different transformer current statuses with different characteristics which are dependable on many factors i.e. system components, transformer characteristics, etc. provide a challenging task.

The different states of the transformer include normal operation, faulty operation, external fault, magnetizing inrush, and over-excitation. Most conditions can supply differential current, only internal fault required to be cleared amongst other conditions.

Traditional protection uses a harmonic restrain method to suppress unwanted protection operation, the second harmonic used to discriminate inrush from other statuses. Although the improvement of core material introduces a decreased level of second harmonic, this may be seen by a protection as a faulty condition. On the other hand, CT saturation and shunt capacitor or distributive capacitance is in a long extra high voltage line increases second the harmonic level in a faulty condition.

2.4 State of Art

Since traditional harmonic restrain have a limitation with new cores and insertion of capacitance that motivate many pieces of research to discriminate transformer current status. They used a mixture of three steps to reduce the number of processed data rather than processing a larger number of differential current samples with unnecessary information. Subsequently, the system is less complicated. A wide variety of researches are based on:

- Optimization techniques such as genetic and swarm-based techniques such as Ant colony optimization, Gravitation search optimization, and particle Swarm optimization which has some focus.
- Signal analysis using time-frequency analysis such as wavelet transform, s-transform, and Huang Hilbert transform.
- Classification algorithm such as artificial intelligence, and tree based algorithm.

Oliveira, Bretas and Ferreira (2014) use Discrete Wavelet Transform to extract features, where it uses spectrum energy criteria to compare with the adaptive threshold, as well as, different mother wavelet and fault resistance. It showed that during the initial fault, the second harmonic increased to 70% of fundamental before returning to pre-fault value in about 2 cycles. In this algorithm, two blocks are used: detection and classification or disturbance identification. However, as fault resistance increased similar energy variation to energization, which weakening the system. It was used to detect internal, external, and inrush current, while CT saturation has been mentioned in one case.

In (Gethanjali, Raja Slochanal & Bhavani 2008) two ANN architectures have used, one for internal fault detection which gives states of: internal fault or other disturbances, and the other for monitoring which in turn gives normal, inrush, over-excitation and CT saturation. They used Multi-Layered Feed Forward Neural Network (MLFFNN) trained with Back Propagation Neural Network (BPNN), since this technique suffer from stuck in local optima, other technique used Particle Swarm Optimization (PSO). The training algorithms require bounded-differentiable activation function, which applies to sigmoid. Moreover, PSO performs better than BPN. Yazadani-Asrami et al. (2015) Classify internal fault, inrush, CT saturation, over-excitation and normal condition using Bayesian classifier (BC) with normal distribution by means and variance for single dimension and means, variance and co-variance for multi-dimension. Each class will have normal distribution, and following this distribution most probable class is considered. The BC gives two states: internal and others. Other disturbances are classified using Improved GSA and PSO optimized FFNN. Similar to BPN, GSA may be stuck in local optima hence the use of IGSA, as the IGSA perform better than PSO.

Weng et al. (2019) Treat internal fault, external fault, CT saturation with fault, deep saturation, and data sampling by means of algorithm similarity, where phase information difference extracted using Discrete Frechet distance rather than continuous FD because difficulty to extract similarity information using CFD. External fault will have low similarity while fault with CT saturation will have higher similarity, with internal fault have highest similarity.

In (Samantary & Dash 2011) differential current samples (16 samples/cycle) fed to decision tree and output give rise to two states: inrush and internal fault. For verification, CT saturation and external fault have been included in addition to inrush in one class, and internal in other classes, which rise the best classification rate (92.5%). The authors found this technique suitable for power transformer protection.

Barhate, Thakre, and Deshmukh (2016) Generates three membership functions which are: differential, restrain currents, and inrush detector. Inrush detector is dedicated using primary dead angle detection to set inrush current flag. These three inputs fed to fuzzy inference engine rule base code. If detector activated, then the algorithm act as inrush detector. While if not activated, the detector relay issues trip/ no trip based on restrain and differential membership function.

Peres and Silva (2019) have taken high and low side voltage currents and normalized it. Then phasor is extracted, where fundamental, harmonic, and negative sequence phasors extracted (where fundamental phasor used to extract negative sequence). Then harmonic and negative sequence instantaneous signal in discrete-time extracted. The instantaneous harmonic signal used to generate harmonic restrain, which is used to reinforce and support restrain current. Disturbance detection is available using energy, where flag change status, whether disturbance happens, then integration ratio- operation and pick up- calculated, and after (1ms). Either internal or external disturbance is detected. As a drawback, this technique requires to be blocked during energization since it will issue a false trip.

(Angrisani et al. 2018) Applied to over-current protection of high voltage side that considers two states: inrush and fault current. Huang Hilbert Transform used to extract Intrinsic Mode Functions (IMF) by means of Empirical Mode Decomposition (EMD) and sifting process, Form Factor Deviation of instantaneous frequency evaluated as a decision criterion. Instantaneous frequency trajectory is examined using Hilbert Transform of real-valued signal. Over-current due to inrush will have an oscillating frequency trajectory, while fault will give approximately constant frequency trajectory.

In (Behvadni, Seifossadat & Saffarian 2019), differential current of any of three phases checked to exceeds the limit using percentage differential restrain. One cycle of fundamental frequency samples used to find zero, negative and positive sequence by means of Clarks Transformation. Then Modified Hyperbolic S-Transform (MHST) matrix calculated using zero and positive sequence, wherein this matrix rows and columns correspond to frequency and time respectively, important information can be achieved: magnitude, phase, and frequency. The previously mentioned information used to extract decision features which are: the energy of first level, variance index, mixed energy-amplitude index, phase standard deviation of second harmonic frequency. One of the most important features is the energy of the first level.

Saravanan and Rathinam (2017) use pseudo-characteristic to extract core operating regions and thus linearity check by using orthogonal polynomial representation to model inrush/ fault detector. This algorithm consists of three sequential steps: amplitude check, harmonic check, pseudo-characteristic check.

Tripathy, Maheshwari, and Verma (2010) suggested digital differential protection using Optimal Probabilistic Neural Network (PNN) optimized by PSO technique. Spread or smoothing parameter defined using PSO as trial and error used in homoscedastic PNN, while PSO used in heteroscedastic PNN. In this algorithm, external and normal discriminated by comparing two consecutive peaks. Whereas, over-excitation discriminated by comparing voltage to frequency ratio with rated voltage to frequency ratio. If these conditions are not available, the algorithm will run optimized PNN to discriminate between inrush and internal fault. The PNN performs better than FFBPNN.

Shah and Bhalja (2013) use Wavelet Analysis to extract features include the standard deviation of detail 1coefficient, the Support Vector Machine (SVM), and the Radial Basis Function (RBF) kernel to classify standard deviation of detail 1 coefficient. Daubichies (db4) mother wavelet used due to its characteristic closeness to fault signal of study it has been reported in the literature to be best suitable.

In (Hasheminejad & Esmaeili 2013) S-Transform Amplitude Matrix (STA) which a is result of S-Transform Complex Matrix used to extract Maximum Amplitude Curve (MAC), which is a maximum amplitude versus frequency, as well as Standard Deviation Curve (SDC), which is standard deviation of time values evaluated versus frequency. The two independent routes MAC and SDC will have a separate Hidden Markov Model (HMM) blocks include two for internal fault, two for external fault, two for inrush current, and six blocks for six different curves. Each block will have a transition matrix and an emission matrix. The total Likelihood Probability (LP) calculated by adding LP for the MAC route and SDC route, and the most probable will be the result. The MAC and SDC compensate accuracy of each other especially in noisy conditions, SDC has low accuracy, meanwhile, MAC improves accuracy.

In (Fernandes, Costa & Medeiros 2016) Maximal Overlap Discrete Wavelet Transform (MODWT) which does not use down sampling. The Spectral energy of sliding window calculated, which decomposed to scaling and wavelet coefficient energy, where this energy compared to some threshold as a disturbance detector. If disturbance detected, ANN will be initiated with a sliding window of some interval that will generate pairs -differential and restrain- energy vectors of 15 samples, and the first sample will be taken before event occurrence with 12 quantities each. For external and transformer energization warning will be issued, while internal fault will issue trip and another ANN classify the fault in 10 cases same as previous ANN technique. Rasoulpoor and Banejad (2013) used DWT also to extract features. The total energy of sliding window and detail coefficient at five levels are also calculated as used to calculate the energy percentage vector of 5 values. Then the correlation coefficients (Cf) between sliding windows calculated, as well as the detection criteria achieved by counting the number of dips in correlation coefficients (Cf) values versus time for three phases of differential current. The internal fault percentage energy vectors are correlated and one dip at disturbance occurrence and no dip after, while inrush current change in correlation coefficient (Cf) appears in all three phases with oscillation.

In (Sheng & Rovnyak 2002) decision Tree (DT) used to classify different data composition and compare between them; differential current, restrain current and percentage differential current. In addition, second and fifth harmonics added to previous data that perform better. While, wavelet detail coefficients d_3 added to the first case which perform better than both cases. Above all, the combination of all currents, harmonics and wavelet coefficients has superior performance than all, which perform better 3% than currents and harmonic case.

In (Shi et al. 2011) Mathematical Morphology (MM) is used to extract shape features and fed to FFNN trained by Levenberg-Marquardt backpropagation for pattern recognition; inrush or internal fault current. The vertical distances between structuring elements (SEs) are the shape feature corresponding to d2, d5, d8, d11, d13 which are distances between each structuring element and SE15 corresponds to 2, 5, 8, 11, 13, 15 samples for base of 32 samples per period. Theses distances are normalized to d13 to eliminate amplitude information.

Ali et al. (2018) proposes the current ratio and voltage ratio differential protection. In this technique, traditional differential protection is checked, which if satisfied other three criteria also checked before declare the transformer and phase’s status: current ratio, voltage ratio and current direction. The current ratio and voltage ratio are the percentage of absolute difference of absolute RMS fundamental currents/voltages to sum of absolute RMS fundamental currents/voltages. Specified current ratio range in conjunction with instantaneous current direction is used to distinguish internal, external, and loaded energization (sympathetic inrush current). While another current ratio in conjunction with voltage ratio ranges used to distinguish no-load energization and the internal fault with no-load energization. Ali et al. (2019) provide a more specific explanation and practical evaluation of the proposed method as it studied sympathetic inrush current, and trip signal delayed for 1ms for security reasons.

Ozgonenel & Karagol (2014) perform differential protection as a wavelet-neuro system. Rather running a random number of features, minimum description length (MDL) used to optimally choose mother wavelet and Shannon entropy to optimally choose the resolution (number of detail coefficients). Thus feature vector dimension; the standard deviation of detail coefficient will be optimal, which will be fed to FFNN for two status decisions: internal and inrush current. The authors find that the optimal combination is bi-orthogonal3.3 (bior3.3) mother wavelet with three detail coefficients. However, they declare that the feature selection algorithm is highly sensitive to signal employed.

2.5 Research Gap

Traditional protection adaptability is limited, it does not have conditional monitoring features and its protection is compromised. Yet, many pieces of research have been done in this field, most of them did not distinguish between all statuses of the transformer, and only some of them have processed this issue, rather more complex and less interpretable systems suggested.

However, new methods of ensemble techniques tested for the field of differential protection. The random forest and boosting provide more accurate, less complex, and interpretable results. Its internal measures provide a powerful tool for model validating and features evaluation.

Chapter Three

Approach

3.1 Introduction

Different trees based technique has been developed over the years including single tree, Bagging, random forest, and boosting. The random forest and boosting perform good results that are competitive with each other, while single trees and bagging are retarded versions. Trees are good candidate classifier for random forest technique, as it reduces variance since trees are weak classifier that performs better than chance. Even a random forest is an ensemble technique that uses a tree-based algorithm, which is an extension to the bagging algorithm that is considered as a predecessor technique. This technique has internal measures can be used to judge the algorithm including error, strength, correlation, variable importance.

3.2 Forward: General Information about Current Classification

Modern power transformers put to several dissimilar conditions due to transformers nature. These dissimilar conditions have become a challenging task to address. These transformer current statuses may introduce differential current that may trigger the protection. A Single signal used to classify all these statuses, which can take different values depending on: transformer ratings, and type of disturbance. Hence, we need to classify these statuses: normal, inrush, over-excitation, ct-saturation and internal fault.

Traditional protection takes raw samples of differential current and pick-up above some threshold, leaving the classification task to power system personnel to define their threshold, which may be compromised and set to danger. Traditional protection uses figure 3.1 to define increasing characteristics that divide the space into two regions: operating region, and restrain region (Ali et al. 2019). The operating region handles internal faults, while the restrain region handles external through current. Rather over-excitation and inrush are handled using harmonics to fundamental ratios.

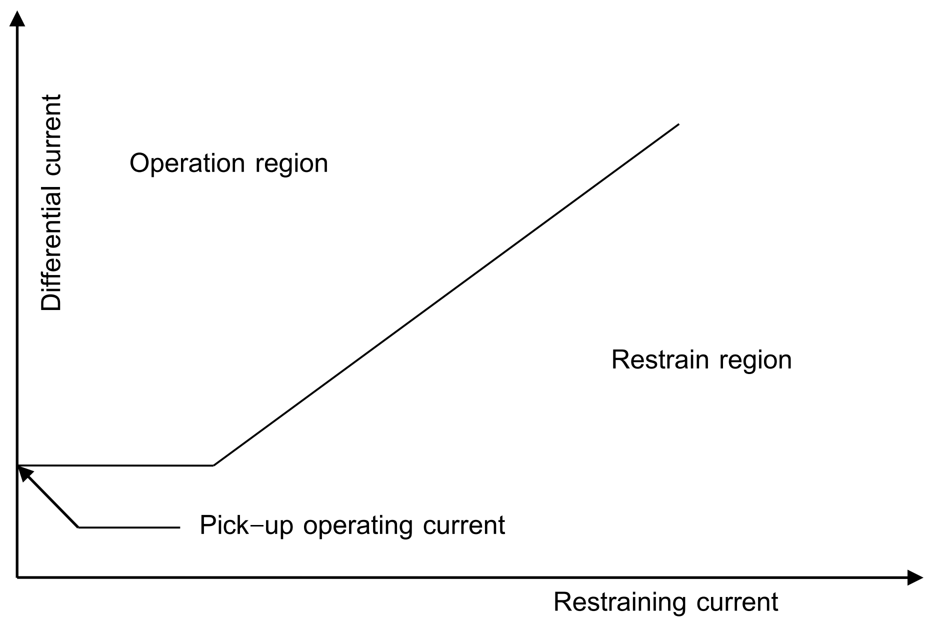


Figure 3.1: Traditional differential protection characteristic (Ali et al. 2018)

3.3 Transformer Model

The transformer model has two parts: winding and core. Since the transformer model needs to include transients unlike steady-state analysis as it has two main aspects, which are nonlinearity and frequency dependency, these feature present difficulties. The nonlinearity arises from magnetic core saturation region, whereas frequency effects on winding and core sections.

The steady-state transformer model based on matrix representation can be given in Ohm's law as:

$$V = ZI \tag{3.17}$$

where V is voltage, Z is impedance and I is current.

For modelling transient phenomena including the inductance effect:

$$\left[\frac{di}{dt}\right] = [L]^{-1}[v] + [L]^{-1}[R][i] \tag{3.18}$$

Where:

$\left[\frac{di}{dt}\right]$: current rate of change vector.

$[L]^{-1}$: inductance inverse matrix.

[R]: resistance matrix.

[i]: current vector.

This representation is valid for frequencies to coincide with the name plate for frequencies up to 1 KHz. For further discussion, the transformer core and winding topology affect the transformer model, i. e: Three-legged transformer zero-sequence flux passes through the air, while five-legged transformer provides bath through the core.

The transformer model three-phase, three-legged, and two winding can be represented, (Tokic et al. 2015) where a simplified model is available. In this model, which is based on self and mutual inductances, this technique has the problem of close values for self and mutual inductances (de Leon & Semlyen 1994). However, the transformer model has been solved by state-space matrices as follows:

$$\left[\frac{dX}{dt}\right] = [A][X(t)] + [B][U(t)] \tag{3.19}$$

Where:

$X(t)$: state variables = $[i_{L1} \quad i_{L2} \quad i_{L0} \quad \phi_{1_j} \quad \phi_{2_k} \quad \phi_{3_l}]$

$[U(t)]$: input vector= $[e_1 \quad e_2 \quad e_3 \quad S_{h_{1j}} \quad S_{h_{2k}} \quad S_{h_{3l}}]$

A: coefficient matrix

B: coefficient matrix

And: i_{L1}, i_{L2} and i_{L0}: are inductors currents

ϕ_{1_j}, ϕ_{2_k} and ϕ_{3_l}: hysteric inductors model flux

e_1, e_2 and e_3: are system voltages.

$S_{h_{1j}}$, $S_{h_{2k}}$ and $S_{h_{3l}}$: current sources from nonlinear hysteric inductors

model

One of the main concerns in the transient analysis is stiffness, since the problem may swing from non-stiff to extremely stiff situation which may affect the algorithm used and the number of steps (Tokic et al. 2015; de Leon & Semlyen 1994). Since some numerical methods require a small step size to ensure stability, the explicate numerical method needs to be avoided stiff system (Tokic et al. 2015).

Further development of the model can be achieved by taking into account winding and core topology, frequency dependency, and capacitor effect. Winding resistance can be approximated using (Martinez et al. 2005):

$$R = R_{dc} + \Delta R_{ac} \left(\frac{f}{f_0}\right)^m \tag{3.20}$$

Where:

R_{dc}: DC resistance.

ΔR_{ac}: slope of the curve at specified value of m.

$\frac{f}{f_0}$: ratio of required frequency AC resistance to fundamental frequency.

m= 1.2- 2.

Some more accurate representation can be adopted using the Foster series circuit (Martinez et al. 2005; de Leon & Semlyen 1994). Also, to have accurate results regarding iron core losses and nonlinearity can be achieved using the Cauer circuit (de Leon & Semlyen 1994).

The general model needs nameplate data and tests to estimate values, simulations are performed for the practical power transformer ratings obtained from Tamilnadu Electricity Board (TNEB), India, so data have been used from (Gethanjali, Raja Slochanal & Bhavani 2008). SIMULINK was used for modelling 20 different transformers, each transformer take one of the following power ratings: 16MVA, 25MVA, 5MVA, 3MVA, 2MVA, and one of the following voltage ratings: 110/33KV, 110/11KV, 66/33KV and 66/22KV. The transformer model is considered as a functional approximation which exhibits terminal behaviour. The output samples of this model were extracted using Fast Fourier Transform:

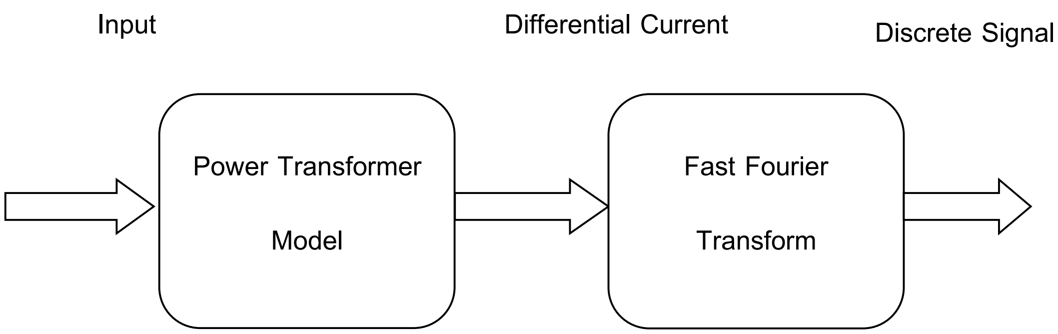


Figure 3.2: Protection system model

3.4 Current Signal Sampling

Regular protection systems use instrument transformer to supply analogue signals, which in turn supplied to relays. Protections have been evolved over the years from mechanical to electronic to digital. While old relay handled analogue signal, digital protection as the name implies deals with discrete signals rather continuous. Thus sampling is necessary for modern protection systems.

In sampled signal limited numbers of samples (4 – 20) are used rather than infinite like an analogue signal, so less information treated and a faster operating system. However, in traditional differential protection samples are used to calculate RMS fundamental differential, and restrain current as follows:

$$\mathrm{I_d(n)} = \text{Fundamintal of } (|\mathrm{i_p} - \mathrm{i_s}|) \tag{3.15}$$

$$\mathrm{I_r(n)} = \text{Fundamintal of } (|\mathrm{i_p} + \mathrm{i_s}|)/2 \tag{3.16}$$

The use of n indicates the use of samples. Equations 3.15 and 3.16 need the fundamental component of the signal to be calculated to use in operating and restraining action of protection. So, Discrete Fourier Transform is used to extract fundamental, second, and fifth harmonics.

3.5 Data Set Utilized

Differential current samples extracted in (Gethanjali, Raja Slochanal & Bhavani 2008). Hence, the discrete signal was sampled with 16 samples/cycle, each sample in the resulting data will be denoted by (P1, P2, P3 ...P16) where the symbol (P) denotes to Point. And the output will be denoted as (Type), which will take values (1, 2, 3, 4, 5). Full data can be found in appendix A, while the sample data shown in the following table 3.1:

Table 3.1: sample data used

Type	1	2	3	4	5
P16	606.682	0	-1.5673	-153	6590.6
P15	-0.4749	-4133.9	-3.1534	-14.4	-9634.2
P14	330.63	0	10.3537	44.9	5628.4
P13	329.63	4771.8	-6.7013	58.2	2704
P12	0.2936	0	-1.2971	14.4	-8834.5
P11	635.56	0	9.7868	-150.7	8575.1
P10	-1.0443	0	-10.207	-14.4	-2103.4
P9	1.1148	3433.9	-0.5718	47.2	-5827.8
P8	788.35	2999.6	6.372	60.6	9379
P7	-1.0845	-2675.4	-11.2082	14.4	-6417.2
P6	1.7889	-3735.2	6.2082	-148.2	-1361.4
P5	755.2	1699.9	1.2331	-14.4	7994.6
P4	-0.727	608.6	-6.7359	49.8	-9125.2
P3	207.3747	0	24.3207	63.5	3441.7
P2	-0.1481	-670.6	-0.0004	14.3	4716.7
P1	0	0	0	2524.8	-9626.9

Five Numbers (1, 2, 3, 4, and 5) assign to the five different cases of the transformer, as shown in the following figure 3.3:

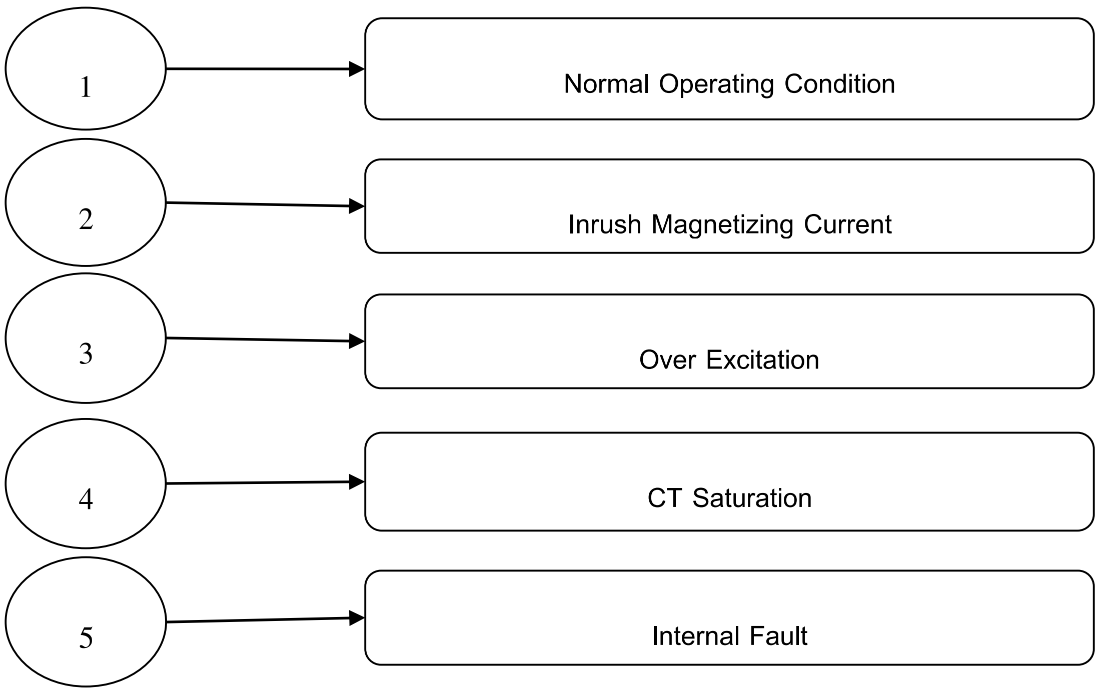


Figure 3.3: Current status sequence

3.6 Development of Ensemble Techniques for Current Classification

3.6.1 Tree Based Algorithm

These techniques partition the problem space into a separate domain, each domain or region is a rectangle, and each space divided using a recursive binary partition. Two elements are needed to perform this operation:

- Variable (feature) to split.
- Point to split on.

Different measures are used to guide tree-building algorithms based on trees building goals which are regression, classification, and purpose of using including growing or pruning trees. Regard regression, the sum of squares is a good impurity measure. While in classification, other measures of impurities are used as Gini index, cross-entropy, and misclassification errors. However, the Gini index and cross-entropy are more sensitive than misclassification errors. Consequently, they are more suitable for growing trees.

The tree is interpretable since the whole space is described by some inequalities. On the other hand, trees are weak learners:

$$P_\theta(h(x, \theta) = Y) > 0.5 \tag{3.1}$$

$$\text{And} \quad \text{error rate} < 50\% \tag{3.2}$$

Where:

P_θ: probability for this specific random vector.

h: classifier.

θ: random vector and x: input vector.

That means there is a better chance as the error rate is less than 50%. It also has a high variance, which is common since the change in input data can change the whole split process. Different tree algorithm is available, including CART, ID3, C4.5, and C5.0 (Hastie, Friedman and Tibshirani 2017).

3.6.2 Bagging

The Bagging technique uses bootstrap samples to build trees that are averaged over the ensemble to reduce variance, leaving bias unchanged. For classification, every tree cast a vote and the majority of trees vote to present a result, as shown in figure 3.4 while, figure 3.5 shows regression:

$$f_{Pred} = \frac{1}{N} \sum_{1}^{N} f_n \qquad (3.3)$$

Where:

f_n: prediction of nth predictor.

f_{Pred}: overall prediction for all predictors.

N: number of bootstrap samples (trees).

The main idea of bagging is to have an independent identical distribution, which implies zero correlation between pairs of trees in the ensemble and the same bias. Thereby it is suitable for high variance low bias examples. Unfortunately, bagged sampled vectors are not independent (Hastie, Friedman and Tibshirani 2017) and (Alfaro, Gámez and García 2013).

3.6.3 Random Forest

Random forest construction is similar to bagging, but it differs as it introduces more randomness to the model using different methods, such as:

- Random feature selection
- Random linear combination of input
- Random noise in the output

This technique commonly uses random feature selection where the number of features selected is less or equal to the total number of features, where these features used to split on by selecting the best split. More randomness implies less correlation, and more strength.

With the increasing number of trees generalization error will have an upper limit:

$$\text{Generalization error} = \frac{\rho(1-s^2)}{s^2} \quad (3.4)$$

Where:

ρ: average correlation between trees vector conditioned to training data

s: strength of classifiers given by the margin function

It is evident that error is combination of two trade-off values; increased correlation will increase error hence low classifier performance and vice versa. While increased strength will reduce error hence better classifier performance. As an example lower number of randomly selected input will reduce the correlation (Breiman 2001).

Another way to understand the principle in terms of variance; since independent identical distributed random variable with variance σ^2 will have variance for the average given by:

$$\text{Variance} = \frac{\sigma^2}{N} \qquad (3.5)$$

Where N is the number of trees in the ensemble.

Nevertheless, the variables are not independent rather identical distributed, and the variance of the average is given by:

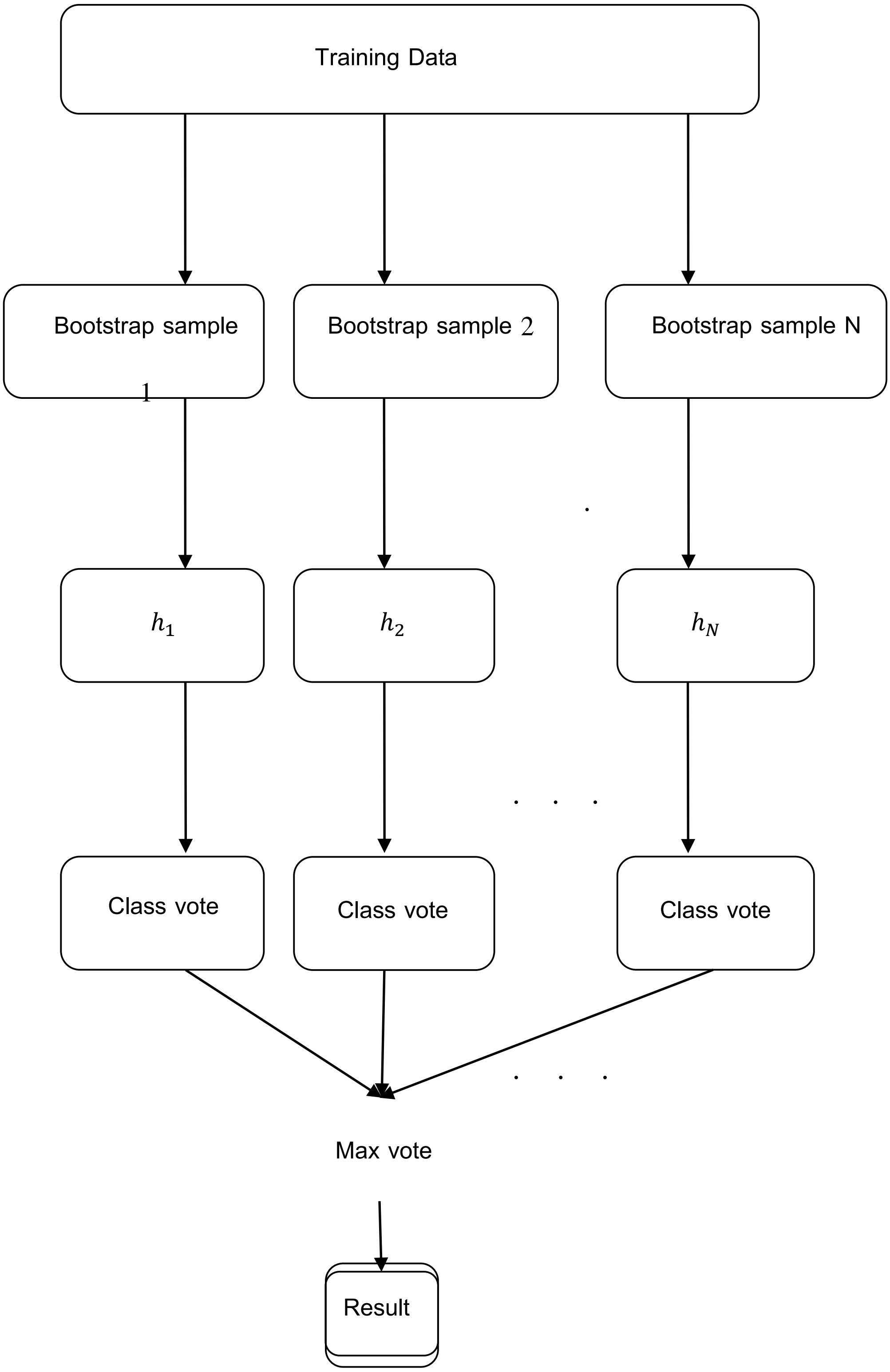


Figure 3.4: Classification bagging

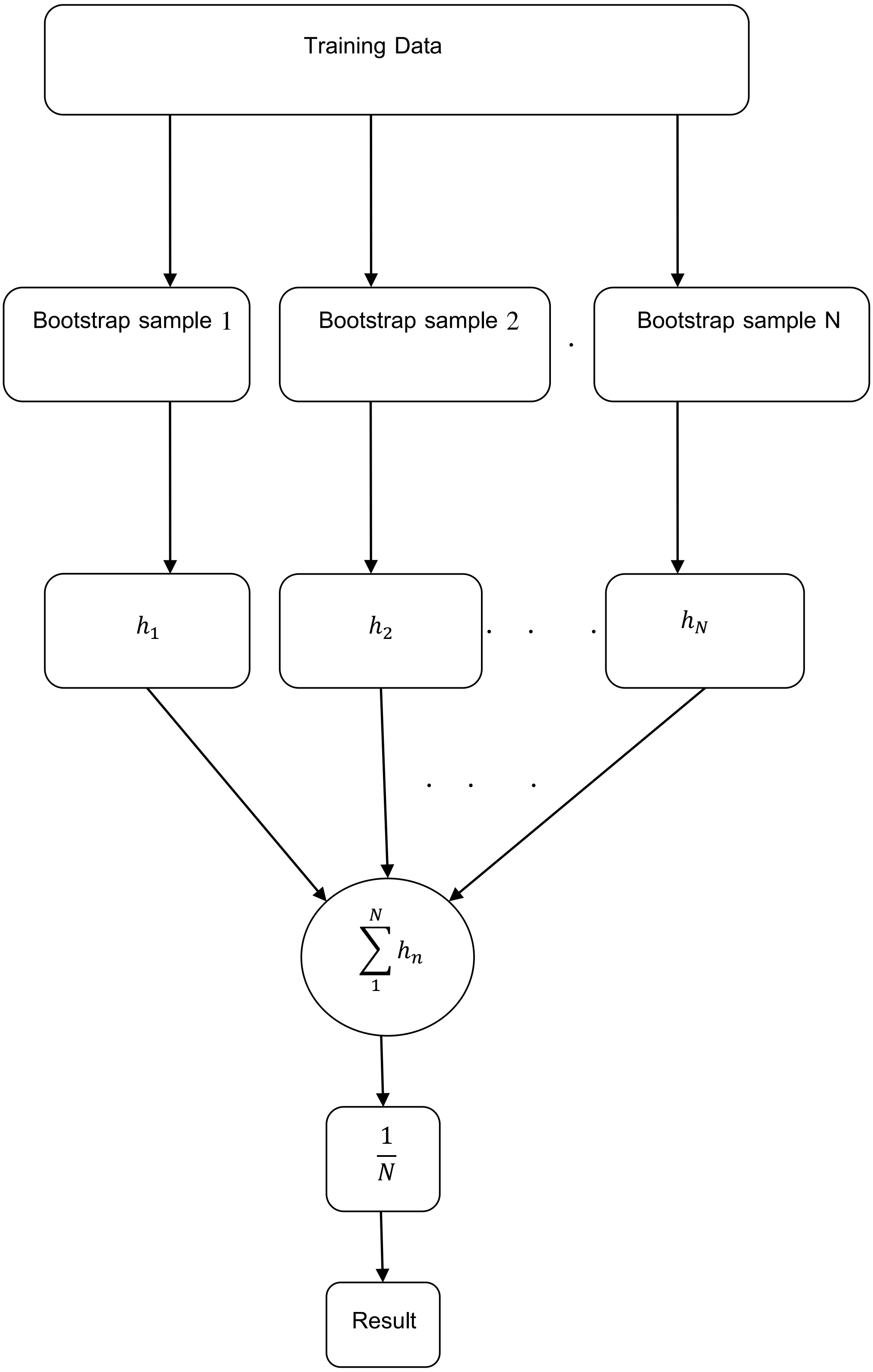


Figure 3.5: Regression bagging

$$\text{Variance} = \rho\sigma^2 + \frac{(1-\rho)\sigma^2}{N} \qquad (3.6)$$

, Where:

ρ: sampling correlation between pair of trees rather average

σ: sampling variance of single tree

The above equation clearly states that the second term vanishes with an increasing number of trees, limited to the value given by the first term. As stated earlier, random forest reduces variance by keeping bias unchanged hence variance is limited to the multiplication of correlation and tree variance. So, it reduces correlation without affecting variance too much by inserting more randomness like random selection (Hastie, Friedman and Tibshirani 2017).

The different number of input selected can be taken, but the default number of input selected is mentioned in the following table:

Table 3.2: Default values for number of feature to split and minimum node size

	Classification	Regression
Number of feature selected	$\sqrt{\text{number of inputs}}$	$\frac{\text{number of inputs}}{3}$
Minimum node size	1	5

3.6.4 Out of Bag

3.6.5 Random Forest Model

Random Forest Model is based on the R language, which uses the randomForest package. The data were separated into two groups, namely: training and testing with 80% and 20%, respectively. To validate the model, internal estimates and testing are used.

Now 16 features are discrete samples of differential current with typically 4 variables to split, and the response of 5 classes available. Therefore, the setting these data into the randomForest algorithm and bootstrap datasets generated from training.

The default values in the algorithm as follows:

ntree: number of trees to be grown = 500

mtry: number of variables to be selected randomly as candidate for splitting which differ for classification = $\sqrt{p}$ and for regression $p/3$, p : number of variables

nodesize: minimum size of terminal node for classification (1) for regression (5)

maxnodes: maximum number of terminal nodes subject to limits by nodesize.

The model was run and OOB error investigated for the best number of trees (ntree) and variables to split (mtry). Furthermore, changing the number of nodes by tuning (nodesize) and (maxnodes) , the result confirmed by extracting trees and observe its parameters.

Figure 3.6 shows the random forest model with input and output data. The output data clarified to be five cases.

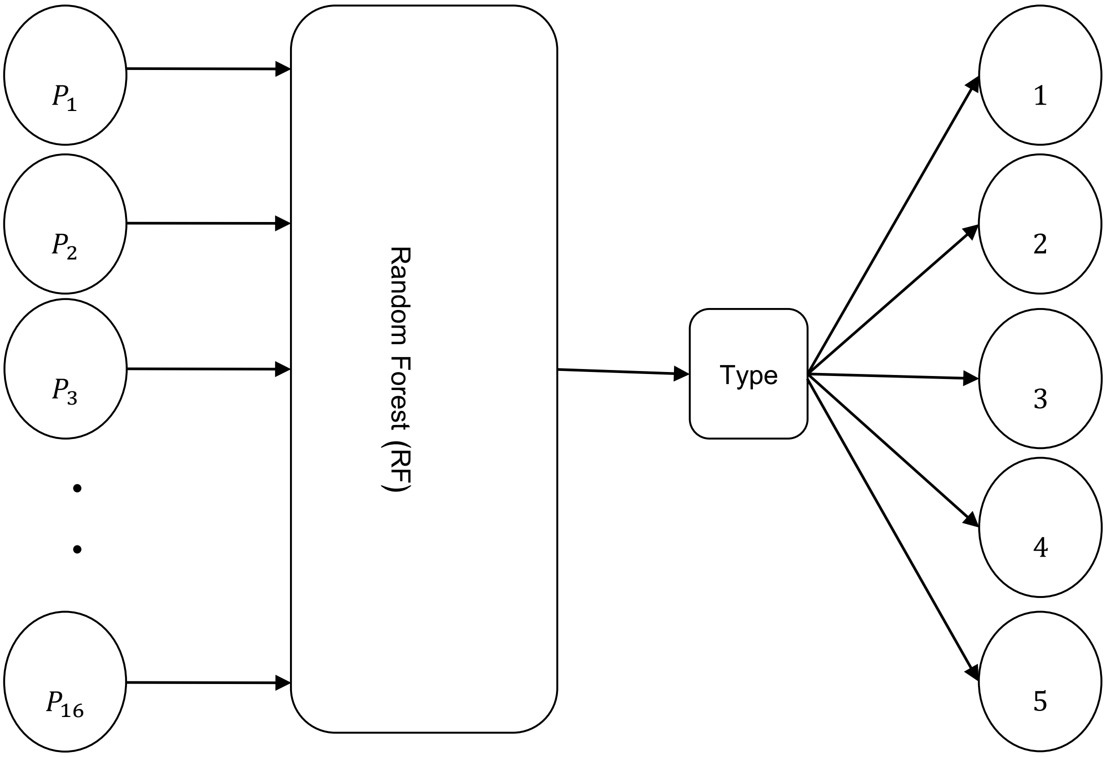


Figure 3.6: Random forest model

3.6.6 Boosting

Boosting is another ensemble technique. Nevertheless, it has a superficial similarity with bagging and random forest. Since the previous techniques build classifiers in a parallel way, the Boosting technique builds them in serial granting weights in two different steps.

Boosting represents the family of the algorithm: AdaBoost, AdaBoost.M1, SAMME, and others. Adaboost is a bi-class boosting technique. Moving to multiclass using forward stage-wise additive model, and fitting additive model (Hastie, Friedman and Tibshirani 2017):

$$\text{predection} = \sum \beta\, b(x;\gamma) \tag{3.8}$$

Where:

β: expansion coefficient.

$b(x;\gamma)$: function of multivariate argument characterized by a set of parameters γ.

By using a loss function that uses exponential loss for Adaboost.M1 and SAMME:

$$L\big(y, f(x)\big) = e^{-yf(x)} \tag{3.9}$$

Where:

y: function true value.

f(x): predicted value of function at each step.

Adaboost.M1 algorithm in figure 3.7 start by initializes the weight to (Alfaro, Gámez and García 2013):

$$w_i = \frac{1}{N} \tag{3.10}$$

Where i indicates the weight for each respective observation, and N is the number of observations. The following is the weighted error calculated by:

$$\text{error} = \sum\nolimits_{\text{observations}} w_i\, I(h(x) \neq y) \tag{3.11}$$

Where I( ) is the indicator function which outputs 1 if the argument is true and 0 otherwise. The training rate is calculated based on the error. It represents the contribution of each classifier in the final result since it is inserted in the final summation also, but in this stage, it represents learning rate α as follows:

$$\alpha = \ln \frac{1-\text{error}}{\text{error}} \tag{3.12}$$

The previous equation is based on Freund and Schapire, while the other boosting multiply of the previous equation by half is suggested by Breiman. Hereafter, weights are adjusted by updating then normalizing them, and wrongly classified examples have more attention, hence more weight. Whereas correctly classified examples have fewer weights using equation (3.13):

$$w_{i_{new}} = w_{i_{old}}\, e^{\alpha\, I(h(x) \neq y)} \tag{3.13}$$

The new training set used to train the new classifier, the steps are repeated where each classifier gives weighted vote using α.

The SAMME only differ from adaboost.M1 as it takes into account the number of classes by the meaning of training rate modifying it to:

$$\alpha = \ln \frac{1-\text{error}}{\text{error}} + \ln k - 1 \tag{3.14}$$

Where:

k: number of classes.

Which further subject random guess to be $\frac{1}{k}$ rather than $\frac{1}{2}$.

3.6.7 Boosting Model

The R function gbm used to get insight into this algorithm. The main features of gbm will be: distribution= multinomial because of multi-class classification problem, ntree= 701, bag.fraction= 0.5 to perform OOB estimation, shrinkage =0.1, cv. folds= 5 to perform cross-validation. Different ratios of training and testing will be done. The training to testing ratios will be 80/20, 60/40 respectively.

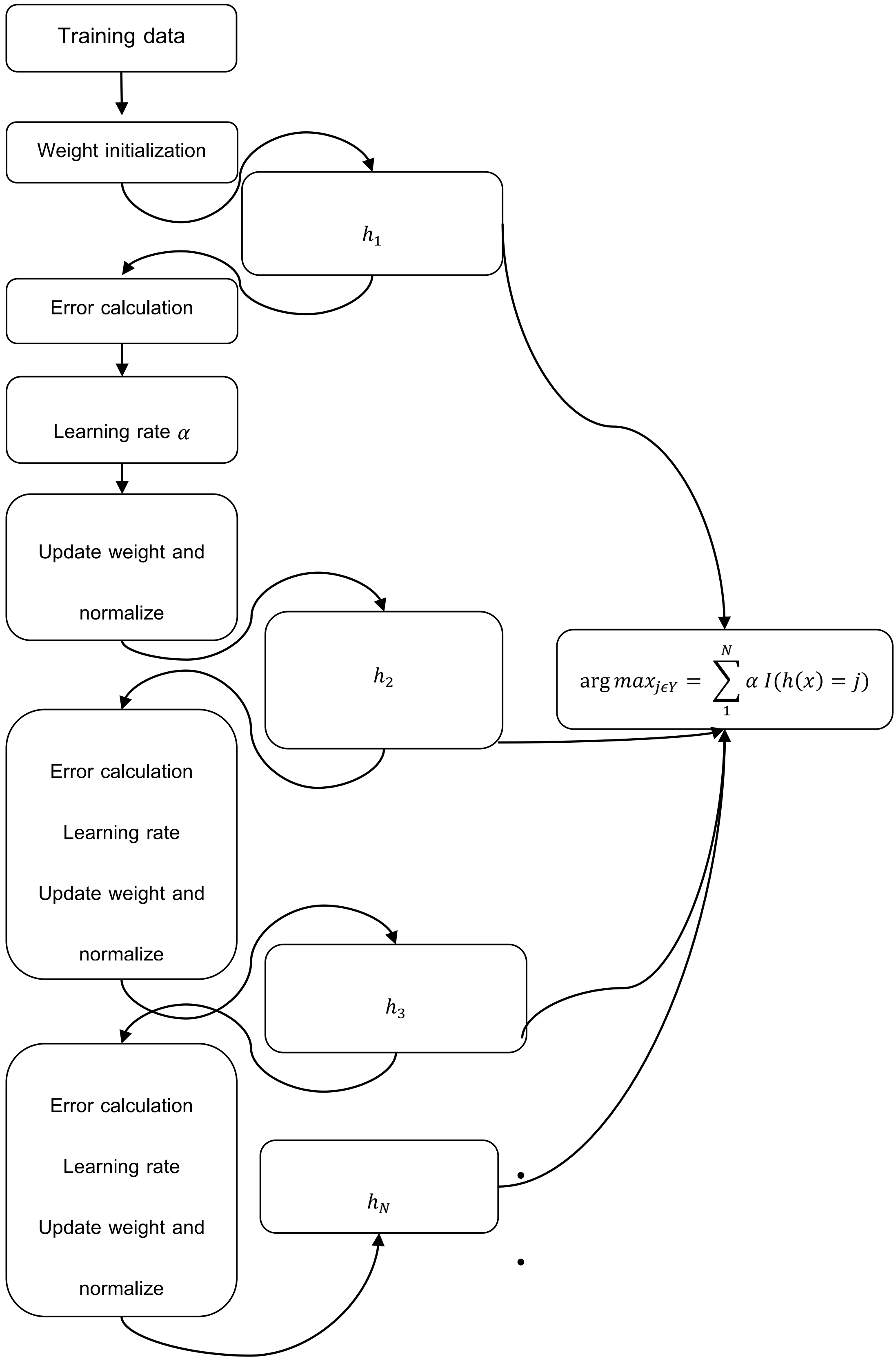


Figure 3.7: Boosting algorithm

Chapter Four

Analysis & Results

4.1 Classification Methodology

The random forest model is applied to differential current samples, where 16 features data columns are needed to be investigated, and one class output column is the Type (1, 2, 3, 4,

and 5). Random forest runs with a single candidate to be split and 2001 trees with 100% of data used in training, and minimum error produced by 33%. The following graph captures OOB error development with respect to number of trees:

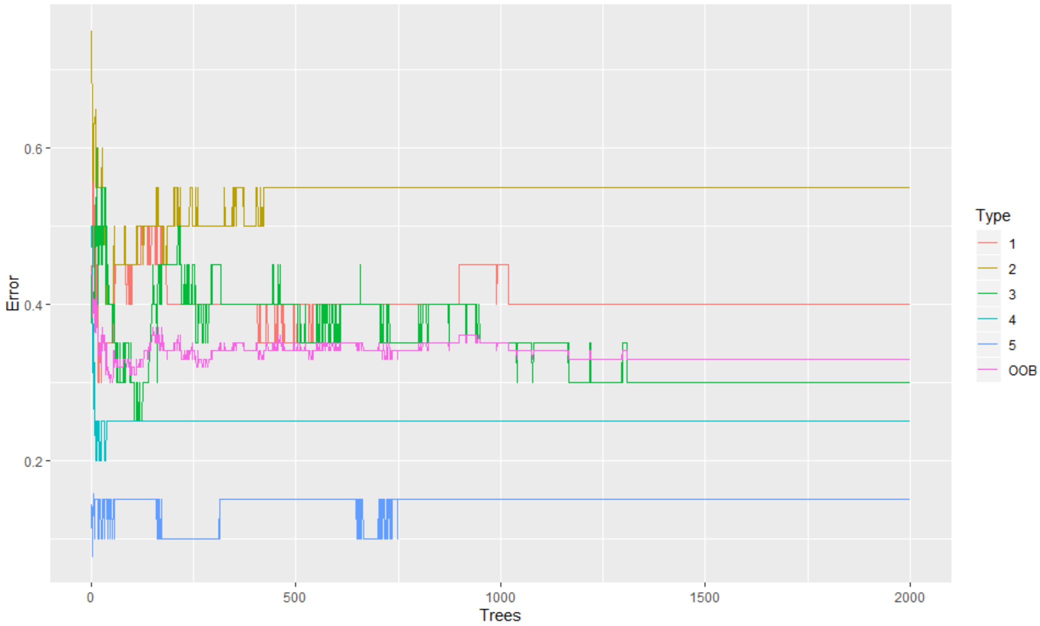


Figure 4.1: OOB error vs. number of trees for original data

These errors indicate poor performance with around one-third data misclassified. It has a normal operating condition, inrush current, over-excitation, CT-saturation, and an internal fault error rate of 40%, 55%, 30%, 25%, and 15% respectively. To have a deep insight into these errors, have an overview at confusion matrix:

Table 4.1: Confusion matrix for original data

		Prediction					
		1	2	3	4	5	Class error
True	1	12	6	0	1	1	0.4
	2	5	9	6	0	0	0.55
	3	0	5	14	1	0	0.3
	4	0	0	5	15	0	0.25
	5	3	0	0	0	17	0.15

The matrix confusion clarifies that 3 internal fault examples classified as a normal status, which will have huge destruction to the apparatus and continuity of supply if that happens and fail to trip occur. Other misclassified examples as 6, 1, and 1 normal are cases classified as inrush condition, CT-saturation and internal fault respectively imply false tripping case. While 5, and 1 over-excitation cases classified as inrush, and CT-saturation. The 5 CT-saturation examples are classified as over-excitation.

Further investigation is needed to explore places of model weaknesses. One to start is outliers, as indicated in graph 4.2 the y-axes correspond to the discrete differential current point, and the x-axis indicates the value of that current. Despite some outliers have been removed, there was not a generalized and powerful model. However, many pieces of literature reported random forest insensitivity to outliers.

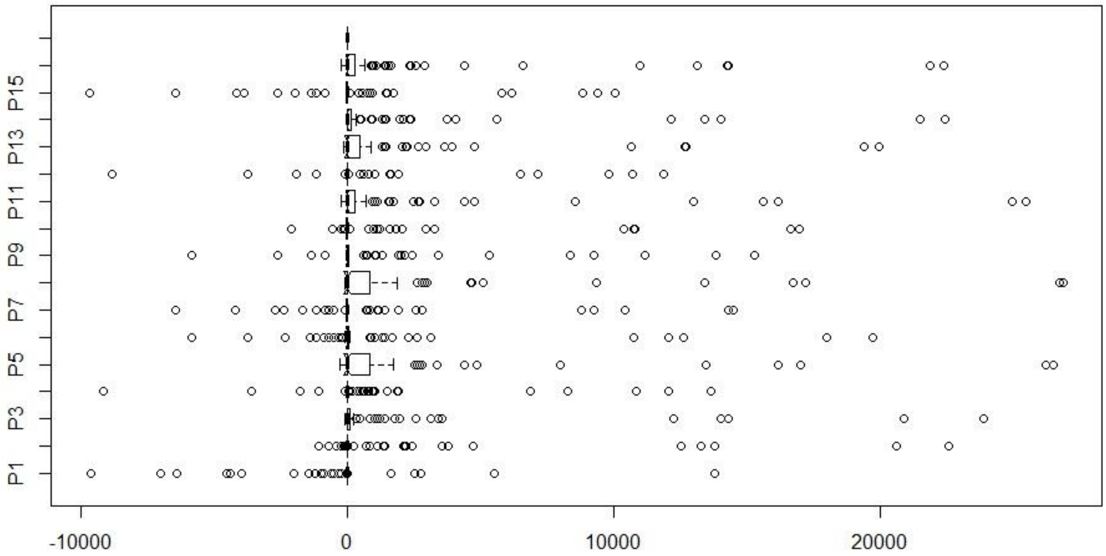


Figure 4.2: Outlier box-plot

To get an overall perspective for the data, the margin used for searching predictor confidence and reliability. As noticed in figure 4.3 margin vs. examples, around one-third of the data below zero, which indicates wrongly classified values, and around half of truly classified are below 0.5 with less reliability.

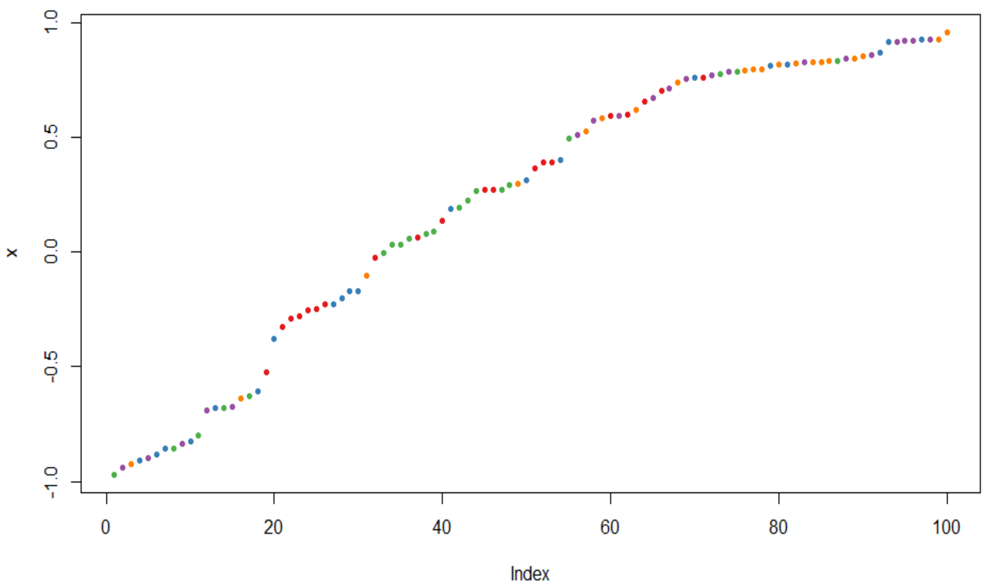


Figure 4.3: Margin vs. examples plot for original data

The above data have been treated all the same while they have different transformer ratings- apparent power, primary, and a secondary voltage-, with a different limit for each case. To overcome this issue, reweight current samples with different structures and relation i.e.: ${}^{S}/_{\sqrt{3}V_p}$, ${}^{S}/_{\sqrt{3}V_s}$, ${}^{S}/_{\sqrt{3}V_{average}}$ and different functions of them which still not good enough with OOB error of around 42%, 24%, and 27% respectively. Also, mixed data sets performed as a combination of those data sets and original data set, which also report poor performance with OOB error similar to the above limits. Also still suffering from failing to trip and false trip.

Finally, the apparent power of five power transformers has been replicated with the same distribution over five cases overall data set with different structures. The exponential relation of ${}^{S}/_{\sqrt{3}V_{average}}$ gives powerful random forest performance that reaches 0%. As a result, the random forest run with a single candidate to be split and 701 trees with 100% of data used in training, and minimum error produced by 0%. %. The following graphs in figure 4.4 capture OOB error development concerning the number of trees:

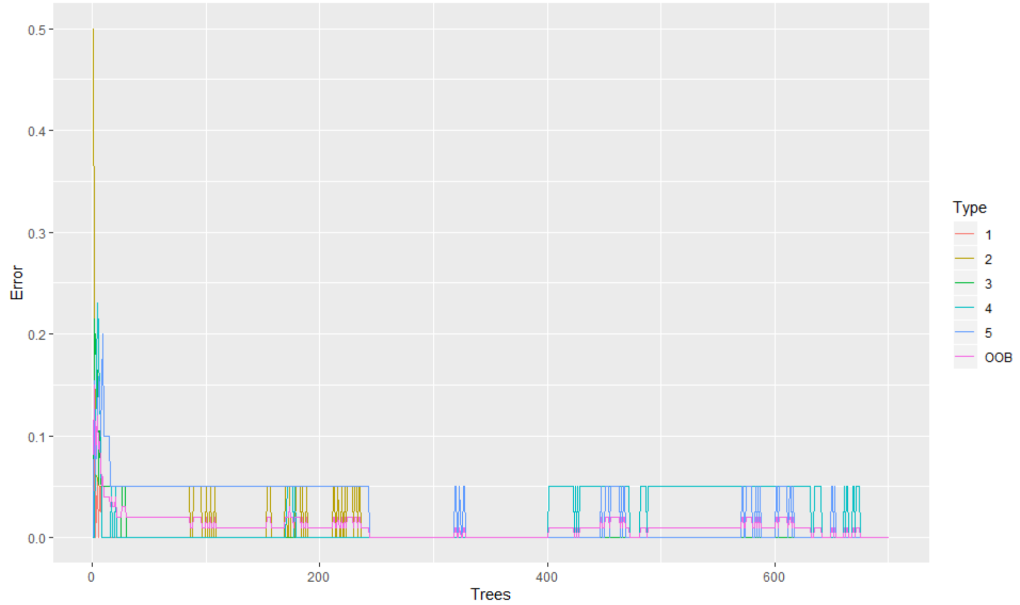


Figure 4.4: OOB error vs. number of trees for modified data set

Meanwhile, the confusion matrix gives a perfect situation as all numbers are diagonal, and the off-diagonal are zeros. The error rate of all cases (normal, inrush, over-excitation, CT-saturation, and internal fault) is 0%. The following table explains this principle:

Table 4.2: Confusion matrix for modified data set

		Prediction					
		1	2	3	4	5	Class error
True	1	20	0	0	0	0	0.0
	2	0	20	0	0	0	0.0
	3	0	0	20	0	0	0.0
	4	0	0	0	20	0	0.0
	5	0	0	0	0	20	0.0

To review how performance has been improved, the margin is investigated. In figure 4.5, some examples have a low margin with a minimum of 0.003703704. However, all data have a positive margin, which implies truly classified examples. Most examples are above 0.5 with a lot of the examples near 1 or 1, which means perfect prediction; this implies predictor confidence and reliability. In contrast to the previous case, half of the examples are below 0.5 with negative values.

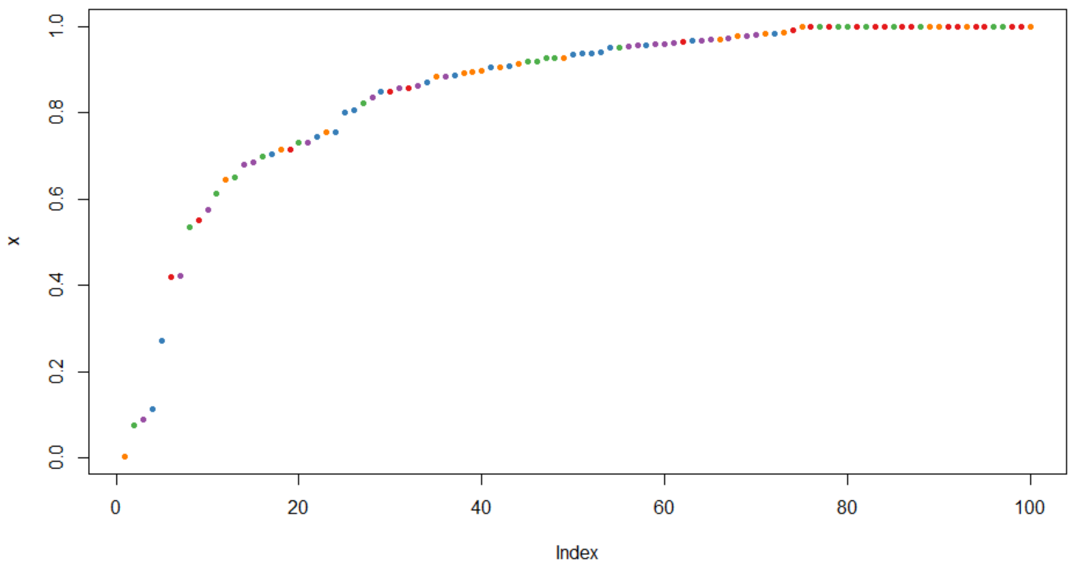


Figure 4.5: Margin vs. examples plot for weighted samples and identically distributed apparent power

The variable importance, which is one of the random forest strong features, gives insight over the significance and contribution of each feature to system accuracy. Moreover, it reflects the error decrease contribution of each variable to the overall accuracy. Thus, two measures of variable importance shown, which are the mean decrease accuracy (MDA), as well as the mean decrease Gini (MDG). Those factors are captured in the following figure:

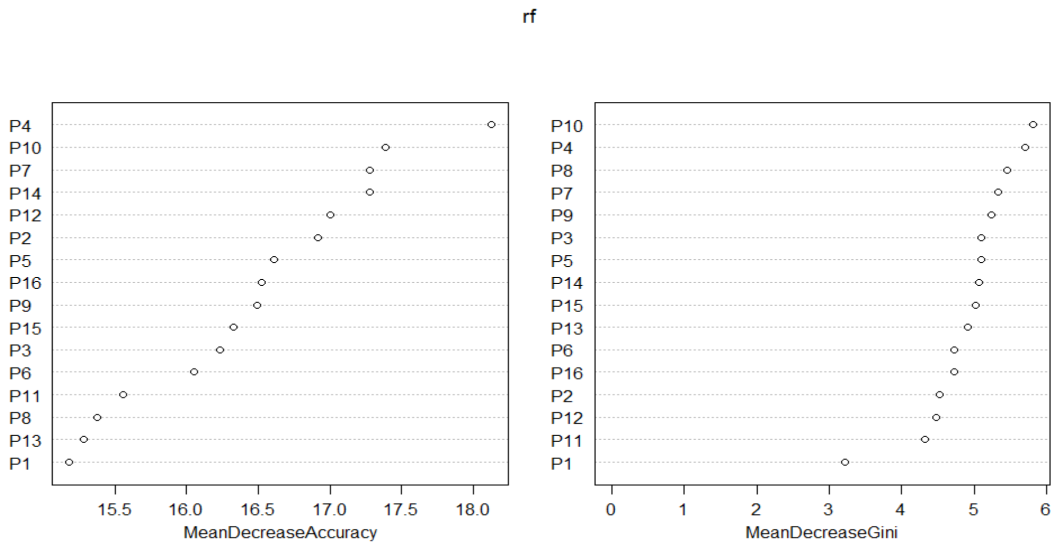


Figure 4.6: Variable importance

The MDA measured using OOB samples by permuting these samples for each tree. Based on OOB prediction error recorded, and OOB error recorded after permuting the variable, a sum of differences between them averaged over all trees in the ensemble. As the larger MDA, the more important is the variable. While MDG measures the sum decrease of node impurity by splitting on that variable, and the larger MDG, the purer is the variable.

Surprisingly, the importance of the samples is approximately the same for all samples. However, some early moments of the signal have higher importance than some late moments of signal i.e.: P_4 has higher importance than P_{16}. This observation indicates that early moments contain important information regarding signal behaviour, thus it is good points to be judge upon. Accordingly, different data windows will be investigated in the next section.

4.2 Data Window

Until now, a highly accurate system that requires one cycle to operate has been achieved. As stated previously, different moments of signal hold a different amount of information, thus different importance. This fact could be used to reduce the data window and ensure faster performance. Consequently, different data windows will be investigated with $3/4$, $1/2$, and $1/4$. Therefore, random forest run with a single candidate to be split and 701 trees with 100% of data used in training, and minimum error produced 0%. The following graphs in figure 4.7 capture OOB error development concerning several trees for ¾ data window:

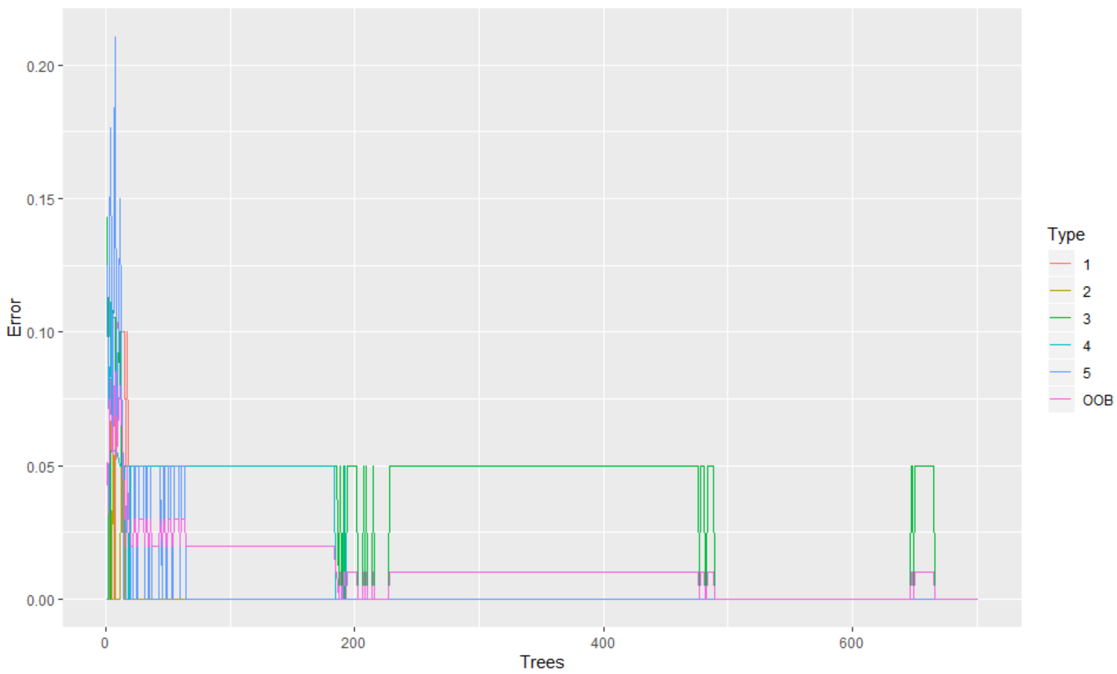


Figure 4.7: OOB error vs. number of trees for modified data set with ¾ data window

Running random forest with a single candidate to split and 701 trees, with 100% of data used in training, the minimum error produced is 1%. Nevertheless, reducing the number of trees to 101 minimum drops the error to 0%. The following graphs in figure 4.8 capture the OOB error development concerning several trees for ½ data window:

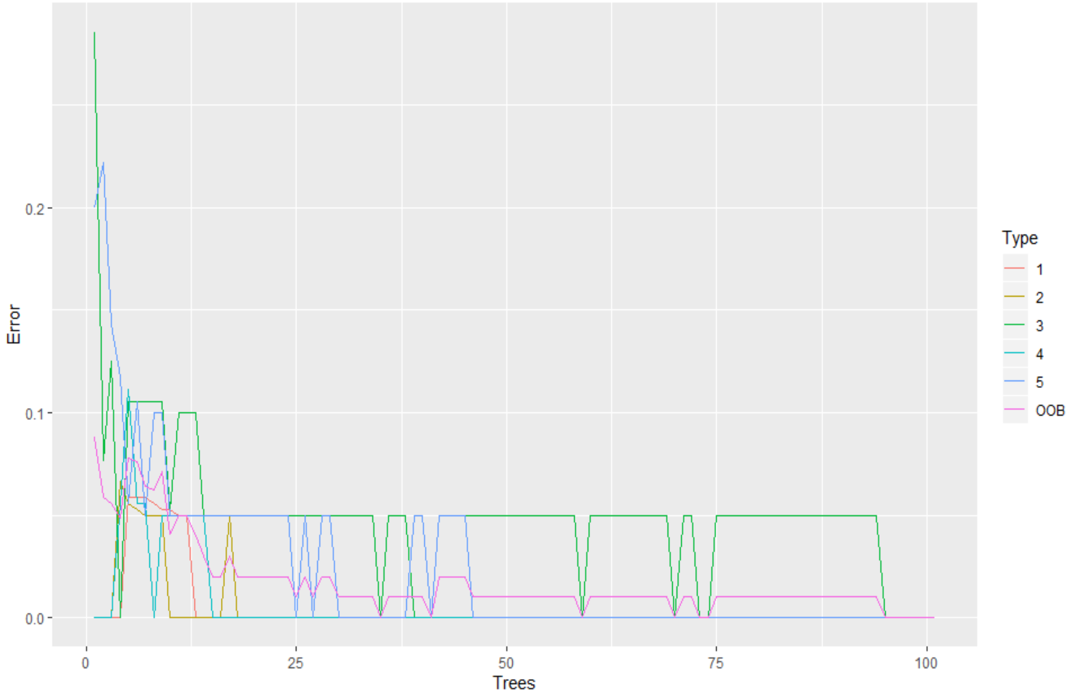


Figure 4.8: OOB error vs. number of trees for modified data set with ½ data window

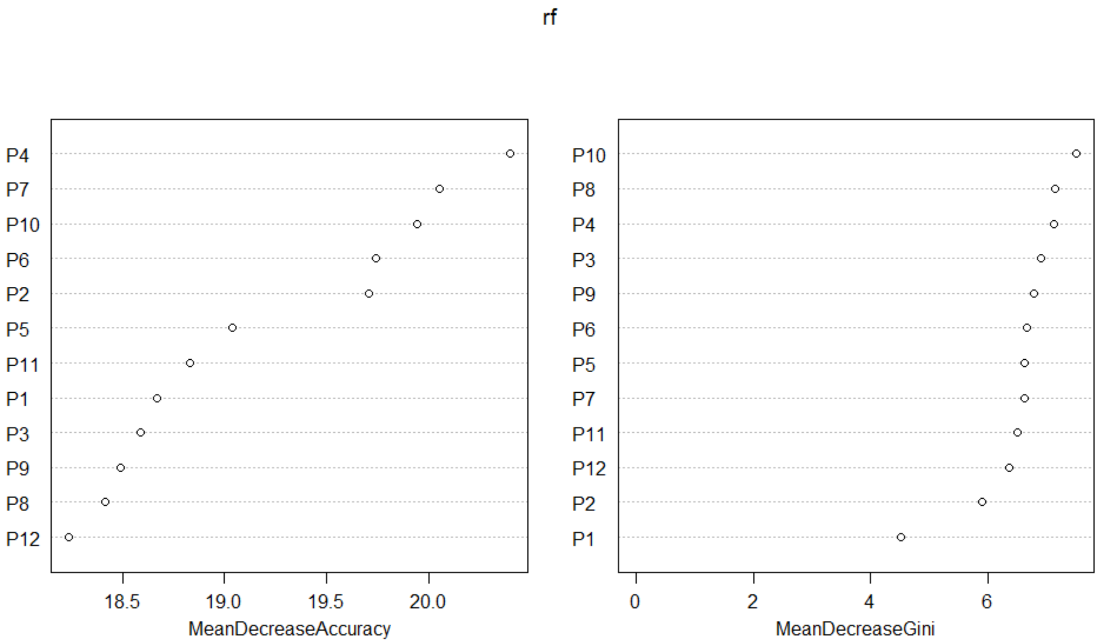


Figure 4.9: Variable importance for ¾ data set

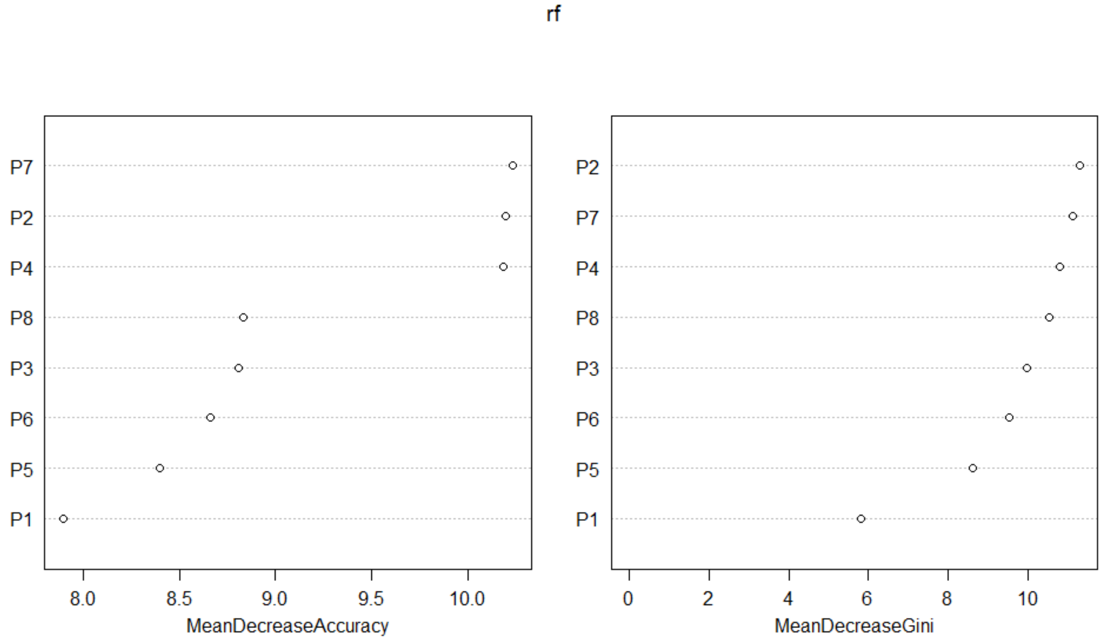


Figure 4.10: Variable importance for ½ data set

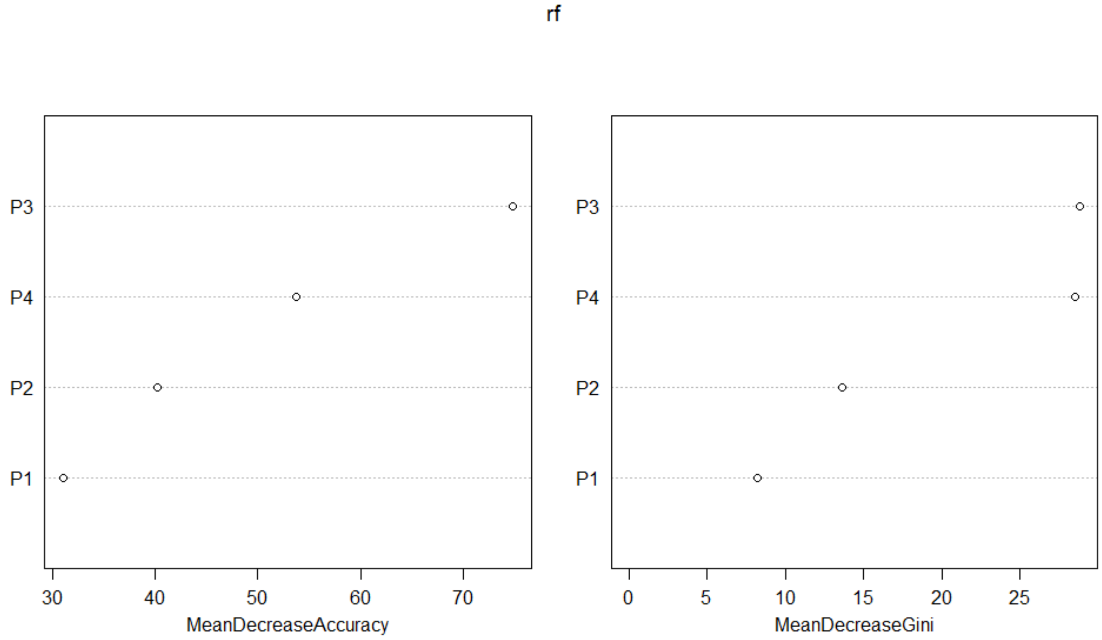


Figure 4.11: Variable importance for ¼ data set

The previous figures, from 4.9 to 4.11, reflect the variable importance of different data windows. While running random forest with a single candidate to be split and 701 trees with 100% of data used in training, the minimum error produced is 2%. However, by changing candidates to split to 4, minimum error drops to 1%. The following graph in figure 4.12 capture OOB error development concerning several trees for ¼ data window:

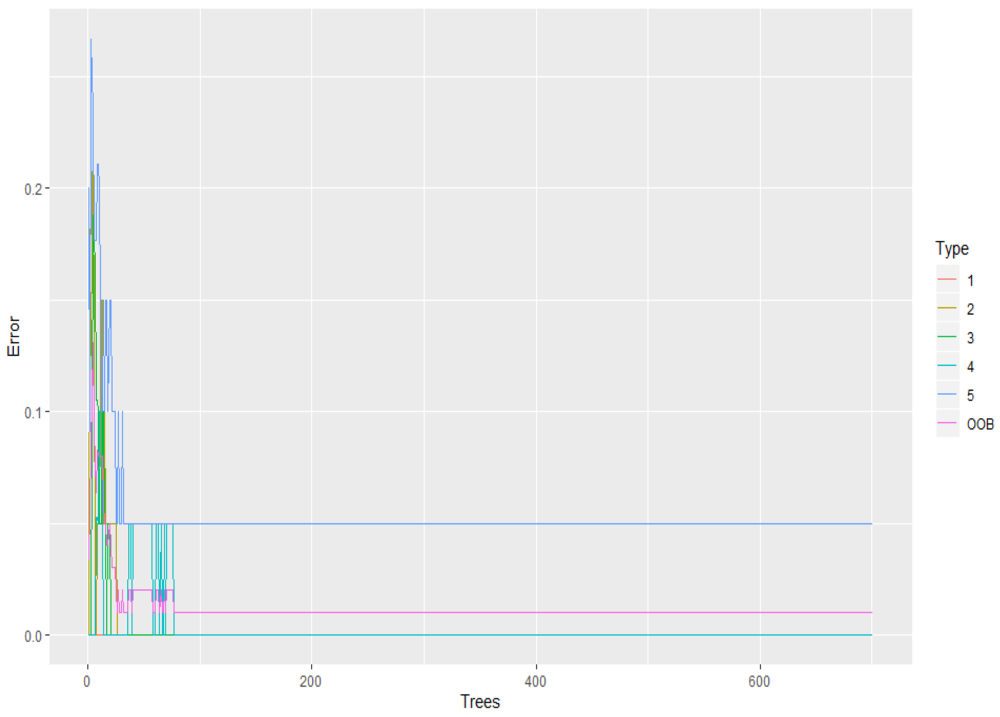


Figure 4.12: OOB error vs. number of trees for modified data set with ¼ data window

4.3 Model Testing

The previous analysis was based on OOB error which is good enough to make a comparison between different models. In order to increase test model credibility, different ratios of training and testing will be done, and training to testing ratios will be 80/20, 60/40 respectively.

First, random forest runs with a single candidate to be split and 701 trees with 80% of data used in training and 20% in testing. Furthermore, a minimum error produced 1.25% as indicated in the next confusion matrix:

Table 4.3: Confusion matrix for train to test ratio 80/20

		Prediction					
		1	2	3	4	5	Class error
True	1	16	0	0	0	0	0.0
	2	0	16	0	0	0	0.0
	3	1	0	11	0	0	0.08333333
	4	0	0	0	18	0	0.0
	5	0	0	0	0	18	0.0

The following graph capture OOB error development with respect to a number of trees:

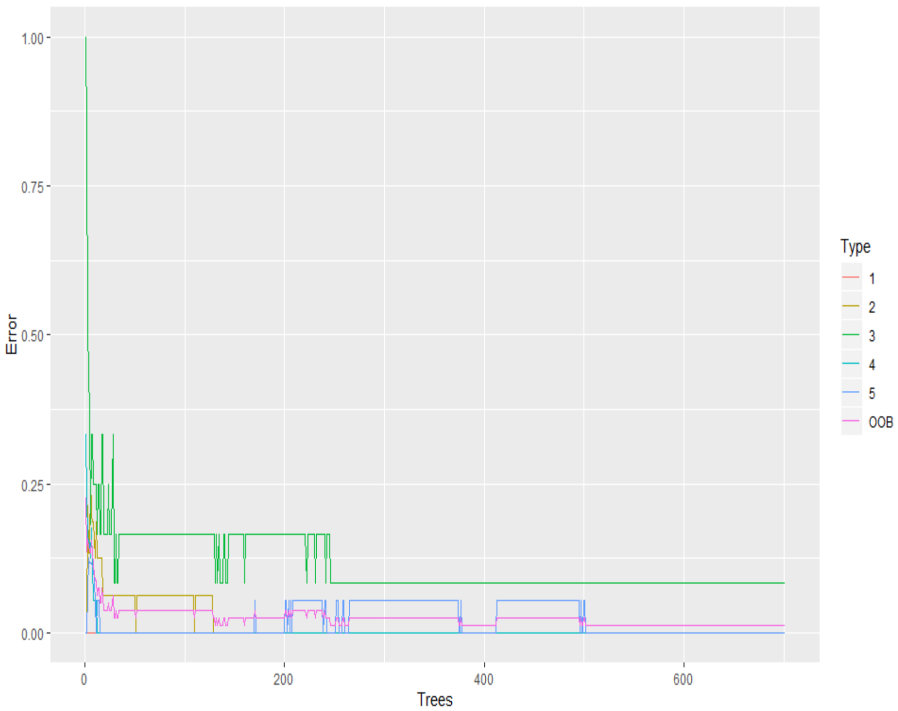


Figure 4.13: OOB error vs. number of trees for modified data set with 80/20 ratio

The following table shows comparison between true class value and prediction value. As can be seen, all examples are identical except example number 73 which predicts over-excitation as a normal condition. By reviewing the confusion matrix, it can be also seen that it predict over-excitation as normal condition. Even those were wrongly classified, it is not critical since both cases: normal and over-excitation are both no trip cases, hence no trip issued. Meaning, if there are two classes; no trip include: normal, inrush, over-excitation, and CT-saturation and trip includes: internal fault then there will be 100% accurate protection system. Moreover, in this case model works perfectly as protection system, rather less accurate condition monitoring.

Table 4.4: Testing versus prediction for 20% testing

	Testing	Prediction
7	2	2
18	3	3
20	5	5
21	1	1
23	3	3
31	1	1
32	2	2
33	3	3
34	4	4
58	3	3
62	2	2
64	4	4
68	3	3
73	3	1
78	3	3
87	2	2
88	3	3
91	1	1
96	1	1
100	5	5

Secondly, random forest runs with single candidate to be split and 701 trees with 60% of data used in training and 40% in testing. Furthermore, minimum error produced 5% as indicated in the next confusion matrix:

Table 4.5: Confusion matrix for train to test ratio 60/40

		Prediction					
		1	2	3	4	5	Class error
True	1	12	0	0	0	0	0.0
	2	0	12	1	0	0	0.07692308
	3	1	0	10	0	0	0.09090909
	4	0	1	0	10	0	0.09090909
	5	0	0	0	0	13	0.0

The following graph capture OOB error development with respect to number of trees:

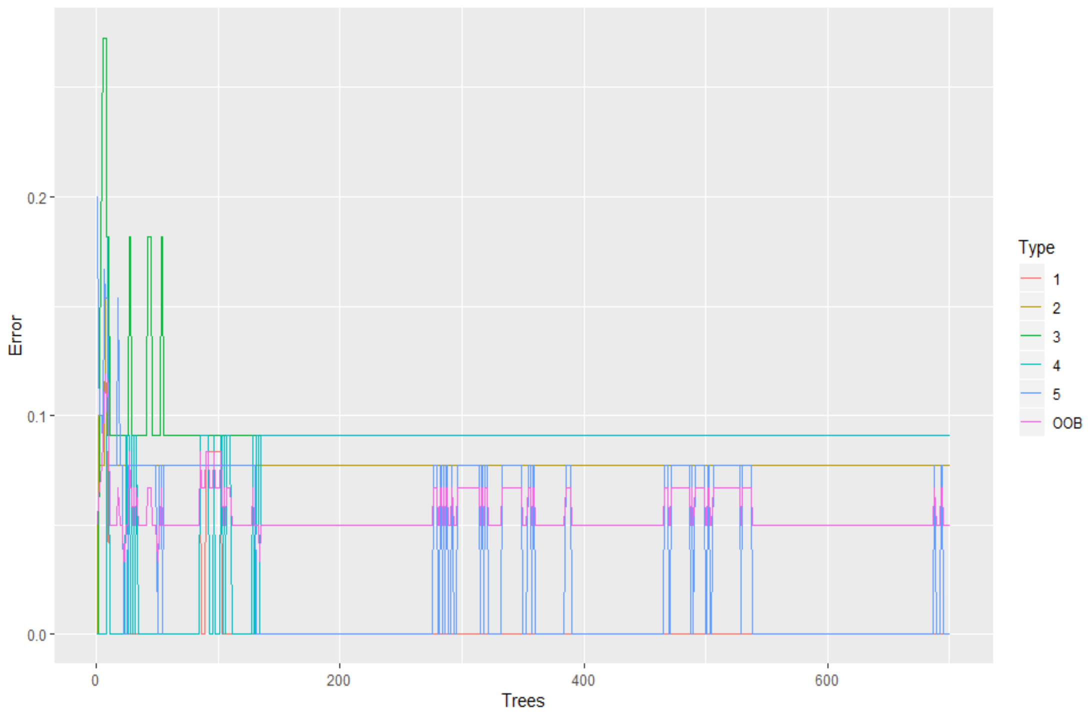


Figure 4.14: OOB error vs. number of trees for modified data set with 60/40 ratio

As previously, next table shows comparison between true class value and prediction value. As can be seen, all examples are identical. Taking another look on confusion matrix reveals that no internal fault has been confused with another class. Even other classes are wrongly classified among them; inrush detected as over excitation, over-excitation detected as normal, and CT-saturation detected as inrush. Again this is not critical; neither fails to trip nor false tripping. Same argument applied here, as if there are two classes: trip, and no-trip. There will be 100% accurate protection element. Hence, in this case, the model works perfectly as a protection system, rather less accurate condition monitoring.

Table 4.6: Testing versus prediction for 40% testing

	Testing	Prediction
1	1	1
3	3	3
7	2	2
8	3	3
10	5	5
13	3	3
14	4	4
18	3	3
19	4	4
20	5	5
21	1	1
24	4	4
26	1	1
28	3	3
33	3	3
35	5	5
37	2	2
39	4	4
42	2	2
47	2	2
49	4	4
52	2	2
53	3	3
56	1	1
58	3	3
59	4	4
60	5	5
65	5	5
66	1	1
68	3	3
74	4	4
76	1	1
80	5	5
84	4	4
86	1	1
87	2	2
89	4	4
92	2	2
95	5	5
96	1	1

4.4 Greedy vs. Limited Size

By default, random forest grows individual trees to maximum size such that each terminal node will have single feature. To investigate effect of limited size trees in the ensemble, trees maximum terminal node will be conditioned to suitable value. To start, investigate in single tree for single candidate to be split and 701 trees, with 100% of data used in training, which produce minimum error 0%. Next table give single tree for that case, where first column indicate node number while next two columns represent daughter node from that parent node which if zero it is terminal node. Fourth column gives feature to split which if NA then it is terminal, next fifth column indicate value of split for the feature. Sixth column labelled with -1 for terminal node and 1 if not. Final column indicate prediction class which if NA then it is not terminal.

Table 4.7: Single tree from the ensemble for 100% data subject to train and greedy case

	left daughter	right daughter	split var	split point	status	prediction
1	2	3	P3	5.53E-30	1	#N/A
2	4	5	P5	5.79E-89	1	#N/A
3	6	7	P6	-3.6E-23	1	#N/A
4	0	0	#N/A	0	-1	1
5	8	9	P8	5.33E-19	1	#N/A
6	10	11	P10	-5.8E-06	1	#N/A
7	12	13	P13	-4.5E-18	1	#N/A
8	14	15	P4	-3.7E-84	1	#N/A
9	0	0	#N/A	0	-1	2
10	0	0	#N/A	0	-1	5
11	16	17	P14	-2.5E-18	1	#N/A
12	18	19	P10	-2.7E-09	1	#N/A
13	0	0	#N/A	0	-1	2
14	0	0	#N/A	0	-1	3
15	20	21	P7	4.1E-90	1	#N/A
16	0	0	#N/A	0	-1	5
17	0	0	#N/A	0	-1	4
18	0	0	#N/A	0	-1	5
19	0	0	#N/A	0	-1	4
20	22	23	P2	-4E-06	1	#N/A
21	0	0	#N/A	0	-1	3
22	0	0	#N/A	0	-1	5
23	0	0	#N/A	0	-1	1

Terminal nodes in this tree, which differ from one tree to the other easily captured to be 12, i.e.: the second tree has 15 terminal nodes, and the third tree has 13 terminal nodes. This work investigates the effect of limiting it to half i.e.: six terminal nodes.

Table 4.8: Single tree from extracted for 100% data subject to train and 6 nodes case

	left daughter	right daughter	split var	split point	status	prediction
1	2	3	P4	-5.7E-19	1	#N/A
2	0	0	#N/A	0	-1	5
3	4	5	P6	4.58E-30	1	#N/A
4	6	7	P4	-1E-22	1	#N/A
5	8	9	P9	-2.5E-19	1	#N/A
6	0	0	#N/A	0	-1	4
7	10	11	P15	-2.2E-83	1	#N/A
8	0	0	#N/A	0	-1	4
9	0	0	#N/A	0	-1	2
10	0	0	#N/A	0	-1	3
11	0	0	#N/A	0	-1	1

By forcing this restrain, it is obvious that there will be 6 terminal nodes with minimum error raised to 9% as can be seen in the next confusion matrix:

Table 4.9: Confusion matrix for 100% data subject to train and 6 nodes case

		Prediction					
		1	2	3	4	5	Class error
True	1	20	0	0	0	0	0.0
	2	0	19	1	0	0	0.07692308
	3	5	0	15	0	0	0.09090909
	4	0	0	0	20	0	0.09090909
	5	0	1	0	2	17	0.0

4.5 Boosting

Boosting is another powerful ensemble technique, this techniques performance is highly comparable to random forest that surpass in some cases. Furthermore it is used to rise the performance of the investigate system.

Different boosting models run using different percentage ratios of train to test. Thus, different optimal number of trees yields. Next graph captures reduction in learning error (black line) and testing error (green line), while the blue line indicates the optimal number of trees.

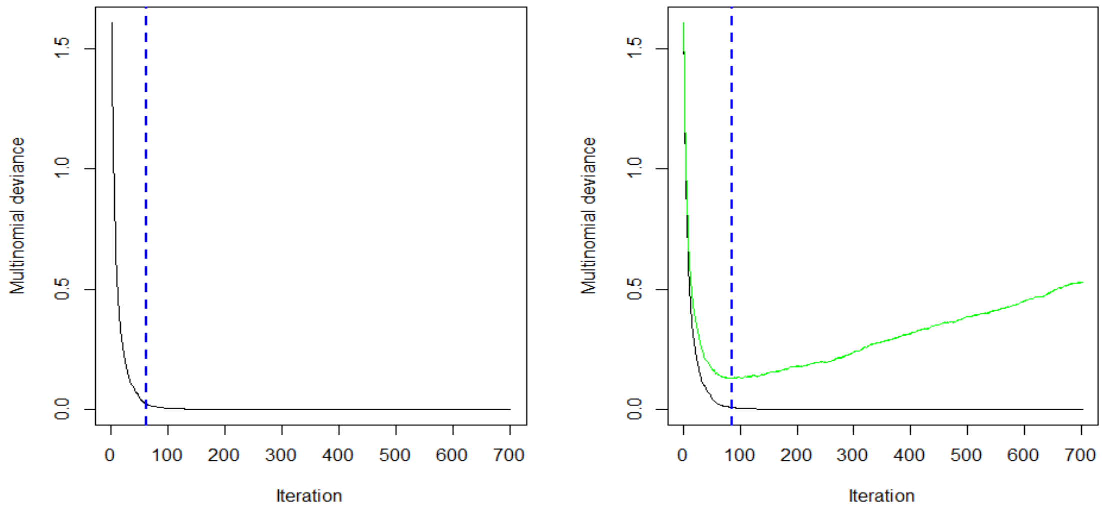


Figure 4.15: Optimal number of trees using; OOB to left, and cross validation to right for modified data set with 100% train data

The importance of optimal number is to achieve faster model, and not to over-fit data. Different optimal number of trees issued by two different techniques, data set ratios and different runs can be seen in the next table:

Table 4.10: Optimum number of trees for different data sets and datasets

	100% original dataset	100% modified dataset	80/20 % modified dataset	60/40 % modified dataset
C.V.	58	87	89	70
OOB	15	63	42	40

Even so, optimal number changes from single run to another, using optimal number of trees each case at a time. System will be trained same as previous parameters with original data set used. At the same time prediction is achieved by optimal number of trees with train to testing ratio 60/40. Next table shows comparison between prediction and testing for this case:

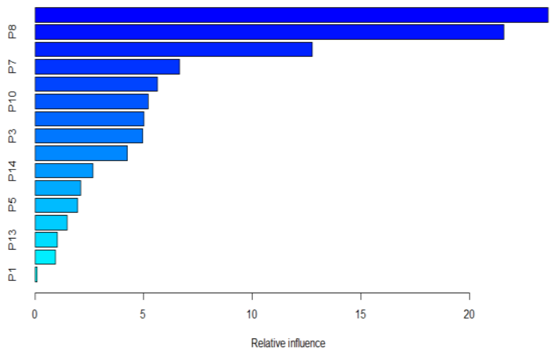
Table 4.11: Testing versus prediction for 40% testing with original dataset

	Testing	Prediction		Testing	Prediction
1	1	1	21	5	5
2	2	2	22	3	3
3	3	3	23	4	5
4	4	4	24	3	3
5	3	3	25	1	1
6	1	1	26	3	3
7	3	3	27	4	4
8	4	4	28	2	2
9	1	1	29	4	4
10	4	4	30	3	3
11	1	1	31	4	3
12	4	5	32	4	3
13	5	5	33	1	1
14	1	1	34	2	1
15	2	2	35	4	3
16	4	5	36	5	5
17	1	1	37	3	2
18	2	2	38	3	2
19	3	3	39	4	3
20	4	5	40	5	5

It is obvious from this table that examples 12, 16, 20, and 23 are all false tripping cases. Other miss-classified examples are 31, 32, 34, 35, 37, 38, and 39 which do not contain false tripping nor fail to trip cases. As a result, 11 out of 40 examples are wrongly classified which yield error rate of 27.5%.

Even so, other boosting models were conducted with same previous parameters with 80/20 and 60/40 train to testing ratio, which yield correct classification for all examples.

As in random forest, feature importance can be investigated. Thus, boosting model was run with all data used to train model with same previous parameters. Next graph gives indication to variable importance captured by boosting:



4	23.6066992	P_9	5.0184874		
8	21.5621208	P_3	4.9696931	P_6	1.4749698
1	12.7421081	P_2	4.2631377	P_{13}	1.0332702
7	6.6609434	P_1	2.6838868	P_{15}	0.9585620
1	5.6409609	P_1	2.1043767	P_1	0.1048687
1	5.2193576	P_5	1.9565575		

Figure 4.16: Variable impotence with optimal number of trees for 100% train data

4.6 Comparison between Proposed Algorithm and ANN

Traditional ANN does not provide reliable and accurate classification, i.e.: Feed Forward Neural Network Back Propagation (FFNN), Cascade Forward Back Propagation (CFBPNN) and Radial Basis Function Neural Network (RBFNN) (Yazadani-Asrami et al. 2015). Hence, other algorithms or combination of algorithms need to be applied. As a result, combinations of optimisation and ANN have been tested. PSO optimized ANN and IGSA optimized ANN was applied in (Gethanjali, Raja Slochanal & Bhavani 2008) and (Yazadani-Asrami et al. 2015) respectively.

Gethanjali, Raja Slochanal & Bhavani (2008) use train to test ratio of 80-to-20. Yazadani-Asrami et al. (2015) have used cross validation for testing Bayesian classifier and one transformer shown to test conditional monitoring. Both have reported accurate result as high as 100%.

Proposed technique has used strict validating regimes; validated using internal measures: out-of-bag error, variable importance, and margin. In addition, different train to test ratio of: 80-to-20 and 60-to-40 are used to validate the model.

Table 4.12 shows comparison between different techniques. Proposed technique will take into account confusion matrix and testing data set in random forest and testing data set in boosting.

Table 4.12: Comparison between different techniques

<table>
<tr><th></th><th>PSO-ANN</th><th>IGSA-ANN</th><th>RF 100% training</th><th>RF 80-to-20</th><th>RF 60-to-40</th><th>Boosting 80-to-20</th><th>Boosting 60-to-40</th></tr>
<tr><td>Protection element</td><td rowspan="2">100%</td><td rowspan="2">100%</td><td rowspan="2">100%</td><td>100%</td><td>100%</td><td>100%</td><td>100%</td></tr>
<tr><td>Conditional monitoring element</td><td>98%</td><td>97%</td><td>100%</td><td>100%</td></tr>
</table>

Previous results were captured for complete cycle of 16 samples/ cycles. Also different data windows considered of: ¾, ½ and ¼ cycles corresponding to 12, 8 and 4 samples/ cycles have been summarized in table 4.13. All cases in table 4.13 are for random forest with 100% data used in training, only internal measures used to validate the model.

Table 4.13: Different data window

	¾ Cycle	½ Cycle	¼ Cycle
Protection element	100%	100%	95%
Conditional monitoring element	100%	100%	100%

Table 4.12 and Table 4.13 indicate competitive accuracy and faster response. So, ensemble techniques provide fast response hence more compact system; less processing and memory needed. Besides, these techniques provide measures that could be used to understand some features more deeply. In addition, these models are compatible to existing hardware, hence only changing the algorithm rather than outside equipment or provisions.

Chapter Five

Conclusion and Future Work

5.1 Conclusion

This thesis is mainly intended to provide differential protection and conditional monitoring through subjecting sampled differential signal to ensemble techniques; random forest and boosting were used. Both techniques were applied using R language; random Forest and gbm packages. At first, practical transformer rating model was adopted and sampled differential current of this model was extracted. Then, these samples are fed to ensemble techniques to provide classification function. This technique is intended to issue tripping status for internal fault and no trip status otherwise. Meanwhile, all cases are classified including: normal, inrush, over-excitation, ct-saturation and internal fault to form conditional monitoring system.

This was achieved by subjecting random forest to different data sets weighting regimes and mixture of them. Even low performance was achieved with original data set with one third examples misclassified with negative margin values, different data set weighting scenarios were tested which also reported low performance. Two step data modifying procedure has given powerful performance, these two steps are:

- Weighted samples of Exponential relation of ${}^{\mathrm{S}}\!/\!_{\sqrt{3}\mathrm{V}_{\mathrm{average}}}$.
- Identically distributed apparent power.

New modified data has shown high accuracy with zero percent out-of-bag error for 100% of data used in training. Rather, different train to testing ratios for validating the model 80-to-20% and 60-to-40% used, which result perfect 100% accurate protection element rather less accurate conditional monitoring element. Data window has been changed to test model under faster performance. Different windows of: 3/4, 1/2, 1/4 of original data have shown high accuracy under faster responses.

Effectiveness of random forest growth deepness have been captured by condition tress terminal node to about half that of the greedy case. Hence error has reached 9%in compare to 0%.

Boosting also has been subjected to original and modified data set conditioned to optimum number of trees. And so, error has been monitored using testing data set for 80-to-20% and 60-to-40%. As a result, error has been decreased from 27.5% in original data set with 60-to-40% to 0% with modified data set in both 80-to-20% and 60-to-40%.

This new technique is used for the first time; it will contribute to provide a new scope of understanding differential signal different cases. Two measures of variable importance used: Mean Decrease Accuracy (MDA), and Mean Decrease Gini (MDG). Thus, it will provide how signal moments differ in importance’s from each other, giving surprisingly result that approximately all moment have similar importance’s even so some early moments have higher importance’s than late ones, which is contradiction to common sense of the more late instances the more the importance. In addition, variables importance’s have been captured for different data window size. Whereas, variable importance using boosting have showed similar result, rather higher variance between higher and lower variables importance’s.

New technique has showed accurate, fast differential element. In addition, external measures have proven that it is important to study model and signal used.

5.2 Future Work

Random forest and boosting are group of families that can be explored amongst to further find best version of random forest and boosting that are most adaptable to differential current signal statues i.e.: different random forest injected randomness forms, and different boosting methods. On the other side, different data representation and mixture of those representations may be tested in conjunction with different model version to find best pair of data representation and model.

Different data sets with power transformer rating variance to study effect of transformer rating. Furthermore, identical apparent power vector feature effectiveness under different data set and transformer ratings could be testing.

Also, Different data sets could be used to validate the model to further extent and proves its ability to real world application. This conducive to further apply this model to real transformer, even low size transformer can be used to validate technique ability in real world application.

References

[1] Oliveira, M., Bretas, A. and Ferreira, G. (2014). ***Adaptive differential protection of three-phase power transformers based on transient signal analysis***. **International Journal of Electrical Power & Energy Systems**, 57, pp.366-374.

[2] Geethanjali, M., Raja Slochanal, S. and Bhavani, R. (2008). ***PSO trained ANN-based differential protection scheme for power transformers***. **Neurocomputing,** 71(4-6), pp.904-918.

[3] Yazdani-Asrami, M., Taghipour-Gorjikolaie, M., Mohammad Razavi, S. and Asghar Gholamian, S. (2015). ***A novel intelligent protection system for power transformers considering possible electrical faults, inrush current, CT saturation and over-excitation***. **International Journal of Electrical Power & Energy Systems**, 64, pp.1129-1140.

[4] Žarkovic, M. and Stojkovic, Z. (2017). ***Analysis of artificial intelligence expert systems for power transformer condition monitoring and diagnostics.*** **Electric Power Systems Research**, 149, pp.125-136.

[5] Breiman L. (2001) Random forests. ***Machine Learning*** , 45,5-32

[6] Ibrahim, I. and Khatib, T. (2017). ***A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm***. **Energy Conversion and Management**, 138, pp.413-425.

[7] Hastie, T., Friedman, J. and Tibshirani, R. (2017). **The elements of statistical learning**. 2nd ed. New York: Springer.

[8] Weng, H., Wang, S., Wan, Y., Lin, X., Li, Z. and Huang, J. (2019). ***Discrete Fréchet distance algorithm based criterion of transformer differential protection with the immunity to saturation of current transformer***. **International Journal of Electrical Power & Energy Systems,** 115, p.105449.

[9] Samantaray, S. and Dash, P. (2011). ***Decision Tree based discrimination between inrush currents and internal faults in power transformer***. **International Journal of Electrical Power & Energy Systems**, 33(4), pp.1043-1048.

[10] Shi, D., Buse, J., Wu, Q., Jiang, L., Xue, Y. (2011). ***Fast Identification of Power Transformer Magnetizing Inrush Currents Based On Mathematical Morphology and ANN*. IEEE Power and Energy Society General Meeting**, pp. 1-6.

[11] Tokic, A., Milardic, V., Uglešic, I. and Jukan, A. (2015). ***Simulation of three-phase transformer inrush currents by using backward and numerical differentiation formulae***. **Electric Power Systems Research**, 127, pp.177-185.

[12] Martinez, J., Walling, R., Mork, B., Martin-Arnedo, J. and Durbak, D. (2005). ***Parameter Determination for Modeling System Transients—Part III: Transformers IEEE PES Task Force on Data for Modeling System Transients of IEEE PES Working Group on Modeling and Analysis of System Transients Using Digital Simulation (General Systems Subcommittee)***. **IEEE Transactions on Power Delivery**, 20(3), pp.2051-2062.

[13] de Leon, F. and Semlyen, A. (1994). ***Complete transformer model for electromagnetic transients***. **IEEE Transactions on Power Delivery**, 9(1), pp.231-239.

[14] Brandwajn, V., Dommel, H. and Dommel, I. (1982). ***Matrix Representation of Three-Phase N-Winding Transformers for Steady-State and Transient Studies*. IEEE Power Engineering Review**, PER-2(6), pp.27-28.

[15] Barhate, V., Thakre, K. and Deshmukh, M. (2016) ***Adaptable Differential Relay using Fuzzy Logic Code in Digital Signal Controller for Transformer Protection*. 57th International Scientific Conference on Power and Electrical Engineering**.

[16] PROTECTION APPLICATION HANDBOOK. ***ABB BA THS / BU Transmission Systems and Substations***, pp.163-174.

[17] Peres, L. and Silva, K. (2019). ***Power transformer protection using an instantaneous-current-value negative sequence differential element*. International Journal of Electrical Power & Energy Systems,** 108, pp.96-106.

[18] Cardelli, E., Della Torre, E., Esposito, V. and Faba, A. (2009). ***Theoretical Considerations of Magnetic Hysteresis and Transformer Inrush Current***. **IEEE Transactions on Magnetics,** 45(11), pp.5247-5250.

[19] Angrisani, L., Bonavolontà, F., Liccardo, A. and Schiano Lo Moriello, R. (2018). ***Identification and classification of transformers current***

transients through Huang Hilbert Transform. **Measurement**, 125, pp.123-132.

[20] Behvandi, A., Seifossadat, S. and Saffarian, A. (2019). ***A new method for discrimination of internal fault from other transient states in power transformer using Clarke's transform and modified hyperbolic S-transform***. **Electric Power Systems Research**, 178, p.106023.

[21] Saravanan, B. and Rathinam, A. (2017). ***Inrush Blocking Scheme in Transformer Differential Protection***. **Energy Procedia**, 117, pp.1165-1171.

[22] Tripathy, M., Maheshwari, R. and Verma, H. (2010). ***Power Transformer Differential Protection Based On Optimal Probabilistic Neural Network***. **IEEE Transactions on Power Delivery,** 25(1), pp.102-112.

[23] Shah, A. and Bhalja, B. (2013). ***Discrimination Between Internal Faults and Other Disturbances in Transformer Using the Support Vector Machine-Based Protection Scheme.*** **IEEE Transactions on Power Delivery,** 28(3), pp.1508-1515.

[24] Hasheminejad, S. and Esmaeili, S. (2013). ***Transient actions analysis of power transformers based on S-transform and hidden Markov model.*** **International Transactions on Electrical Energy Systems,** 24(6), pp.826-841.

[25] Fernandes, J., Costa, F. and Medeiros, R. (2016). ***Power Transformer Disturbance Classification Based on the Wavelet Transform and***

***Artificial Neural Networks*. International Joint Conference on Neural Networks**, 2016, pp 640-646.

[26] Rasoulpoor, M. and Banejad, M. (2013). ***A correlation based method for discrimination between inrush and short circuit currents in differential protection of power transformer using Discrete Wavelet Transform: Theory, simulation and experimental validation*. International Journal of Electrical Power & Energy Systems**, 51, pp.168-177.

[27] Sheng, Y. and Rovnyak, S. (2002). ***Decision trees and wavelet analysis for power transformer protection*. IEEE Transactions on Power Delivery**, 17(2), pp.429-433.

[28] Shi, D., Buse, J., Wu, Q., Jiange, L. and Xue, Y. (2011). **Fast Identification of Power Transformer Magnetizing Inrush Currents Based On Mathematical Morphology and ANN**.

[29] Ali, E., Helal, A., Desouki, H., Shebl, K., Abdelkader, S. and Malik, O. (2018). ***Power transformer differential protection using current and voltage ratios*. Electric Power Systems Research**, 154, pp.140-150.

[30] Ali, E., Malik, O., Abdelkader, S., Helal, A. and Desouki, H. (2019). ***Experimental results of ratios-based transformer differential protection scheme*. International Transactions on Electrical Energy Systems,** 29(11).

[31] Ozgonenel, O. and Karagol, S. (2014). ***Transformer differential protection using wavelet transform*. Electric Power Systems Research**, 114, pp.60-67.

[32] Martínez-Muñoz, G. and Suárez, A. (2010). ***Out-of-bag estimation of the optimal sample size in bagging***. **Pattern Recognition**, 43(1), pp.143-152.

[33] Alfaro, E., Gámez, M. and García, N. (2013). ***adabag: AnRPackage for Classification with Boosting and Bagging*. Journal of Statistical Software,** 54(2).

Appendices

Appendix A: R Coding for Random Forest and Boosting

This code was built based on R which is a statistical and data analysis special program and language. Please be aware that other graphs and features which are introduced in this report need different coding lines. The following code represents the lines used to apply and run Random Forest and Boosting:

```
install.packages("randomForest")
install.packages("gbm")
library(gbm)
library(randomForest)
install.packages("xlsx")
library(xlsx)
path <- "C:\\Users\\user\\Desktop\\AI\\Book4NEW.xlsx"
data <- read.xlsx(path, sheetIndex = "Sheet1")
head(data)
str(data)
data$Type = as.factor(data$Type)
data_set_size= floor(nrow(data)*0.80)
index <- sample(1:nrow(data),size = data_set_size)
training <- data[index,]
testing <- data[-index,]
set.seed(1)
```

```
rf <- randomForest(Type ~ .,data = training , mtry = 1,maxnodes = 5, ntree = 5001, replace= TRUE ,  importance = TRUE)
result <- data.frame(testing$Type, predict(rf, testing[,1:16], type = "response"))
boost.rf=gbm(Type ~ . , data=training , distribution = "multinomial" ,shrinkage = 0.1,bag.fraction =0.5,keep.data = FALSE,cv.folds = 5, n.trees=701)
result <- data.frame(testing$Type, apply(predict(boost.rf, testing[,1:16], type = "response",ntree=gbm.perf(boost.rf, method = "cv")), 1, which.max))
summary(boost.rf, n.trees = gbm.perf(boost.rf, method = "cv"))
```

Appendix B: Data set Used

Type	1	2	3	4	5	1	2	3	4	5
P16	606.682	0	-1.5673	-153	6590.6	982.5	0	-2.456	-238	2903.8
P15	-0.4749	-4133.9	-3.1534	-14.4	-9634.2	-0.7	-6419.2	-5.0487	-23	-3862.7
P14	330.63	0	10.3537	44.9	5628.4	552.2	0	16.157	70	1989.9
P13	329.63	4771.8	-6.7013	58.2	2704	550.8	7.4973	-10.476	91	1413.8
P12	0.2936	0	-1.2971	14.4	-8834.5	0.5	0	-2.0131	23	-3703.1
P11	635.56	0	9.7868	-150.7	8575.1	1029.8	0	15.2832	-235	3295.5
P10	-1.0443	0	-10.207	-14.4	-2103.4	-1.2	0	-15.95	-23	-539.7
P9	1.1148	3433.9	-0.5718	47.2	-5827.8	1.8	5365.4	-0.8912	74	-2621.7
P8	788.35	2999.6	6.372	60.6	9379	1269.7	4686.7	9.9493	95	3.7907
P7	-1.0845	-2675.4	-11.2082	14.4	-6417.2	-1.7	-4180.4	-17.515	23	-2353.8
P6	1.7889	-3735.2	6.2082	-148.2	-1361.4	2.8	-5836.3	9.7149	-231	-887.9
P5	755.2	1699.9	1.2331	-14.4	7994.6	1219.2	2656.1	1.9187	-23	3396.3
P4	-0.727	608.6	-6.7359	49.8	-9125.2	-1.1	950.9	-11.0349	78	-3583.3
P3	207.3747	0	24.3207	63.5	3441.7	364.3	0	38.14	99	1064.6
P2	-0.1481	-670.6	-0.0004	14.3	4716.7	-0.2	-1047.7	-0.0004	23	2200.9
P1	0	0	0	2524.8	-9626.9	0	0	0	13806	-3933

1	2	3	4	5	1	2	3	4	5	1	2
358.96	0	-0.8067	-95.4	1489.6	159.79	0	-0.403	-47.7	901.5	86.9021	-13.8125
-0.3132	-2609.7	-1.526	-9	-1928.5	-0.1725	-	-0.7476	-4.5	-1156.6	-0.1108	-
185.92	0	5.3885	28.1	944.9	72.791	0	2.6855	14	553.9	34.4909	0
185.2025	2955.5	-3.4809	36.5	755.9	72.3585	1444.3	-1.7399	18.2	464.4	34.2047	841.409
0.1666	0	-0.66	9	-1876.6	0.0669	0	-0.3224	4.5	-1132.2	0.0326	0
375.87	0	5.0964	-94.1	1615.5	167.1961	0	2.545	-47.1	960.2	90.8906	0
-0.669	0	-5.3086	-9	-219.4	-0.3515	0	-2.6538	-4.5	-120.1	-0.2185	0
0.6793	2146.2	-0.2827	29.4	-1358.6	0.3227	1073.1	-0.1374	14.7	-826.9	0.1859	643.847
470.7	1874.8	3.3241	37.8	1.8917	214.02	937.4	1.6613	18.9	1132.9	118.7332	562.444
-0.6957	-1672.1	-5.8315	9	-1138.7	-0.3652	-836	-2.8582	4.5	-674.2	-0.2269	-501.6
1.1	-2334.5	3.0759	-92.6	-503.3	0.5325	-	2.3317	-46.3	-316.5	0.3116	-700.35
499.3	1062.4	1.0634	-9	1723.3	202.702	531.2	10.4725	-4.5	1039.2	111.6787	318.7327
-0.4727	380.4	-0.0742	31.1	-1775.5	-0.2541	190.2	-0.0367	15.6	-1061.1	-0.1605	114.116
106.31	0	12.067	39.8	476.5	30.6566	0	5.9292	19.8	272.2	8.2131	0
-0.134	-419.1	-0.0004	9	1149.7	-0.105	-209.5	-0.0004	4.5	701.3	-0.079	-125.7
0	0	0	5523.3	-1974.3	0	0	0	2761.8	-1185.9	0	0

3	4	5	1	2	3	4	5	1	2	3
-0.2418	2.86E+01	32.6272	-0.0483	-13.8125	-0.2418	-28.6	-38.1322	18.5161	11012	-8.8744
-0.4341	-2.70E+00	102.94	-0.175	-827	-0.434	-27	-17.1254	-0.3044	8843	-28.055
1.6036	8.40E+00	30.4002	36.8288	0	1.6036	8.4	60.0064	1.6408	12175	60.3019
-1.0438	1.09E+01	-84.3393	38.5323	841.4094	-1.0438	10.9	-59.3677	-0.415	10675	-37.8619
-0.1867	2.70E+00	-82.1	0.1893	0	0.187	2.7	15.6546	0.2616	9844	-11.4064
1.524	-2.83E+01	32.6273	-0.0429	0	1.524	-28.3	39.3768	21.7399	13002	54.3972
-1.5886	-2.70E+00	102.94	-0.17	0	-1.5886	-2.7	-65.8865	-0.7421	10415	-57.4566
-0.0634	8.80E+00	30.399	43.9968	643.848	-0.0634	8.8	44.5885	0.8667	11173	-2.4515
6.2623	1.13E+01	-84.3393	46.006	562.444	6.2623	11.3	9.0061	126.2338	13422	35.3118
0.3273	2.70E+00	-82.045	0.1955	-501.61	0.3273	2.7	-56.0994	-0.7817	10436	-62.976
1.3247	-2.78E+01	32.6272	-0.0365	-700.352	1.3247	-27.8	62.5081	1.3615	12643	34.9377
6.2078	-2.70E+00	102.9351	-0.1633	318.7327	6.2078	-270	-23.5514	91.4455	13476	6.4733
-0.0217	9.30E+00	30.4001	52.719	114.1165	-0.0217	9.3	-32.4327	-0.5238	10879	-42.0836
3.4775	1.19E+01	-84.3393	55.127	0	3.4775	119	64.9404	1.6238	14040	122.8565
-0.0004	3.00E+00	-82.045	0.2031	-125.73	-0.0004	273.3	-50.3396	-0.0754	13283	-0.0002
0	1.67E+03	33.755	0	0	0	1657.1	-0.7446	0	-946	0

4	5	1	2	3	4	5	1	2	3	4
6.80E-08	-12.8394	0.0083	2365.1	-1.7479	-6.80E-08	-12.7923	0.0011	1425.7	-1.0434	-6.80E-08
-3.66E-08	-5.5604	-0.0764	1495.3	-4.3436	-3.66E-08	-5.6151	-0.0497	874.8	-2.5009	-3.66E-08
-3.59E-07	19.9424	0.3126	2396.9	11.7741	-3.59E-07	19.9647	0.1837	1406.3	7.0827	-3.59E-07
2.55E-07	-19.8605	-0.0986	2226.4	-7.508	2.55E-07	-19.8345	-0.063	1342.9	-4.5003	2.55E-06
-1.50E-07	5.3685	0.0366	1653.1	-1.5313	-1.50E-07	5.3136	0.0181	56.4	-0.9726	-1.50E-07
4.53E-08	13.0042	0.2754	2667.2	11.0248	4.53E-08	13.0489	0.1614	1577.69	6.6229	4.53E-08
-1.91E-07	-21.9585	-0.1642	2062.2	-11.4404	-1.91E-07	-21.9598	-0.1024	1238	-6.8608	-1.91E-07
8.63E-08	14.9795	0.1575	1935.7	-0.7011	8.63E-08	14.9371	0.0906	1116	-0.4259	8.63E-08
1.83E-08	2.8495	5.4066	2820.1	7.1622	1.83E-08	2.9056	0.1083	1676.9	4.3031	1.83E-08
1.27E-07	-18.6222	-0.1723	1938.6	-12.5631	1.27E-07	-18.6507	-0.1073	1147.4	-7.5356	1.27E-07
-2.27E-08	20.8892	0.2563	2297.6	7.3462	-2.27E-08	20.8698	0.1499	1319.1	4.4175	-2.27E-08
4.31E-08	-7.996	0.0693	2837	1.3907	4.31E-08	-7.9425	0.0377	1673.1	0.8408	4.31E-08
-6.36E-08	-10.6785	-0.1209	1887.4	-8.4724	-6.36E-08	-10.7272	-0.0764	1061.4	-4.1425	-6.36E-08
2.16E-08	21.6238	0.3086	2587.2	28.382	2.16E-08	21.6327	0.1812	1427.5	16.9908	2.16E-08
-1.08E-08	-16.8828	-0.0502	2435	-0.0002	-1.08E-08	-16.8453	-0.0378	1357.6	-0.0002	-1.08E-08
0	-0.094	0	-478.2	0	0	-1.51E-01	0	-300.2	0	0

5	1	2	3	4	5	1	2	3	4	5
-12.7333	-0.0007	945.8	-0.6938	-6.80E-08	-63.5206	678.8184	1.43E+04	-0.6283	-3.40E-07	-63.8843
-5.6833	-0.0346	570.2	-1.639	-3.66E-08	-28.4991	-0.567	5.81E+03	-0.7358	-1.83E-07	-28.0837
19.9921	0.121	918.5	4.7264	-3.59E-07	100.014	309.1715	1.34E+04	3.7431	-1.80E-06	99.847
-19.8017	-0.0435	890.2	-2.9987	2.55E-07	-98.9486	307.3426	1.27E+04	-2.6646	1.27E-06	-99.149
5.244	0.0106	618.4	-0.6539	-1.50E-07	26.0906	0.4511	6.53E+03	-0.3105	-7.49E-07	26.5169
13.1045	0.1061	1033.6	4.4174	4.53E-08	65.5917	710.7748	1.56E+04	3.7581	2.27E-07	65.2528
-21.9615	-0.0697	816	-4.5726	-1.91E-07	- [illegible]	-1.3268	1.08E+04	-4.0589	-9.54E-07	- [illegible]
14.8826	0.0589	716.9	-0.2846	8.63E-08	74.2798	1.5387	8.37E+03	-0.1277	4.31E-07	74.6051
2.9759	0.0708	1096.9	2.8704	1.83E-08	14.9824	910.103	1.68E+04	2.4933	9.20E-08	14.5581
-18.6859	-0.073	742.6	-5.0229	1.27E-07	-93.4955	-1.3855	8.82E+03	-3.5265	6.36E-07	-93.2792
20.8451	0.0984	836.2	2.9427	-2.27E-08	104.1811	2.4309	1.08E+04	1.0649	-1.13E-07	104.3319
-7.8765	0.0236	1076.8	1.5193	4.31E-08	-39.2501	862.2454	1.62E+04	13.6891	2.16E-07	-39.656
-10.7879	-0.0524	653.5	-0.0682	-6.36E-08	-54.0007	-0.9142	6.91E+03	-0.0455	-3.18E-07	-53.6302
21.6428	0.1194	879.4	11.2935	2.16E-08	108.24	131.0846	1.22E+04	6.6611	1.08E-07	108.1736
-16.7977	-0.028	854.5	-0.0002	-1.08E-08	-83.886	-0.0793	1.26E+04	-0.0002	-5.40E-08	-84.1757
-0.2216	0	-203.3	0	0	-1.241	0	-4.53E+03	0	0	-0.8103

1	2	3	4	5	1	2	3	4	5
1.08E+0	2.24E+04	-9.84E-01	-3.40E-07	-6.13E+01	1.91E+02	1.99E+02	-1.95E-01	-3.40E-07	-5.91E+01
-9.00E-	9.42E+03	-1.15E+00	-1.83E-07	-3.10E+01	-1.94E-01	1.48E+03	-2.24E-01	-1.83E-07	-3.34E+01
5.02E+0	2.15E+04	5.84E+00	-1.80E-06	1.01E+02	7.51E+01	3.76E+03	1.18E+00	-1.80E-06	1.02E+02
5.00E+0	2.00E+04	-4.17E+00	1.27E-06	-9.77E+01	7.45E+01	3.66E+03	6.65E-01	1.27E-06	-9.63E+01
7.00E+0	1.07E+04	-4.84E-01	-7.49E-07	2.35E+01	1.23E-01	1.59E+03	-6.13E-02	-7.49E-07	2.10E+01
1.13E+0	2.50E+04	5.87E+00	2.27E-07	6.76E+01	2.00E+02	4.40E+03	3.61E+00	2.27E-07	6.95E+01
-	1.70E+04	-6.34E+00	-9.54E-07	-1.10E+02	-4.32E-01	2.94E+03	9.02E-01	-9.54E-07	-1.10E+02
2.40E+0	1.39E+04	-2.00E-01	4.31E-07	7.22E+01	4.63E-01	2.02E+03	-3.81E-02	4.31E-07	7.02E+01
1.44E+0	2.67E+04	3.89E+00	9.20E-08	1.76E+01	2.62E+02	4.66E+03	4.29E+00	9.20E-08	2.00E+01
-	1.43E+04	-5.53E+00	6.36E-07	-9.47E+01	-4.51E-01	2.81E+03	-3.69E-02	6.36E-07	-9.58E+01
3.80E+0	1.80E+04	1.62E+00	-1.13E-07	1.03E+02	7.41E-01	2.63E+03	3.32E-01	-1.13E-07	1.02E+02
1.37E+0	2.62E+04	2.23E+00	-2.16E-07	-3.67E+01	2.46E+02	4.41E+03	4.22E+00	2.16E-07	-3.43E+01
-	1.21E+04	-7.12E-02	-3.18E-07	-5.62E+01	-3.05E-01	1.50E+03	-1.40E-02	-3.18E-07	-5.83E+01
2.27E+0	2.09E+04	1.05E+01	1.08E-07	1.09E+02	1.69E+01	3.15E+03	2.01E+00	1.08E-07	1.09E+02
-	2.06E+04	-2.00E-04	-5.40E-08	-8.21E+01	-6.90E-02	3.58E+03	-2.00E-04	-5.40E-08	-8.02E+01
0	-6.98E+03	0	0	-3.8658	0	-	0	0	-6.3924

1	2	3	4	5	1	2	3	4	5
1.05E+02	2.40E+03	2.01E+00	-3.40E-07	-5.62E+01	6.36E+01	1.54E+03	1.32E+04	-3.40E-07	-4.27E+01
-1.24E-01	7.59E+02	-4.89E-02	-1.83E-07	-3.62E+01	-8.81E-02	4.43E+02	-3.18E-02	-1.83E-07	-1.86E+01
3.50E+01	2.12E+03	1.35E+00	-1.80E-06	1.03E+02	1.69E+01	1.34E+03	8.76E-01	-1.80E-06	6.65E+01
3.46E+01	2.07E+03	1.35E+00	1.27E-06	-9.45E+01	1.66E+01	1.31E+03	8.78E-01	1.27E-06	-6.62E+01
6.59E-02	8.06E+02	-3.60E-02	-7.49E-07	1.78E+01	3.86E-02	4.70E+02	-2.34E-02	-7.49E-07	1.78E+01
1.10E+02	2.49E+03	2.13E+00	2.27E-07	7.17E+01	6.65E+01	1.59E+03	1.40E+00	2.27E-07	4.34E+01
-2.68E-01	1.62E+03	5.08E-01	-9.54E-07	-1.09E+02	-1.84E-01	1.02E+03	3.13E-01	-9.54E-07	-7.32E+01
2.69E-01	1.05E+03	-2.23E-02	4.31E-07	6.75E+01	1.74E-01	6.35E+02	-1.45E-02	4.31E-07	4.99E+01
1.46E+02	2.64E+03	2.54E+00	9.20E-08	2.30E+01	9.08E+01	1.70E+03	1.67E+00	9.20E-08	9.55E+00
-2.79E-01	1.17E+03	-2.18E-02	6.36E-07	-9.71E+01	-1.91E-01	7.20E+02	-1.42E-02	6.36E-07	-6.21E+01
4.36E-01	1.43E+03	1.63E-01	-1.13E-07	1.01E+02	2.85E-01	8.91E+02	8.00E-02	-1.13E-07	6.96E+01
1.37E+02	2.52E+03	2.49E+00	2.16E-07	-3.12E+01	8.42E+01	1.63E+03	1.63E+00	2.16E-07	-2.66E+01
-1.92E-01	7.89E+02	-8.30E-03	-3.18E-07	-6.08E+01	-1.33E-01	4.84E+02	-5.40E-03	-3.18E-07	-3.56E+01
5.25E-01	1.79E+03	1.17E+00	1.08E-07	1.09E+02	3.44E-01	1.16E+03	7.49E-01	1.08E-07	7.21E+01
-6.09E-02	2.10E+03	-2.00E-04	-5.40E-08	-7.78E+01	-5.26E-02	1.38E+03	-2.00E-04	-5.40E-08	-5.62E+01
0	-8.59E+02	0	0	-9.5131	0	-5.73E+02	0	0	-0.3842

1	2	3	4	5	1	2	3	4	5
-4.27E+01	6.79E+02	1.43E+04	-1.41E+00	-2.27E-07	-4.28E+01	1.08E+03	2.19E+04	-2.21E+00	-2.27E-07
-1.86E+01	-5.67E-01	6.19E+03	-3.19E+00	-1.22E-07	-1.85E+01	-9.00E-01	1.01E+04	-3.34E+00	-1.22E-07
6.65E+01	3.09E+02	1.40E+04	8.99E+00	-1.20E-06	6.65E+01	5.02E+02	2.25E+04	1.40E+01	-1.20E-06
-6.62E+01	3.07E+02	1.27E+04	-6.09E+00	8.48E-07	-6.62E+01	5.00E+02	1.94E+04	-9.39E+00	8.48E-07
1.78E+01	4.51E-01	7.16E+03	-9.47E-01	-5.00E-07	1.80E+01	7.00E-01	1.19E+04	-1.45E+00	-5.00E-07
4.34E+01	7.11E+02	1.62E+04	8.67E+00	1.51E-07	4.33E+01	1.13E+03	2.55E+04	1.35E+01	1.51E-07
-7.32E+01	-1.33E+00	1.08E+04	-9.14E+00	-6.36E-07	-7.32E+01	-2.10E+00	1.67E+04	-1.43E+01	-6.36E-07
4.99E+01	1.54E+00	9.25E+03	-4.03E-01	2.88E-07	5.00E+01	2.40E+00	1.53E+04	-6.22E-01	2.88E-07
9.55E+00	9.10E+02	1.72E+04	5.68E+00	6.10E-08	9.43E+00	1.44E+03	2.69E+04	8.87E+00	6.10E-08
-6.21E+01	-1.39E+00	9.27E+03	-1.09E+01	4.24E-07	-6.20E+01	-2.10E+00	1.45E+04	-1.57E+01	4.24E-07
6.96E+01	2.43E+00	1.21E+04	4.46E+00	-7.60E-08	6.97E+01	3.80E+00	1.98E+04	6.93E+00	-7.60E-08
-2.66E+01	8.62E+02	1.70E+04	1.11E+00	1.44E-07	-2.67E+01	1.37E+03	2.65E+04	1.71E+00	1.44E-07
-3.56E+01	-9.14E-01	8.28E+03	-4.17E+00	-2.12E-07	-3.55E+01	-1.40E+00	1.36E+04	-7.36E+00	-2.12E-07
7.21E+01	1.31E+02	1.43E+04	1.89E+01	7.20E-08	7.21E+01	2.27E+02	2.39E+04	2.96E+01	7.20E-08
-5.62E+01	-7.93E-02	1.38E+04	-2.00E-04	-3.60E-08	-5.63E+01	-1.00E-01	2.26E+04	-2.00E-04	-3.60E-08
-0.3842	0	-4.36E+03	0	0	-0.2564	0	-6.40E+03	0	0

1	2	3	4	5	1	2	3	4
-4.21E+01	1.91E+02	4.42E+03	-4.36E-01	-2.27E-07	-4.14E+01	1.05E+02	2.58E+03	-2.61E-01
-1.93E+01	-1.94E-01	1.74E+03	-6.96E-01	-1.22E-07	-2.00E+01	-1.25E-01	9.55E+02	-4.13E-01
6.68E+01	7.51E+01	4.09E+03	2.82E+00	-1.20E-06	6.71E+01	3.50E+01	2.34E+03	1.69E+00
-6.58E+01	7.45E+01	3.93E+03	-1.87E+00	8.48E-07	-6.54E+01	3.46E+01	2.27E+03	-1.12E+00
1.71E+01	1.23E-01	1.93E+03	-2.98E-01	-5.00E-07	1.63E+01	6.59E-02	1.04E+03	-1.77E-01
4.40E+01	2.00E+02	4.78E+03	2.71E+00	1.51E-07	4.46E+01	1.10E+02	2.73E+03	1.63E+00
-7.32E+01	-4.32E-01	3.28E+03	-2.86E+00	-6.36E-07	-7.32E+01	-2.68E-01	1.85E+03	-1.71E+00
4.93E+01	4.63E-01	2.46E+03	-1.26E-01	2.88E-07	4.87E+01	2.69E-01	1.31E+03	-6.33E-02
1.03E+01	2.62E+02	5.09E+03	1.78E+00	6.10E-08	1.11E+01	1.46E+02	2.89E+03	6.48E+00
-6.25E+01	-4.51E-01	2.61E+03	-3.07E+00	4.24E-07	-6.29E+01	-2.79E-01	1.41E+03	8.81E-02
6.93E+01	7.41E-01	3.14E+03	1.99E+00	-7.60E-08	6.91E+01	4.36E-01	1.68E+03	1.16E+00
-2.58E+01	2.46E+02	4.86E+03	1.07E+01	1.44E-07	-2.51E+01	1.37E+02	2.74E+03	6.41E+00
-3.63E+01	-3.05E-01	1.94E+03	-3.69E-02	-2.12E-07	-3.69E+01	-1.92E-01	9.87E+02	-2.20E-02
7.22E+01	1.69E+01	3.57E+03	5.84E+00	7.20E-08	7.23E+01	5.25E-01	1.97E+03	3.46E+00
-5.57E+01	-6.90E-02	3.79E+03	-2.00E-04	-3.60E-08	-5.52E+01	-6.09E-02	2.18E+03	-2.00E-04
-1.1639	0	-1.42E+03	0	0	-1.9185	0	-8.57E+02	0

5	1	2	3	4	5
-2.27E-07	-4.06E+01	6.36E+01	1.66E+03	-1.74E-01	-2.27E-07
-1.22E-07	-2.09E+01	-8.81E-02	5.74E+02	-2.71E-01	-1.22E-07
-1.20E-06	6.74E+01	1.69E+01	1.49E+03	1.13E+00	-1.20E-06
8.48E-07	-6.50E+01	1.66E+01	1.45E+03	-6.27E-01	8.48E-07
-5.00E-07	1.54E+01	3.86E-02	6.14E+02	-6.59E-02	-5.00E-07
1.51E-07	4.53E+01	6.65E+01	1.74E+03	3.66E+00	1.51E-07
-6.36E-07	-7.32E+01	-1.84E-01	1.16E+03	1.22E+00	-6.36E-07
2.88E-07	4.79E+01	1.74E-01	7.85E+02	-4.17E-02	2.88E-07
6.10E-08	1.20E+01	9.08E+01	1.84E+03	4.29E+00	6.10E-08
4.24E-07	-6.33E+01	-1.92E-01	8.52E+02	2.99E-02	4.24E-07
-7.60E-08	6.87E+01	2.85E-01	1.03E+03	7.41E-01	-7.60E-08
1.44E-07	-2.42E+01	8.42E+01	1.75E+03	4.24E+00	1.44E-07
-2.12E-07	-3.77E+01	-1.33E-01	5.83E+02	-1.46E-02	-2.12E-07
7.20E-08	7.24E+01	3.44E-01	1.24E+03	2.27E+00	7.20E-08
-3.60E-08	-5.45E+01	-5.26E-02	1.42E+03	-2.00E-04	-3.60E-08
0	-2.8589	0	-5.72E+02	0.00E+00	0

جامعة النجاح الوطنية

كلية الدراسات العليا

تطوير نظام خبير لتشخيص أعطال محولات القوى الكهربائية باستخدام تقنية الغابة العشوائية

إعداد

غازي عرار

إشراف

د. تامر خطيب

قدمت هذه الأطروحة استكمالا لمتطلبات الحصول على درجة الماجستير في هندسة القوى الكهربائية، بكلية الدراسات العليا، في جامعة النجاح الوطنية، نابلس، فلسطين.

2020

تطوير نظام خبير لتشخيص أعطال محولات القوى الكهربائية باستخدام تقنية الغابة العشوائية

إعداد

غازي عرار

إشراف

د. تامر خطيب

الملخص

هذه الرسالة سوف تساهم في مجال أنظمة الحماية الكهربائية، حيث أن أنظمة الحماية التقليدية فشلت في تجاوز محدداتها، ظهرت الحاجة لإيجاد تقنيات جديدة لحل هذه المعضلة. في هذه الرسالة تقنيات المجموع استخدمت لحل هذه المشكلة. وبالتالي نظام حماية تفاضلي تم بناؤه من أجل توفير عنصر حماية: فصل، عدم فصل وعنصر مراقبة وعرض الحالة للتمييز بين حالات المحول الكهربائي المختلفة: طبيعي، التدفق المغناطيسي، فرط الاستثارة، تشبع محول التيار، عطل داخلي.

نماذج 20 محول بمواصفات حقيقية مع الخمس حالات تشغيلية للمحول تنتج 100 مثال من أجل تدريب وفحص التقنية المقترحة باستخدام 1600 قراءة. الغابة العشوائية دربت باستخدام 100 مثال أوليات بدون معالجة، وفحصت باستخدام مقاييس داخلية: خطأ خارج الحقيبة، الهامش، مصفوفة الخلط والقيم المتطرفة. ثم تم تدريب وفحص الغابة العشوائية باستخدام مجموعة بيانات معدلة معادة التوزين. خطأ خرج الحقيبة والهمش استخدما للمقارنة بين مجموعة البيانات الأولية والمجموعة المعدلة.

استخدمت نسب تدريب الى فحص مختلفة: 80–الى–20 و 60–الى–40 للتحقق من قوة وموثوقية النموذج المستخدم. كما تم انشاء نماذج أسرع من الغابة العشوائية باستخدام نوافذ بيانات مختلفة: $\frac{3}{4}$، $\frac{1}{2}$ و$\frac{1}{4}$ دورة وكان الناتج عنصر حماية دقيق وعنصر مراقبة وعرض الحالة عالي الدقة. بالإضافة لذلك، تم مقارنة وفحص نسختين من الغابة العشوائية حسب عمق الشجر فيها: عمق كامل وعمق محدود.

تطبيق تقنية التعزيز على مجموعة البيانات الأولية والمجموعة المعدلة باستخدام نسب تدريب الى فحص مختلفة: 80–الى–20 و 60–الى–40 للتحقق من التقنية. فحصت التقنية مشروطة بعدد الأشجار الأمثل باستخدام: خطأ خارج الحقيبة، التحقق المتقاطع. بالمثل، فان أهمية المتغير مشروطة بعدد الأشجار الأمثل.

بالإضافة لما سبق، تم التعامل مع أهمية المتغيرات باستخدام تقنيتا المجموع، ولاحقا تم الوصول الى استنتاج بخصوص أهمية الاشارة عند لحظات مختلفة.

تقنيتا الغابة العشوائية والتعزيز أظهرت نتائج واعدة وقدرة على التصنيف للمشكلة المطروحة. حيث انهما تستطيعان توفير نتائج دقيقة، سريعة وذات موثوقية واعتمادية.