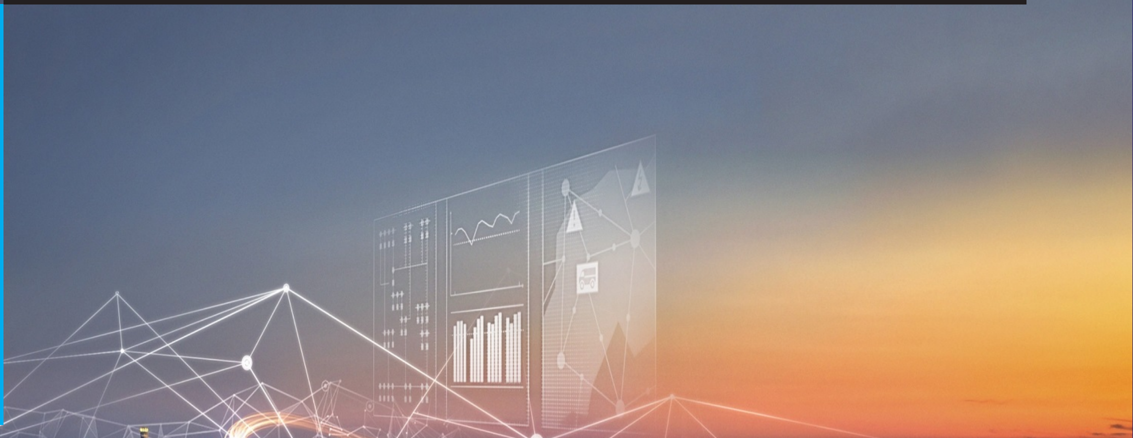
Detection and Classification of Disturbances in the Grid Using Discrete Wavelet Analysis and Machine Learning

Master Thesis

Kiran Kumar

Technische Universiteit Delft



Delft University of Technology

Challenge the future

DETECTION AND CLASSIFICATION OF DISTURBANCES IN THE GRID USING DISCRETE WAVELET ANALYSIS AND MACHINE LEARNING
Master Thesis

in partial fulfillment of the requirements for the degree of

Master of Science
in Electrical Engineering

at the Delft University of Technology,
to be defended publicly on Wednesday Oct 30, 2019 at 14:00.

by

Kiran KUMAR

Thesis Committee:

Prof. dr. Peter Palensky,	Thesis Chair, IEPG, TU Delft
Dr. Miloš Cvetković,	Daily Supervisor, IEPG, TU Delft
Dr. Mohamad Ghaffarian Niasar,	External Member, DCES, TU Delft
Rishabh Bhandia, M. Sc. ,	PhD. Candidate, IEPG, TU Delft

An electronic version of this thesis is available at http://repository.tudelft.nl/.

TU Delft Delft University of Technology

Acknowledgements

The foremost, expression of my gratitude, is extended to the best pair of professors, one could beseech, as their guiding light, Prof. Dr. Peter Palensky and Dr. Miloš Cvetković. I, sincerely thank you, for your continuous support, invaluable insights and the abundant patience, exhibited during our numerous feedback meetings, without which, this thesis wouldn't be a reality.

Analogous to the electromotive forces, pushing the charged particles in an electric circuit, through the various hurdles of resistances, potential wells and air-gaps, always trying to complete the circuit, my PhD. Supervisor, Rishabh Bhandia, has been my motive force, through the entire process of developing this thesis. Thank you, for inspiring and constantly supporting me.

I am also grateful to Dr. Mohamad Ghaffarian Niasar, from the DCES group at TU Delft, for providing his valuable time, to be part of my thesis committee and evaluating my work. I thank you, for guiding me through the various steps of my M.Sc. program.

I gratefully acknowledge the collective help, that I have received from the various members of the IEPG group, in the form of resource and knowledge sharing. The excellent administrative support from the IEPG secretary, Ellen, is worthy of appreciation. My friends and colleagues, Mohit, Devendra and Vetrivel, have been of great help in helping me maintain my sanity, while also helping me document my work. Thank you! I thank all my colleagues and friends, of the past two years, for being part of this wonderful journey, not limited to academics.

Finally, I thank my wife, Anns, for tolerating me and supporting me, all the way from Kerala, India, while also helping me with my queries in protection of the electrical power system.

Kiran Kumar
Delft, October 2019

Contents

Acknowledgements v

List of Figures ix

List of Tables xi

Abstract xiii

1 Introduction 1
- 1.1 Background and Motivation . . . 1
- 1.2 Research Questions . . . 5
- 1.3 Methodology . . . 5
- 1.4 Thesis Contributions . . . 5
- 1.5 Outline . . . 6
- References . . . 6

2 DWT and Selection of the Mother Wavelet 9
- 2.1 Discrete Wavelet Transforms (DWT) . . . 9
- 2.2 Comparison of DWT Using Various Wavelet Families . . . 11
- References . . . 13

3 Simulation of Disturbances in PSCAD 17
- 3.1 Disturbances . . . 17
- 3.2 Disturbances Simulated in PSCAD . . . 18
- References . . . 22

4 Feature Extraction and Training the ML Module 27
- 4.1 Feature Extraction . . . 27
- 4.2 ML Module Training . . . 30
- References . . . 35

5 Results and Conclusion 37
- 5.1 Results . . . 37
  - 5.1.1 Decision Tree v.s. Gradient Boosting . . . 37
  - 5.1.2 Comparison of features . . . 41
  - 5.1.3 Implementation of ML . . . 42
- 5.2 Conclusion . . . 44

List of Figures

1.1 System Frequency of National Grid, UK (Time in GMT) 2

2.1 Examples of mother wavelets 10
2.2 DWT representation as tree of high and low pass filters. 10
2.3 Input Test Signal for Wavelet Analysis. 12
2.4 Analysis of Test Signal using Daubechies Wavelet Family (db). 13
2.5 Comparison of decomposition of test signal during normal behaviour. . . 14
2.6 Comparison of decomposition of test signal during high frequency disturbance. 15
2.7 Comparison of decomposition of test signal during low frequency disturbance. 16

3.1 IEEE 9 Bus System 19
3.2 IEEE 14 Bus System 20
3.3 Circuit Diagram of the HIF model implemented 21
3.4 V-I Characteristics for HIF in 30 Bus System (33 kV) 22
3.5 V-I Characteristics for HIF in 34 Bus System (24.9 kV) 23
3.6 Current waveform for HIF in 34 Bus System (24.9 kV) at Bus 848 on Phase A 23
3.7 Current waveform for Grounded faults in 9 Bus system. 24
3.8 Voltage waveform for Grounded faults in 9 Bus system. 24

4.1 Shannon Entropy of DWT Details 29
4.2 Shannon Entropy of DWT Details (Level 1 to 4) 30
4.3 Example of DT predicting the chances of survival of passengers on Titanic, based on sex and age [5]. 31
4.4 Flow Chart for training of ML Module 33
4.5 Flow Chart for Classification Program using trained ML Module 34

5.1 Prediction of Events in 34 Bus system using Gradient Boosting (GB) Classifier in voltage signals 38
5.2 Prediction of Events in 34 Bus system using DT Classifier in voltage signals 38
5.3 Prediction of Events in 34 Bus system using Gradient Boosting (GB) Classifier in current signals 39
5.4 Prediction of Events in 34 Bus system using DT Classifier in current signals 39
5.5 Prediction of Events in 30 Bus system using Gradient Boosting (GB) Classifier in voltage signals 39
5.6 Prediction of Events in 30 Bus system using DT Classifier in voltage signals 40
5.7 Prediction of Events in 30 Bus system using Gradient Boosting (GB) Classifier in current signals 40

5.8 Prediction of Events in 30 Bus system using DT Classifier in current signals 40
5.9 Accuracy of ML module trained using 34 Bus system signals over 25 runs . 41
5.10 Accuracy of ML module trained using 30 Bus system signals over 25 runs . 42
5.11 Comparison of accuracy while removing features 43

List of Tables

3.1 Simulation Summary 21

5.1 Comparison of accuracy of GB and DT classifiers, pre and post normalisation 38

ABSTRACT

The growing energy demands and the global shift towards Renewable Energy resources (RES), has resulted in rapid growth of the electrical power grid. With the integration of further RES into the power grid, the need for a resilient and smart power grid has never been greater. For meeting this necessity, the electrical power grid is undergoing a massive overhaul in terms of infrastructure, catalysed by the advent of Information and Communication Technology (ICT). Integration of ICT layers onto the power grid, has furthered the use of data analytic and processing tools, in the field of electrical power systems. Utilizing the plethora of data being collected from the power grids, and the modern ICT developments, such as, in the field of Machine Learning (ML), the quest to make the grid smarter and stronger continues. However, this being the infant stages of research, in ML implemented power grid protection, there is a vast variety of problems, yet to be discovered. This thesis, explores the efficacy of ML, used in tandem with Discrete Wavelet Transforms (DWT) and trained using basic statistical features of current and voltage signals in detection of a variety of disturbances in the power grid.

The thesis can be summarised as follows:

- **Identification of signal processing tool**

  The current and voltage signals, collected from the power system, can be utilised directly, for analysis of the characteristics of disturbances in it. However, by utilising signal processing tools, a greater insight into the signal can be obtained. DWT is one such tool, used for decomposing the signal, based on frequency, while also retaining the time domain data of the signal. The decomposition utilises special functions, called mother wavelets, and an infinite number these can be developed. A study, comparing a variety of family of these mother wavelets, yielded db11, of Daubechies family, suited for detection of disturbances in electrical signals.

- **Disturbance signature generation**

  Disturbance signatures for commonly occurring events in a power system, along with rarer and harder to detect events such as HIF were simulated using PSCAD for multiple system voltages. Further, data for equipment failure related incipient faults were obtained from the PES work group of IEEE. These disturbance signals were decomposed using DWT, for training the ML module.

- **ML implementation for detection and classification of disturbances**

  Using the signature of disturbances generated and obtained, supervised ML modules were trained to detect and classify the disturbances by analysing the basic statistical features, mean, standard deviation, variance, RMS, number of zero and mean crossings, Shannon entropy and energy, extracted from the DWT decomposition of the current and voltage signals. A comparison of the accuracy of, two of the common classification models, Decision Tree and Gradient Boosting, yielded Gradient boosting as more accurate at classification of the disturbance. A comparative analysis of the statistical features, also provided an insight into their contribution towards the accuracy of ML.

1

Introduction

The electric grid has been the corner stone for modern technology, by being the primary medium for transfer of energy across the globe. With the push for cleaner sources of energy, due to the increasing awareness of climate change, the dependency on electricity has increased substantially. This in turn has increased the burden on the electric grid, due to both increased energy being pumped in and increased energy being drawn from the grid.

With the electrical grids being interconnected across continents, there are chances of unmitigated disturbances in the grid, for e.g. line to line faults, line to ground faults, high impedance faults (HIFs), equipment failures etc. resulting in cascading failures spanning vast geographical expanses. As a result it is very crucial to mitigate such disturbances as early as possible, however, the first step for mitigation is to detect the disturbances and then to classify them.

With the advancement in computing and processing techniques, Machine Learning (ML) has emerged as the stepping stone to real-world artificial intelligence applications. Implementation of smart algorithms combined with ML techniques can aid in the detection and classification of disturbances, named in the previous paragraph, onto which further mitigation techniques can be applied, to limit the effect of these disturbances on the grid. This in turn helps in achieving smarter and resilient grids [1] [2].

1.1. Background and Motivation

The interconnection of electric grids makes it vulnerable to cascading failures with an origin that can be located at a geographically far away location. These failures can result in disruption of power supply or failure to meet basic power quality requirements, resulting in large penalties to the Distribution System Operator (DSO), loss of revenue to the industries contracted to this DSO, and disruption of public utility services like transportation, communication networks, domestic power supply etc. Hence, it is crucial for the DSO to mitigate the effects of such failures, or at least limiting the number of affected entities.

In the recent years, there have been a few such incidents, where an unmitigated cascading failure resulted in city-wide to nation-wide blackouts. One such notable blackout was the Manhattan blackout on 13^{th} of July, 2019 [3]. The Manhattan blackout left 72,000 customers of Con Edison, a major DSO in the city of New York, without power supply and affected a part of city's metro, while disrupting over 200 traffic and street lights. Although fatalities were avoided, the 5 hours of power disruption caused a tense and chaotic situation in the city; requiring rescue operations to evacuate people stuck in elevators and metros. Con Edison attributed the blackout to a burning 13kV cable, which further resulted in a cascading failure [3]. This was deemed as a critical equipment failure. This blackout has affected the DSO, as calls for revoking their license and large penalties are being raised. The DSO was penalised an amount of 18 million USDs in 2007 for a nine day blackout in 2006. Similar penalties are expected as an outcome of the Manhattan blackout of 2019 [4].

Another such blackout was the Great Britain blackout on 9^{th} of August, 2019, which was the country's largest in over a decade [5]. This blackout affected the Newcastle airport and nearly a million customers of UK's National Grid DSO. Although, the blackout only lasted for over an hour; however, the aftereffects were evident the subsequent day as well, with delayed trains and flights. The outage was caused due to disconnection of a 730 MW gas power plant in Cambridgeshire and 2x 406 MW wind farm units in Hornsea, while the demand hit 23 GW. This sudden disconnection of over 1.5 GW supply, separated by over 150 km, resulted in automated load shedding. This is part of the mitigation protocol for fall in frequency, visualised in figure 1.1. Although, the resultant blackout was the result of mitigation actions, the effects of the loss of power was widespread and has resulted in an enquiry to ascertain whether the National Grid is to be penalised for poor planning and mitigation actions [6].

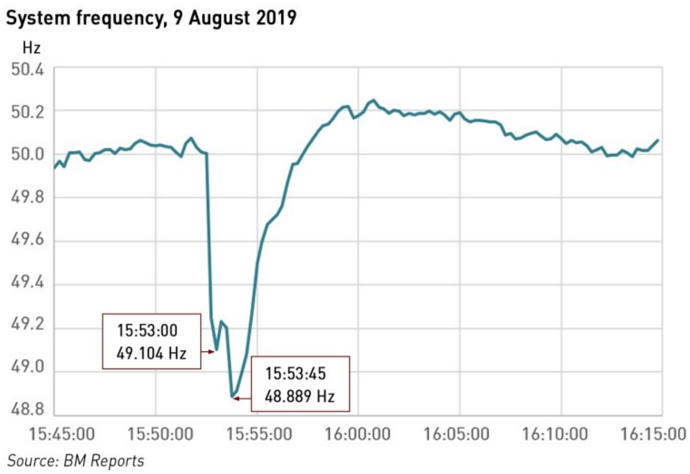


Figure 1.1: System Frequency of National Grid, UK (Time in GMT)

It is seen that power outages is a persistent problem, and can be the result of disturbances in the grid, causing a cascading failure across the system. Hence, it is of critical importance to mitigate such disturbances at the earliest to avoid widespread power outages.

For decades a wide variety of techniques have been used to detect and distinguish the disturbances in the grid. Traditional circuit breakers and relays, measuring the signals from current and voltage transformers, have been employed for protection of the grid. However, with the advent of ICT and stringent grid regulations, the need for more accurate and faster methods can be met using the plethora of computational tools developed in the recent times [2] [7].

Mathematical tools like Fourier Transforms and Fast Fourier Transforms have been used extensively in frequency analysis of disturbances in the grid. However, Wavelet Transforms (WT) have been found to have better suitability in analysis of transients in electrical power systems [8] [9]. As a result, WT analysis shall be mathematical tool used in this thesis, specifically, Discrete Wavelet Transform (DWT) analysis, as the signals analysed are discrete time signals. Further, detection and classification comprise the initial steps of mitigation of system failures, and shall be focused on, in this thesis.

The existent electrical grid is in the process of transformation into smarter grids, with layers of Information and Communication Technology (ICT) being added to an expanding and interconnected network. Smart meters, as a part of ICT layer, collects a vast amount of data to be processed and analysed [10]. This raw data, can be analysed using the various data processing tools available and can be broken down further, using data analysis tools, for drawing conclusive information. These meters also add significant processing capability to the monitoring system. Utilising the existing signal processing techniques, the measured currents and voltages can be used to detect disturbances in the grid.

Classification of these disturbances, however, forms one of the most challenging tasks. This is due to the difficulty in determination of distinguishing characteristics, of the various disturbances. ML modules can be trained to identify the distinguishing characteristics and use them to classify these disturbances. The accuracy of the classification is dependent on the selection of signal features, as well as the amount of training data provided to the ML module.

Various research has employed DWT for detecting faults in the power system. K. Rajesh and N. Yadaiah explored a technique to detect, locate and identify line to line faults and line to ground faults in single phase and three phase AC system, using Wavelet Modulus Maxima (WMM) in their research [11]. This paper depicts the clear superiority of wavelet transforms in detection of the said faults, however, it is an early work, which needs to be developed further, to encompass larger networks, with multiple voltage grades and different disturbances. T. S. Sidhu and Z. Xu utilised wavelet based analysis techniques, in detection of incipient faults in underground cables [12]. The method developed was capable of detecting and identifying sub-cycle incipient faults, while being able to distinguish it from capacitor and load switching events. However, the method was best suited for detecting single line to ground faults, even with high impedance. Another method was developed using DWT based analysis to detect HIF in distribution networks and distinguish it from solid faults by W. C. Santos et al. in their research [13]. The research was able to distinguish HIF from other faults, however, it did not focus on classifying what the detected solid fault was.

As seen above, each of the research that developed methods to detect disturbances in the power system/ grid, focused mainly on a set of disturbances that have similar characteristics. Basically, this is due to the proportional increase in complexity of classification algorithms, and subsequent decrease in distinguishable characteristics of disturbances, with respect to the increase in types of disturbances to be classified. However, ML proves to be a strong candidate for use in classification, due to the inherent capability of ML in pattern recognition.

R. A Sowah et al. deal with development of fault data collector for detection, location and classification of faults in power grid using ML [7]. However, it is limited open circuit and short circuit faults. The research showcases the superiority of Decision Tree (DT) based ML module over k-Nearest Neighbour (kNN) and Support Vector Machine (SVM), in classification problems. Ferhat et al. use DWT and ML to classify power quality events in their research [14]. The research focuses on classifying power quality events like voltage sag, voltage swell, harmonics, momentary interruptions and flicker, however, it does so detailing the feature selection and extraction process. B. Bhattacharya and A. Sinha also utilise ML for classification of faults in the power grid in their research [15]. This, again focuses only on classification of line to line and line to ground faults.

In order to move towards a smarter grid, the addition of intelligent features can be in different phases, where, each new addition moves us closer to achieving the goal of smart grids [2]. As the literature, mentioned in the above paragraphs, deal with a particular set of disturbances and methodology to classify them, this thesis shall attempt to cover a larger set of disturbances, while adapting some of the methodologies put forth. This is to move closer to developing a single system for detection and classification of every disturbances in the power grids of the future.

This thesis shall utilise DWT analysis technique to detect and ML techniques to classify the disturbances listed below:

- Disruptive
  - Line to Line (LL) Faults
  - Line to Line to Line (LLL) Faults
  - Line to Ground (LG) Faults
  - Line to Line to Ground (LLG) Faults
  - Line to Line to Line to Ground (LLLG) Faults
  - HIFs
  - Possible Equipment Failure Signature
- Non-disruptive
  - Load Switching
  - Capacitor Bank Switching

The disruptive disturbances are the ones, that are considered harmful to the system, and can end up disrupting the normal operation of the power grid. Whereas, non-disruptive disturbances like switching of load, capacitor banks, which do not result in disruption of the normal operation of the power grid.

1.2. RESEARCH QUESTIONS

THE research questions, this thesis aims to answer are as follows:

- What mother wavelet is to be utilised for detection and classification of disturbances in grid? Why?
  - Comparison of the different mother wavelets available, using test signal.
- How effective are basic statistical features, extracted from the discrete wavelet decomposition, for classifying the disturbances? How accurate are the methodologies used for training ML module?
  - Analysis of the decomposition of the signals and training the ML module using basic statistical features.
  - Analysis of popular ML training methodologies by accuracy comparison.
- How can the detection and classification module be implemented? What language/ software to use? Why?

1.3. METHODOLOGY

THE thesis methodology is as listed below:

1. **Selection of mother Wavelet**: As a variety of families of wavelets are available and an infinite amount of new ones can be made using them, it is of significant interest to select an appropriate mother wavelet for the DWT analysis of current and voltage measurements. A comprehensive comparative analysis is carried out to select the same.
2. **Simulation of disturbances**: To identify the features of the decomposition of the DWT analysis at different voltage levels, PSCAD is used to simulate various events in IEEE 9, 14, 30 and 34 bus systems. The measurements are saved as COMTRADE files.
3. **Pre-processing and training of ML module**: A set of measurements from the PSCAD simulations, are pre-processed for feature extraction and is used for training the ML module using Python language.
4. **Detection and Classification**: A set of measurements from the PSCAD simulations, is pre-processed and fed to the ML module to verify the accuracy of classification.

1.4. THESIS CONTRIBUTIONS

THE following contributions have been made through this thesis:

- Comparative analysis of DWT of electrical wave forms, using multiple mother wavelets and selection of the best suited mother wavelet.

- Python based DWT analysis, pre-processing of data and ML implementation.
- Investigation of the accuracy of detection and classification using ML, based on extracted basic statistical features.

1.5. OUTLINE

THIS thesis document follows the structure as listed below:

- **Chapter 1**: Provides an introduction to the thesis, while summarising the intended outcomes.
- **Chapter 2**: Summarises DWT, detailing the comparative study of DWT analysis using various mother wavelets and selection of the best suited wavelet.
- **Chapter 3**: Details the disturbances simulated and analysed.
- **Chapter 4**: Deals with the feature extraction, training and implementation of ML module.
- **Chapter 5**: Draws conclusions of the thesis and puts forth the future work to be undertaken.

REFERENCES

[1] A. Ipakchi and F. Albuyeh, *Grid of the future,* IEEE Power and Energy Magazine **7**, 52 (2009).

[2] H. Gharavi and R. Ghafurian, *Smart grid: The electric energy system of the future [scanning the issue],* Proceedings of the IEEE **99**, 917 (2011).

[3] P. McGeehan, *A burning 13,000-volt cable touched off manhattan blackout, con edison says,* The New York Times (2019).

[4] P. McGeehan and A. Southall, *Blackout: Con edison apologizes, but offers few clues about 'root cause',* The New York Times (2019).

[5] P. Karasz, *U.k. seeks answers after biggest power failure in a decade,* The New York Times (2019).

[6] W. Mathis and M. Carr, *London blackout blamed on drop in wind and natural-gas power,* Bloomberg (2019).

[7] R. A. Sowah, N. A. Dzabeng, A. R. Ofoli, A. Acakpovi, K. M. Koumadi, J. Ocrah, and D. Martin, *Design of power distribution network fault data collector for fault detection, location and classification using machine learning,* in *2018 IEEE 7th International Conference on Adaptive Science Technology (ICAST)* (2018) pp. 1–8.

[8] D. C. Robertson, O. I. Camps, J. S. Mayer, and W. B. Gish, *Wavelets and electromagnetic power system transients,* IEEE Transactions on Power Delivery **11**, 1050 (1996).

[9] Y. Kihira, N. Yamamura, and M. Ishida, *Proposal of evaluation method of electric power fluctuation using wavelet transform,* in *2009 International Conference on Electrical Machines and Systems* (2009) pp. 1–4.

[10] R. Morello, C. De Capua, G. Fulco, and S. C. Mukhopadhyay, *A smart power meter to monitor energy flow in smart grids: The role of advanced sensing and iot in the electric grid of the future,* IEEE Sensors Journal **17**, 7828 (2017).

[11] K. Rajesh and N. Yadaiah, *Fault identification using wavelet transform,* in *2005 IEEE/PES Transmission Distribution Conference Exposition: Asia and Pacific* (2005) pp. 1–6.

[12] T. S. Sidhu and Z. Xu, *Detection of incipient faults in distribution underground cables,* IEEE Transactions on Power Delivery **25**, 1363 (2010).

[13] W. C. Santos, F. V. Lopes, N. S. D. Brito, and B. A. Souza, *High-impedance fault identification on distribution networks,* IEEE Transactions on Power Delivery **32**, 23 (2017).

[14] F. Uçar, F. Alçin, B. Dandil, and F. Ata, *Machine learning based power quality event classification using wavelet — entropy and basic statistical features,* in *2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR)* (2016) pp. 414–419.

[15] B. Bhattacharya and A. Sinha, *Intelligent fault analysis in electrical power grids,* in *2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)* (2017) pp. 985–990.

2

DWT AND SELECTION OF THE MOTHER WAVELET

This chapter provides a basic introduction to DWT and the various wavelet families. It further, analyses comparatively the ability of the different wavelet families to be utilised in this thesis. The best suited wavelet is then selected for use throughout this thesis.

2.1. DISCRETE WAVELET TRANSFORMS (DWT)

THE basic theory of wavelets provides a decomposition of a function $x[k]$, into a family of time-frequency localised functions called wavelets. In discrete functions, discrete wavelet transforms are obtained using the mathematical equation shown below:

$$W_{\psi}[m,n] = \frac{1}{\sqrt{a_0^m}} \sum_{k=-\infty}^{+\infty} x[k]\psi\left[\frac{k-a_0^m nb_0}{a_0^m}\right] \tag{2.1}$$

Where $\psi[k]$ is called the mother wavelet, and is the analysing function used to decompose the input signal $x[k]$. The mother wavelet's only restriction is to be short and oscillatory, i.e, the average of the function must be zero and it should decay at both ends quickly. This leads to generation of infinite numbers of mother wavelets. Wavelet transforms have two parameters called scaling parameter (a) and translation parameter (b), and are functions of integers m and n, where, $a = a_0^m, b = a_0^m nb_0; a_0 > 1$ and $b_0 \neq 0$[1].

As an infinite number of mother wavelets can be generated, it becomes tedious to generate one for each use case. A family of mother wavelets have been generated and analysed for the use cases in power signals. Some of the available mother wavelets, have been shown in figure 2.1.

It is seen in the figure 2.1, that the mother wavelets are localised in time and have finite energy. Each of the different mother wavelets shown, has a different shape, and was developed to excel in specific applications. As this thesis deals with sinusoidal wave forms, with high frequency disturbances, it would be of importance to see, how each of

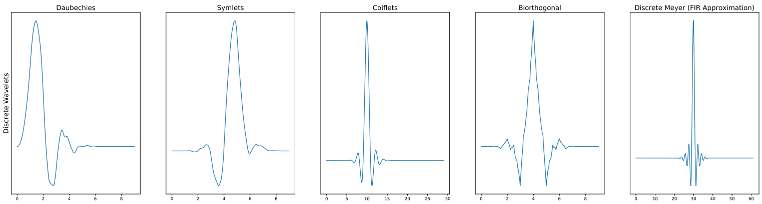


Figure 2.1: Examples of mother wavelets

the wavelet family, decomposes a test signal, representing the current or voltage waveform with disturbances.

Advantages of DWT over Fourier Transforms (FT) and Short Time Fourier Transforms (STFT) have been showcased in previous research for power systems [1] [2]. Essentially, DWT, due to the time-frequency localisation characteristics, and logarithmic frequency coverage, provides better resolution of fast transient disturbances, that are experienced in power systems.

DWT can be represented as a tree of high and low pass filters, as shown in figure 2.2. Here, the input signal is s, while the outputs d_i and a_i are the DWT decomposition of input signal s. d_i corresponds to the output of the i^{th} High Pass filter (HPF) and is called the detail, whereas, a_i corresponds to the output of the i^{th} Low Pass filter (LPF) and is called the approximation.

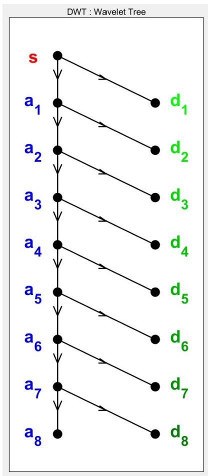


Figure 2.2: DWT representation as tree of high and low pass filters.

The signal s is split into two samples of equal length, sampled at a rate half of the sampling rate of signal s. One of the samples is passed through the LPF, obtaining the approximation at level 1, a_1 and the other sample is passed through the HPF to obtain the detail at level 1, d_1. This process is iterated again over the approximations, to obtain the approximations and details at the next level. The process iterates till the approximations cannot be decomposed further.

Thus it is seen that the number of details and approximations, that can be extracted or the max level to which a signal can be decomposed using DWT is determined by the sampling rate of the input signal s. It is to be observed that, the range of frequencies in the HPF at each level decreases as the level increases, thus d_1 contains higher frequency components as compared to d_2, which in turn contains higher frequency components as compared to d_3 [1]. Thus, the high frequency transients will be filtered out in the initial levels of decomposition, depending on the sampling rate of the input signal, and shall be visible in the details.

2

It is this property of the DWT, that shall be used to evaluate the suitability of the mother wavelets available, for the disturbances in this thesis.

2.2. Comparison of DWT Using Various Wavelet Families

An infinite number of mother wavelets can be generated, as observed in previous section. However, in this thesis, the family of mother wavelets, to be analysed, shall be limited to the ones available in two of the popular tools used in power system analysis, MATLAB and Python. The wavelet toolbox available in MATLAB, is utilised for this study, due to the GUI advantage of MATLAB over Python. This makes it easier to observe the signal decomposition using DWT.

For this comparative study, a signal, representing the current or voltage waveform in a power system, is generated. A sine wave of 50 Hz was generated as the base signal, with an amplitude of 400. The signal was generated for the time period of a second, i.e., 50 cycles. To emulate the noises and harmonics present in the grid's electrical parameters, a noise in the form of random number was added to the base signal. Further, to emulate a high frequency disturbance, a 550 Hz cosine wave with 200 amplitude, is added to the base signal from time period 0.30 seconds to 0.36 seconds, i.e., 3 cycles of the base signal. A low frequency disturbance is added in the form of a 25 Hz sine wave of 200 amplitude from time period 0.60 seconds to 0.66 seconds, i.e. 3 cycles of the base signal. This signal is referred to as the test signal and is shown in figure 2.3.

This test signal was decomposed using the following mother wavelets:

- Daubechies Family (db)
- Symlets Family (sym)
- Coiflets Family (coif)
- BiorSplines Family (bior)
- DMeyer Family (dmey)
- Fejer-Korovkin Family (fk)

Some of the families of mother wavelets listed, have a set of mother wavelets in their family. For e.g. in Daubechies family, there are mother wavelets named db1, db2, db3

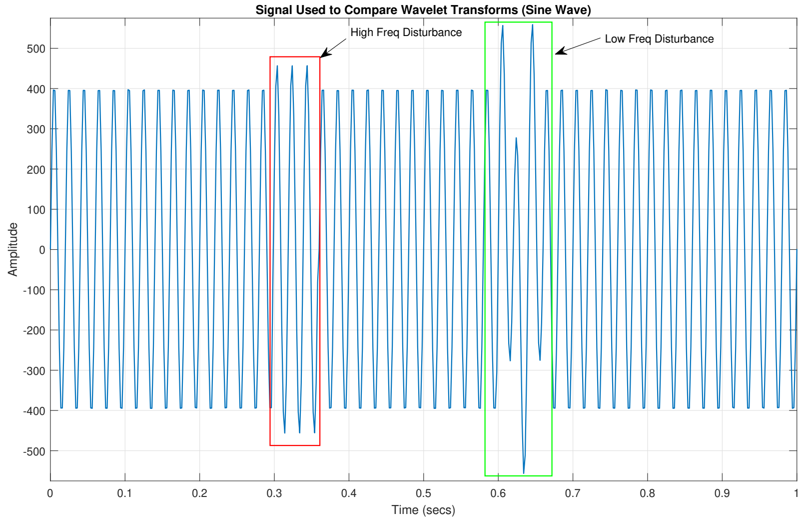


Figure 2.3: Input Test Signal for Wavelet Analysis.

etc. Of this db1 is the most basic and primitive mother wavelet, also called Haar wavelet, and is omitted in this analysis, as it yields no structured information from the test signal. For the test signal, that was sampled at 500Hz, i.e. 500 samples are available for 1 second, the maximum level of decomposition possible is 8. Figure 2.4, shows details at level 1, d_1 for the various mother wavelets belonging to the Daubechies family.

From the figure, it is noticeable that for mother wavelets db7 onward, the start and end of the low frequency and high frequency disturbances is clearly seen in the detail. It has been known that a minimum of db4 is required for analysis of transients in power system signals [3], however, it is observed here that the db4 fails to detect the start and end of the disturbances adequately.

A similar visual analysis of the decomposition of the test signal using the different families of mother wavelets yielded a result as follows:

- Symlets Family wavelets displayed similar properties as that of Daubechies.
- From Coiflets Family wavelets coif4 and coif5 showcased the decomposition same as that of db11.
- BiorSplines Family wavelets provided decomposition with higher noise during the cycles in test signal with no disturbance, and hence was found to be unsuitable.
- DMeyer Family consists of a single wavelet and the decomposition of the test signal yielded detail with large oscillations during the period of high frequency disturbance, which is unsuitable for the purpose of this thesis.

As it was observed that the decomposition of daubechies family and symlets family of wavelets are similar and db7/sym7 onward provide insight into the start and end of the

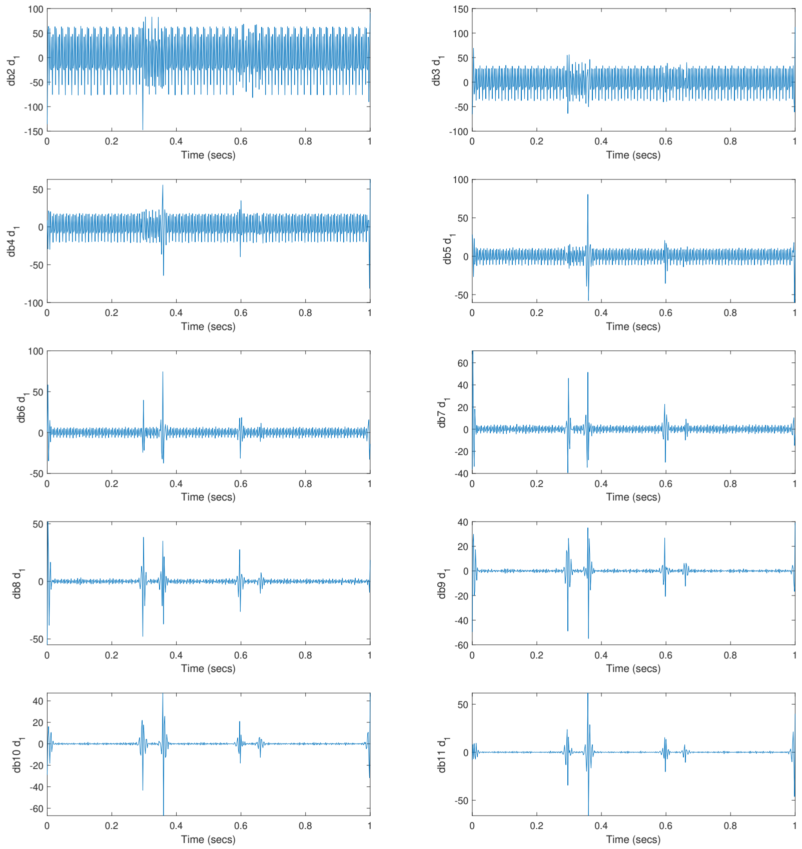


Figure 2.4: Analysis of Test Signal using Daubechies Wavelet Family (db).

disturbances in test signal same as coif4, coif5, we are free to choose the mother wavelet from one amongst these. However, research in detection of fast transients and harmonics in power systems, have showcased the advantage of selecting daubechies family of wavelets, especially db10 and above[4]. This is in agreement with the rudimentary comparative exercise carried out and mother wavelet db11 is selected for the use in this thesis..

REFERENCES

[1] D. C. Robertson, O. I. Camps, J. S. Mayer, and W. B. Gish, *Wavelets and electromagnetic power system transients,* IEEE Transactions on Power Delivery **11**, 1050 (1996).

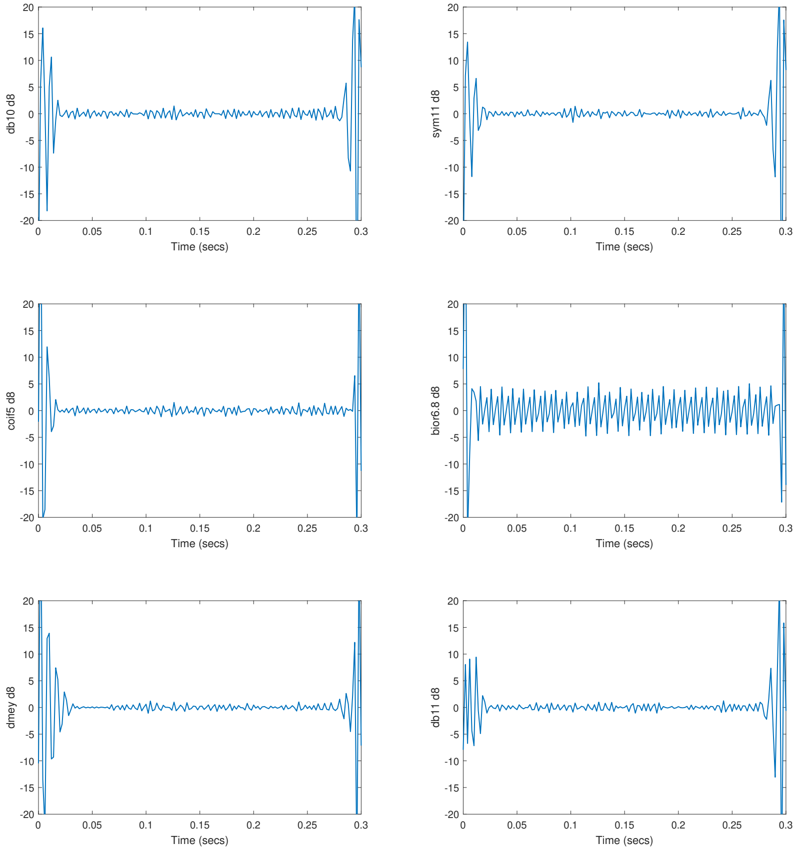


Figure 2.5: Comparison of decomposition of test signal during normal behaviour.

[2] K. Rajesh and N. Yadaiah, *Fault identification using wavelet transform,* in *2005 IEEE/PES Transmission Distribution Conference Exposition: Asia and Pacific* (2005) pp. 1–6.

[3] N. S. D. Brito, B. A. Souza, and F. A. C. Pires, *Daubechies wavelets in quality of electrical power,* in *8th International Conference on Harmonics and Quality of Power. Proceedings (Cat. No.98EX227)*, Vol. 1 (1998) pp. 511–515 vol.1.

[4] S. Wu, Z. Piao, J. Ji, X. Li, B. Hu, and P. Yang, *Harmonics detection in electric power systems based on wavelet packet transform,* in *2010 International Conference on Intelligent Computation Technology and Automation*, Vol. 2 (2010) pp. 425–427.

2

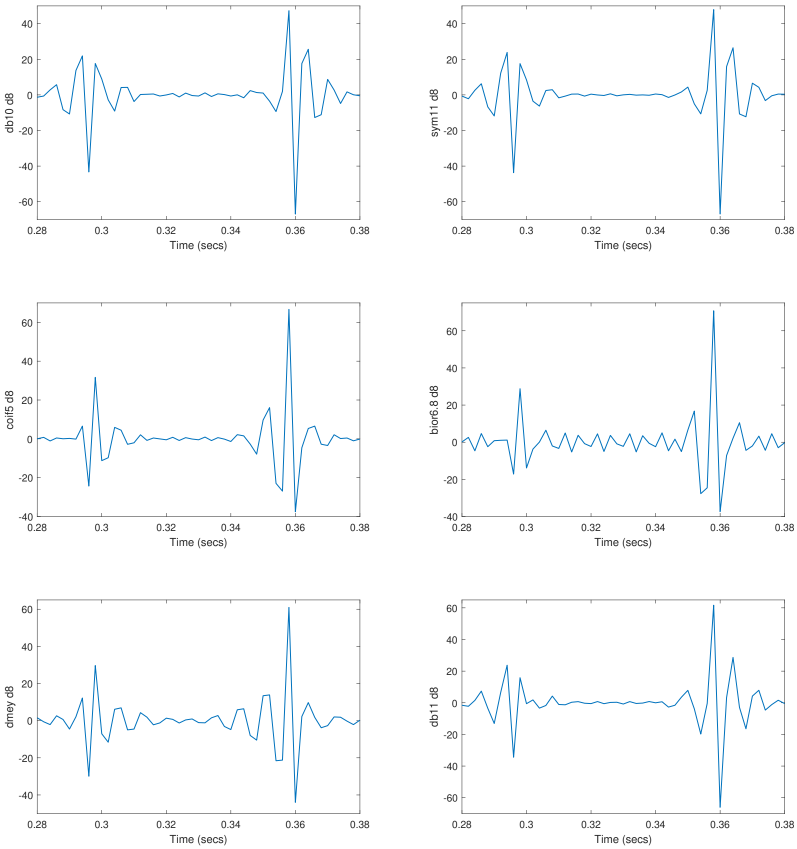


Figure 2.6: Comparison of decomposition of test signal during high frequency disturbance.

2

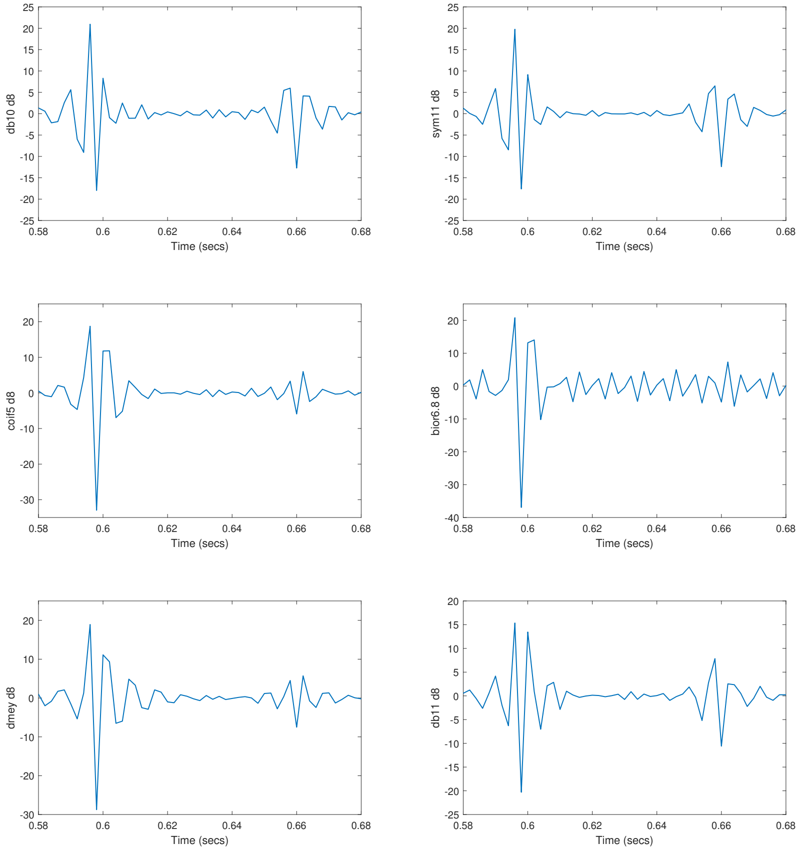


Figure 2.7: Comparison of decomposition of test signal during low frequency disturbance.

3

Simulation of Disturbances in PSCAD

The power grids comprise of various voltage levels, and hence, the magnitude of disturbances can vary over a large range, according to the voltage level it occurs at. To analyse the characteristics of these disturbances, it is necessary to have the data at various voltage grades. The data can be collected from actual field cases or can be simulated using various tools available. For this thesis, PSCAD is used for simulation of different events in systems, with different voltage levels, resulting in a wide variety of disturbances. This simulation data is used to train the ML module and to test the ML module in further chapters.

3.1. Disturbances

Disturbances in a power grid is a fairly common occurrence, with the vast amount of equipment connected and disconnected to and from the grid at any given time and the exposure of the transmission system to the nature. These disturbances are not an issue, until the normal operation of the grid is affected. Faults such as 2ϕ, 3ϕ, 1ϕ to ground, 2ϕ to ground and 3ϕ to ground have been the staple faults, that a conventional circuit breaker and relay protect the power grid against. However, faults with high impedance, tend to go undetected and uncleared by the conventional protective systems, due to the low fault currents and lack of definitive characteristic of the fault [1]. Detection and classification of HIF has been a hot topic for the past few decades, with new and innovative techniques being developed every year.

Researches using DWT and ML for detection and classification of HIF have emerged in the last few year; where different feature sets are utilised to better the accuracy and time required to detect and classify [2]. However, these researches focus mainly on the detection and classification, while letting the conventional relays and circuit breakers handle the conventional faults. This is a good approach as a start, however, in the future of smart grids, it would be beneficial to use a single module for detection and classification of any disturbances in the grid, while keeping the conventional protection devices as a contingency plan to fallback on.

Another such disturbance, which may go undetected by the existent protection devices, or might be ignored as a conventional fault, is the incipient faults, which can be early signs of equipment damage [3] [4]. These faults are highly difficult to detect, due to following reasons:

- Intermittent in nature.
- Very short fault duration (1/4 of a cycle)
- Low fault currents

Detecting these early signs of equipment damage is of great importance to the DSO, as an unanticipated damage of equipment can lead to unplanned downtime, which in turn will result in penalties. The Manhattan blackout is one such example of equipment damage, that resulted in large scale blackout. However, it is to be noted that the underlying cause of the blackout has not been made public yet.

This thesis aims to try and detect as well as classify these hard to detect disturbances, while also being able to detect the conventional faults, and distinguish them from non-disruptive disturbances.

3.2. DISTURBANCES SIMULATED IN PSCAD

To understand the effect of the different voltage grades used in the electrical networks, test systems ranging from 20 kV to 230 kV were used. Various events were simulated to bring in disturbances in the test systems and a 1-D db11 wavelet transform was applied on the measured current and voltage waveform. The details of the simulated events, wavelet transforms and the inferences have been detailed in the following sections.

This also helps to analyse the scaling of the results of the DWT with various voltage and current measurements, to further determine a classifier differentiating normal operation from disturbances.

To simulate various events, IEEE 9 bus, 14 bus, 30 bus and 34 bus (without wind load) [5] transmission systems were used in PSCAD software. The simulation was run for 6 seconds, however the measured data were only recorded from 1 second to 6 seconds. This was done so to avoid the start-up disturbances of the system. The following events were simulated in each of the test systems:

- Load Switching
- Capacitor Switching (not simulated in 9 Bus system)
- Breaker Switching (disconnecting parts of the system)
- Internal Faults (Fault On resistance of 1 Ω)
    1. Phase A to Ground Fault
    2. Phase B to Ground Fault

3. Phase C to Ground Fault
4. Phase A-B to Ground Fault
5. Phase A-C to Ground Fault
6. Phase B-C to Ground Fault
7. Phase A-B-C to Ground Fault
8. Phase A to Phase B Fault
9. Phase A to Phase C Fault
10. Phase B to Phase C Fault
11. Phase A to Phase B to Phase C Fault

- HIF (not simulated in 9 and 14 Bus systems)

The voltage and current were measured between different buses and the data was recorded using COMTRADE 99 recorder module in PSCAD, with a 5 μs time step and 50 μs sampling time. The COMTRADE 99 format is the most widely used file format for recording and exchange of power system transients. The format is used as per recommendation of IEEE Power and Energy Society.

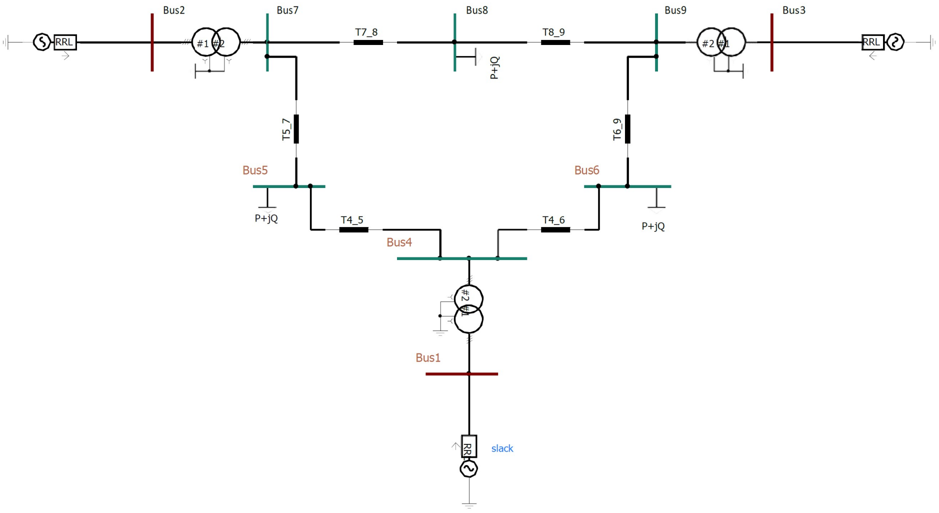


Figure 3.1: IEEE 9 Bus System

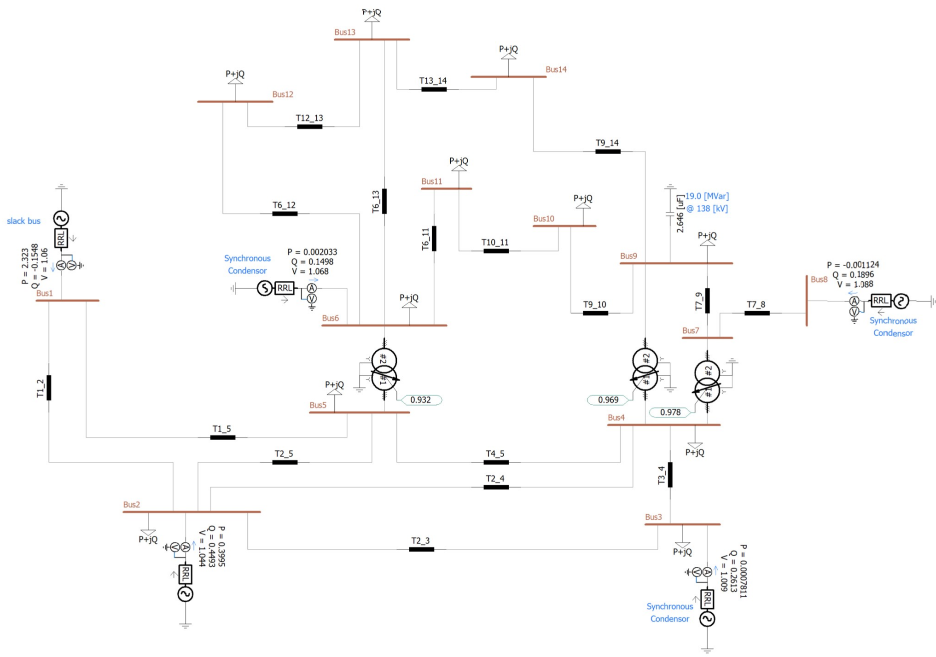


Figure 3.2: IEEE 14 Bus System

For the conventional faults, the simulations were run with 1 Ω fault resistance. The faults were simulated at multiple locations in the systems, while voltage and current measurements for each phase were recorded. However, for HIF, a module was designed to emulate the unpredictability of the faults observed in real world measurements, and is shown in figure 3.3. This module, was utilised in the IEEE 30 and 34 bus systems, as HIFs are a rarity in higher voltage systems. The HIF module was designed based on the model developed in [1] which was developed on the base of [6]. The values of the DC voltage sources V_p and V_n, and resistances R_p and R_n, were recalculated for the IEEE 30 and 34 bus systems, i.e., 33 kV and 24.9 kV systems respectively. The values of resistances of R_p and R_n as well as the DC voltage sources V_p and V_n, were varied using random number generators, while limiting it within a range, determined by the voltage level of the system.

The V-I characteristics of the HIF modelled for 33kV and 24.9kV are shown in figures 3.4 and 3.5, respectively. To emulate the variable arcing resistances observed in the HIF, variable resistances were used, with random number generator for assigning random resistance values to the fault path. The breakdown voltages, for the initiation of the HIF, were also varied withing a range of values by varying the DC sources. Too incorporate an asymmetry in the HIF characteristics, the DC voltage sources were ensured to have a different range limit. This is observed as the multiple lines in V-I characteristics. Figure 3.6 shows the current waveform of phase A in 34 Bus system, measured between buses 834 and 842, when the HIF module is switched on at bus 848 at 1 second, however, the

Bus System	No. of Disturbance Signal Data Recorded	No. of Disturbances Simulated
34	1860	31
30	4080	34
14	2436	29
9	864	16
Total	9240	110

Table 3.1: Simulation Summary

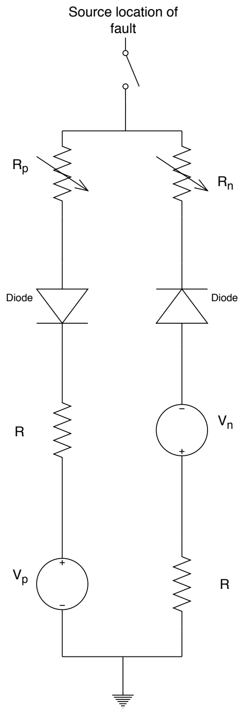


Figure 3.3: Circuit Diagram of the HIF model implemented

arcing starts at 1.0025 seconds, when the voltage of the system is able to forward bias the forward conduction diode in the HID module. A similar pattern is observed in negative half-cycle, wherein the reverse conduction diode is forward biased. The current waveform is characterised by the irregularity in current during fault, caused due to the varying fault resistance, as well as the changing breakdown voltages. The non-symmetrical nature is also observed in the figure, as modelled by the diodes and DC voltage sources. The current and voltage are recorded at multiple buses, thus providing different waveform for the same disturbance.

Along with disruptive disturbances, non-disruptive disturbances like load switching

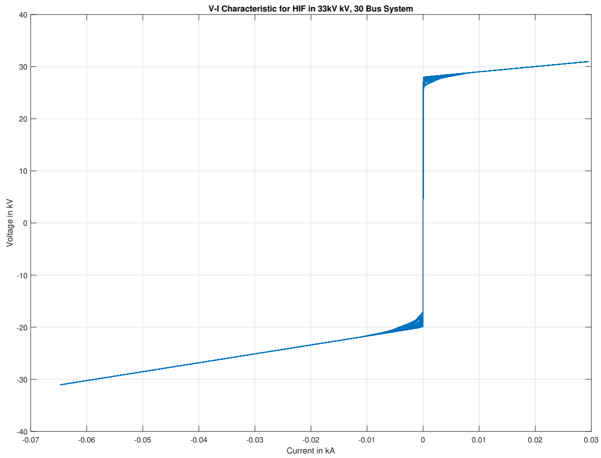


Figure 3.4: V-I Characteristics for HIF in 30 Bus System (33 kV)

and capacitor bank switching is simulated as well. These voltage and current recordings were used to both train and test the ML module in the next chapter. Current and voltage signatures of power equipment failures were obtained from the IEEE Power and Energy Society [7]. The gap-less data for 2 days of current and voltage waveform were downloaded and used for this thesis. The data was not available in COMTRADE 99 format and was available in .mat format.

Figures 3.7 and 3.8 show the current and voltage signals measured in 9 bus system, for bus 1 phase A, bus 2 phase B, bus 3 phase C and between buses 4 and 5 for phase A. In these signals, following disturbances simulated, can be observed:

- Phase A to Ground fault at Bus 8 from 2 to 2.1 seconds.
- Phase B to Ground fault at Bus 8 from 2.6 to 2.7 seconds.
- Phase C to Ground fault at Bus 8 from 3.2 to 3.3 seconds.
- Phases AB to Ground fault at Bus 8 from 3.8 to 3.9 seconds.
- Phases AC to Ground fault at Bus 8 from 4.4 to 4.5 seconds.
- Phases BC to Ground fault at Bus 8 from 5 to 5.1 seconds.
- Phases ABC to Ground fault at Bus 8 from 5.6 to 5.7 seconds.

REFERENCES

[1] R. Bhandia, J. J. Chavez, M. Cvetkovic, and P. Palensky, *High impedance fault detection in real-time and evaluation using hardware-in-loop testing,* in *IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society* (2018) pp. 182–187.

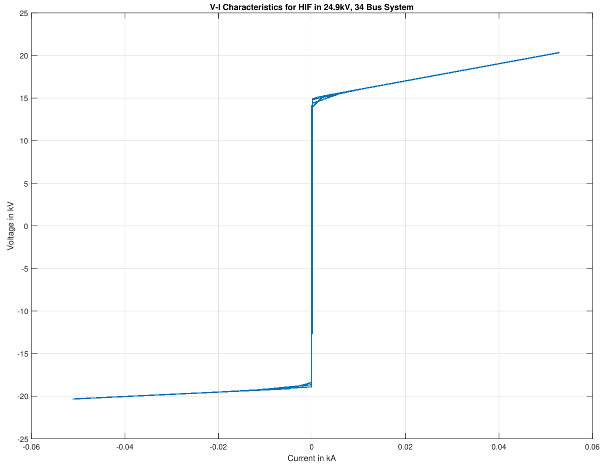


Figure 3.5: V-I Characteristics for HIF in 34 Bus System (24.9 kV)

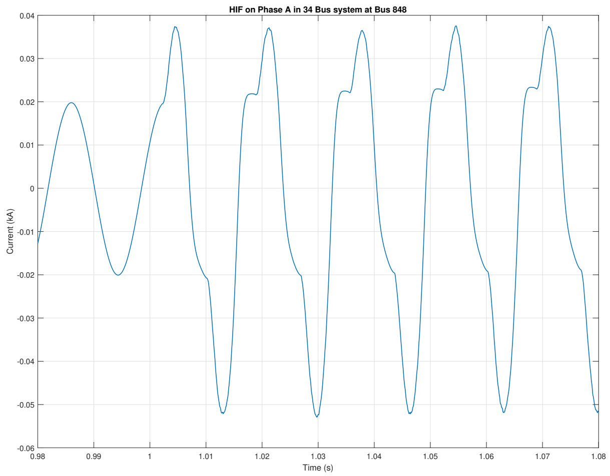


Figure 3.6: Current waveform for HIF in 34 Bus System (24.9 kV) at Bus 848 on Phase A

3

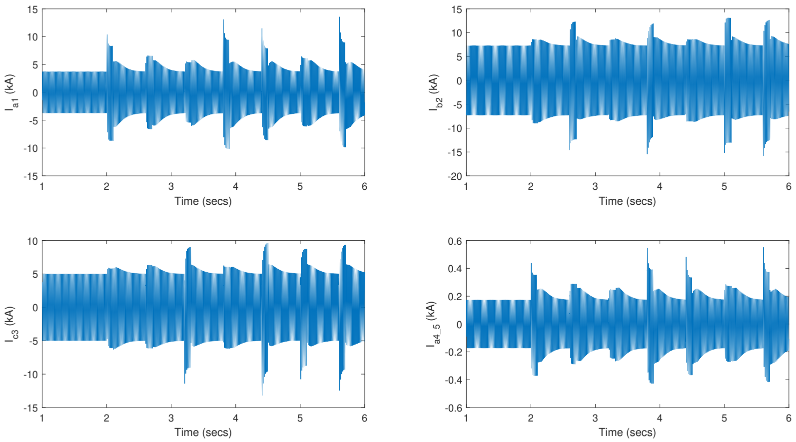


Figure 3.7: Current waveform for Grounded faults in 9 Bus system.

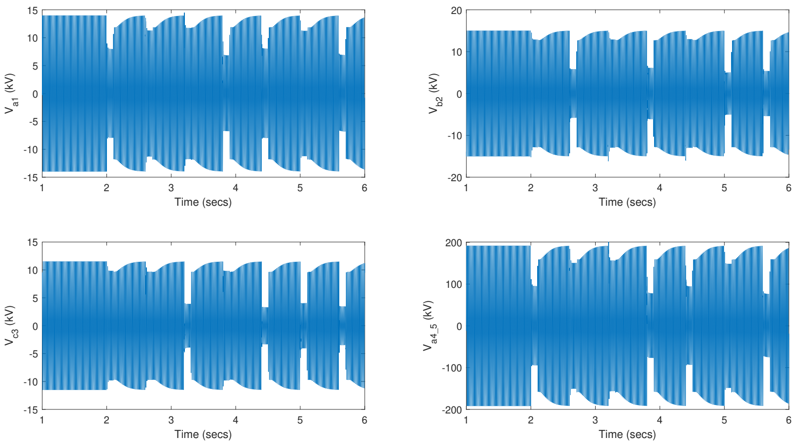


Figure 3.8: Voltage waveform for Grounded faults in 9 Bus system.

[2] A. . Sedighi, M. . Haghifam, O. P. Malik, and M. . Ghassemian, *High impedance fault detection based on wavelet transform and statistical pattern recognition,* IEEE Transactions on Power Delivery **20**, 2414 (2005).

[3] T. S. Sidhu and Z. Xu, *Detection of incipient faults in distribution underground cables,* IEEE Transactions on Power Delivery **25**, 1363 (2010).

[4] S. Kulkarni, A. J. Allen, S. Chopra, S. Santoso, and T. A. Short, *Waveform characteristics of underground cable failures,* in *IEEE PES General Meeting* (2010) pp. 1–8.

[5] *IEEE benchmark models and examples [online],* `https://hvdc.ca/knowledge-base/topic:51/v:`.

[6] R. E. Lee and M. T. Bishop, *A comparison of measured high impedance fault data to digital computer modeling results,* IEEE Power Engineering Review **PER-5**, 35 (1985).

[7] *IEEE working group on power quality data analytics,* `http://grouper.ieee.org/groups/td/pq/data/`.

3

4

Feature Extraction and Training the ML Module

The advance in computational technology, and processing power of residential computers, have made Machine Learning (ML) an easy to access subset of Artificial Intelligence (AI). ML has been available in two of the popular tools used by electrical engineers, MATLAB and Python, along with wavelet modules. This ease of availability and added benefit of complex pattern recognition, is utilised in classification problems of power systems and is explored in this chapter.

4.1. Feature Extraction

ML modules can be segregated into supervised, unsupervised and reinforcement learning modules [1]. As the name suggests, in supervised learning modules, the ML module is trained to recognise patterns in a set of data, using available set of data and the patterns labeled in it, called the training data set. In case of unsupervised ML modules, the underlying patterns in a set of data is recognised and utilised by the module for prediction and classification of newer data. Finally, in reinforcement ML modules, the module keeps dynamically learning patterns in the training data, optimising the prediction or classification results. In this thesis, the method of supervised learning is utilised to develop a ML module for classification of the disturbances explored in the previous chapter. Supervised learning is preferred here due to the amount of classifications to be made. With higher the number of labels/ patterns to be recognised, the unsupervised and reinforcement learning based ML modules, require higher training data and time for achieving accuracy comparable to supervised ML module.

For training the ML module, certain features from the voltage and current signals can be extracted, for a better distinguishing of disturbances in the power system. This is preferred over the use of current and voltage signals directly for training the ML module, as the underlying pattern during a disturbance, is more distinguishable using signal processing and statistical mathematical tools [2]. As mentioned in chapter 2, DWT is a powerful method to process the voltage and current signals, resulting in distinguishable disturbance patterns. This property is utilised to process the signals from chapter 3. However, this only provides a pattern for distinguishing the start and stop of a disturbance in the power system as seen in figure . To have a better classification of the disturbances, further features, based on statistical and signal analysis tools are extracted, namely:

- Root Mean Square (RMS),
- Energy,
- Shannon Entropy,
- Mean,
- Standard Deviation,
- Variance
- Number of zero crossings,
- Number of mean crossings

These features are extracted for the details and approximations at each level of decomposition of the voltage and current signals using DWT. Thus for a signal, with 8 levels of decomposition, a total of 8 x 8 [Details] + 8 x 8 [Approximation] = 128 features are extracted. As explained in the previous section, the maximum level of decomposition depends on the sampling rate of the signal to which DWT is applied.

RMS, Energy, Mean, Standard Deviation, Variance, Number of Zero Crossings and Number of Mean Crossings are basic statistical criteria, used for summarising a signal. Another such criteria is the Shannon entropy of a signal. In classical sense, entropy is the measure of disorder in a system. In information theory, pioneered by Claude Shannon, the entropy is defined analogous to the entropy defined in classical sense (thermodynamics) [3]. It can be defined as a measure of the unpredictability of the average information content, and is quantified as below:

$$H = -\sum p_i log(p_i) \tag{4.1}$$

where, H is the entropy and p_i is the probability of the information i occurring in the information set. This is useful for the analysis of a signal, as the entropy of a low probability value, i.e. disturbances in a sinusoidal waveform of current and voltage, in a signal carries more "information" as compared to a high probability value.

Figure 4.1, shows a comparison of the Shannon entropy values of each detail of the DWT of current measurement between buses 834 and 842 in IEEE 34 bus system for different events. The entropy values for each event, is comparable with respect to the detail, however, for the LLLG and LLL faults, the entropy values appear to be similar at level 8 of decomposition, where majority of the entropy is concentrated. If the Shannon entropy of the first 4 levels of decomposition are zoomed into, we see that the values differ for different event. This is shown in figure 4.2.

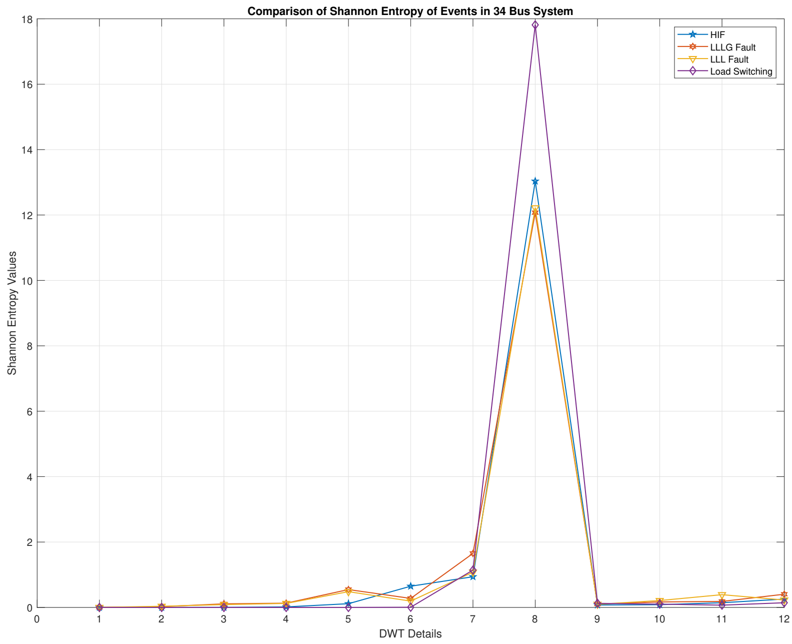


Figure 4.1: Shannon Entropy of DWT Details

Each of the DWT decomposition of the voltage and current signal is summarised by these features and are used by the ML module for classification of the disturbance. These features are extracted from the training data, and are then grouped as per the disturbance they belong to, and the event labels are grouped similarly. This can be seen as a matrix, x, with features of a signal in each column, and a vector, y, with the label of the event, to which each feature belongs to, as shown in equation 4.2

$$\begin{aligned} Event\ Labels,\ y = & \quad \{y_1 \quad y_2 \quad . \quad . \quad . \quad y_m\} \\ Features\ of\ each\ event,\ x = & \left\{ \begin{array}{cccccc} x_{11} & x_{12} & . & . & . & x_{1m} \\ x_{21} & x_{22} & . & . & . & x_{2m} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ x_{n1} & . & . & . & . & x_{nm} \end{array} \right\} \end{aligned} \tag{4.2}$$

where, y_i, is the event label corresponding to the features in column i of x, i.e., x_{1i} to x_{ni}. These are used to train the ML module for classification of events as per the features to the corresponding labels.

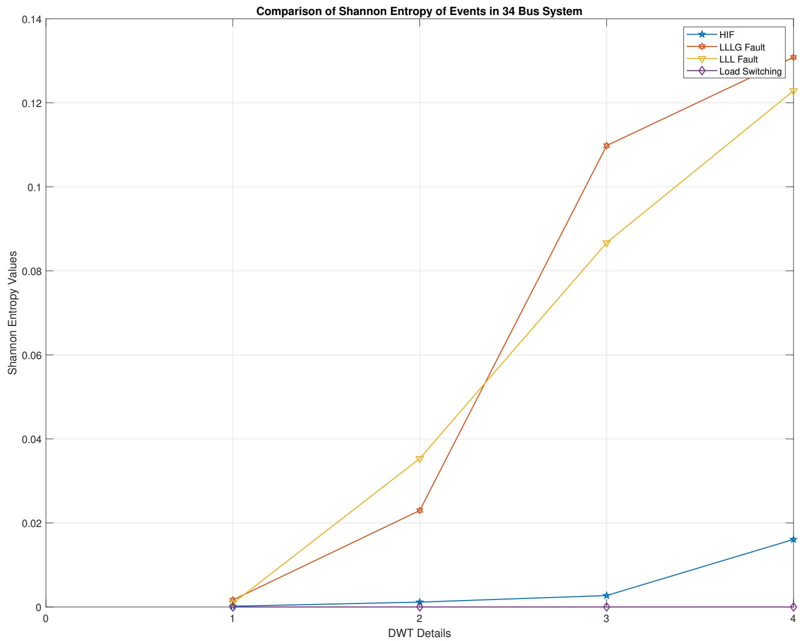


Figure 4.2: Shannon Entropy of DWT Details (Level 1 to 4)

4.2. ML MODULE TRAINING

As mentioned in the previous section, there are different ML methodologies available for use in supervised training. Some of the methods available in MATLAB and Python for classification problems are k-Nearest Neighbour (kNN), Support Vector Machines (SVM), Decision Trees (DT), Gradient Boosting (GB) etc. Of these, (DT), is known to be the simplest, with a high accuracy rate in classification problems using supervised learning, along with the shortest training time [4]. The DT works very similar to human brain's decision making process. Every flowchart can be classified as a decision tree and hence it is one of the easiest ML modules to comprehend. Figure 4.3 shows an example of a small DT algorithm, that predicts the chances of survival of passengers on the Titanic, based on the age and sex of the passengers. The accuracy of prediction can be observed to be varying with each level, i.e. leaf, of the tree branch.

As seen in figure 4.3, DT algorithms assign logical rules to each leaf, where a certain feature is equated to a value, resulting in two categories, True or False, which is further broken down based on the further logical rules assigned to the leaves. The further down the branch we go, the smaller the data-sets become. This may lead to classification modules, that provide high accuracy when classifying the test data, however result in low accuracy when classifying test data, as the logical rules assigned by the DT algorithm is over-fitted to the specific features of the training data. This is undesirable in a robust classifier and can be avoided by increasing the variety in features of the training set. However, there can be a trade-off in the accuracy with large feature sets, due similarity in features of different events to be classified. Hence, it is desirable to select features than can be used to distinguish the events to be classified. The ML modules are equipped with the abilities to reduce the number of features to be used for classification, by removing the features that are highly correlated; to counter the over-fitting of the ML module for a particular training set, the ML module can optimise the depth of the branch of tree needed for better classification, also reducing training time, and can be trained to cross-validate the predictions.

Another popular method is GB, which is highly utilised in use cases, where the ML model has a weak performance, i.e., the model is unable to classify accurately. It utilises mathematical modeling to boost the weak model, for better accuracy. The mathematical concept of the same is complex and is left out of the scope of this thesis, however, it can be simplified as, initially making multiple weak models, and then, building a model based on these weaker models to achieve a model with higher accuracy of classification. The underlying weaker models include DT based models as well. The benefit of this methodology, is to achieve a better accuracy of classification, using features that are, seemingly insufficient for the classification problem. This, however, requires a longer time period for training the ML module, as the multiple weaker modules are built first and then the strong model is utilised for training the ML module. The ML modules in this thesis were trained using both DT and GB, and a comparison of the results were made.

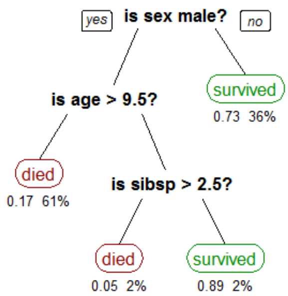


Figure 4.3: Example of DT predicting the chances of survival of passengers on Titanic, based on sex and age [5].

For training the ML module, the data from the PSCAD simulations as well as the data from the PES workgroup is utilised. As the data is stored in two different formats, simulations in COMTRADE 99 and 2-day gap-less equipment failure signature in .mat format, the first step would be to read this data and convert into the format beneficial in training the ML module. However, to decide this, a tool for implementation of ML module was to be zeroed in. From the two common tools, MATLAB and Python, the following were observed:

- MATLAB:

  - The visualisation of data and the breakdown of signals using DWT is very simple.
  - ML modules along with features can be easily visualised.
  - Is not open source and needs license for usage.
  - The analysis of the PES work group data is highly resource consuming and requires large processing power, unavailable in normal computers.

- Python:
  - Open source tool with a variety of open source Integrated Development Environments (IDEs).
  - The analysis of large data, such as the PES work group data, is easy, with lower resource consumption and lower processing power requirements.
  - The visualisation of data is not as elegant and simple as MATLAB.
  - Requires a strong knowledge in writing program codes.

Python's drawbacks were minuscule when compared to those of MATLAB, as high resource requirement of MATLAB and the need for license purchase to use the tool, outweighs the benefits of MATLAB. For this reason, Python was utilised for implementing the ML module in this thesis.

A simple representation of activity flow, followed in this thesis, for training the ML module is shown in the figure 4.4.

As the flow-chart shows, the data is initially read into the Python's IDE, in this theses, Jupyter Notebooks were used. The voltage and current signals are then pre-processed, readying it for feature extraction. The pre-processing of data from COMTRADE 99 files involve conversion of data to raw signal data, by processing the .dat and .cfg files. The signals are then separated into voltage and current signals. The part of the signal, corresponding to the disturbances, is extracted based on time of event simulation, and each signal is labelled as per the disturbance. After this, the program applies DWT on each disturbance signal, and extracts the features, mentioned in the previous section, for each level of detail and the final level of approximation. These features are grouped as per the event label they represent. This grouping can be visualised as the matrices in equation 4.2. The grouped data is then split into training data and test data, usual norm of 70:30 split, i.e., 70% of data is used to train the ML module and 30% is used to test the accuracy of the ML module. The training data, containing of x matrix with features and y matrix with event labels, is used to train the DT ML module. The accuracy of the ML module is explored by using the test data. The trained ML module is then extracted and saved as a pickle file (common file format for saving data between python programs).

This ML module, can then be used by any other python program, for classifying the disturbances in the signals provided. However, the features extracted from the voltage and current signals in this classification program must be same as the features extracted in the training program. Such a program was developed for this thesis and the flowchart for same is shown in figure 4.5

The trained ML module is initially loaded into the python program and then the current and voltage signals are read. The signals are pre-processed same as in the training

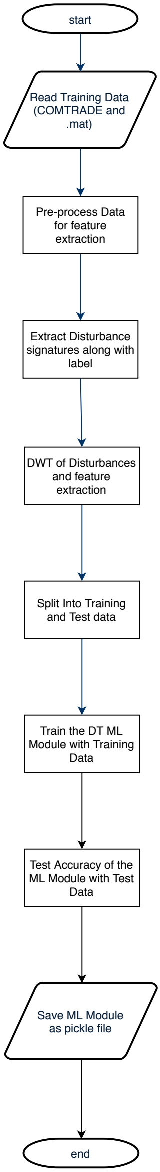


Figure 4.4: Flow Chart for training of ML Module

program, wherein the current and voltage signals are separately grouped. To emulate the conditions in the grid, as moving window is utilised to sample the signals. This window is limited by the number of data points used in training the ML module, which was 4000 data points, corresponding to 0.2 seconds of signal data. The data within this window is then subjected to DWT and feature extraction, same as in the training program. The data in this window is then classified using the trained ML module. If the classification is not normal, then signal in the time period of this window, is classified as disruptive or non disruptive based on the label assigned by ML module.

If the classification is as normal, then the program checks if it is the end of the signal data, and moves the window to the next 2 cycles, i.e., the window is moved by 670 data points, if it is not the end of the signal data. The process of applying DWT, extraction of features, classification by ML module and shifting of window is repeated, until the end of signal is reached. The program stops once the end of signal is reached.

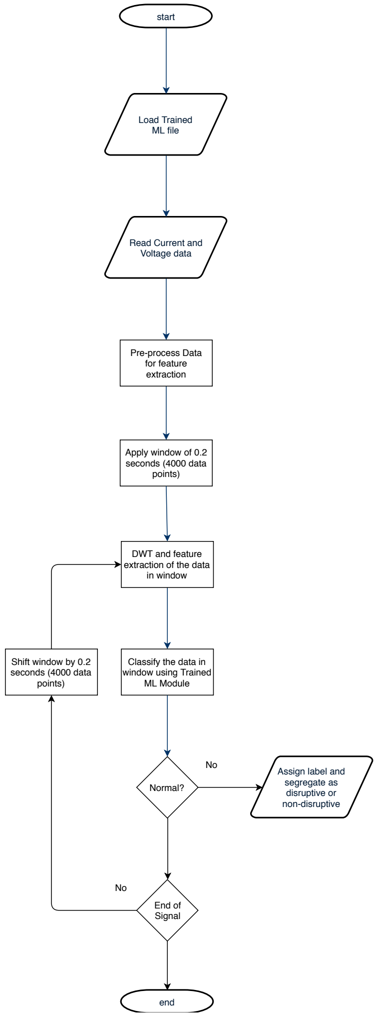


Figure 4.5: Flow Chart for Classification Program using trained ML Module

Hence, this program, essentially checks for a disturbance every 2 cycles, i.e., 0.0335 seconds, and classifies the disturbance detected as disruptive or non-disruptive, along with the possible cause for disturbance.

REFERENCES

[1] P. Domingos, *A few useful things to know about machine learning,* Commun. ACM **55**, 78 (2012).

[2] F. Uçar, F. Alçin, B. Dandil, and F. Ata, *Machine learning based power quality event classification using wavelet — entropy and basic statistical features,* in *2016 21st International Conference on Methods and Models in Automation and Robotics (MMAR)* (2016) pp. 414–419.

[3] C. E. Shannon, *A mathematical theory of communication,* The Bell System Technical Journal **27**, 379 (1948).

[4] Y. Kashnitsky, *Topic 3. classification, decision trees and k nearest neighbors,* `https://mlcourse.ai/articles/topic3-dt-knn/`.

[5] S. Milborrow, *An example of a cart classification tree,* `https://upload.wikimedia.org/wikipedia/commons/f/f3/CART_tree_titanic_survivors.png` (2011).

5

Results and Conclusion

The program developed in the previous chapter, was tested within multiple disturbance event scenario and the results drawn from these tests are visualised and explained in this chapter. This chapter also serves as the conclusion of this thesis, listing the achievements with respect to the objectives and research questions set at the beginning. It also delves into the shortcomings of this thesis and the future work, that can be embarked on, to overcome these shortcomings.

5.1. Results

The previous chapter explored the necessities of training a ML module and developed a program to utilise the ML module to detect and classify disturbances in the grid. The training program, provided an insight into the accuracy of the trained ML module for classification of disturbances. In the following comparisons, the ML module trained and tested using 34 Bus system is mainly focused on, however, similar results were observed in the other bus systems as well.

5.1.1. Decision Tree v.s. Gradient Boosting

For the ML module, a comparison between one of the simplest learning modules, Decision Tree and Gradient Boosting (GB), one of the best suited learning module for developing a strong model, using features that might not yield high accuracy. A significantly higher accuracy was noticed for ML Module using GB as compared to DT, for the same training and test data of simulation recording.

The figures 5.1, 5.3, 5.5 and 5.7 represent the number of classifications, grouped by events, made by GB classifier trained ML Module on the test data set of voltage and current signals, respectively. As listed previously, the accuracy of the classifications, based on the basic statistical features, is above 75% for GB classier based ML module. The figures showcase the ability of more accurate classification of the events using voltage signals, over current signals. As the test data set was randomly selected, the number of test signals, corresponding to each event, is unevenly distributed. Voltage signal based

Type of Classifier	For Voltage Signals		For Current Signals	
	Pre-normalisation	**Post-normalisation**	**Pre-normalisation**	**Post-normalisation**
GB	87.57%	88.24%	76.14%	76.14%
DT	79.41%	79.08%	63.73%	66.01%

Table 5.1: Comparison of accuracy of GB and DT classifiers, pre and post normalisation

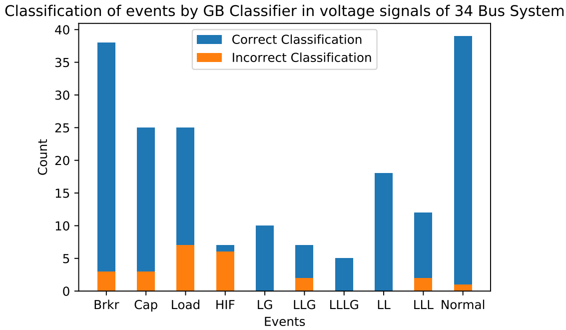


Figure 5.1: Prediction of Events in 34 Bus system using Gradient Boosting (GB) Classifier in voltage signals

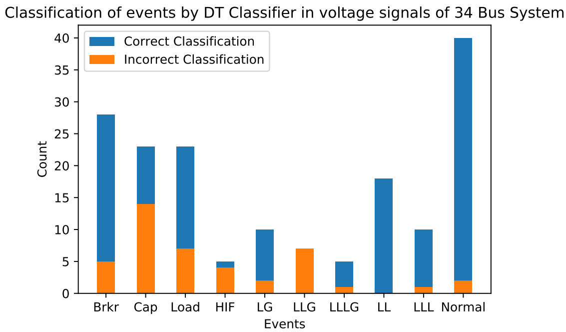


Figure 5.2: Prediction of Events in 34 Bus system using DT Classifier in voltage signals

classification of HIF, yielded 15 accurate classifications and 2 inaccurate classifications, as opposed to 12 accurate and 6 inaccurate classification based on current signal. Similarly the figures 5.2, 5.4,5.6 and 5.8, represent the number of classifications made by DT Classifier trained ML module, using voltage and current signals, respectively. Although, the DT classifier trained ML module is able to achieve an accuracy above 75%, in voltage signals, the accuracy in current signal, is fairly low. The figures show that the LLLG , HIF and Cap switching events are highly likely to be classified inaccurately when DT is used with current signals.

This thesis explored the possibilities to increase this accuracy by introducing normalisation of the current and voltage signals, wherein the current and voltages were normalised over the base of peak current and peak voltages encountered in the system,

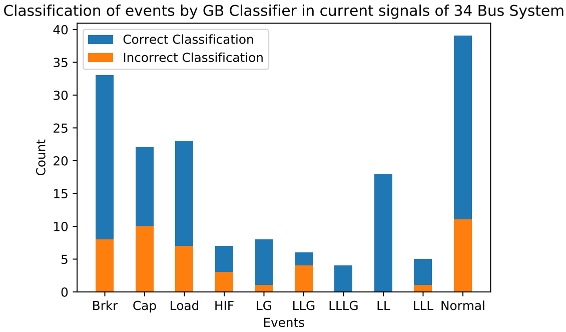


Figure 5.3: Prediction of Events in 34 Bus system using Gradient Boosting (GB) Classifier in current signals

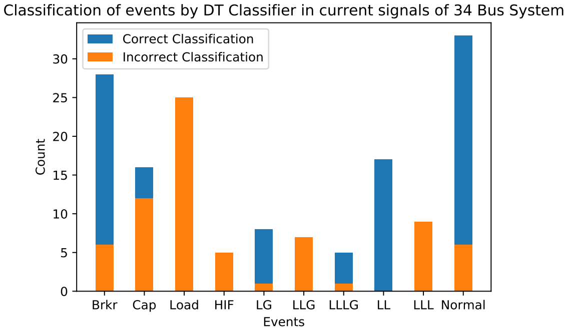


Figure 5.4: Prediction of Events in 34 Bus system using DT Classifier in current signals

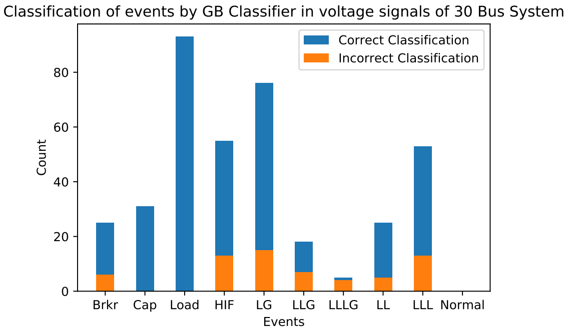


Figure 5.5: Prediction of Events in 30 Bus system using Gradient Boosting (GB) Classifier in voltage signals

when there is no disturbance, i.e., normal operation. This was done so to avoid having to deal with different magnitudes of voltage and current, when the voltage levels are changed. However, the improvement in accuracy was only marginal in the case of GB Classifier trained ML module, although, improvements were noticed in the case of DT

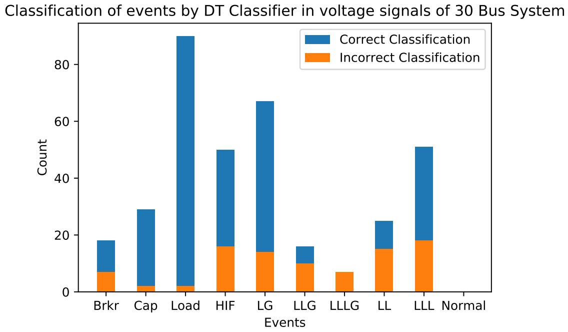


Figure 5.6: Prediction of Events in 30 Bus system using DT Classifier in voltage signals

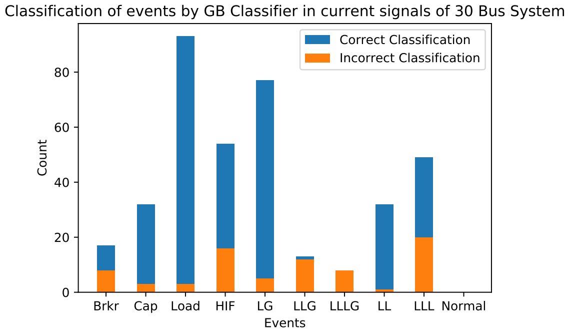


Figure 5.7: Prediction of Events in 30 Bus system using Gradient Boosting (GB) Classifier in current signals

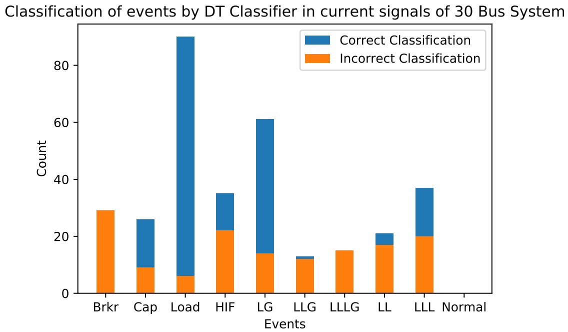


Figure 5.8: Prediction of Events in 30 Bus system using DT Classifier in current signals

Classifier trained ML module.

The improvements in accuracy were only marginal, when using the same module for classification of events in 34 bus system. However, there was an increase in accuracy when using this normalised trained ML module for classification of events, compared to non-normalised trained ML module in a system with different voltage level, 30 bus system. This improvement in accuracy, though, was of no significance, as the accuracy was below 15%. Hence, for each voltage grade tested, a separate module was trained, which had similar accuracy characteristics as in 34 Bus system, i.e., voltage signals provided better accuracy for classifying the events both in GB and DT classifier trained ML modules, as compared to current signals. The accuracy was higher for GB classifier trained ML module in these cases as well.

As the accuracy of the ML module is calculated by randomly splitting the simulation data by 70:30 ration, wherein the 70% of data is used to train the ML module and 30% is used to test the ML module accuracy, the training and test data can make a difference in the achieved accuracy. This accuracy also changes with respect to the amount of data used to train and test the ML modules. This can be observed in the figures 5.9 and 5.10, wherein the accuracy scores over a series of 25 runs are plotted.

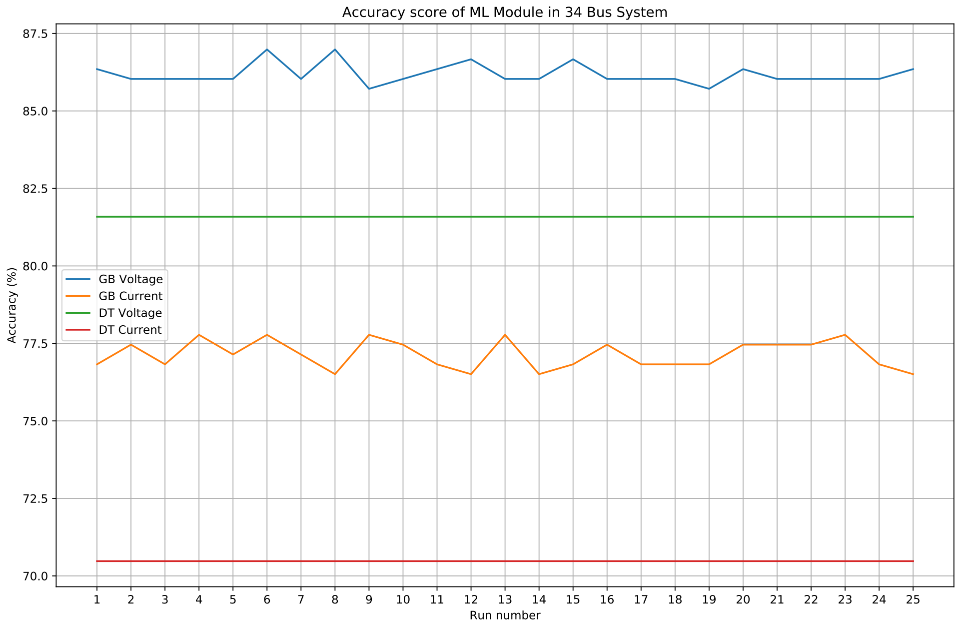


Figure 5.9: Accuracy of ML module trained using 34 Bus system signals over 25 runs

5.1.2. COMPARISON OF FEATURES

To further analyse the effect of each feature, described in the previous chapter, on the accuracy of the ML module, an elimination method was utilised. Each of the features, one at a time, were left unused, while training the ML module, to analyse the impact of the feature on accuracy of the module. The accuracy of the ML module, on the test data is utilised for this comparative study. Each statistical feature, amounts to sixteen features of a signal, as explained in the previous chapter. Figure 5.11 provides an insight into how a feature, or a set of features, on elimination can affect the accuracy of classification.

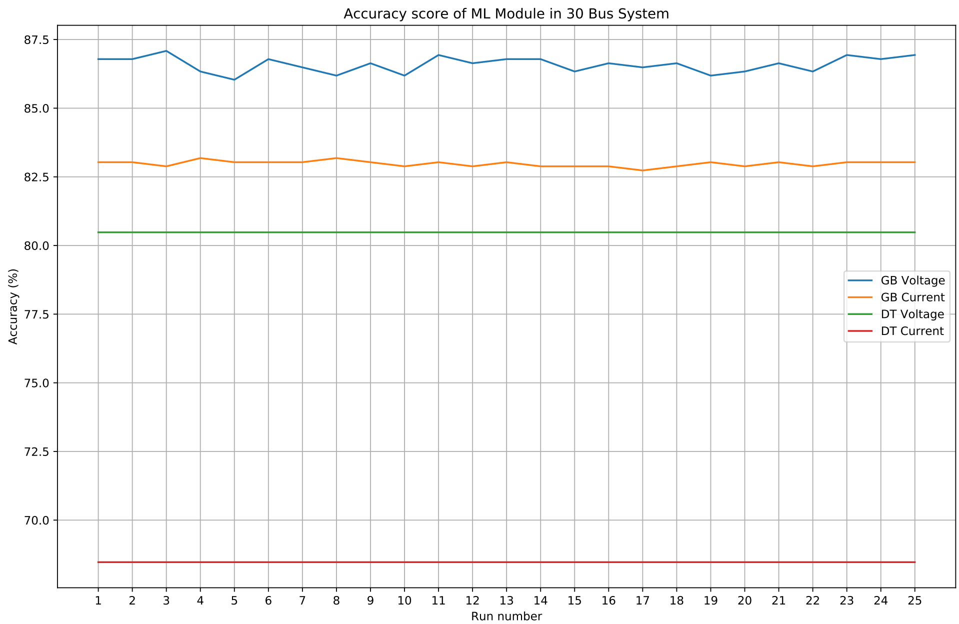


Figure 5.10: Accuracy of ML module trained using 30 Bus system signals over 25 runs

It was observed that the elimination of Shannon entropy, results in the minimum accuracy in calculation, when voltage signals are considered. This is irrespective of the classification trainer use. The fall is roughly 6-7% . However, this only causes a minor fall in accuracy, when current signals are utilised, causing a fall of 1-2%. It was also observed that elimination of variance, results in minimum impact on the accuracy when voltage signals are involved, even resulting in the highest accuracy when DT classifier is used to train the ML module. It can be understood from this activity that, Shannon Entropy is an important feature in classification of voltage signals, whereas, the number of mean and zero crossings are important for classification current signal. It is also understood that none of the features acts solely as the most important feature, i.e., being the sole feature, that can be used to obtain an accuracy rating as high as the combination of the features. Thus, making the combination of these statistical features, a better candidate as opposed to individual features, which is the conclusion drawn initially, however, was verified in this thesis.

5.1.3. IMPLEMENTATION OF ML

As with any technological research, the final aim is to implement it in real world scenarios. Such an attempt was made in this thesis, wherein, the trained ML module, was utilised to detect and classify disturbances simulated in 34 Bus system using PSCAD. The software algorithm developed, and shown in 4.5, attempted to use just the ML module as an agent for detection and classification of disturbances in the signal. However, without an added block for detection of a disturbance, the ML module trained and developed in

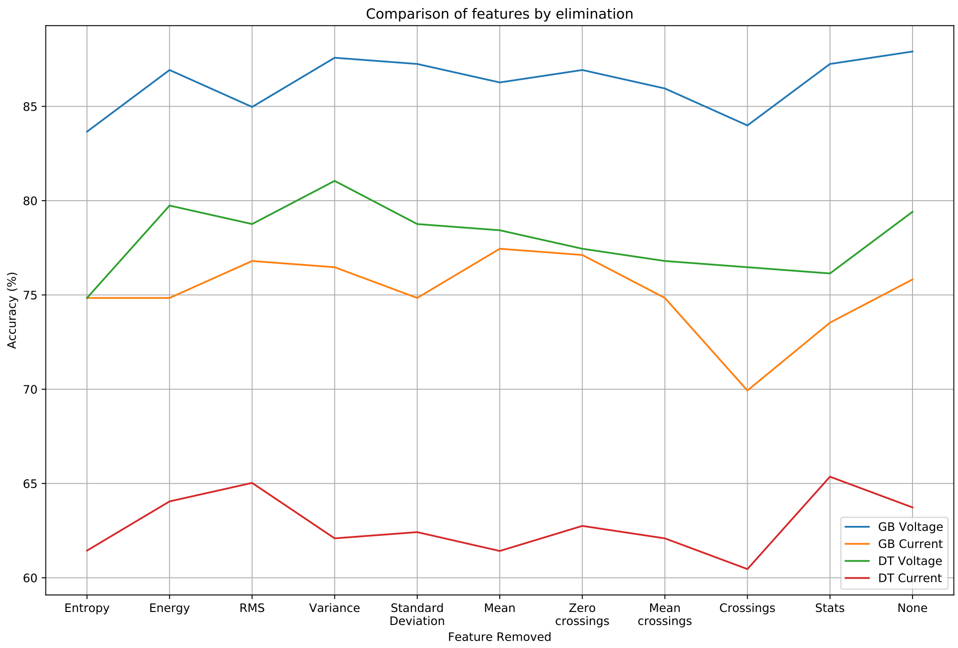


Figure 5.11: Comparison of accuracy while removing features

this thesis, fails to detect and classify the disturbances or normal operation in the signals simulated. This can be attributed to the following:

- The algorithm performs DWT on the signal inside the moving window, irrespective of the presence or absence of a disturbance. This leads to false flagging, even during normal operation of the system.
- The basic statistical features focused on, in this thesis, performs well on isolated disturbance signals, however when the disturbance signatures cross over a window, the ML module is unable to classify it accurately.

5.2. CONCLUSION

THIS thesis aimed at tackling the issue of detection and classification of multiple disturbances in the power grid using a single program, while utilising the basic statistical features of a signal, to distinguish between the disturbances. The same was achieved by developing a detection and classification program utilising the DWT and ML modules in Python. A comprehensive analysis of different mother wavelets in chapter 2, provided db11, of Daubechies family, as the suitable mother wavelet for the use-case in this thesis. Some of the important and hard to detect disturbances, namely HIF and incipient faults resulting from power equipment failures, were identified and the data for the same were obtained through PSCAD simulations and IEEE PES work-group, as detailed in chapter 3. A ML module based on DT and GB was trained for classification of the disturbances in the grid and a program was developed to detect and classify disturbances in the grid, based on basic statistical features of a power signal, as detailed in chapter 4.

To understand the performance of two of the most popular ML classification algorithms, Gradient Boosting and Decision Tree, a comparative analysis of the accuracy and training time was conducted. While the DT classifier is immensely fast at being trained as compared to GB classifier, while using the same training data, the performance of GB classifiers are better in terms of accuracy, especially while using the basic statistical features of a signal for classifying a disturbance.

Another such comparative analysis of these statistical features in the previous section, yielded the importance of Shannon Entropy in accurate classification of voltage signals, while also highlighting the importance of a very basic signal statistics, number of times a signal decomposition crosses the mean and zero, in the case of classifying disturbances in current signal. This analysis also further verified the common notion of using a combination of features for training the ML.

While a program for implementing the ML module for detection and classification of various disturbances in a system was developed, the proposed algorithm provides a low detection and classification accuracy. However the same can be improved upon with further research in pre-classification blocks, which can again be implemented using ML modules, or rudimentary mathematical blocks.

This thesis, hence, achieved its task of detection and classification of disturbances in the grid, using DWT and ML. However, the performance rating of the developed program can be improved by exploring the following:

- Adding to the variety of disturbance simulations for training the ML module, for e.g., varying the fault resistances and duration.
- Using larger and accurate distribution system models.
- Adding more features, based on recent research in detection of HIF and incipient faults.
- Adding more kinds of disturbances for detection and classification.
- Varying the sampling rate of the signal to match the sampling rate of existent grid monitoring devices.

The above can be included in the future work of this thesis, along with, improving on the detection algorithm, by adding detection blocks into the program, before the ML module comes into play. This is in turn reduces the processing time of the data, and can improve the speed of the program when dealing with real-time data from live grid. The disturbance data simulated in this thesis, shall be made public, for utilisation of other engineers.

With the increase in implementation of ICT layer over the physical layer of the grid, the sampling rate of data, and processing power of the grid monitoring devices will increase by many manifolds. This thesis, has attempted to explore this future, of a smart grid, while contributing to the smartness of the grid.