`

Intelligent fault detection technique for distribution network with interconnected distributed generation sources

by

SIPHO PELICAN LAFLENI

Submitted in accordance with the requirements

For the degree of

Master of Engineering (90121)

at the

UNIVERSITY OF SOUTH AFRICA

SUPERVISOR: PROF M SUMBWANYAMBE

CO-SUPERVISOR: DR TS HLALELE

JUNE 2024

Declaration

Name: Sipho Pelican Lafleni

Student number: 54337852

Degree: Master of Engineering (90121) Electrical

Dissertation: Intelligent fault detection technique for distribution network with distributed generation sources.

I hereby affirm that the aforementioned dissertation is the result of my individual effort, scholarly inquiry, and diligent research. All sources that have contributed to this research or have been utilized and quoted have been appropriately acknowledged and referenced with complete citations.

Furthermore, I affirm that I have not previously presented this work, either in its entirety or in part, for assessment at UNISA for any other degree or at any other institution of higher learning.

05.06.2024

SIGNATURE **DATE**

Abstract

Machine Learning (ML) or Artificial Intelligence (AI) approaches in a distribution protection system are proposed to detect and categorize distribution network (DN) issues with Distribution Generation. This study highlights an important but generally overlooked distribution generation feature in energy transformation. Distributed generators (DGs) integrated into Distribution Network (DN) capability improve power quality, system dependability, voltage sags, and emergency backup during protracted grid outages. A technical and global analysis of DG technology's increased penetration is revealing its effects such as the utility systems undergoing elevated fault current and load flow changes, which will affect current protective relaying, particularly overcurrent relays. A protective system that can respond to the new dynamic DN is required to avoid consequences. Intelligent protection system application is a suitable distribution network solution for the explained challenge.

The study conducts a thorough assessment of various methods used in detection and diagnostic systems inside distribution networks that have interconnected distributed generators (DGs). The study specifically emphasizes the implementation of intelligence-based approaches. This evaluation is conducted through a thorough literature review. A comprehensive model of a distribution network, encompassing all essential components, was constructed and subjected to simulation. This model incorporated all pertinent parameters associated with the distribution network. Subsequently, several forms of faults were intentionally introduced at a specific point inside the micro-grid. The purpose of this exercise is to gather voltage and current signals at the busbar and then the collected data is subsequently transformed into numerical values to facilitate machine-learning modelling. The implementation of intelligent approach for fault detection in distribution networks with various machine-learning techniques, allows the approach to form part of the objective to gather related signals that are pre-processed as variable features in order to extract required data that can help identify distribution network and classify faults at most efficient and accurate way. The primary objective of the protection system was to analyze the underlying failure mode, determine the fault type quickly and accurately, and identify the faulty line in the system. .

S shows the micro-grid's successful operation with current and voltage signals evidently shown . The current and voltage signals are transformed to numerical values for feature extraction which is key requirement for machine learning modelling. The derived variable features from feature extraction were trained and tested to validate fault diagnosis and classification to find the best machine learning fault classifier, Support vector machine (SVM) classifiers shows excellent results with 99.9% accuracy in validation and testing. These accuracy results meet the difficult requirements of a micro grid protection systems and SVM's ability to simulate non-linear decision boundaries, which is valuable in many applications.

Keywords: Distributed generators (DGs), Distribution network (DN**)**, Artificial Intelligence (AI), Artificial Neuron Network (ANN), Wavelet transform (WT)

List of Publication

Published

- Sipho P Lafleni, Mbuyu Sumbwanyambe, Tlotlollo S Hlalele, "Intelligent Fault Detection Technique for a Distribution Network with Distributed Generation Sources - Status Review," *International Journal of Engineering Trends and Technology*, vol. 71, no. 8, pp. 333-347, 2023. Crossref, **https://doi.org/10.14445/22315381/IJETT-V71I8P229**

Published

S. P. Lafleni, M. Sumbwanyambe and T. S. Hlalele, "Multi-class Fault Analysis Using Quadratic SVM Classifier Fault Detection Technique for Micro-grid," *2024 32nd Southern African Universities Power Engineering Conference (SAUPEC)*, Stellenbosch, South Africa, 2024, pp, doi: 10.1109/SAUPEC60914.2024.10445052.

Acknowledgement

- Undertaking a Master's research endeavour is a rigorous and significant undertaking that necessitates a strong sense of discipline, diligent effort, self-sacrifice, and support. The abandonment of the journey the loss of hope for completion, and the potential feeling of despair can be attributed to the necessity of setting aside various activities, such as familial and social engagements, in order to maintain a productive study routine. I am deeply grateful for the multitude of individuals who have provided unwavering support during my journey towards the attainment of my master’s degree. While it may be challenging to enumerate all individuals deserving of my appreciation, I would like to really convey my heartfelt gratitude to each and every one of them.

- In gratitude to Almighty God, Lord and Saviour Jesus Christ, and my spiritual home, St Johns Apostolic church "eChibini," I give thanks for bestowing upon me the grace, good health, and lifesustaining resources necessary to complete this research; and for providing me with encouragement and support during my darkest moments.
  I am extremely appreciative.

- To my amazing sister Nombulelo, my brother Coolman and “Young fella’ Sbanisenkosi Lafleni, and for motivating me to strive for more significant objectives and having faith in my capability to achieve them. I express my gratitude towards you.

- In recognition of my marvellous daughter Zanokuhle Siphokazi's instigation of this master's program, I extend my sincere appreciation for your perseverance during the times I was unable to be with you for "Daddy time", and for your unwavering silent encouragement amidst my challenges. Unlike me, you likely have a greater claim to this degree. Also your mom Nonsindiso Mkhize for taking care of you, when I was not around busy trying to complete my work

- To my supervisor and co supervisor, Professor Sumbwanyambe and Dr Hlalele, for the substantial contributions, assistance, and direction provided throughout the entirety of this research endeavour. Your consistent willingness to assist and promptness have been an immense relief. Sincere appreciation very much.

- To the other members of my family, for encouraging me to persevere and pursue my objectives and for being there for me during difficult times.
- To my phenomenal network of support, in times of pleasure and hardship, I appreciate you being my rock solid support system. Regarding my current state of self-isolation, I appreciate your support, encouragement, and understanding, as well as your prayers.
- To my Friend Christopher Shear, I know we do not speak often, but my friend you and your family have always been dear to me, you were the very first person who encouraged me to pursuit this degree. Thank you Mr Shear

To all individuals included in the aforementioned list, including any unintentional omissions, I also extend my wishes for heavenly benefits from the entity known as God, specifically in the form of the almighty Lord Jesus Christ, that exceed expected results. Moreover, it is hoped that this celestial being will satisfy the deep-seated cravings held within your innermost beings, with the ultimate objective of exalting his formidable name.

Dedication

This dissertation is dedicated to the late Simon Funani and Ntombizanele Pretty Lafleni, my Father and Mother, respectively. Despite your absence in celebrating this accomplishment, I consistently sense your influence in my daily existence. As I envision your gaze upon me, I aspire to avoid disappointing you throughout my own odyssey.

Table of Contents

List of Figure

List of Table

List of Word Abbreviation

4IR	Fourth Industrial Revolution
AC	Alternating current
AI	Artificial Intelligence
AMI	Advanced Metering Infrastructure
AMR	Autonomous Moving Robot
ANN	Artificial Neural Network
AR	Augmented-Reality
CBM	Condition-Based Monitoring
CNC	Computer Numerical Control
CNN	Convolutional Neural Network
CNN	Convolutional Neural Networks
CPS	Cyber-Physical Systems
CPU	Central Processing Unit
DC	Direct current

DFIG	Doubly-Fed Induction Generator
DG	Distribution generation
DL	Deep learning
DLGF	Double line ground fault
DN	Distribution Network
DT	Decision Tree
DWT	Discrete Wavelet transform
FPI	Fault Passage Indicators
GA	Genetic algorithym
HIF	High-Impedance Failures
IEEE	Institute of Electrical and Electronics Engineers
KBS	Knowledge base system
KKT	Karush-Kuhn-Tucker
kW	Kilo watt
LLF	Line to Line fault

LLLGF	three-phase to ground fault
MG	Micro-grid
ML	Machine learning
MW	Megawatt
OMS	Outage Management Systems
PR	Pattern recognition
PV	photovoltaic
RNN	Recurrent Neural Networks
RTU	remote terminal units
SCADA	Supervisory Control and Data Acquisition
SCIG	Squirrel-Cage Induction Generator
SLGF	Single line ground fault
SVM	Support Vector Machine

1. Chapter one: Introduction

1.1 Introduction of the research

Due to fast technological and scientific innovation, modern DN electrical protection systems are growing more complex. One of the most important requirements for power system current protection systems is fault diagnosis. When constructing and managing a system, the protection system fault diagnosis becomes the top priority. A protection system fault diagnostic detects, isolates, and identifies faults in operational data. Artificial intelligence (AI), a robust technique for identifying patterns, has attracted the interest of numerous academics and exhibits promise in the realm of detecting faults in electrical distribution networks. Recognizing fault patterns poses a challenge due to the extensive range and heterogeneity of response signals. Therefore, a typical fault diagnostic system comprises two main components: data processing, namely feature extraction, and fault identification.

1.2 Detailed problem statement

The distribution network encounters several challenges, with one of the key challenges being the distribution network protection system's response to the connection of distribution generation (DG) units. The extent to which distributed generation (DG) is included, together with the specific type of DG integration including components like synchronous machines or power electronics, has a significant impact on the functioning of protection systems. Figure 1.1 showcases the protection system being analysed, which has a radial structure. The layout of the protection scheme ensures efficient response to power flow in a single direction. This system is capable of providing reliable protection against fault currents occurring within the network. The incorporation of a distributed generation (DG) into the power distribution system, as depicted in Figure 1.2, introduces the possibility of fault contribution from DGs. This, in turn, alters the short circuit levels within the system. Consequently, it is conceivable that the fault current may exceed the breaking capacity of circuit breakers and fuses, thereby leading to malfunctions in the protection system. Due to the bi-directional fault current contribution, the tripping mechanism can experience a range of malfunctions. This can lead to traditional protection schemes failing to operate effectively, especially when there are distribution generation (DGs) connected to the distribution network (DN).

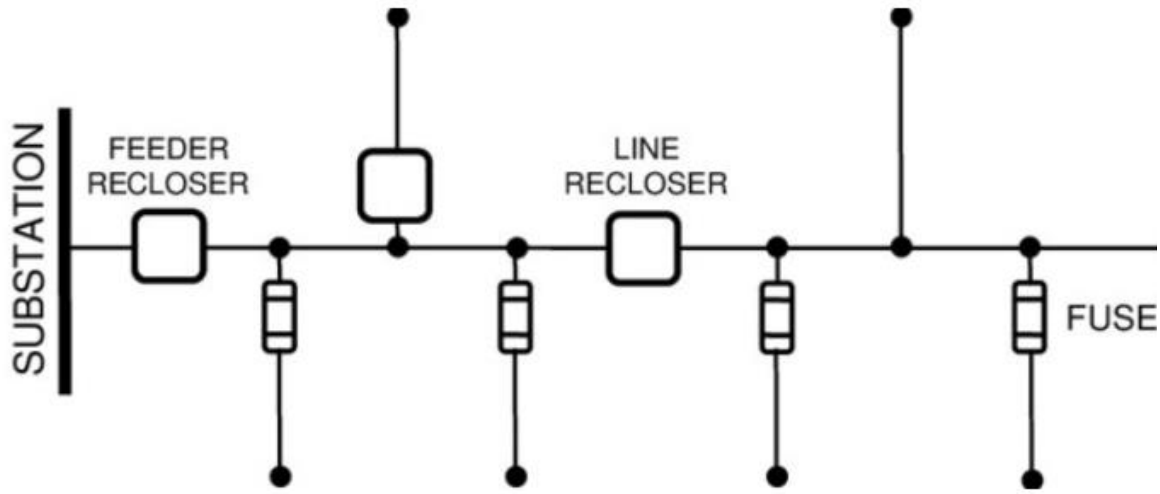


Figure 1.1. Main utility without DG integration[1].

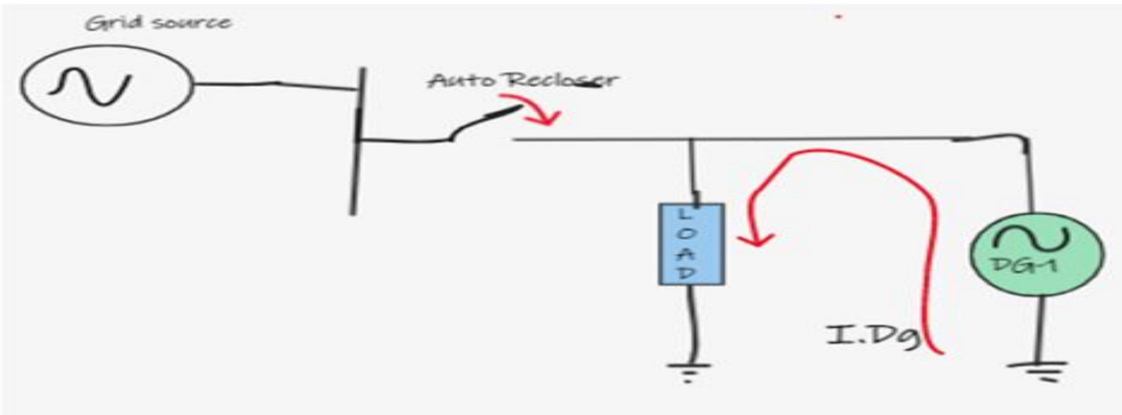


Figure 1.2. Grid source integrated with DGs.

1.3 Context of the problem

The challenge presented today on distribution networks is the urge for a low to free carbon economy globally and locally, this phenomenon can exert a substantial influence on the operations of the distribution system, particularly with regards to the incorporation of distribution generations into the grid. The increasing fascination with utilizing renewable energy sources in conventional electricity distribution networks has presented novel issues for electric power system operators and the network protection system. According to reference[2], the two primary factors that are essential for expeditious restoration of a malfunctioning component are the identification of the fault site and determination of the fault kind. The primary objective of a distribution network scheme is to ensure a consistent and uninterrupted supply of electrical power, even during abnormal conditions. The goal is to minimize any period of power outage to an absolute minimum. Power outages can be caused by Natural events, which can cause short circuits faults, which makes these faults being single phase either to ground or phase to phase or phase to phase to ground or a three-phase fault. The attributes of an electrical protection system are operating by applying and setting up protection devices (relays, fuses, recloser etc...) to ensure optimal sensitivity to malfunction conditions while minimizing the occurrence of false alarms under typical operating circumstances. According to Section 5.1.1.2 of the grid code standard, the prescribed guidelines for grid protection settings specify that protection relays must offer optimal protection against short-circuit. Additionally, these relays should offer remote and/or local backup for faults in the distribution network that have not been resolved. It is important to note that these relays should not be configured to trigger overload tripping [3].

These modern distribution networks with DG integrated provide difficulty to detect and locate the fault in the networks hence leading to regulation of grid codes being challenge to adhere to. Many artificial intelligent techniques can be adapted or can be implemented directly to solve complex power system challenges, which can be very powerful in identifying faulty condition, classification of fault in the distribution network with integrated with DG. Artificial intelligent (AI) techniques form such as knowledge-based systems (KBS) in fuzzy logic, neural networks and genetic algorithms also known as prominent methods are all first introduced and then investigated in terms of their applicability in the protection systems in the distribution network.

1.3.1 Research gap (Machine learning application in Micro-grid protection)

In an electrical distribution network, faults must first be identified, correctly diagnosed, and repaired as quickly as feasible. The relays that protect the electrical system from blackouts can be operational by using the DN protection system. A solid fault detection system offers a relaying function that is efficient, dependable, quick, and secure. In recent decades, The issue of energy usage has emerged as a prominent worry for numerous nations worldwide. This is because certain countries have limited energy supplies and are concerned about the effects that burning fossil fuels has on the environment. As a result, a new strategy has been put into place that aims to minimize the use of fossil fuels at all levels and explore sources of renewable energy. As a result, experts and decision-makers are paying close attention to the topic of renewable energy. Renewable energy sources encompass solar, wind, geothermal, hydroelectric power, biomass, and biofuels. These resources are readily accessible on a global scale and have the potential for renewal over a period. Renewable energy possesses the potential to function as a feasible alternative to fossil fuels in the realm of power generation, owing to the inherent spatial constraints and limited availability of the latter. Distributed generation (DG) refers to the utilization of technologies that generate electricity from renewable energy sources in close proximity to or directly at the point of consumption. In order to uphold the necessary level of dependability in the power system, it is crucial for these recently developed technologies to furnish the operator with a timely and precise response to any anomalous conditions.

The voltage and current levels of the distribution network are modelled according to normal distribution system levels using Simulink software. The network construction is modelled using Matlab Simulink and parameter data taken from an already successful network, and voltage and frequency level adjusted to local standard. The data signals from the sensors located at the bus-bars are utilized for the purpose of gathering and analysing the current signals in both steady-state (normal condition) and transient state (fault condition). The acquired data is subsequently subjected to pre-processing and labelled for feature analysis, enabling it to undergo testing and training in order to identify the optimal machine learning classifier for multi-class fault classification. The performance of the proposed algorithm will be evaluated by considering various parameters, including fault types, fault location, DG location, and the number of DGs. The evaluation results will be presented in terms of classification accuracy. There are many types of classifiers but it is difficult to determine the best performing. Hence, the performance can have evaluated using various tools such as confusion metrics, scatterplot for the validation of the analysis

1.4 Research objective and questions.

This research work proposes a fault diagnosis method by application of algorithm based on machine learning in a micro-grid with Solar PV, BESS, and non-linear loads in which the performance of the fault classifiers will be evaluated in efficiency and accuracy to observe the best machine model for the Micro-grid protection system. The objective is to use a model that uses DGs interconnected to the grid namely the, Solar PV with inverter and a battery energy storage system unit which will be a MG test system which is implemented as a case study that will allow us to analyse the feasibility of implementation of the ML algorithms for speedy, straightforward and cost-effective means of fault diagnosis in order to increase the distribution network system reliability and security. With the case study model, the signals of current and voltage of the system are very complex, it becomes impossible to determine the correct classifier without first testing all the available classifiers to determine the best classifier model for that micro-grid. The best classifier can only be determined in the micro-grid model by first training and testing before we can conclude the best classifier for fault diagnostic.

1.4.1 Research questions.

The research problems associated with the problem

- What are the different types of Distribution generation sources that can be integrated into the grid and how does distributed generation assist the grid?
- What are the various fault types that manifest within the distribution network?
- What are different types of Machine Learning fault classifiers that can be implemented for fault classifying in micro-grid protection schemes?
- What data signals for fault classification are required?
- What is the difference between the Binary and Multi-class fault classification method? And how flexible and responsive is the artificial neural network techniques to changing conditions in the distribution network.

1.4.2 Research Objective.

- To simulate Distributed Generation (DG) interconnected into the grid which are referred to small-scale power generation technologies and define the different types of DG sources with minimum of two implemented into the micro-grid.
- Define types of faults that are possible on the Micro-grid and then create symmetrical fault at one location of the micro-grid and collect the different types fault simulated current and voltage waveforms at the Point of common coupling (PCC) as by theoretical definition, of which these waves forms will be converted into numerical values which is required for machine learning modelling.
- Defining classifiers that can be implemented for Fault diagnosis in Micro-grid and the practicing their ability by Train and test all Machine Learning Classifiers available in Matlab Simulink to determine the best fault classifier model for multi-class fault classification.
- Collection of faults and normal waveforms signals of the simulated micro-grid at different bus-bar locations after collection pre-process and label these of large data set of current and voltage into smaller samples.
- By generating faulty and healthy data, and then to implement a multi-class classification by using different classifier to detect different combinations of faults which are 6(six) different types of faults which a twelve total faults rather then and a combination of 2(two) which is Binary classification.
- Determine the most efficient classifier in terms of accuracy using confusion metrics and scatterplot tools in classification learner.

1.4.3 Goal of the research topic

The goal is to simulate a hybrid micro-grid distribution network with DGs integrated into the grid implement and show the effectiveness of the Machine learning tools for fault detection and diagnosis for the required outcomes in classification,. This study's findings will help the network operators in guarantee more precise methods of fault detection and diagnosis as well as minimize damages and reduce workforce waste throughout the process of fault localization in these futuristic distribution networks.

1.5 Research methodology

- A substantial literature review is conducted in the initial stages of the research with technical papers nevertheless dominating the literature survey. The literature review is done on by confirmed previous work that has been completed on the similar topic of distribution generation intergrade on the distribution network protection system, which has enlightened the author on general aspects of the topic.
- A plan and procedure are created using the research technique, including full assumptions, detailed modelling, data collecting, analysis, and interpretation. We employ abstract investigation and simulations to solve the problem via logical research
- The second stage of the research is the modelling of the micro-grid which will be used as the test bench and hence the calculations of different parameters of the models and control circuits inverters and energy storage circuits will be required. This stage also includes specification of different characteristics of a South African like distribution network especially the voltage levels which will conclude the ideal model of a Micro-grid test system that would be suitable for carrying out studies for the research.
- In the third stage, the hybrid AC micro-grid distribution power system network after being modelled and successfully operating using Matlab software at supply frequency of 50Hz then data collection of three phase voltage and current signals under normal conditions and fault conditions will be collected at chosen locations of bus-bars and then which also, fault analysis for Single Line to ground fault, Line-to-Line fault, Double Line to ground fault and three-phase fault is done.
- This data is pre-processed and labelled accordingly in a tabular form for feature extraction using Matlab feature diagnosis tool to find the best feature for the model in terms of accuracy.
- After feature extraction then we use classification learner for multi-class fault detection and classification to see the best classifier for the modelled micro-grid by training and testing and observing the classifier with the highest accuracy.

1.6 Dissertation outline

The outline for the complete dissertation will contribute five chapters in the following order and format. Chapter three elaborate on ML methodology process of accuracy validation, with chapter two explaining the different types of methodologies which guides the chapter three methodology.

Chapter 1: Introduction

This chapter is written to signify the problem; to set the scene for the research and give an overview of the main components. The introduction will be written in such a way that it is both instructive and self-contained.

Chapter 2: Literature Review

This chapter situates the subject in the context of past research and academic literature on artificial intelligence methodologies for fault detection, location, and categorization. The purpose of Chapter 2 is to give a critical mix of experimental literature organized by key topics or variables. By establishing the theoretical and conceptual framework of the study, the thesis aims to demonstrate how the issue solves a gap or problem in the literature.

Chapter 3: Methodology

The study sets the approach inside a certain methodological method in this chapter. They will include a description of the model and sample, as well as the data gathering method used and the analysis methodologies and tools employed. The Methodology phase will provide a comprehensive overview of the study's design and methodology.

Chapter 4: Analysis and results

This chapter integrates and discusses the results from the simulation of the model based on the research questions, literature review, and conceptual framework of the study spotting patterns and themes, as well as finding discrepancies in the results, will be one of the approaches employed in the analysis. This will provide opportunity to think deeply about the study's findings as well as their practical and theoretical consequences.

Chapter 5: Conclusions and Recommendations

Conclusion and recommendations will be presented in this chapter of the dissertation as a set of concluding remarks and recommendations. That is, they will be claims based on evidence, which must be supported by the evidence. Its goal is to reflect an integration of the study findings, analysis, interpretation, and synthesis for each discovery

2. Chapter Two: Literature Review

2.1 Introduction to the chapter

This chapter includes a literature overview of the application of artificial intelligence (AI) technologies in a distribution protection system with the goals of identifying, categorizing, and detecting problems in the distribution network. This chapter also covers fault detection strategies in micro-grid and fault detection techniques in power systems.

The three components of an electric power system are generation, transmission, and distribution. Electricity is generated at the electricity generating facility and then delivered to distribution stations via transmission lines. The transmitted line is used to supply power to users via distribution networks, which provide the necessary power, with the basic goal of a power system networks being to provide clients with uninterrupted power. The electrical power system can be susceptible to occurrence of fault at any time, which cannot be avoided due to natural or uncontrolled events. When a fault develops in the system, it is necessary to identify the faulty line and remove it from the grid in order to keep the power flowing and the service running smoothly for customers. The extent of the breakdown is determined by several factors, including the fault current, the position of the short circuit, and the voltage level. The challenge of recognizing and classifying distribution network failures has long existed, and it will only get more problematic with the addition of DG to the DN, which will have a detrimental impact on the DN's protection strategy.

2.2 Types of Fault

Many literatures have mentioned in writing of the four kinds of fault that can transpire in distribution systems which are important in protection systems of the distribution network which are namely: single line to ground fault (SLGF)[4], 2)double line to ground fault (DLGF)[4], 3)line to line fault (LLF) and 4) three-phase to ground fault (LLLGF)[5]. General knowledge explains occurrence of the single line-to-ground fault in the distribution network line as a situation when one conductor drops to the ground or comes in contact with the neutral conductor.

The electrical distribution network's fault detection system monitors one or more electrical conductors and produces an input signal to represent the multitude of electrical signal settings in the section of the network to be monitored for fault isolation and identification. Fault diagnosis can be explained as in [6]where the type of fault, together with the magnitude ,location and time period of the detection of a fault in the system is found.

2.3 Types of Distribution generation and fault current contributions

2.3.1 Squirrel-Cage Induction Generator (SCIG) and Doubly-Fed Induction Generator (DFIG)

In electrical power systems, there are several types of distribution generation and fault current contributions. These can include the following: Historically and presently, induction and synchronous machines have been another type of distribution generation being used for application in thermal, hydro, and wind generation plants.

Squirrel-Cage Induction Generator (SCIG) and Doubly-Fed Induction Generator (DFIG) are the most common induction generators used in wind power generation. Wind turbines and certain smaller hydropower systems generally incorporate induction generators into their design. In order to produce power and induce voltage, these machines are asynchronous and rely on the relative motion of the magnetic fields in the rotor and the stator[7]. Wind power is a renewable energy source that generates clean electricity. The turbine nacelle is the main component of a wind power system, housing all the generating components in a wind turbine, as depicted in figure 2.1 in which the turbine rotor captures mechanical energy, which is converted into electrical energy by the generator. The generated electrical energy must be regulated and conditioned before being integrated into the power grid for interconnection use.

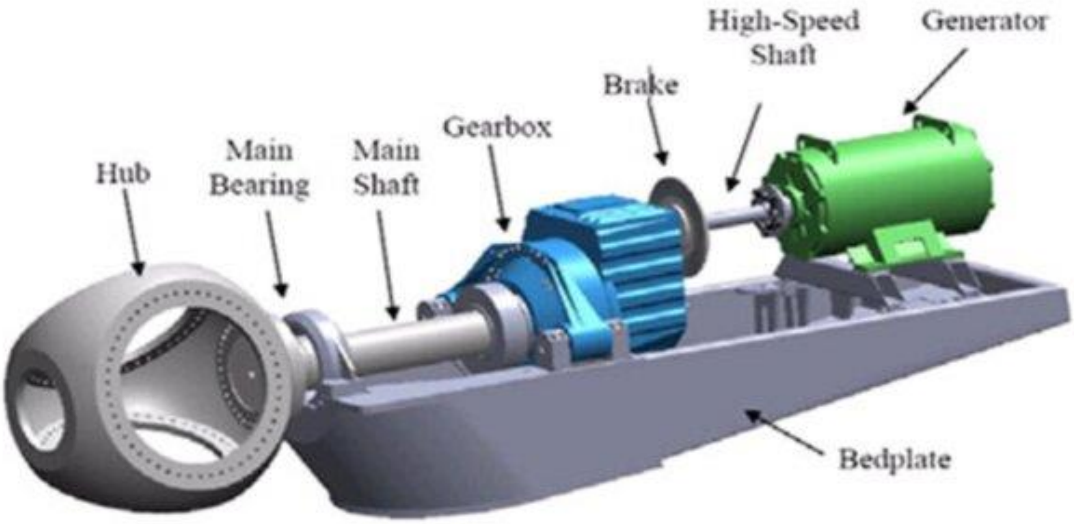


Figure 2.1. The general layout of the Wind power nacelle[8]

The main element is the power electronics, which allows for very effective and flexible regulation and control of power, voltage, and frequency. The power electronics interface between energy generation, storage, and transmission, which serves as a link between wind turbines and the grid, is playing an increasingly important role in distributed power systems as the amount of wind generating increases.

2.3.2 Solar PV Distribution Generation

Photovoltaic (PV) technology has a lengthy history of bringing the concept of solar energy to life. The utilization of solar energy for powering satellites in outer space stands as one of the pioneering applications of this technology. In the year 1973, the esteemed University of Delaware undertook the commendable task of erecting the pioneering solar edifice, aptly named "Solar One"[9]. The inaugural case of "distributed generation" was realized with the successful completion of the initial grid-connected solar photovoltaic (PV) system in Kerman, California during the 1990s.

PV-DG may be broadly categorised into three sorts. The first being the Utility-Scale PV-DG which are sized in Megawatt (MW) (1-10MW)[10]. In their integration to the grid, the requirement is that they have to have nominal capacities which are compatible to conventional feeders or to distribution substations substation ratings with the grid. In order to effectively handle the conversion of energy from direct current (DC) to alternating current (AC), it is necessary to incorporate multiple power-electronic inverter modules that operate in parallel. These modules are equipped with both internal and external protection systems to ensure the safe and reliable functioning of the system. The protection systems include fast overcurrent protection, safeguards against under and overvoltage and frequency deviations, as well as active anti-island protection schemes. The purpose of the anti-island protection is to prevent the photovoltaic-distributed generation (PV-DG) plant from supplying power to the grid in the event of a loss of utility grid connection.

The second category is Medium-Scale PV-DG (10–1,000 kW), which has installations on small to big structures such as apartment complexes, traditional retail stores, government building rooftops, or industrial parks. Their usual interconnection structure may resemble that of utility-scale PV-DG plants, with separate interconnection transformers and smaller plants with PV-DG capacity equal to the demand they supply. Small-Scale PV-DG Capacity ranges from 1 to 7 kW, with solar panels put on the roofs of customer houses and connected to the customer side grid, which is typically 230V and single phase, producing more or less electricity than necessary by the customer's load.

2.3.3 Synchronous DG's Impact on Fault current contribution

The operation and design of a synchronous generator can be studied in many information sources including [11] where synchronous based DG technology has been used in distribution power systems for a long time due to it benefit of ability to increase the reliability in the power system and decrease environmental impact concerns. In [12] the abstract highlights local interconnection of generators in distribution grids has consequences for the protection systems in which the emphasis is placed on the connection of generators to the radial distribution system, as well as considerations such as loss of selectivity, overcurrent protection level, earth-leakage protection, disconnection of generators, islanding, and single-phase connections [12]. When DG is used as a micro grid, it provides reactive power correction, increases system stability, and reduces voltage sag. Synchronous generators are included in distributed generators; however, they contribute to the most fault current compared to the other types of DGs.

2.3.4 Summary of DG's Impact on Fault current contribution

Contributions to fault current can originate from a variety of sources, including synchronous generators, induction generators, distributed generation sources, and even large industrial loads. It is important to note that the features of the power system, the installed generating capacity, and the type of fault being evaluated can all affect the specific types of generation and fault current contributions. Furthermore, new types of generating sources and their related fault current contributions can be introduced due to improvements in technology and changes in power system designs. Distributed generators also include synchronous generators; nevertheless, in comparison to the other kinds of DGs, synchronous generators are responsible for the greatest amount of fault current [13].

2.4 Conventional fault detection techniques in distribution network

Power system fault diagnosis has always been broken down into two distinct groups in their methodology: namely a) based on post-fault line impedance measurement, and b) method that collect data by monitoring the signal a malfunction causes. Therefore, constant monitoring of voltage, current, impedance, etc. is essential for speedy fault-finding and repair in power distribution systems, and it increases power networks' dependability. The distribution network does not need to be overhauled drastically to accommodate hybrid micro grids, a new technology that combines the best features of AC and DC micro grids. However, due to the incorporation of numerous forms of Distributed Generation (DG), its protection confronts a number of issues. The quantity and direction of short circuit current are both affected by the incorporation of DGs into the power system. Therefore, the complexity of a micro grid's overcurrent protection design increases. The distribution network continues to face various problems that arise for a variety of reasons, affecting distribution network reliability and potentially resulting in costly repairs, lost productivity on operations, and power loss to customers. Traditional networks and modern networks will face these realities as the grid becomes more diverse in terms of power sources.

In distribution networks, various conventional fault detection techniques have been used for the purpose of identifying and locating faults. Below are several frequently employed techniques:

Overcurrent Protection: overcurrent protection is a commonly employed method for identifying and addressing faults that arise due to excessive currents. The relays are designed to activate and interrupt the current in a specific section of the distribution network when it surpasses a predetermined threshold.

Directional overcurrent: protection is a type of relay that enhances discrimination capabilities by taking into account the direction of current flow. Selective isolation of faulted sections is achieved by utilizing relays that are specifically designed to trip only in the direction of the fault.

Distance protection: is a type of relay that calculates the impedance between the location of the relay and the point where a fault occurs. The detection and isolation of faults can be achieved by comparing the measured impedance with a pre-set impedance characteristic.

Differential protection: involves the use of relays to analyse and compare the currents that flow into and out of a specific area of protection. The presence of a significant difference suggests a potential fault within the designated area. In response, the relay will activate and disconnect the faulty section to prevent further issues.

Fault Passage Indicators (FPIs): are devices that are strategically installed on distribution feeders with the purpose of promptly detecting and indicating the occurrence of faults. Visual or audible indicators are useful in pinpointing the location of a fault by reducing the search area.

Fault location algorithms: are employed to determine the distance to the fault point after a fault has been detected. The algorithms make use of data obtained from measurement devices like current and voltage sensors, along with information about the network topology.

Outage Management Systems (OMS): refer to software systems that aim to enhance fault management by integrating fault detection, location, and other distribution network data. The system automates the process of fault analysis, allowing for quicker fault restoration by providing real-time information to operators.

Supervisory Control and Data Acquisition (SCADA): systems are utilized to effectively monitor and control the distribution network. Data is collected from remote terminal units (RTUs), enabling operators to identify abnormalities and potential faults in real-time.

The conventional techniques discussed in this chapter serve as the fundamental basis for detecting faults in distribution networks. It is important to mention that there are newer technologies, like advanced metering infrastructure (AMI) and distribution automation systems, which are being implemented to improve fault detection and response capabilities.

2.5 Theory background of AI Fault detection methods

In recent times, there has been rapid progress in fault techniques for diagnosing, detecting, and classifying faults. This progress can be attributed to advancements in signal processing hardware, artificial intelligence, Global Positioning System (GPS), and communication technology. These advancements have created new opportunities for researchers to study the limitations of traditional distribution networks when connected to the grid. Intelligent electronic devices are used instead of traditional measuring devices, such as current transformers and voltage transformers, to overcome challenges in data acquisition and gather a substantial amount of data. Researchers can utilize big data to create advanced and intelligent methods for detecting faults.

2.5.1 Logic flow

In the absence of machine learning techniques, a study in [14] has suggested the use of combinational logic and fuzzy logic methods at the component level for fault identification. In order to assess performance [15] has introduced an a fault detection system that relies on data obtained from the IEEE 9 bus model, using a three-step logical process. In [16] the data is extracted at both ends of a transmission line, and a three-phase fault implementation is carried out to detect, classify, and locate the fault. If a method for detecting high impedance faults (HIF) in distribution systems is utilized, [17] provides a comprehensive description of the logical sequence to resolve HIF systems.

During the time frame of 2000-2010, the concept of grid diversity in the energy mix was considered as a forward-thinking approach. In fact, researchers such as [18] suggested the use of fault detection approaches that rely on obtaining measurement parameters using current transformers. The integration of Wireless Sensor Networks (WSNs) in Smart Grid as a tool for fault detection in the logic flow process has emerged as a prominent opportunity. This is because traditional technologies have been perceived as a disadvantage due to their high cost when implemented in a smart grid. A study,[19] has summarized the opportunities and challenges in wireless communication, highlighting the need for a robust data transmission method for effective functioning. The discussion in [20] concluded that the Smart Grid's distribution network was unable to function without a strong and reliable data transmission infrastructure. The Smart Grid has the capability to support substantial volumes of Distributed Generation (DG), although it will necessitate the development of new DG connectivity standards. In addition, by utilizing data transfer and distributed processing, DGs of any size can be effectively monitored, controlled, and evaluated.

2.5.2 Fuzzy Logic

For the regulation of intricate and non-linear systems, fuzzy logic is an effective modelling technique. In the realm of AI system implementation, fuzzy logic is leveraged to emulate the cognitive processes and reasoning abilities exhibited by humans. Fuzzy logic is an algorithmic approach utilized for decision-making, wherein a meticulously constructed rule base is employed to accommodate a multitude of conflicting and indeterminate factors. The concept is derived from notions that acknowledge and exploit the intermediate region between two polar opposites, and is governed by rational equations as opposed to intricate differential equations [21]. A simplified explanation is provided, as depicted in Figure 2.2, illustrating a variant of the intelligent fault detection process cycle.

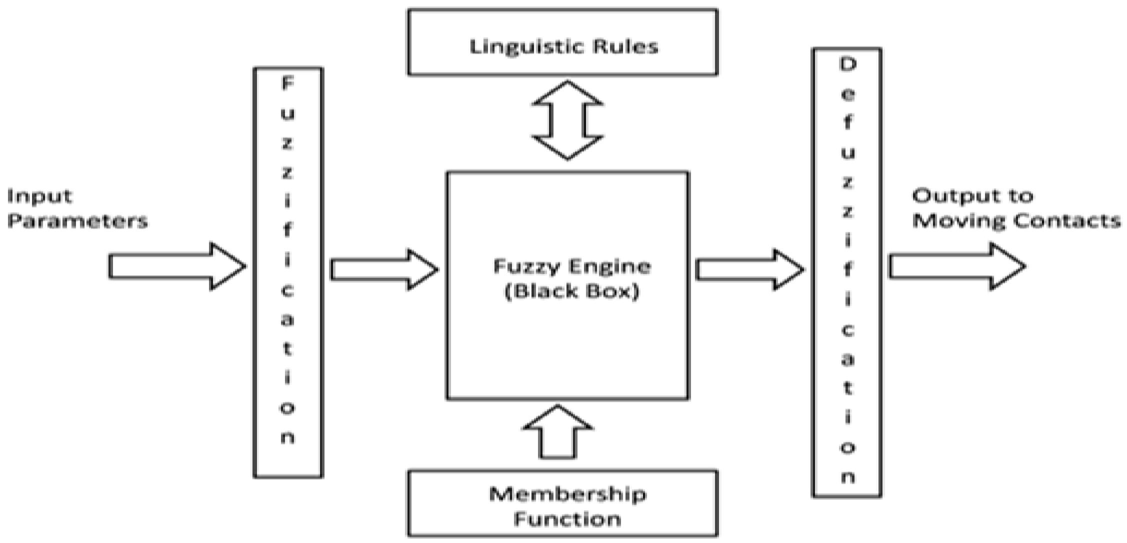


Figure 2.2. Fuzzy logic Intelligent fault detection process [22].

2.5.3 Artificial neural network

One of the ingenious techniques utilized for identifying issues in distribution systems involves the utilization of artificial neural networks (ANNs). As a result of the artificial neural network's advanced pattern recognition capabilities, the identification of errors can be accomplished effortlessly. However, in order to identify the underlying issue in the relationship between the input data and the predicted output, it is necessary to undergo a training process. A classification-oriented neural network comprises four layers, namely the input layer, hidden layer, softmax layer, and output layer. The voltage and current of the three-phase neural network are trained by feeding the input values into a processing unit, which, in this case, consists of six inputs. The neural network in the distribution line produces four outputs, namely the line voltages and currents of the three phases A, B, and C, along with the ground. The utilization of pattern recognition technique in the context of electrical power distribution system involves a systematic approach to identify and distinguish between abnormal and normal conditions. This approach draws inspiration from the functioning principles of the brain and nervous system, as depicted in Figure 2.3.

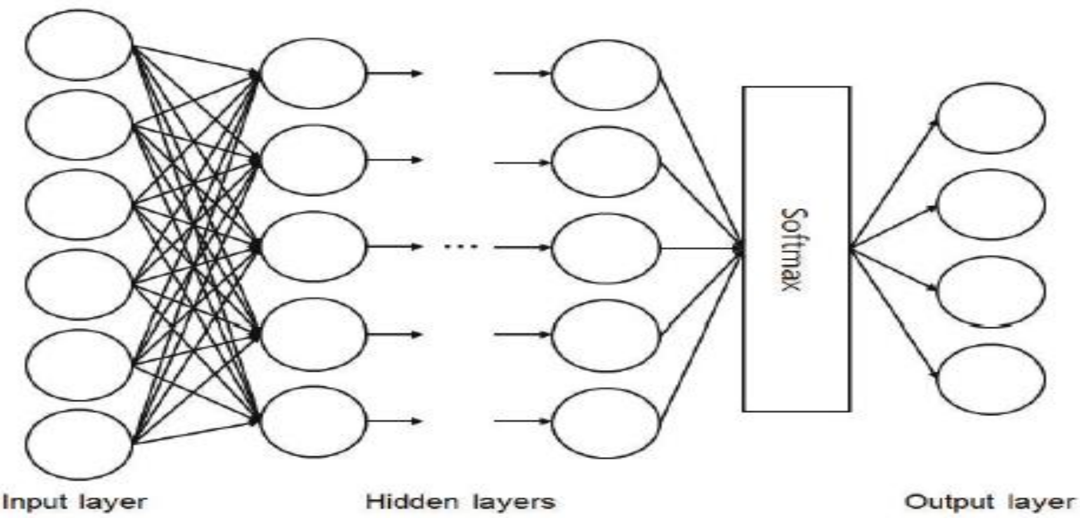


Figure 2.3 The typical topology of a neural network for doing classification[23].

Feature scaling of the aforementioned nature can be employed to normalize input data before being fed into the input layer, as is customary for data-driven fault diagnostic methodologies, guaranteeing that all values are confined within the specified range [24]:

$$X' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (2,1)$$

The input data undergoes a series of nonlinear transformations in the hidden layers, resulting in progressively more sophisticated representations (features):

$$h_1 = \sigma(\omega_{1x} + b_1) \quad (2,2)$$

$$h_l = \sigma(\omega_l h_l + b_l), l = \{2 \ldots d\} \quad (2,3)$$

Where X ϵ R^{n_x} is the input vector h_l ϵ R^{nhl} is the hidden representation, W_l ϵ $R^{nhl\ X\ nhl-1}$ is the weight matrix, b_l ϵ R^{nhl} is the bias vector and d is the number of hidden layers. It is important to remember that n_{hl} (the amount of neurons in each hidden layer) with d are hyper parameters whose values must be established before neural networks can be trained. To make the aforementioned nonlinear transformation is used in this work, and the rectified linear unit (ReLU), which is defined as:

$$\sigma(x) = \max(0, x) \quad (2,4)$$

The output of the preceding hidden layer experiences the transformation as illustrated in h_1 by equation (2) without the activation function:

$$h_s = \omega_s h_d + b_s \quad (2,5)$$

Furthermore, the softmax layer employs the subsequent softmax function to ascertain the values of every output neuron.

$$y_j = \frac{esp(h_s, j)}{\sum_{j=1}^{n_{hs}} esp(h_s, j)} \quad (2,6)$$

2.5.4 Support Vector Machine (SVM)

The Support Vector Machine (SVM) class of learning machines has demonstrated effectiveness in addressing Pattern Recognition (PR) problems [25]. The maximization of the margin in classification problems can be seen as an equivalent approach to the conventional framework. However, by focusing on the minimization of the weight vector, engineers can achieve enhanced control over the generalization of the problem. The solution can be obtained by utilizing a set of support vectors that are sparsely distributed. In order for classification to occur, it is imperative that the separation surface, regardless of whether it is linear or non-linear, be implemented within the input space. The utilization of linear combinations of kernels associated with Support Vectors can be employed for the purpose of defining the separation function in Support Vector classification, as exemplified in the context below:

$$f(x) = \sum_{zj \in s} \alpha_j y_j k\left(x_j, x\right) + b \quad (2.7)$$

Training patterns are denoted by x_i class labels by $y_i \in \{+1, -1\}$ and the collection of Support Vectors is indicated by S. Therefore, the dual formulation results in:

$$\min_{0 \le \alpha_i \le C} = \sum_{ij} \alpha_j Q_{ij} \alpha_j - \sum_i \alpha_i + b + \sum_i y_i \alpha_i \quad (2.8)$$

Where a_i the respective coefficients, b are is an offset matrix, and a_i are related coefficients. In the inseparable situation, the symmetric positive definite kernel is denoted as $Q_1 = y_1 y_j \mathrm{K}\left(x_i, x_j\right)$. Karush-Kuhn-Tucker (KKT) dual conditions can be written as

$$g_i = \frac{\partial w}{\partial w} = \sum_i Q_{ij} \alpha_j + y_i b - 1 = y_i f(x_i) - 1 \quad (2.9)$$

And

$$\frac{\partial w}{\partial w} = \sum_i y_{ij} \alpha_j = 0 \tag{2.10}$$

This partitions the training set into S the support vector set (0<a_i<C, g_i=0), E the error set (a_i=C, g_i<0) and R the well classified set (a_i=0, g_i>0)

2.5.5 Discrete Wavelet Transform (DWT)

The utilization of the wavelet transform (WT) is employed in order to disintegrate a signal into a collection of "wavelet" basis functions across both temporal and frequency domains [26]. The wavelet transform utilizes the expansion and contraction of fundamental functions to discern distinct frequency components within a signal. Utilizing the wavelet transform methodology, it is possible to decompose a given signal into its constituent frequencies. The utilization of the mother wavelet as the fundamental basis function is due to its inherent properties of dilation and translation. In this scenario, the low-frequency component of the signal is obtained through the utilization of wide windows, while small windows are employed to capture and represent any discontinuities present in the signal.

$$Wf(m,n) = 2^{m/2} \int f(t) \varphi\left(2^{-m} t - n\right) dt \qquad (2,11)$$

To where m corresponds to the frequency and n to the duration. The typical formula for a wavelet series is

$$f(t) = \sum_{k=\infty}^{k=\infty} c_k \phi(t-k) + \sum_{k=-\infty}^{\infty} d_{ik} \phi\left(2^i t - k\right) \tag{2.12}$$

$$\phi(x) = \sqrt{2} \sum_n h_o \Phi(2x - n) \tag{2.13}$$

Where h_1 $f(t) = \sum_{k=\infty}^{k=\infty} c_k \phi(t-k) + \sum_{k=-\infty}^{\infty} d_{ik} \phi\left(2^i t - k\right)$ the high pass is filter coefficient and Φ(x) is the wavelet function.

2.6 Evolution of intelligent techniques in power system protection

Significant progress has been made in the application of artificial intelligence (AI) approaches for fault detection in electrical distribution networks that are equipped with distribution generation because of the growth of intelligent techniques used in power system protection. Some of the most well-known AI methods of the time are discussed below.

Reference [27] discusses supervised machine learning architecture-based machine-learning algorithms, in which there is demonstrated performance analysis in fault categorization and location identification. Historically Fault detection using ML algorithms is common. ML techniques like SVM, Random Forests, and Neural Networks can now detect distribution network issues. Based on past data, ML methods have classified, located, and predicted faults.

Simultaneously, there has been an upsurge in the utilization of Deep Learning techniques for the purpose of fault detection, specifically Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). In the realm of problem diagnosis and classification, deep learning models exhibit exceptional performance owing to their remarkable capability to extract intricate features from vast volumes of data. The utilization of deep learning techniques for the detection of distribution network failures has exhibited promising outcomes.

While more sophisticated methods such as machine learning and deep learning have taken precedence in defect detection due to their ability to manage intricate patterns and vast volumes of data, fuzzy logic has been utilized in fault detection to address imprecisions and uncertainties. The application is based on the assessment and diagnosis of fault conditions using linguistic factors and fuzzy rules, utilizing measurable data. Fuzzy logic has demonstrated its utility in addressing the inherent uncertainties linked to the detection of faults in distribution networks.

2.6.1 Artificial Neural Network

Artificial Neural Networks (ANNs), being one of the pioneering methodologies in the realm of artificial intelligence, have been extensively employed in the domain of power engineering research for a considerable duration, in fault detection and classification resulting highly innovative outcomes. The verification of its application is duly referenced [28, 29]. Studies by [30] proposes a novel method for detecting and classifying problems in transmission lines by analyzing oscillography data. In [31] the oscillography data is utilized as instantaneous current/voltage samples, which are then utilized as inputs for artificial neural networks. According to empirical observations, the utilization of neural networks and genetic algorithms (GAs) in practical applications is accompanied by inherent challenges pertaining to the training process[32] The author proposed a solution for power systems fault diagnosis by utilizing a Bayesian networks-based approach. The method utilizing a Bayesian network exhibited enhanced outcomes in comparison to the preceding approach, as depicted in [33]and [34]. The research proposes the use of an Artificial Neural Network (ANN) to improve problem diagnostics in a deregulated distribution system that experiences high impedance faults (HIF) which is considered by [35]. This new environment presents novel challenges that impact the operation, control, and protection of the system.

Artificial neural networks (ANNs) are employed as the preferred method for identifying defects in PV arrays. The operation and definition are authored by [36] additionally, it involves the establishment of a certain approach for carrying out its operations. Regarding power system engineering [37], the neural network architecture is employed to detect faults in a power transmission line system. Its primary objective is to establish a comprehensive structure for distance protection, with individual zones being segregated by neural networks. The faults being examined include single phase, two phases, and both phases to ground.

The author [38] provides a comprehensive explanation through a range of scenarios a novel method for detecting defective segments and precisely locating faults in power system distribution networks is proposed, which combines the Discrete Wavelet Transform (DWT) with artificial neural networks (ANN).

Studies have also been conducted on the topic of Fault detection and Localization in DC micro-grids using Artificial Neural Networks, as mentioned in the article [39] which presents an innovative approach utilizing artificial neural networks (ANNs) for the purpose of fault detection and localization in a micro-grid featuring a low voltage DC bus. According to the provided information, the implementation of the suggested technique can lead to the achievement of a more dependable DC micro-grid [39] due to its remarkable capability of swiftly identifying issues on the direct current (DC) bus and effectively isolating them without the need to de-energize the entire system. In order to ensure utmost precision, the neural network undergoes training utilizing real-world instances of DC bus short circuit scenarios.

The utilization of online condition monitoring and protection technology is imperative owing to the challenges associated with identifying and detecting high impedance faults in power distribution networks. This has had a notable and favorable impact, particularly in systems that employ dispersed generation. This fact is confirmed and demonstrated in [40], where a technique is introduced for identifying and describing High Impedance Faults (HIFs) in Distribution Generation (DG) systems. The implementation of this technique displays nonlinearity. A study was done to examine the influence of using various neural network topologies and data input modalities on fault detection, in order to highlight the wide range of approaches applied.

An intelligent fault classification technique is being developed by leveraging the capabilities of an artificial neural network with [41] utilizing the isolation model as the foundation for artificial intelligence and advanced signal processing methodologies.

Recent presentations have demonstrated the utilization of artificial intelligence (AI) models for devising solutions to issues pertaining to adaptive system protection. Esteemed authors, for instance, have extensively discussed this subject matter[42] and assess the methodology employed in the identification and categorization of anomalous states within a system. It is imperative to emphasize that for training and evaluating AI models, these methodologies leverage the existing data. When validating ML efficiency in application, it is important to consider authors who have conducted extensive research and analysis in the field such as [43] who use DNNs, short for deep neural networks, are a variant of artificial neural networks (ANNs) characterized by their multi-layered architecture. These networks consist of a multitude of hidden layers, composed of interconnected neurons, which are strategically positioned between the input and output layers. An intelligent fault detection method is described in Reference [44], which offers quick information on the fault's kind, phase, and position for the sake of micro-grid protection. The fault detector that is used in the solution is developed with the help of the discrete wavelet transform and deep neural networks (NNs).

2.6.2 Support Vector Machine:

The strategy based on artificial neural networks and support vector machines for finding defects in radial distribution systems is described by [45] with SVM being clearly defined. You will find Studies conducted after big power outages can be throughout the world found that faults in the protection system were to blame or miss-co-ordination being another cause. This has highlighted the requirement for further post-fault and remedial research utilizing intelligent/knowledge-based systems to enhance protective coordination. After a fault in a surrounding line that feeds into the same substation can be more accurately diagnosed using this method. Possible benefit includes enhanced fault monitoring and diagnostics for safer power system operation, this is studied with success in [46]. By giving attention to development of technology with a viewpoint of preventing catastrophe before any breakdown, [47] in which the system expedites the provision of data pertaining to the fault's type, phase, and location, thereby enabling the facilitation of micro-grid protection. The employed defect detector in the solution is formulated by leveraging the discrete wavelet transform and deep neural networks (NNs).The process of identifying electrical equipment involves the utilization of a support vector machine (SVM) as a classifier and the Zernike moment as an image feature. The utilization of Support Vector Machines (SVM) in conjunction with other intelligent techniques can be feasibly employed in the domain of fault detection and classification, as exemplified by [45] for a Radial Distribution systems. This reference in [48] uses Utilization of the Support Vector Machine (SVM) algorithm for the purpose of identifying the defective line within a substation, as well as determining the precise distance to said faulty line. Support Vector Machines (SVMs) utilizing Radial Basis Function (RBF) kernels possess the capability to acquire knowledge pertaining to the relationship between voltages and currents, as well as the identification of fault source and its corresponding location. While the data collected during a failure event in real-world power systems may have certain limitations, it is worth noting that such measurements can provide valuable insights into the condition of substations and aid in pinpointing the location of faults. Author [49] article elucidates the procedural steps required for the processing of said data, with the ultimate objective of fabricating a fault diagnostic system that exhibits optimal efficacy.

In the preceding decade (2010-2020), Support Vector Machines (SVM) have emerged as an advanced means for precisely identifying the segment of a defective line. SVM, a contemporary machine learning technique, has been developed based on the fundamental principles of statistical learning theory. In [50, 51] shows the empirical evidence suggests that the proposed strategy exhibits potential for practical implementation in real-world scenarios pertaining to smart grid distribution fault diagnosis. It is shown in [52] that Intelligent systems heavily depend on the fundamental principles of knowledge discovery in data techniques, as they have the potential to greatly enhance their overall performance. Despite the fact that the Support

Vector Machine (SVM) detection method has demonstrated remarkable performance capabilities in power application systems that are distinct from Distribution Networks (DN) with Distributed Generators (DGs), studies such as [53] demonstrate a possibility in integrating these detection methods power systems network like transmission line. By implementation of a support vector machine (SVM) classifier and Wavelet Transformation, reference [54] introduces a technique for fault classification in power systems. In[50], the author also presents a methodology for the identification and localization of malfunctions in power distribution networks that are integrated with Distributed Generation (DG) systems. The proposed methodology leverages Support Vector Machines to enhance the fault diagnostic procedure. At this stage protection of the electrical grid in modern DN, the limitations of conventional protection schemes are still present and hence authors like [55, 56] write about an approach on how to Evaluate the Performance of the Support Vector Machine (SVM) Approach for Distributed Generation fault Detection. Then authors like [57] demonstrate or introduce a novel pattern recognition-based methodology that has been devised to detect high-impedance failures (HIFs) within distribution networks. These HIFs encompass a range of issues, such as damaged conductors and arcs, among others. Several methodologies have been previously introduced to address the issue of high impedance faults (HIFs) such as presented in [58] and showed many disadvantages but in the 2010-2020 era authors such as [59] had presented ways to counter the disadvantage in [57, 58]. [60]suggested that a technique be developed that makes use of an algorithm that is based on machine learning (ML), namely support vector machine (SVM). The proposed SVM-based approach is designed to solve the important challenge of discriminating between islanding and grid fault occurrences. This allows the system to achieve better precision in islanding identification while simultaneously detecting grid failures in a more realistic manner. In this article [61], The development of Modified Multi-Class Support Vector Machines (MMC-SVM) technology aims to enable the simultaneous identification and categorization of various types of open-circuit failures that may potentially arise within power distribution systems. The strategy developed by [62] is based on PSO methods and aims to optimize both input characteristics and SVM characteristics simultaneously in order to classify the different fault types observed in the distribution network. The fault types can be categorized into ten distinct categories, including single phase-to-ground faults (AG, BG, and CG), line-to-line faults (AB, AC, and BC), double line-to-ground faults (ABG, ACG, and BCG), and three-phase short-circuit faults (ABC).

The most recent publication presents an intriguing methodology proposed by[63] wherein a suggested technique is employed to detect Single-line, double-line, and triple-line Harmonic Impedance Faults (HIFs) in a typical distributed generating system. During the occurrence of faults, an analysis is conducted on the current signals of the remaining phases in order to observe their respective variations. In this study, we will assess multiple intelligent fault detection techniques in networks that have distributed generators (DGs) integrated. We will thoroughly evaluate the advantages and disadvantages of these techniques. Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy Logic, Genetic Algorithm (GA), and matching method are all machine learning techniques that can be implemented for diversified networks. When summarizing the evolution, it is evident that within distribution networks (DN), there exists a distinct category of faults known as high impedance faults (HIF). These faults pose a challenge in terms of detection, primarily due to their low current magnitude, which renders them undetectable by conventional overcurrent-based protection devices[64].

2.6.3 Wavelet transform (WT)

In [65] the discrete wavelet transformation is a mathematical operation that allows for the decomposition and analysis of signals or data into different frequency components. It involves the use of wavelet functions, which are small, localized waveforms that can be scaled and translated to capture different features of the signal. By applying the discrete wavelet transformation, one can obtain a prominent system in power system applications, by the utilization of WT, a robust novel mathematical tool, can serve as a rapid and highly efficient alternative to the conventional Fourier transform for the analysis of power system transient waveforms[66]. The challenges encountered in the distribution system are highly complex and require meticulous problem-solving. In the study referenced as [67] a fault diagnostic technique is introduced specifically for power distribution systems. This technique aims to identify and classify faults in a systematic manner. The wavelet transform has been introduced in numerous published papers as a valuable technique for analyzing disturbances in power systems. Moreover, the authors in reference [68] put forth a technique that relies on the dyadic wavelet transform as a means of identifying and localizing faults occurring in transmission lines. Ref [69] yields an index which utilizes numerous method's description of transmission line fault detection and faulted-phase selection which shows the coefficient of discrete approximation of the dyadic wavelet transform with the Haar wavelet approach. The utilization of wavelet analysis in power systems has exhibited a steady increase in the quantity of published works, as evidenced by reference [70] Furthermore, it has been observed that the domains of protection and power quality have yielded the highest number of research contributions.

Given that dispersed generators are connected to a low or medium voltage distribution system, it was postulated [71] that the most efficient approach to identify and isolate problematic segments in a distribution network employing distributed generators would involve a network of cooperative relay agents. Decentralized multi-agent protection for DG systems, including HIF detection, fault identification, and coordination of DG system protection systems, is proposed in the study referenced as [72] makes a proposal the new digital protection schemes implemented in condition monitor units are relay agents capable of interacting with apparatus, interacting with other agents to obtain information, and protecting both autonomously and cooperatively. In order to prevent the normal load current effect, wavelet packet analysis was utilized to extract fault characteristics for multi-agent network protection. Or a decentralized multi-agent protection for the DG systems that is capable of HIF detection, fault identification, and coordinate DG system protection systems. In condition monitor units, the new digital protection schemes are relay agents that can seek for information from other agents, interact with equipment, and protect autonomously and cooperatively. The wavelet packet analysis was used to extract fault characteristics in multi-agent network protection to prevent normal load current effect.

Despite the greater reliability of subterranean cables compared to overhead lines, a method for locating faults in underground distribution systems is presented in reference[73] The Discrete Wavelet Transform (DWT), which operates on the principle of traveling waves, is employed to identify the problematic areas of the subterranean distribution system by locating its high-frequency components. A newly developed method for identifying phase-to-ground faults in main distribution networks is detailed in the study referenced as[74] which describes a novel f detection method for unbalanced distribution systems is presented in this article..

In the Latter decade of 2009-2020 , current practice implemented then is to disconnect DG units as soon as feasible to allow traditional protection devices like as fuses, sectionalizes, and reclosers to work correctly, with the goal of avoiding difficulties with the protection system. A problem detection system and a directed comparison scheme have been developed by using high frequency transient signals. [75], It utilizes the DWT based on the Daubechies wavelet function to extract differentiating properties from fault-generated high-frequency current transient signals. The work proposes by [76] shows a protective strategy for the identification of faults with and without DGs based on Wavelet Transforms. Each bus implements this, as faults are simulated, and the fault currents is analysed with Haar wavelet to yield single-level decomposition detail coefficients. The method for fault categorization stated by [77] consists of two distinct logics, depending on whether or not the fault includes ground. This is critical, as the features of a defect affecting the ground differ significantly from those of a fault that does not involve the ground, which must be addressed independently. Other authors such as [2, 78-82] presents a precise method for fault detection and classification of fault kinds and faulted phase(s) in distribution networks.

The author [83] delineates a methodology for the extraction, detection, and classification of brief circuit fault features in a power distribution network in this study. Characteristics are extracted from transient fault currents measured at the source terminal of the network via DWT analysis. For the identification of high impedance faults (HIFs) in active distribution networks, a data mining-driven solution based on discrete wavelet transform (DWT) is proposed [84]. With the IEEE 13-Bus and IEEE 34-Bus systems serving as proved platforms. DWT derives time-frequency characteristics by decomposing two fault current signal cycles and extracting statistical features from the detail coefficients at each level. Correlation coefficients have replaced the decomposition of defect signals. This study investigates the classification and defect detection of series compensated long transmission lines utilizing a wavelet transform and artificial neural networks[85]. The proposed approach detects the ground current signal through the sampling of three phase current signals at least one cycle prior to the fault and one cycle afterward.

The development of strategies capable of handling a multitude of complex or unpredictable scenarios involving a diversified load demand and grid, the production of renewable energy, and ambiguous fault information has emerged as a formidable obstacle for distribution network protection systems to surmount. By utilizing signal processing strategies and machine learning instruments, this article[86] describes a procedure for identifying defects in the dynamic distribution grid. Feedforward neural networks (FFNN) are machine-learning tools, whereas the Hilbert-Huang transform (HHT) and the discrete wavelet transform (DWT) are signal processing algorithms. The author[87] employs sophisticated signal processing and machine learning methodologies to identify and categorize defects by utilizing pre-existing data. The instrument is supplied with a three-phase current measurement as input, and the classification and location of the fault are disclosed as output. The authors of[88] have devised an algorithm. A model is developed during the course of his research, employing the MATLAB/Simulink environment and incorporating artificial neural networks, multi-scale approximation analysis, and discrete wavelet transform. By employing this model, defect detection and classification processes in distribution networks can be executed with remarkable efficiency, velocity, and reliability. Articles of a similar nature, such as [89-92] address the impact of high-impedance failures (HIFs) on the safeguarding of distribution systems. This assertion holds particular validity in the context of micro-grids and distribution networks that incorporate distributed generators (DGs) and employ variable operation strategies.

2.6.4 Fuzzy logic

Particularly concerning control-related issues, the number of papers devoted to the application of fuzzy logic in power systems increased dramatically in the last decade of the twenty-first century. However, it is stated in reference[93] that electrical systems were protected using a family of fuzzy-based algorithms at the start of the previous decade. The authors of this study[94], propose an efficient malfunction detection approach for power systems that employs fuzzy logic to identify the problematic circuit. In CE-Nets, the values of CB and relay are specified by the binary and coordinative relations established by the system. The author employs sagittal diagrams to diagnose fault sections. Subsequently, the alarm data and the estimated fault section are utilized to analyze the failure or erroneous alarms of relays and circuit breakers, as proposed in this work, in order to account for the uncertainties associated with fault section diagnostics of power systems [95, 96]. The proposed system provides candidates for the fault section with regard to both membership and malfunction or false alarm. Using a fuzzy-logic-based multi-criteria approach, [97] demonstrates a novel method for real-time defect classification in power transmission systems. A fuzzy logic-based fault-type identification technique has been proposed by [98] in which this technique requires only three observations, or three measurements of line current, to accurately identify each of the ten categories of short-circuit defects. The defect detection method is described and verified through the utilization of bus bars, transformers, and line sections in a medium voltage test system examination[99]. A combination of fuzzy logic and genetic algorithms (GAs) is utilized to detect errors in these networks. It proved to be efficient, precise, and dependable, and it performed admirably under a wide range of system conditions. The study mentioned above illustrates the potential of fuzzy logic and neural networks in facilitating transmission line defect categorization through advanced pattern recognition. By increasing the algorithm's selectivity for a variety of real-world scenarios that were not consistently predicted during the training process, the method employs[100] .

In light of the escalating difficulties encountered, particularly in the context of islanding, the author presents a novel time-frequency approach for detecting power islands in systems comprising numerous generators in[101, 102] A similar topic is addressed in the article [103, 104] titled Failure Identification in Smart Grids. A state-of-the-art method for identifying power system defects is presented in this related study [105]. [106] investigates and verifies the various fault types and locations that may occur in distribution lines with DGs. A review of recent advancements in fault detection can be found in reference[107]. Using fuzzy logic for fault detection and classification in transmission lines has been demonstrated in [108] to be effective and successful across a wide range of fault scenarios. In addition to identifying and classifying defects, this method offers automated, real-time protection. This research[109] is concerned with the diagnosis of imprecise logic defects. The magnitudes of phase and neutral currents establish fuzzy defect identification criteria. For all substation measurements, the IEEE 34-Node Radial Test model is applied.

Using data collected by sensors and smart meters in the smart grid, this research literature proposes the application of AI algorithms for fault detection, classification, characterization, and localization[110]. This is in light of the escalating volume of data available and the need for greater precision in its utilization. While the author[111] does not implement DG on DN, he does present an artificial intelligence technique known as fuzzy logic-based control strategy in this work. This strategy aims to identify the underlying causes of transmission line problems. This article [112, 113] presents a fuzzy-based intelligent defect detection and classification scheme for DG-integrated distribution lines. In order to replicate the defect detection procedure, two discrete fuzzy inference systems (FIS) are employed. The initial FIS detects the substantial fault current magnitude that is characteristic of typical shunt faults, while the subsequent FIS detects the minimal current magnitude that results from the presence of HIF.

2.7 Discussion, constraints, and emerging trends

The utilization of various AI algorithms for anomaly detection and classification in distribution networks through distribution generation has been explicated. Clearly, they have been employed to identify a multitude of power system malfunctions. In conclusion, each possesses unique assets and weaknesses.

Fuzzy logic rules are defined using simple language, which is advantageous for formalizing human reasoning. The three essential constituents of a fuzzy-logic system are Fuzzification, Fuzzy inference, and Defuzzification. Fuzzy logic can be utilized to differentiate between various fault types, offering advantages over other artificial intelligence techniques. Unlike other methods that necessitate extensive modeling processes, a fuzzy logic system can produce results by specifying only a few rules.

An artificial neural network (ANN) is a computational model that replicates the architecture of the human brain. Similar to the real brain, the model consists of basic processing units that are interconnected in an intricate layered structure. This enables the model to mimic an intricate non-linear function with numerous inputs and multiple outputs. It is evident that the structures of Artificial Neural Networks (ANNs) have achieved significant advancements in the field of defect detection applications. Deep learning techniques, such as deep neural networks and convolutional neural networks, have been employed to automatically extract complex information. Researchers have focused on improving fault detection performance by optimizing network designs, regularization approaches, and training algorithms.

The efficacy of the wavelet transform has been established. The wavelet transform has been demonstrated to be a highly effective method for analyzing and assessing the signals under consideration [114]. The wavelet transform can be used to decompose the distorted signal into several time-frequency domains, as stated in reference [115]. The wavelet transform can be employed to identify different frequency components in a measured signal by manipulating the fundamental functions through expansion and contraction. After undergoing the wavelet transform, the signal is divided into several frequency bands. The mother wavelet, utilizing the properties of dilation and translation, serves as the fundamental basis for the function. The distorted signal can be effectively analyzed and separated into several time-frequency domains using the wavelet transform. The identification of various frequency components in a measured signal can be achieved using wavelet transform, which involves the expansion and contraction of fundamental functions. After undergoing the wavelet transform, the signal is divided into multiple frequency bands. The mother wavelet, utilizing the properties of dilation and translation, acts as the fundamental component of the function.

Using a support vector machine (SVM) for classification necessitates a large amount of training and testing data. As a result, even with a minimal amount of training data, SVM has exhibited exceptional generalization performance. Furthermore, due to the correct nonlinear relationship in mapping utilizing kernel-level functions, data from two or more categories can always be separated by a hyperplane[116]. As a result, it can execute classification jobs with high precision for the purposes of fault identification and classification. Scholarly research has focused on improving the efficacy of Support Vector Machines (SVMs) through the use of various kernel functions, studying ensemble approaches, and addressing issues related with imbalanced datasets. Support vector machines (SVM) have been studied in conjunction with feature selection approaches and hybrid methodologies to improve the efficacy of defect identification.

When dealing with numerous layers of dimensions and continuous features, ANN, DWT, SVM, and deep learning algorithms tend to perform better. Depending on the application, the performance of various algorithms may differ dramatically.

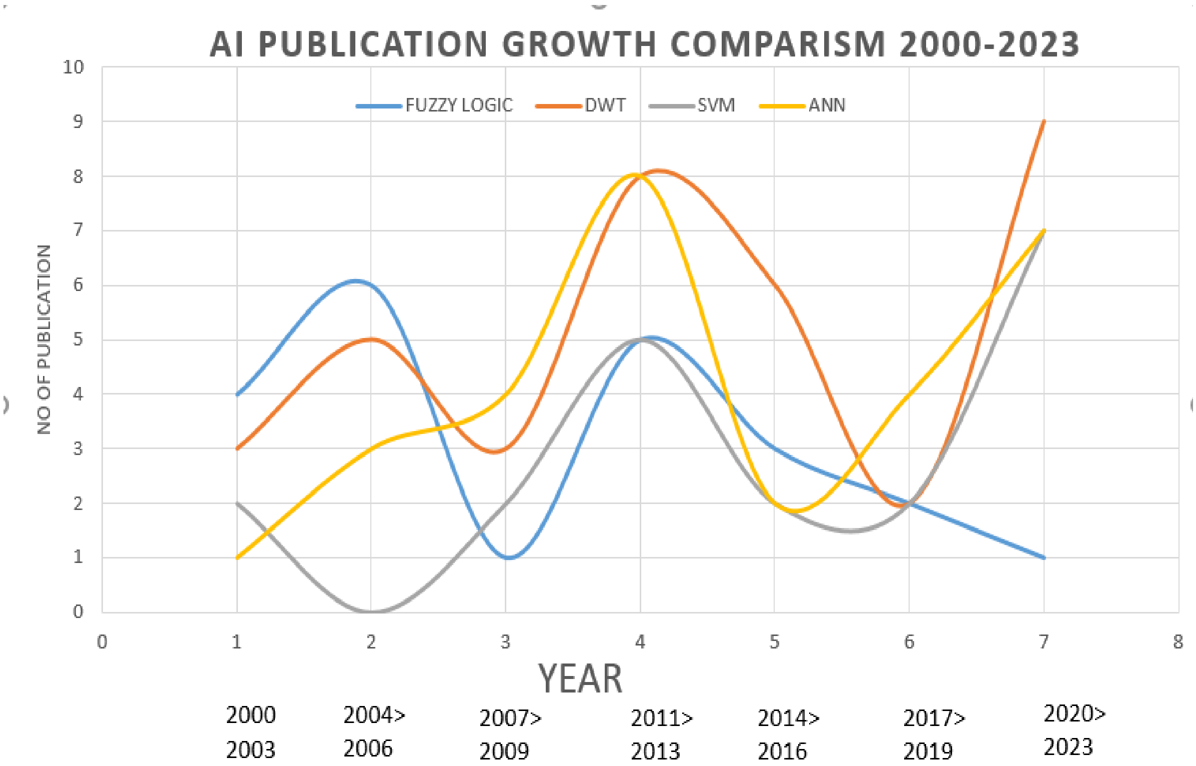


Figure 2.4 Publication of AI techniques for fault detection in with and without DG

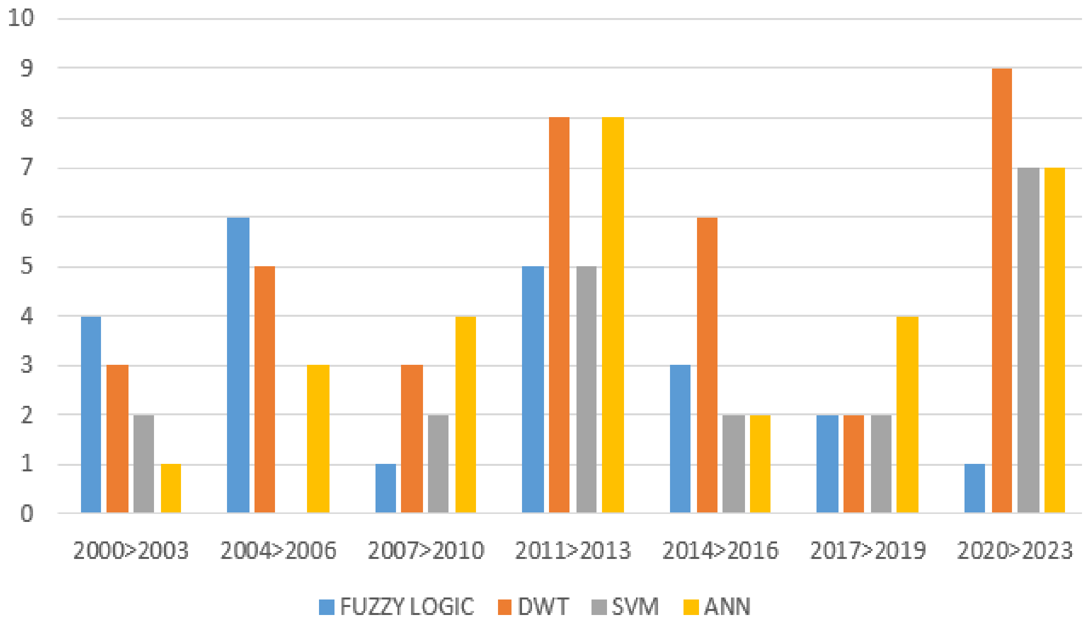


Figure 2.5 Consistency trend of various intelligence techniques in fault detection in power systems.

The graphical representations in Figures 2.4 and 2.5 present a concise overview of publications that conduct a comparative analysis of artificial intelligence (AI) methodologies across the three decades mentioned in section IV. This comparison is made due to the absence of a singular learning algorithm that consistently outperforms other algorithms across all datasets. Hence, the selected publications for the review employ either a singular learning methodology or integrate it with another learning algorithm. This chart has been formulated to facilitate a comparative analysis of the progression of AI technology in the domain of power system protection, specifically in power systems incorporating dispersed generation. The works produced can be classified into two categories: journal articles or conference proceedings. Due to this phenomenon, particularly in recent years, fuzzy algorithms have exhibited notable indications of exhaustion in their utilization within power system protection as a standalone algorithm. However, it is worth noting that the Discrete Wavelet Transform (DWT), Support Vector Machine (SVM), and Artificial Neural Network (ANN) algorithms are currently experiencing a notable surge in the frequency of their appearances within published literature. These strategies do not insinuate the embodiment of the future; instead, in recent times, researchers have been employing combinations of all four of these methodologies in an endeavour to enhance. Based on the examination of Figures 2.4 and 2.5, it can be deduced that multiple methodologies, including Support Vector Machines (SVM), Fuzzy Logic, Discrete Wavelet Transform (DWT), and Artificial Neural Networks (ANN), have been extensively utilized for the purpose of fault detection in electrical distribution networks that encompass distribution generation, starting from the year 2000. The domain of fault detection systems in distribution networks has witnessed significant advancements in multiple facets including performance, feature extraction, hybridization, and network architectures. The aforementioned advancements have led to the development of increasingly accurate and reliable methodologies for fault detection, thereby affirming the observed growth trend.

3. Chapter Three: System Model

3.1 Introduction of the System Model

The primary objective of any electricity supply system is to meet the energy demands of all customers, and distribution carries the energy to the farthest customer using the appropriate voltage level. Associated with distribution networks are a number of auxiliary systems that aid in meeting the demand for supply safety, dependability, and quality. Protection systems, which are implemented to clear faults and limit damage to distribution equipment, are the most essential. Distributed generation, which is explicitly defined in chapter two, can be incorporated into a contemporary distribution network. Consequently, it is now recognized that the protection philosophy of the distribution network can vary from utility to utility based on the nature of each problem and the utility's desired solution.

Utility protection schemes vary depending on the requirements, their methods of operations hence connecting DG units to the grid without negatively impacting safety, reliability and quality of supply [117] remains high on priority. The challenge is that most of the electrical network were not designed to handle the Distribution Generation as grids are currently designed in a top down manner [117] with defined power flows Power can usually flow bi-directionally within a certain voltage level (depending on topology), but unidirectional from higher voltage levels to lower voltage levels [117]. Safe operation of the protection need to be undertaken all the times and the protection system has to be sufficiently selective, in order to optimize reliability and availability of supplied power [118]. Fact is the installation of DG into the grid changes the property of the network and the short-circuit power increases as well as the current pathway. Large impact on the protection system depends upon level of penetration of DG and also on type of interfacing devices with the main aim of protection coordination is to achieve selectivity i.e. to keep healthy part apart from faulty part to achieve stability [119].

3.2 Distributed generation and the distribution system

The artificial intelligence fault classifiers are being utilized in this work to assist in the classification of various sorts of faults for the purpose of multi-class fault classification in a micro-grid. The term "distributed generation" in this context refers to the process of producing electricity at or close to the location where it is consumed, as opposed to producing it at a centralised power station. Solar photovoltaic (PV) and battery energy storage systems (BESS) are examples of renewable energy sources that are incorporated into the micro-grid. The distribution system, on the other hand, is comprised of a network of power lines, transformers, bus-bars, and other pieces of equipment that are responsible for delivering electricity to the loads and other recipients of the electricity.

3.3 Network construction

A grid equivalent of 120 kilovolts (kV), three phase transformers 120kV/11kV, and 120 kilovolt feeders are used in the modelling of the electrical grid. It is possible to create an imbalance on the grid by switching on a load with only one phase. In order to link the micro-grid to the distribution system, a three-phase breaker and a transformer rated at 25 kV/1000 V are utilized.

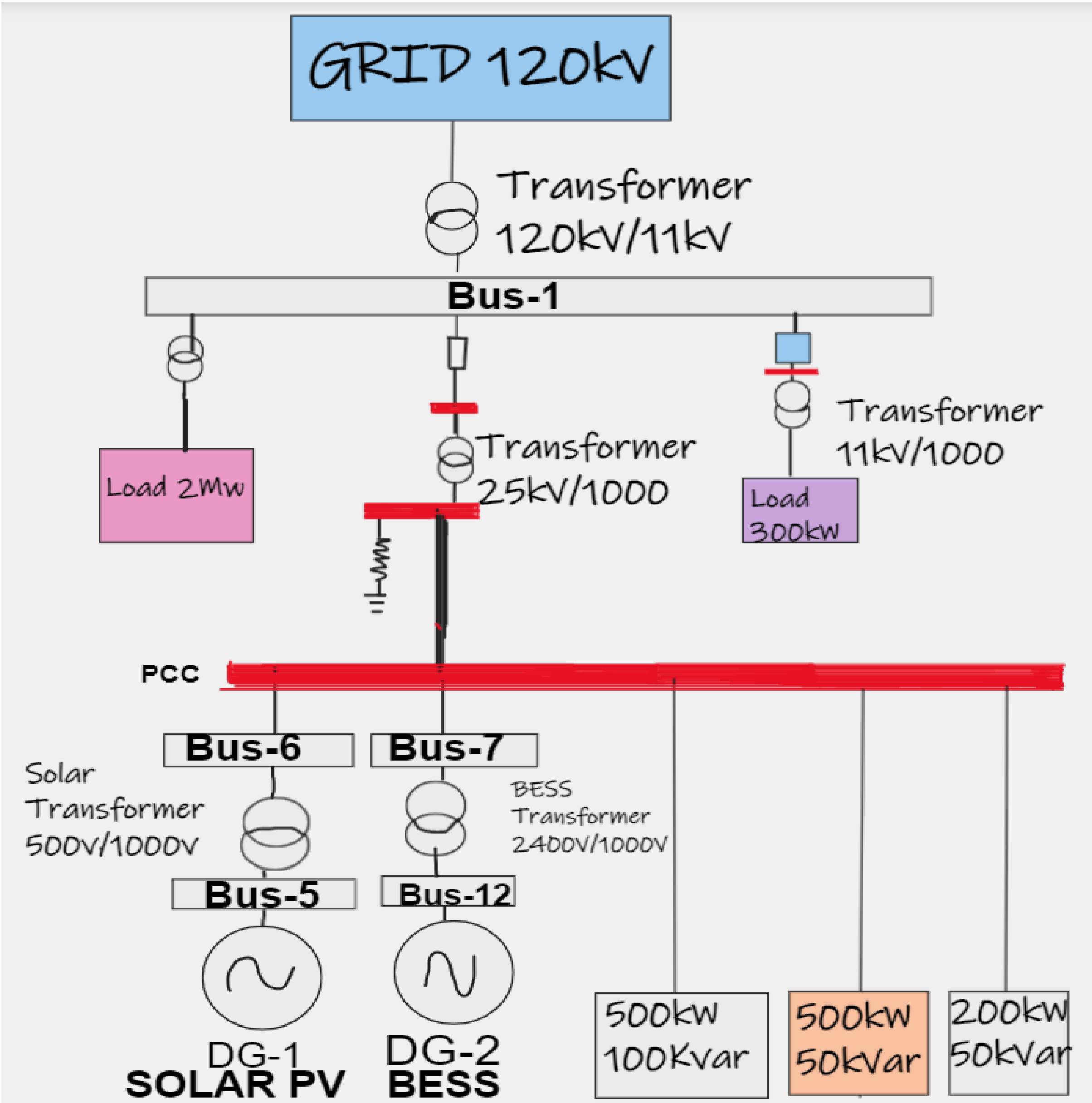


Figure 3.1 Micro-grid schematic drawing

Figure 3.1 refers to the11 kV originating from bus 1 is utilized to establish a connection with a load of 300Kw-2Mvar. This load is accompanied by a grounding transformer that operates in parallel, with the purpose of creating a pathway to ground with a reasonably low impedance. Consequently, this arrangement ensures that the system neutral remains at or close to ground potential. In order to mitigate the impact of re-striking ground faults, it is imperative to impose restrictions on the amplitude of transient overvoltage.

3.4 Feeder line

The line parameters of a distribution network are described in a sequence vector, as outlined in table 3.1. The values of resistance, inductance, and capacitance are evenly spread along the line, and the cable's resistance per unit length, inductance per unit length, and capacitance per unit length determines this distribution. Its corresponding sequence impedance characterizes the system. The determination of the positive and negative sequence impedance equivalents can be obtained through direct calculation from;

$$Z = \frac{V^2}{P} \tag{3.1}$$

Where

Z =equivalent positive and negative sequence impedance

V=nominal phase to phase voltage

P= Three phase short circuit power

Table 3.1 Microgrid feeder parameter data.

Line number	length(Km)	R1 (Ω/km) Ro(Ω/km)	H/km	F/Km
1	2	0,1153 0,413	1,05e-3 3,32e-3	11,33e-009 5,01e-009
2	6	0,1153 0,413	1,05e-3 3,32e-3	11,33e-009 5,01e-009
3	15	0,1153 0,413	1,05e-3 3,32e-3	11,33e-009 5,01e-009
4	30	0,1153 0,413	1,05e-3 3,32e-3	11,33e-009 5,01e-009

3.4 Solar Farm

The photovoltaic array at the solar plant has the capacity to generate (one) 1MW when subjected to sun irradiation of 1000 W/m2 and when the cell temperature is set to 25 degrees Celsius. A Maximum Power Point Tracker (MPPT) system is in charge of regulating the boost's output. In order to harvest the maximum amount of power from the PV array, the 1000 V DC is converted to about 500 V AC via a three-level NPC converter. The purpose of connecting the converter to the micro-grid, an LC filter and a 1-MVA 480V/600V three-phase coupling transformer are utilized. The Structure for mounting the photovoltaic panels include mounting structures that optimize their angle and orientation in order to collect the maximum amount of solar energy possible.

The data in Table 3.2 evaluates the performance out of the solar PV at a specific location and battery performance capability. In data analysis of Solar PV analysis , PVGIS (Photovoltaic geographical information system) is applied by to extract information shown in Table 3.2 for any location on Earth, in our case we have chosen and area in KZN in Ladysmith so that the PVGIS gives data on solar radiation and the efficiency of photovoltaic systems. The amount of electricity a photovoltaics is able to produce at the designated location and how much output get altered throughout the year are two of the most important findings that can be used in the design phase and how effective is a battery at storing all that energy. From the table photovoltaic data is extracted from the given location is shown as Latitude/longitude (Lat/Lon).

Table 3.2 Microgrid feeder parameter data, curtesy of PVGis

Location [Lat/Lon]:	-28.610,29.755
Horizon:	Calculated
Database used:	PVGIS-SARAH2
PV installed [Wp]:	1000000
Battery capacity [Wh]:	600
Discharge cutoff limit [%]:	40
Consumption per day [Wh]:	1000000
Slope angle [Â°]:	35
Azimuth angle [Â°]:	0
Simulation outputs:	
Percentage days with full battery [%]:	100
Percentage days with empty battery [%]:	100
Average energy not captured [Wh]:	1870693.11
Average energy missing [Wh]:	674014.18

Azimuth can be descripted as one of the key factors in determining the efficiency of a solar PV system and Figure 3.2 show the Azimuth, which is essential component for the operation of the solar micro-grid when generating maximum power. Azimuth, when discussing solar photovoltaic (PV) power generation, is the angle between the north and south or east and west orientations of a solar panel or solar array. Because it controls how directly the panels are exposed to sunlight throughout the day, the azimuth angle of a solar panel or array is critical for maximizing its energy generating capacity[120].

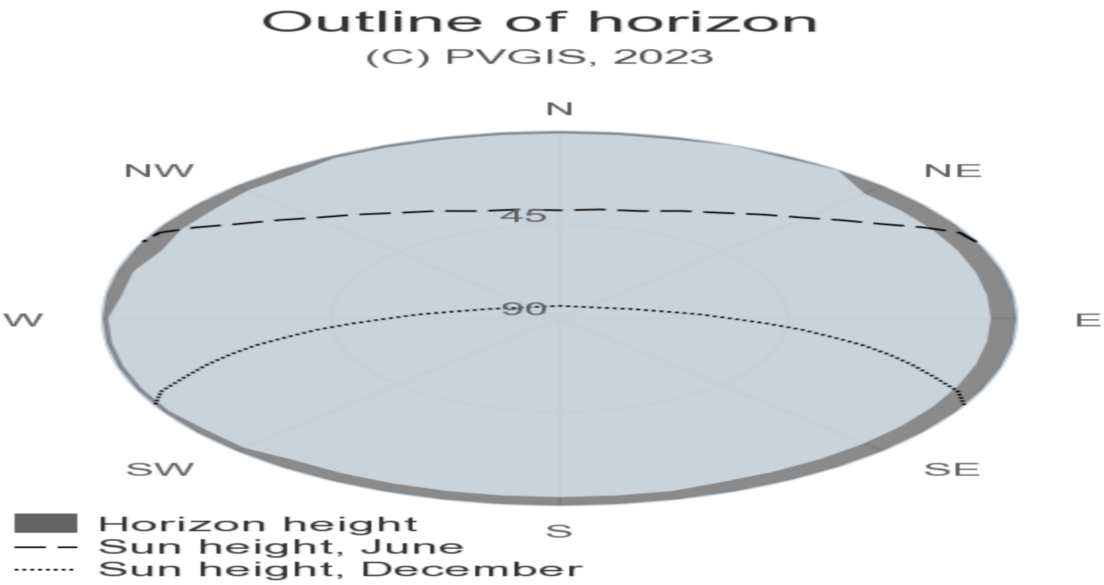


Figure 3.2 Shows the Outline of horizon with angular displacement between the projection of sun rays and a line oriented precisely in the north-south direction.

Further, on Figure 3.3 shows below the location site, which the micro-grid is proposed to be installed. The location benefits from favourable local climate patterns, including optimal sunlight hours, cloud cover, and temperature variations, ensuring consistent performance throughout the year. Therefore, the selection of this location was determined by these characteristics, making it the ideal choice in terms of irradiance in the KZN interior landscape. Unlike coastal areas, which typically have lower irradiance, this location benefits from ample sunlight exposure and minimal shading.

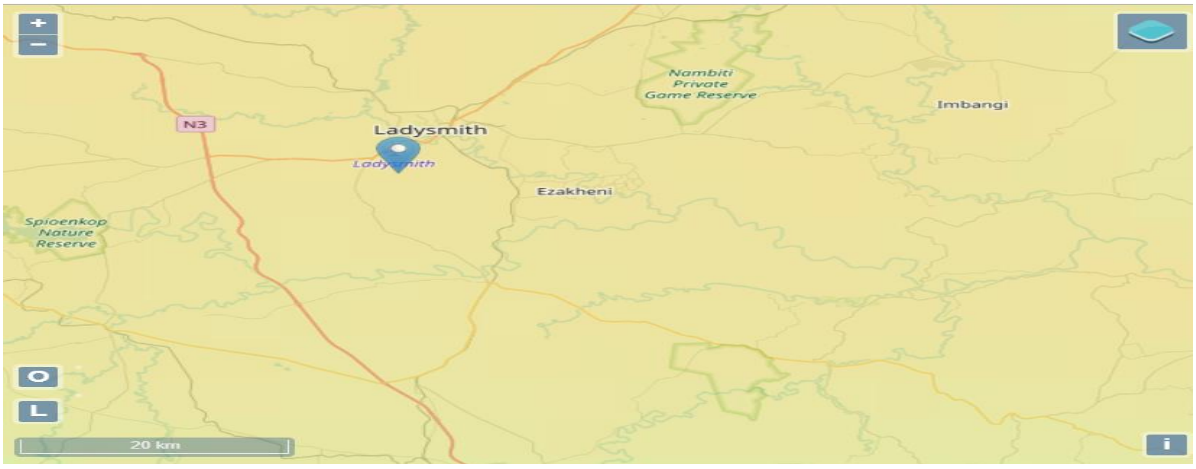


Figure 3.3 Location of 1MW solar plant.

The monthly energy output displayed in Figure 3.4 illustrates the estimation of energy production from various sun-tracking PV systems that are connected to the electrical grid. This data is derived from the tracking photovoltaic (PV) system provided by PVGIS. Figure 3.4 divides the energy output into monthly segment graphically in which the data set comprises the monthly mean values of energy output from the photovoltaic (PV) system and the in-plane irradiation per month, as well as for the entire year. These values have been derived from a multi-year time series of solar radiation and various climatic parameters

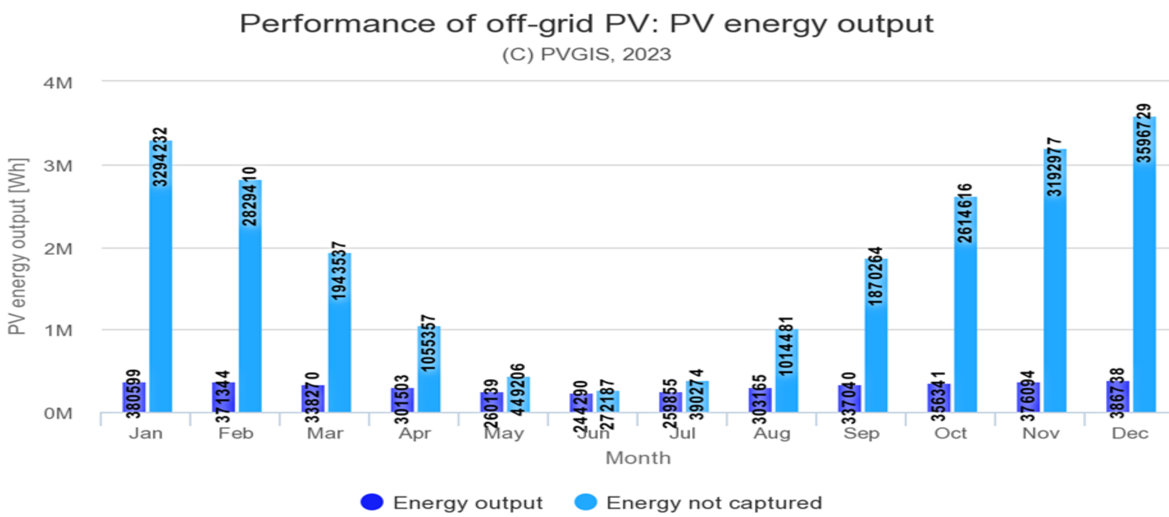


Figure 3.4 Monthly energy output Wh

3.4.1 Solar panel string layout

A total of 14 solar panels are connected together in a series to create a string. Then, 220 of these strings are connected in parallel to generate a maximum power output of 1 MW.

The needed Pv capacity can be calculated using equation 3.2 as shown below.

$$\text{Needed PV capacity} = Needed\_PV\_Capacity = \frac{Daytime\_consumption}{Specific\_yield} \qquad (3.2)$$

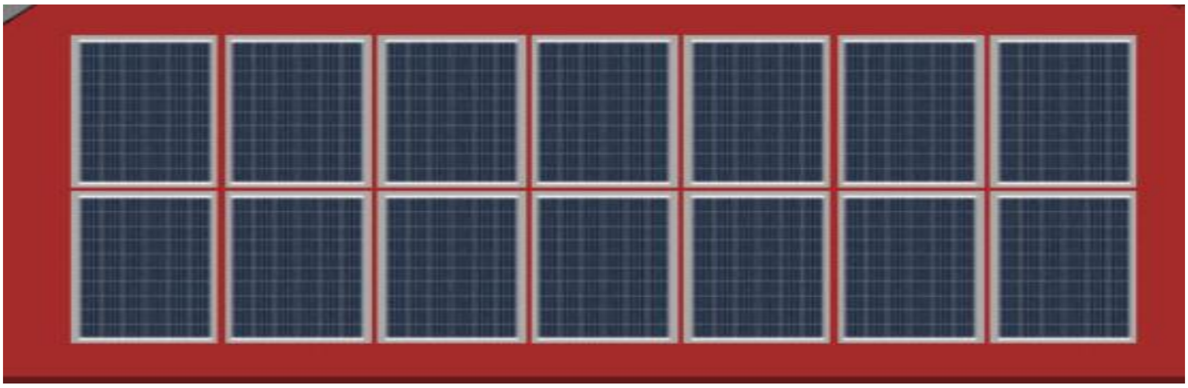


Figure 3.5 Graphical solar panel string layout installation.

3.5 Inverter selection

The solar inverter, which resides within the network, serves as a pivotal element within a solar power system. As previously stated in the literature review, the main purpose of this device is to perform the conversion of direct current (DC) electrical energy, which is produced by solar panels, into alternating current (AC) electrical energy. This converted AC electricity is then suitable for utilization in the electrical grid. The conversion stage of the system necessitates the utilization of an inverter, wherein the careful consideration of power and voltage capability selection becomes of utmost importance.

The selection of the inverter is contingent upon a multitude of factors, encompassing the magnitude of the solar installation, the arrangement of the solar panels, and the precise specifications of the system. The efficiency and reliability of the inverter are crucial factors that heavily influence the overall performance of a solar power system.

3.6 BESS Structure

The Battery Energy Storage System (BESS) encompasses several fundamental components and features that are vital to its functionality and effectiveness. The primary component of the Battery Energy Storage System (BESS) is a battery bank, which typically consists of batteries employing modern chemical compound such as Lithium-ion. These batteries are designed to efficiently store electrical energy. The Battery Energy Storage System (BESS) incorporates an inverter, which facilitates the conversion of the direct current (DC) power stored within the batteries into alternating current (AC) power. This AC power is well-suited for the provision of electrical energy to the micro-grid and for fulfilling diverse electrical loads. The charger function of the Battery Energy Storage System (BESS) allows for the charging of its batteries, particularly at times when the solar plant produces surplus energy or during off-peak hours when electricity costs are lower.

Discharging: In instances of heightened demand or inadequate solar plant output, the Battery Energy Storage System (BESS) releases stored energy to provide power to the micro-grid.

The Energy Management System (EMS) is a sophisticated control system that enhances the efficiency of battery charging and discharging processes by considering several parameters, including solar power, load demand, and energy pricing.

Protection and Control Systems: The Battery Energy Storage System (BESS) is fitted with safety systems that serve to mitigate the risks associated with overcharging, overcurrent, and other potential defects.

The model of the Battery Energy Storage System includes a 480V/1000V transformer, an LCL filter, and a two-level converter. Additionally, the model includes a battery system. The BESS also features a control system that sends voltage references (Vref) to the PWM generator that controls the converter and a control signal (open/close) to the grid breaker. Both of these signals are produced by the control system.

The model of the battery system is composed of Lithium-ion Iron Phosphate (LFP) cells that have 3.2V and 14Ah respectively. They are assembled into a battery string with a voltage of 922V by being organized in multiple modules of cells (72 modules of 4 cells), and then connected in series. In order to create a battery system with a capacity of 1 MWh, our model comprises 80 separate battery strings connected in parallel.

3.6.1 Battery sizing

$$C_{10} = \frac{EnergyConsumption\_\mathrm{X}\_Autonomy}{DOD\_\mathrm{X}\_SystemVoltage} \quad (3.3)$$

- DOD (Depth of discharge) is 50%
- Energy consumption
- Autonomy = 12 Hours
- Vsyst = System Voltage DC = 922 V

3.7 Short circuit calculation

Since short circuits increase current flow, power system elements' current flows can be used to discover defects[121]. Although short-circuit calculations for protection settings will be discussed, they are also used for other purposes.

The fundamental equation for determining short circuit current in a power system is expressed as Isc = V / Z, where Isc represents the short circuit current measured in amperes (A), V denotes the pre-fault voltage measured in volts (V), and Z signifies the total impedance measured in ohms (Ω) between the location of the fault and the power sources[122]

. In order to determine Z, it is necessary to include all the components connected in series between the fault location and the power sources, including transmission lines, transformers, and generators. Each component's impedance can be represented by a complex number consisting of a real part (resistance) and an imaginary part (reactance). Furthermore, it can be represented as a magnitude, known as impedance, and an angle, referred to as the power factor angle.

Three-phase faults and three-phase-to-earth faults with symmetrical impedance can be depicted as single-phase faults because they balance the electrical system. Devoid of this symmetry, asymmetric fault-line-to-earth, line-to-line, and line-to-line-to-earth faults necessitate a practical method of fault analysis.

When considering a three phase system, each vector quantity, voltage or current, is supplanted by three components, resulting in a total of nine vectors that uniquely represent the values of the three phases, which are designated as follows:

1. Positive sequence component
2. Negative sequence component
3. And zero sequence component

With this arrangement, voltage values of any three phase system, Va, Vb , Vc are represented as :

$$V_a = V_{a0} + V_{a1} + V_{a2} \tag{3.4}$$

$$V_b = V_{b0} + V_{b1} + V_{b2} \tag{3.5}$$

$$V_c = V_{c0} + V_{c1} + V_{c2} \tag{3.6}$$

therefore

$$V_b = V_{a0} + \alpha^2 V_{a1} + \alpha V_{a2} \tag{3.7}$$

$$V_c = V_{a0} + \alpha V_{a1} + \alpha^2 V_{a2} \tag{3.8}$$

It can be demonstrated that

$$\begin{matrix} V_a \\ V_b \\ V_c \end{matrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha^2 & \alpha \\ 1 & \alpha & \alpha^2 \end{bmatrix} \times \begin{bmatrix} V_0 \\ V_1 \\ V_2 \end{bmatrix}$$

(3.9)
(3.10)
(3.11)

and

$$\begin{bmatrix} V_{a0} \\ V_{a1} \\ V_{a2} \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \alpha & \alpha^2 \\ 1 & \alpha^2 & \alpha \end{bmatrix} \times \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \tag{3.12}$$

By inverting the matrix co-efficient:

Therefore

$$V_{a0} = \tfrac{1}{3}\left(V_a + V_b + V_c\right) \tag{3.13}$$

$$V_{a1} = \tfrac{1}{3}\left(V_a + \alpha V_b + \alpha^2 V_c\right) \tag{3.14}$$

$$V_{a2} = \tfrac{1}{3}\left(V_a + \alpha^2 V_b + \alpha V_c\right) \tag{3.15}$$

Procedure can also be applied directly to currents.

$$\boldsymbol{I}_a = \boldsymbol{I}_{a0} + \boldsymbol{I}_{a1} + \boldsymbol{I}_{a2} \tag{3.16}$$

$$\boldsymbol{I}_b = \boldsymbol{I}_{a0} + \boldsymbol{I}_{a1} + \boldsymbol{I}_{a2} \tag{3.17}$$

$$\boldsymbol{I}_c = \boldsymbol{I}_{a0} + \boldsymbol{I}_{a1} + \boldsymbol{I}_{a2} \tag{3.18}$$

Therefore

$$I_{a0} = \tfrac{1}{3}\left(I_a + I_b + I_c\right) \tag{3.19}$$

$$I_{a1} = \tfrac{1}{3}\left(I_a + \alpha I_b + \alpha^2 I_c\right) \tag{3.20}$$

$$I_{a2} = \tfrac{1}{3}\left(I_a + \alpha^2 I_b + \alpha I_c\right) \tag{3.21}$$

3.8 Machine learning Methodology

There are many different approaches that can be adopted in the literature for a formulation of the protection scheme challenges when DGs are interconnected. The overall objective of all approaches is to maintain the desired outcome of the performance of protection system when a distribution network becomes diversified by different sources of power by using an intelligent technique. The type of methodological approach used to conduct this experiment is quantitative method. This will involve the mechanism of the methodology which can be summarized into the following main steps as shown in Figure 3.7.

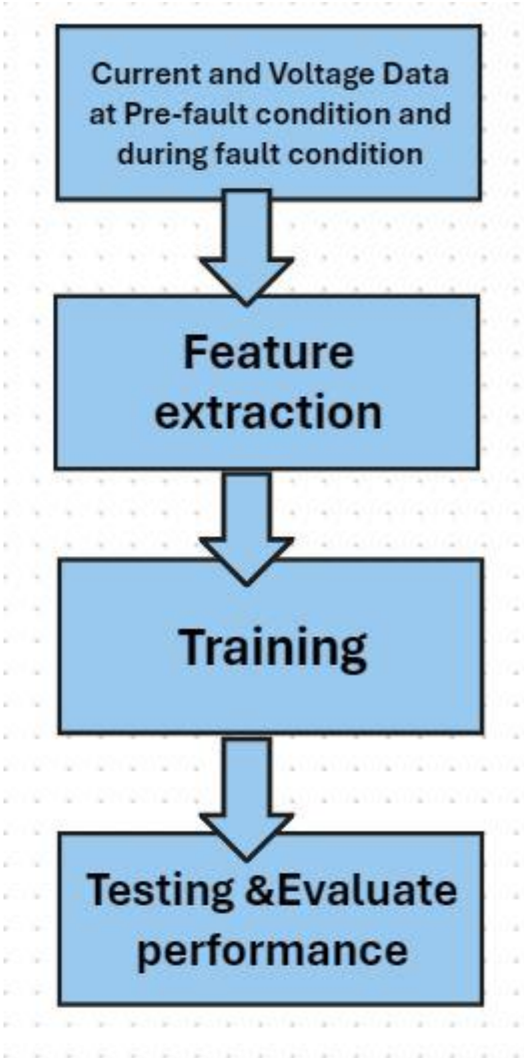


Figure 3.6 Machine learning workflow operation for fault detection.

The use of machine learning is commonplace. Machine learning methods are employed in every aspect of modern life, from medical diagnosis and speech/handwriting recognition to automated trading and movie recommendations. Raw data management, identifying critical properties that affect your model, training numerous models, and performing model assessments can be difficult for any given machine learning application. The model will use common procedures for supervised learning to detect fault in the distribution network with distribution generation.

3.9 Data collection

Data collection is performed on the micro-grid across varied operational conditions, encompassing normal and anomalous scenarios. The table 3.2 below illustrates the depicted faulty scenarios of currents and voltages in the systems at various location. The purpose of this endeavour is to guarantee that the data acquisition encompasses a comprehensive spectrum of circumstances that may arise within a micro grid, thereby facilitating the development of a resilient diagnostic model for constructing purposes.

The collection of voltage and current signals as the required data within the circuit is essential for both the intelligent detection of electrical faults and the comprehensive monitoring of the health of the micro-grid. By evaluating the fluctuations in these signals, which can indicate abnormal circumstances such as short circuits and overloads, it is possible to identify the electrical defects that have occurred.

Table 3.2 Showing data collection points bus-bar rating and their description.

System Bus	Description	Voltage level
Bus-1	Grid bus	11000
Bus-4	Static Load	1000
Bus-5	Inverter Bus(4160V)	5000
V-Pcc	Point of common coupling	1000
Bus-6	Transformer Solar Plant	1000
Bus-7	Transformer Bus(BESS)	1000
Bus-9	Variable Load 3	400
Bus-10	Static Load 4	400
Bus-11	Variable Load 5	400

3.10 Feature Selection and Extraction

Feature selection and extraction improve algorithm performance and efficiency in machine learning, especially when modelling defect detection systems. Selecting or constructing a subset of significant and informative characteristics from the original dataset can improve model performance, reduce overfitting, and speed computation. Selecting the correct features for fault detection algorithms can help find system defects or anomalies. The right feature is chosen by extracting valuable model condition indicators. Condition indicators are crucial for building categorization and prognosis models.

Both feature selection and feature extraction are crucial to building effective fault detection algorithms because they make models more interpretable, generalizable, and robust to noise and outliers in the data.

Diagnostic Feature Designer, a Matlab software, lets one build features and assess probable condition indicators using a multifunction graphical interface with the full workflow for a fault diagnosis program which includes multiple steps that begin with data acquisition and end with deployment and integration of a fault diagnosis algorithm

In the workflow shown in Figure 3.7, first stage is the already modelled Simulink model that will help us generate data. The second stage involved classifying the raw data into "fault codes," "normal current and voltage," and "faulty current and voltage" for a duration of 5 seconds. The data set, which was generated under a variety of fault circumstances, was imported into the program via Matlab's fault diagnosis Simulink. Our Intelligent algorithm for fault diagnosis will derive significantly from this data after we have recorded it and added it to our data ensemble.

In selection of the appropriate feature, it is crucial to identify the feature that is most relevant as they include statistical measures mentioned in the Table 3.3 shown above.

Table 3.3 Description of feature extracted from current and voltage signals.

Signal Features	Description
Clearance factor	Peak value divided by the squared mean of absolute amplitude square roots. For rotating machinery, this feature is highest for healthy bearings and decreases for defective ball, outer race, and inner race. Inner race faults are best separated by the clearance factor.
Crest factor	RMS / peak value. Faults often first appear in signal peakiness before signal root mean squared energy. The crest factor can detect early flaws.
Impulse factor	Analyse the disparity between the elevation of a summit and the average amplitude of the signal.
Kurtosis	Signal distribution tail length, or outlier probability. Developing flaws increases outliers and kurtosis metric. Kurtosis for normal distributions is 3.
Mean	The average or mean value of an array.
Peak Value	The maximum absolute value of the signal shall be determined. Utilized for the computation of alternative impulse metrics.
Rms	Root Mean Square
Sinad	The ratio between the aggregate power of a signal and the combined power of noise and distortion.
SNR	The ratio of signal power to noise power is a fundamental measure in the field of signal processing. It quantifies the relative strength of the desired signal compared to the unwanted noise present in a given system or channel.
Shape factor	RMS divided by absolute value mean. Shape factor depends on signal shape but not size.
Skewness	The signal distribution exhibits an inherent lack of symmetry. Faults have the potential to influence the distribution symmetry, consequently leading to an elevation in the level of skewness.
Std	Standard deviation
THD	The ratio of the power of the total harmonic components to the power of the fundamental component.
Fault code	To depict different types of faults for algorithm understanding

3.11 Fault classification

The task of fault detection and diagnosis in machine learning often involves the common practice of classifying faults using the Classification Learner app in MATLAB. The Classification Learner tool facilitates the execution of prevalent supervised learning operations, including feature selection, validation scheme specification, model training, and result evaluation.

After the feature extraction, by choosing the classification algorithm MATLAB offers a diverse range of algorithms encompassing decision trees, support vector machines, and ensemble methods that are explanatory in table 3.5 shown below. One may also engage in the exploration of various algorithmic alternatives and parameter whose value is used to control the learning process configurations.

Table 3.4 Classifier to be implemented for the fault classification.

FAULT CLASSIFIER	MODEL DESCRIPTION AND ADAPTABILITY
Coarse tree	Predictors are numerical
Medium tree	Model flexibility is medium and easy to interpret fault classes
Fine tree	Flexibility of the model is high with ant leaves to distinct between classes
Linear Discriminant	Ability to create linear boundaries between fault classes
Quadratic Discriminant	Non-linear boundaries between classes, hyperbola or parabola
Binary GLM Logistic Regression	Popular algorithm to use due to easier to interpret results and cannot set any hyper parameter option
Efficient Logistic Regression	Can change the hyper parameter settings for solver, lambda, and beta tolerance.
Gaussian Naive Bayes	Useful for multiclass classification, cannot change any parameter for model flexibility
Kernel Naive Bayes	can change parameter for model flexibility
Linear SVM	Can train SVMs when your data has two or more classes
Quadratic SVM	Ability to separates data points of one class from those of the other class with medium flexibility
Cubic SVM	Data points closest to the separating hyperplane
Fine Gaussian SVM	Ability to make fine detailed separation between classes
Medium Gaussian SVM	Medium distinction between classes
Coarse Gaussian SVM	Make coarse separation between classes, with low flexibility
Efficient Logistic Regression	Easy interpretation with ability to change hyper parameter settings for solver, lambda, and beta tolerant
Efficient linear SVM	Easy interpretation with ability to change hyper parameter settings for solver, lambda, and beta tolerant

Fine KNN	Linear kernel, easiest to interpret
Medium KNN	Medium distinction between classes
Coarse KNN	Coarse distinction between classes, interpretability wise very difficult between classes.
Cosine KNN	Uses Cosine distance metric, with harder interpretability of class
Cubic KNN	Implements cubic distance metric in distinction between classes
Weighted KNN	Uses distance weight to distinct between classes.
SVM Kernel	Increases as the Kernel scale setting decreases, with application flexibility being medium friendly
Logistic Regression Kernel	increases as the Kernel scale setting decreases, with application flexibility being medium friendly
Boosted trees	A powerful ensemble learning method, involves training each new tree to correct the errors made by the previous ones
Bagged trees	Bagged trees are an ensemble method that combines the predictions of multiple decision trees, with each trained on different random subsets of the training data created through bootstrapping.
Subspace Discriminant	It combines principles from discriminant analysis and subspace methods.
Subspace KNN	It an extension of the traditional k-NN algorithm that incorporates the concept of subspace methods to enhance classification performance
Rusboost trees	Is an ensemble learning technique that combines the principles of boosting and random under-sampling to address class imbalance problems in classification tasks
Narrow Neural Network	Typically with good predictive accuracy, with each model being a feedforward, fully connected neural network for classification
Medium Neural Network	Typically with good predictive accuracy, with each model being a feedforward, fully connected neural network for classification
Wide Neural Network	Typically with good predictive accuracy, with each model being a feedforward, fully connected neural network for classification with interpretability of class being difficult
Bi-layered Neural Network	Typically with good predictive accuracy, with each model being a feedforward, fully connected neural network for classification with high model flexibility
Tri-layered Neural Network	Typically with good predictive accuracy, with each model being a feedforward, fully connected neural network for classification with high model flexibility

4. Chapter four: Simulation Setup for Intelligent Protection System

4.1 Introduction

This chapter presents a detailed account of the simulation tests undertaken to study the defect detection system based on the utilization of Machine Learning applications. The simulation will entail the integration of distributed generation (DG) with the distribution network (DN) in order to enhance energy resilience, integrate renewable energy sources, and optimize energy efficiency. The simulation will encompass the micro-grid under both normal and faulty conditions, specifically focusing on various symmetrical faults.

4.2 The Simulation Setup

The objective was to establish a comprehensive micro-grid testing ability that closely replicates real-world conditions in order to assess the efficiency of the fault detection system. It is necessary to identify the essential elements of the micro-grid and their respective measurements, particularly the renewable energy sources like solar panels, energy storage devices such as batteries, line feeders, and the electrical loads that are crucial for the overall functioning of the Micro-grid. Therefore, to create a thorough simulation setup for a fault detection method in a modelled micro-grid, it is necessary to meticulously model the grid, carefully design fault scenarios, integrate detection algorithms, and conduct a comprehensive analysis of the results. The selected simulation platform offered flexibility in simulating various components and defects, as well as giving powerful tools for analysing the performance of the algorithm.

The system test bed used in the study, has two distributed generators namely the Distributed Generator energy source specified as 1500 kVA Solar PV plant with inverter generator converter DC power into Three phase AC power and the second Distributed Generator specified as the 200 kVA Battery energy storage. The Micro-grid feeder lines is modelled using distribution line parameter and feeders in the model.

The predominant configuration for primary distribution systems comprises of three-phase systems. Within the realm of three-phase systems, various fault configurations may manifest, including single-phase faults, two-phase faults, and three-phase faults. In order to adequately mitigate all possible fault scenarios, it is imperative to conduct a comprehensive evaluation of the currents (or voltages) across the entirety of the three phases. Multiple distinct methodologies can be employed to ascertain the existence of high voltage and current components that have been overlaid onto current-voltage signals. In this dissertation, the utilization of artificial intelligence methodology is employed to achieve the desired objective with an initial examination of AI and the utilization of machine learning in the context of a power system simulator model and fault diagnosis for machine conditioning of Micro-grids fault detection algorithms.

4.3 Voltage and Current signal dataset collection.

The collection of voltage and current signals within the circuit is essential for both the intelligent detection of electrical faults and the comprehensive monitoring of the health of the micro-grid. By evaluating the fluctuations in these signals, which can indicate abnormal circumstances such as short circuits and overloads, it is possible to identify the electrical defects that have occurred. The collection of voltage and current signal can be collected at the following bus bars with voltage level depicted as shown in the table Table-4.1. For the purpose of this dissertation, the "Point of Common Coupling" has been chosen as the designated point for data signal collecting in order to conduct a test experiment on the proposed method and analyse the outcomes of the simulation.

4.3.1 Fault current and voltage signals

A fault detection simulation has been conducted on an 11-bus micro-grid system. The system includes a complex network with distinct buses representing various components, such as generators, loads, and storage systems. The micro-grid model, including the network diagram, may be found in appendix-A.

When doing the dissertation and analysis pertaining to the design, protection, and operation of micro grids, one of the most significant parameters to take into consideration is the fault current level in the micro grid. In essence, the expeditious clearance of fault currents holds utmost importance in safeguarding equipment, upholding system stability, reducing operational interruptions, ensuring safety, fortifying grid resilience, and preserving power quality within a microgrid. When it comes to sizing protective devices, estimating fault-clearing times, and assuring the safety and dependability of the micro grid, having a solid understanding of the fault current level is necessary. It is essential to take into account the numerous fault types in the bus that are validated in table 4.1, as individual fault types can have diverse features and ramifications for the operation and protection of the micro-grid. When gathering data for fault analysis in the 11 bus system of the micro-grid, some of the frequent types of faults that should be considered are depicted also in table 4.1 which is shown. A fault detection simulation has been conducted on an 11-bus micro-grid system. The system includes a complex network with distinct buses representing various components, such as generators, loads, and storage systems. The micro-grid model, including the network diagram, may be found in annexure 1.

Table 4.1 Shows the table with bus-bar information and types of faults data that can be collected at the bus-bar.

System Bus	Description	Types of Faults
Bus-1	11Kv	LL, LLG, LLLG, LG,LLL
Bus-4	Load	LL, LLG, LLLG, LG,LLL
Bus-5	Inverter Bus(4160V)	LL, LLG, LLLG, LG,LLL
V-Pcc	1000V	LL, LLG, LLLG, LG,LLL
Bus-6		LL, LLG, LLLG, LG,LLL
Bus-7	Transformer Bus(BESS)	LL, LLG, LLLG, LG,LLL
Bus-9	BESS	LL, LLG, LLLG, LG,LLL
Bus-10	Load	LL, LLG, LLLG, LG,LLL
Bus-11	Load	LL, LLG, LLLG, LG,LLL

4.4 Results of the Micro-grid simulation

In this section, the findings that were obtained from the simulation setup for Intelligent the fault detection and identification technique are presented. This technique involves exciting both normal and faulty condition with signals that have a variety of voltage and current. The findings are presented in the order described below: Voltage and current measurements for a normal and faulty micro-grid.

4.4.1 Final Micro-grid Performance Result

It is necessary to offer a full review of the micro-grid's performance in the context of a research dissertation that is centred on a micro-grid with Machine learning or artificial intelligent-based fault detection applications. The simulation of the 11-bus micro-grid system will yield useful insights into stability analysis. This simulation will facilitate the compilation of data sets for various fault scenarios, which are necessary for machine learning multi-class classification. Therefore, the primary purpose of the research dissertation is to evaluate the performance of the fault detection algorithm. This section will present a synopsis of the micro-grid's performance in terms of its reliability, stability, and effectiveness in defect detection in a way that is succinct and easy to understand. With Table 4.2 characteristics are shown actual simulation results obtained of maximum voltage and current data.

Table 4.2 Show maximum current from normal operation of the Micro-grid model Peak Resultant Maximum Voltage and current signals.

Maximum	I a(Amps)	IbAmps)	Ic(Amps)	Va(Volts)	Vb(Volts)	Vc(Volts)
Bus-1	1,44E+03	1435	1435	8299	7588	7334
Bus-4	1398	1398	1397	6971	6974	6967
Bus-3	1535	1474	1537	70,02	69,29	70,17
Bus-5	269	264,8	270	2442	2384	2375
V-PCC	1569	1593	1619	497,6	403	336
Bus-6	1116	1131	1133	606,6	579,6	565,3
Bus-7	1346	1280	1328	907,1	1001	734
Bus-9	646,2	626,5	531,8	207,3	180,3	184,6
Bus-10	595,1	597,2	598,9	328,4	325,7	314,1
Bus-11	36,86	39,4	39,75	297,5	302,3	232,9

The micro-grid's capacity to sustain a uniform and dependable power provision during typical operational circumstances shall be examined. The notable characteristic of this system lies in its consistent and unwavering performance as shown evidently by Table 4.2 maximum voltage and current align within voltage rating of the model bus-bars where measurement of the datasets are to taken.

4.4.2 Voltage and current per bus-bar signals results

The micro-grid's bus-bar locations will be depicted graphically to display the readings of current and voltage. This study presents a validation of the operational efficacy of the model, allowing for the assessment of parameter ratings for the micro-grid during its functioning.

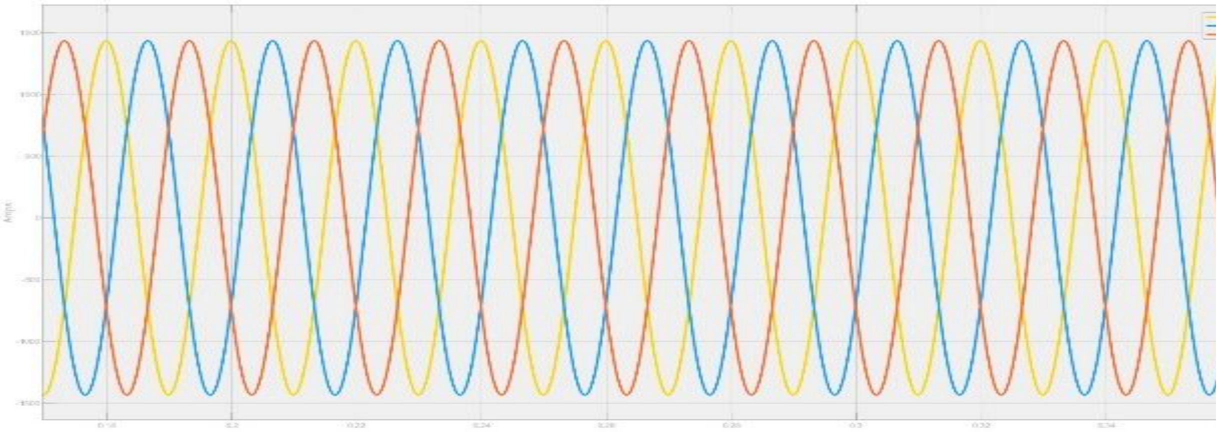


Figure 4.1 Current waveform signal at bus-bar 1 of the micro-grid.

Figure 4.2 Voltage signal at busbar 1

Bus-1, which has an 11kV rating on the micro-grid network, is located along the 120kV supply side of the grid, which has a step down transformer of 120kV/11Kv. The current and voltage signal data set are collected at the bus-bar-1, Figure 4.1, and Figure 4.2 shows these signals in waveforms during normal conditions. For the current signal, the x-Axis is Time and the signal is collected for the period of 0 seconds to 3seconds. The signals legends of the signals as per figures can be depicted as in Table 4.7:

Table 4.3 Shows the legend for the current and voltage signals.

Voltage and current colour coding signals		
Va	Ia	Red
Vb	Ib	Yellow
Vc	Ic	Blue

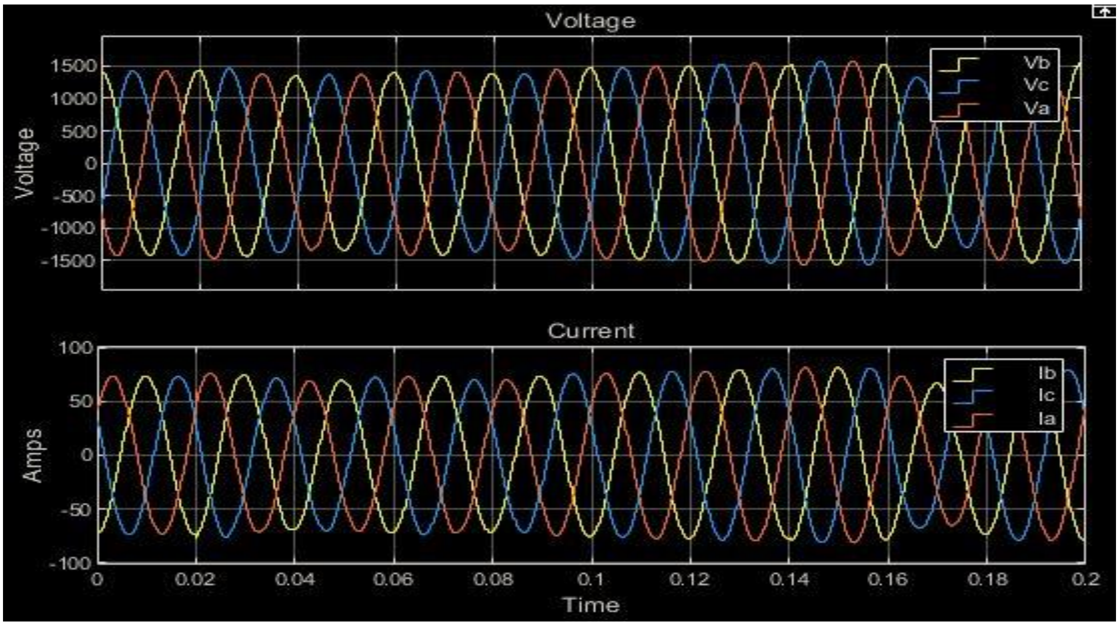


Figure 4.3 Bus-bar 3 current and Voltage signal during normal operation.

Bus-bar- 3 and Busbar-4 can be located at the micro-grid network on the input side of the 11kV/1Kv transformer, and bus-bar 4 after the Figure 4.3, Figure 4.4 and Figure 4.5, shows current and voltage signals under no fault or normal condition with the X-axis showing the period which is from 0 seconds to 3 seconds. Bus three (3) shows the

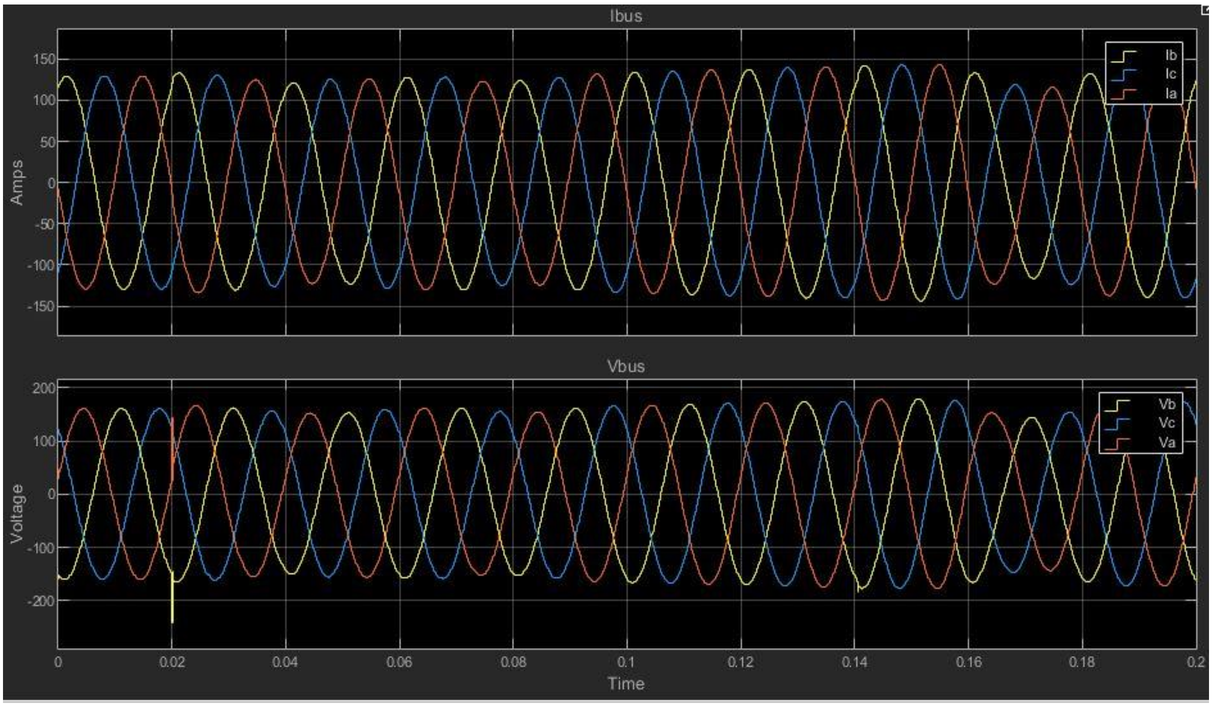


Figure 4.4 Bus-bar 4 Voltage signal during normal operation.

Busbar-4 can be located at the micro-grid network on the output side of the 11kV/1Kv transformer, and bus-bar 4 after the Figure 4.3, Figure 4.4 and Figure 4.5, shows current and voltage signals under no fault or normal condition with the X-axis showing the period which is from 0 seconds to 0.2 seconds.

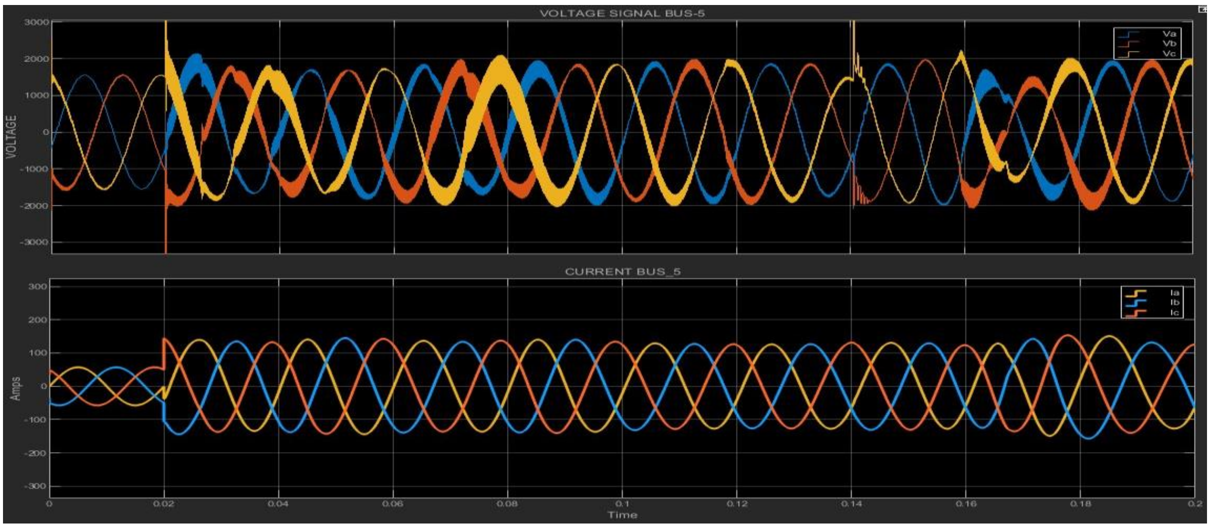


Figure 4.5 Bus-bar 5 Voltage signal during normal operation.

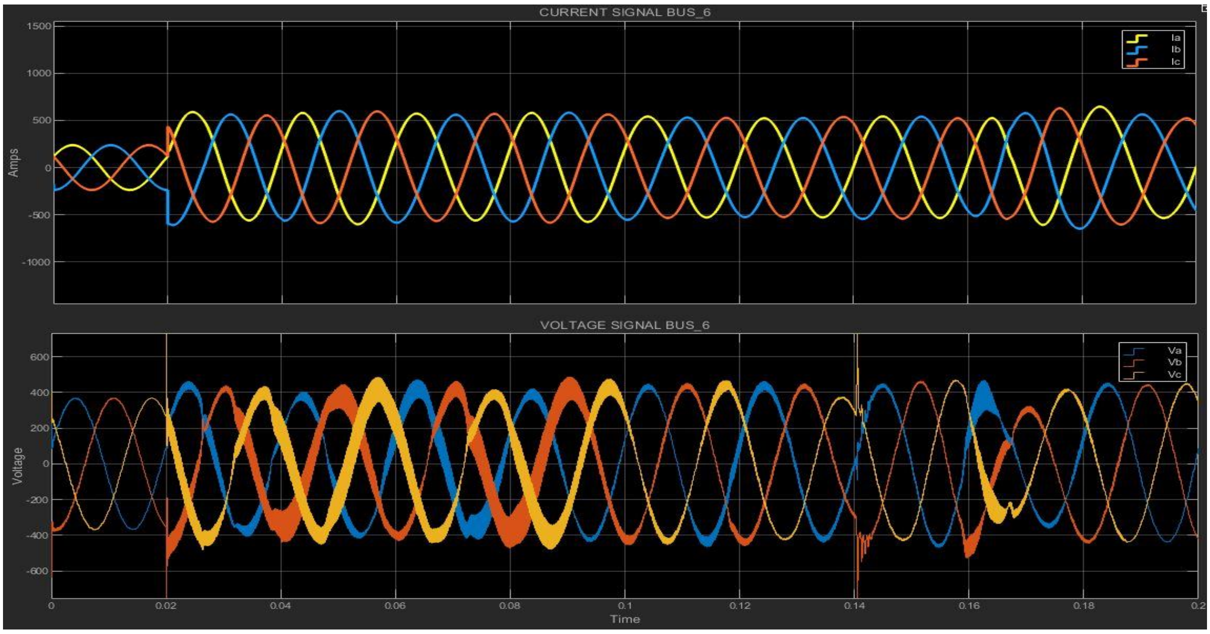


Figure 4.6 Bus bar 6 Voltage and current signal during normal operation.

Bus-bar 5 is located between the solar-PV array and transformer 1 on the network structure, current and voltage signals under no fault or normal condition with the X-axis showing the time period which is from 0 seconds to 3 seconds.

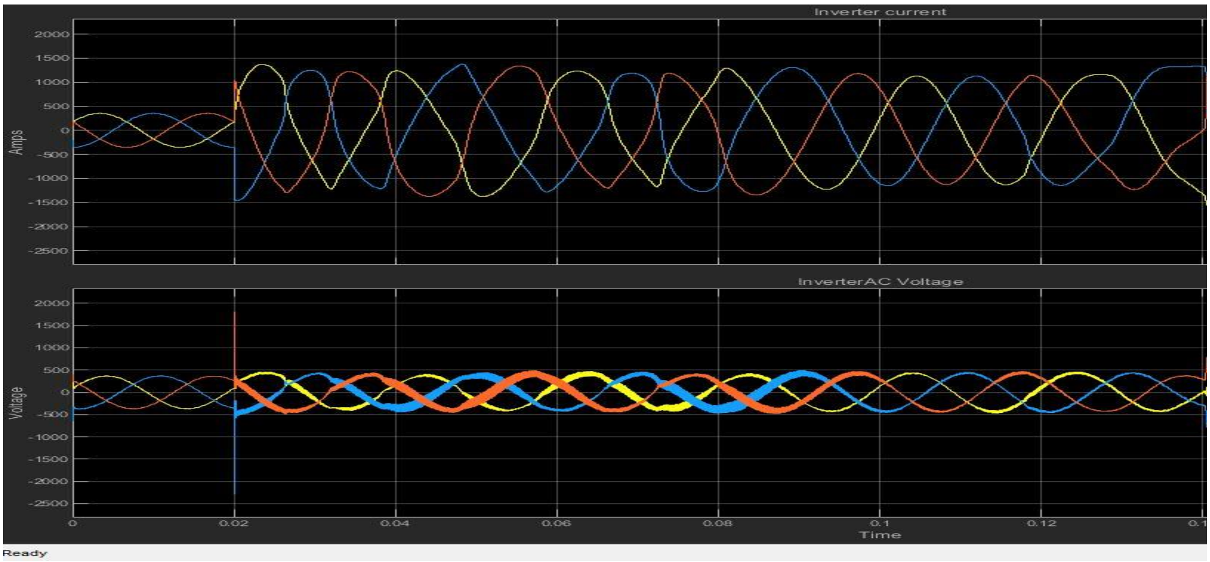


Figure 4.7 Bus-bar 7 Voltage and Current signal during normal operation.

Bus-bar 7 and Bus-bar 12 is located after BESS transformer which is fed from the BESS supply on the network structure, current and voltage signals under no fault or normal condition with the X-axis showing the time period which is from 0 seconds to 0.14 seconds are given in figure.

.

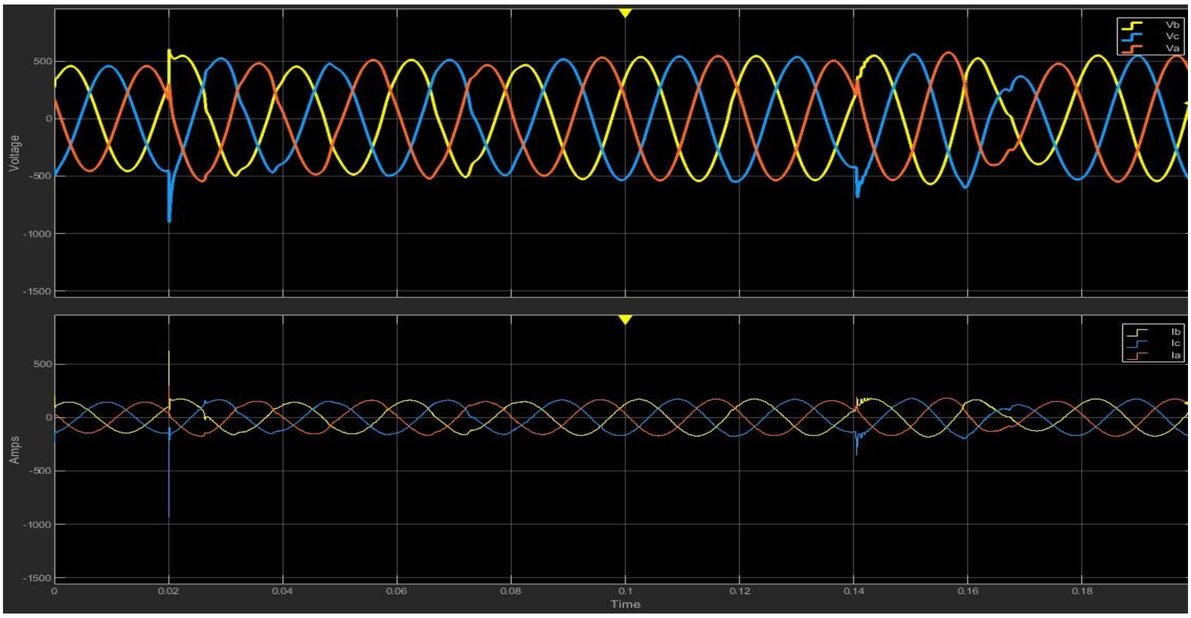


Figure 4.8 Bus-bar 9 current signal during normal operation

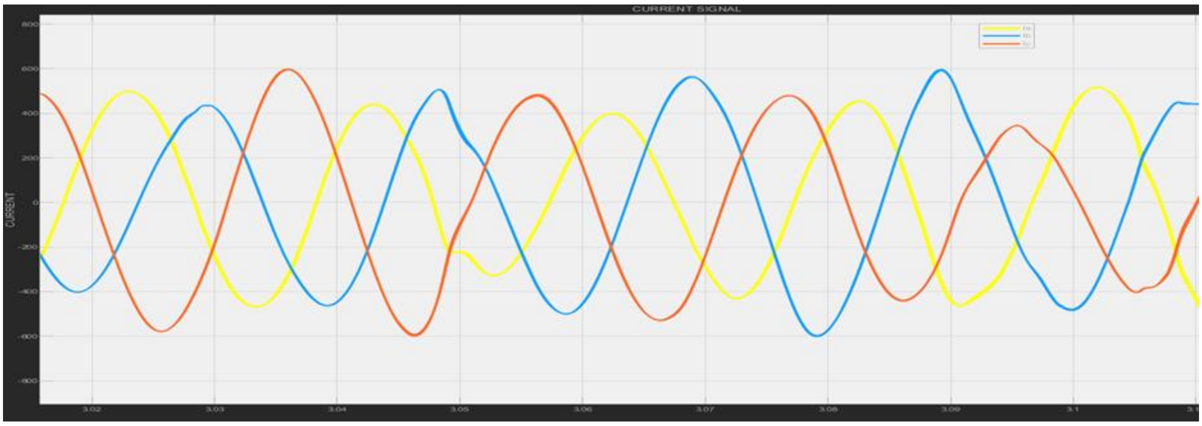


Figure 4.9 Bus-bar 10 Current signal during normal operation.

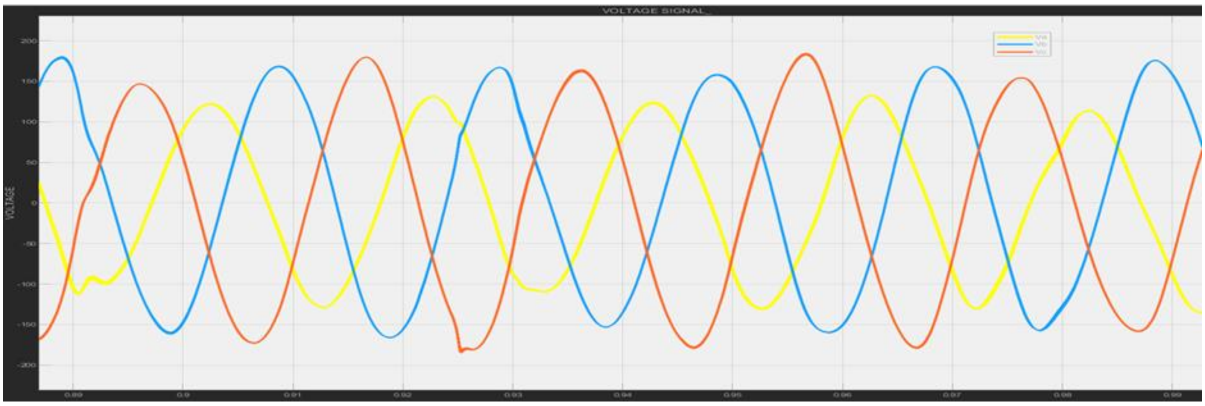


Figure 4.10 Bus-bar 10 Voltage signal during normal operation.

Bus-bar 9 and Bus-bar 10 is located for loads, which is fed from the entire supply on the network structure. The loads are rated at 2x 500kW and single load is at 200kW current and voltage signals under no fault or normal condition with the x-Axis showing the time period which is from 0 seconds to 3 seconds are given in figure.

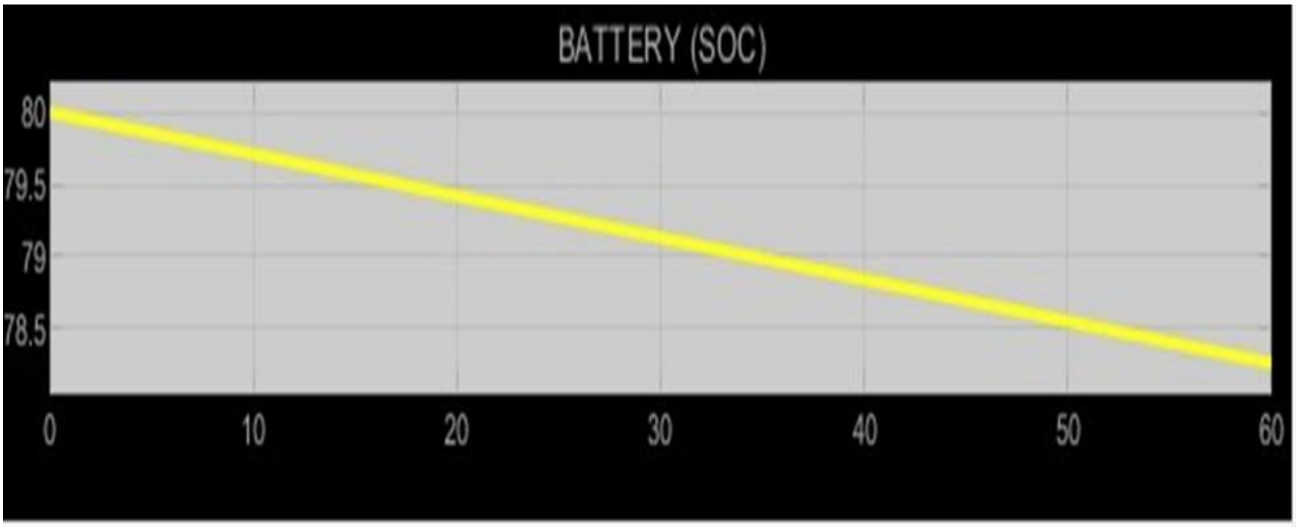


Figure 4.11 The performance of the battery over a period of 60 seconds it rate of discharge.

The performance of the battery system state of charge during a 60-second period of the micro-grid. It shows the battery state of charge drops initially to below 78% during that period. For maximum observation for optimum analysis of the micro-grid performance, the simulation was ran for 60 seconds throughout.

4.5 Fault and signal location

Now that we have witnessed the operation of the network under regular conditions, we are content with the way it operates. We will now proceed to the most important section of the dissertation, by inducing a symmetrical fault in a distribution network with Distributed Generation. This portion deals with the process of producing a fault condition in the distribution network that is symmetrical in nature as shown in Figure. This indicates that the fault currents and voltages are balanced across all three phases of the power system. Choosing the type of fault is done to determine the kind of symmetrical fault that intended to get current and voltage data for machine learning application into the system. Faults that include line-to-line or line-to-neutral connections, as well as three-phase short circuits, are among the most common types to be simulated. The process involves inserting the fault as shown in Figure 4.8 into the distribution network at a location next to PCC. The use of simulation tools that are able to model fault conditions is one way to accomplish this goal. During the time that the fault is active, it is necessary to monitor a variety of characteristics, including voltage, current, and power quality, at a variety of sites throughout the distribution network by incorporating measurement tools into the simulation tools itself in order to collect data.

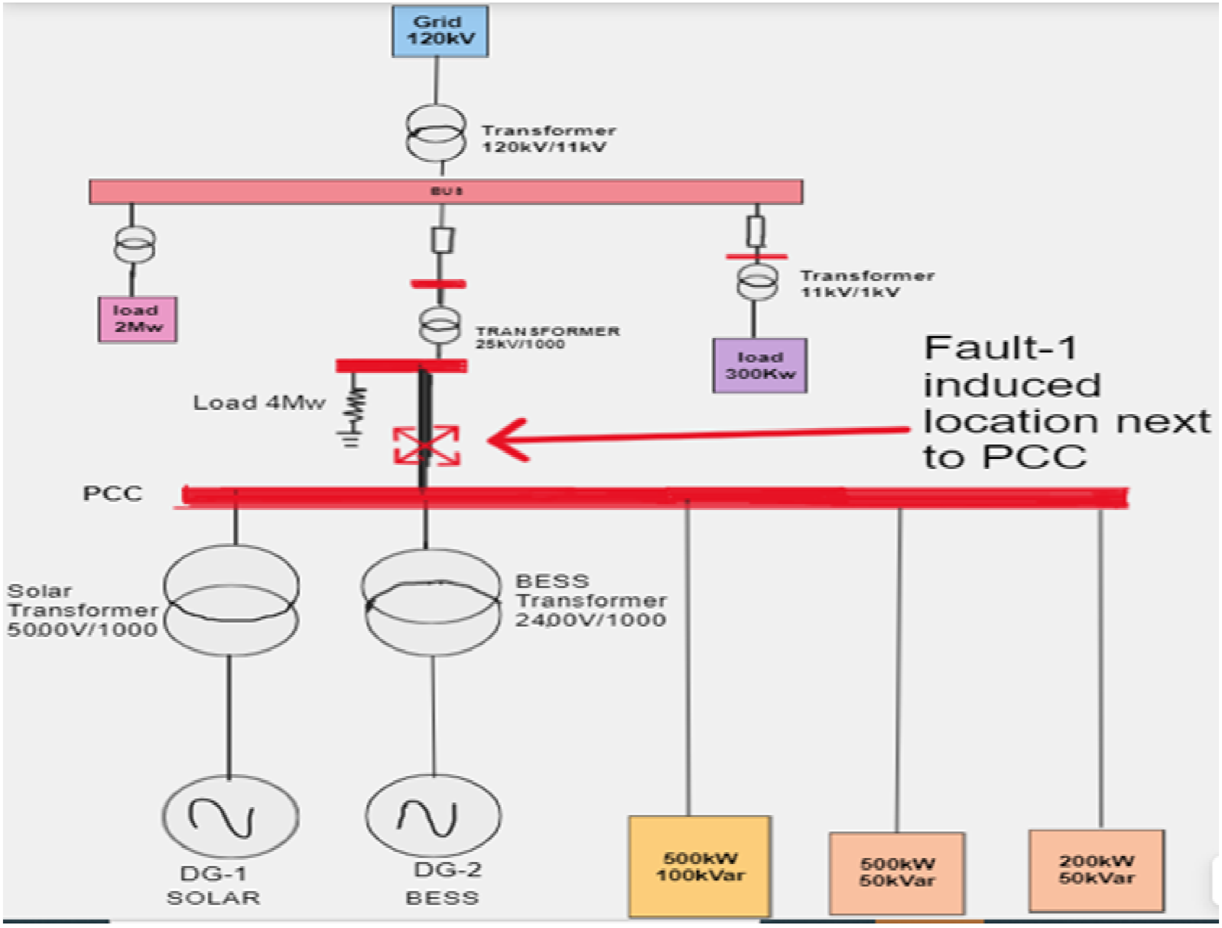


Figure 4.12 Fault induction in the distribution network with Distribution.

Please note the fault could have been induced in any location in the micro-grid, but for the sake of the experimentation the selected location and collection of data signal location (Point of common coupling) has been chosen for investigative and analysis purposes.

4.6 Results after fault signal for data extraction for modelling

Execution of an analysis on the data that have been collected , is done to see if the fault is, in fact, symmetrical, meaning that the currents and voltages are balanced across all three phases. This analysis is essential for establishing whether any corrective actions are required in order to analyse the impact of the fault on the system and to establish and train and test the fault detection algorithm. The below graphical data is collected from the signal connected at the PCC busbar.

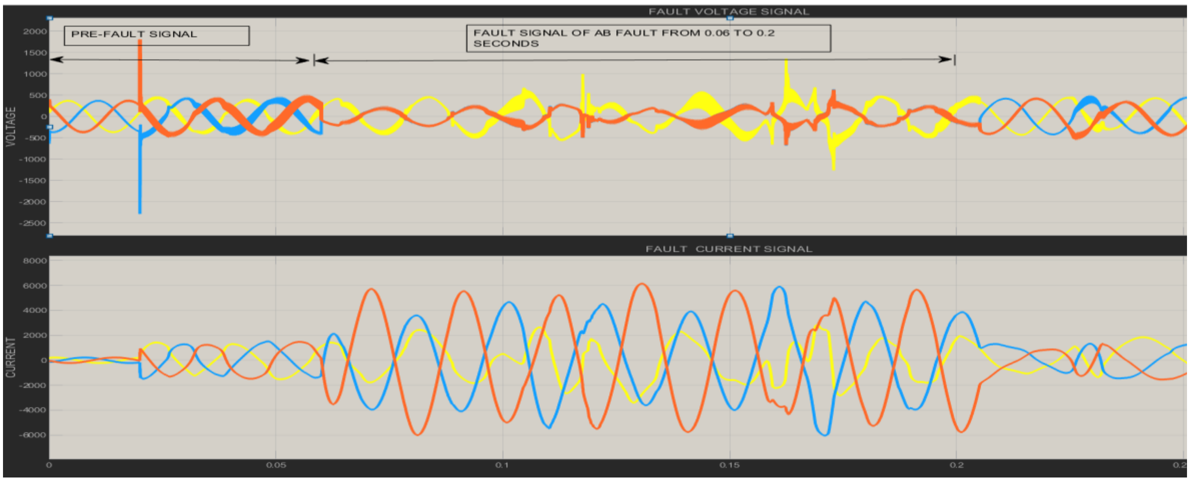


Figure 4.13 shows fault signal from line AB current and voltage.

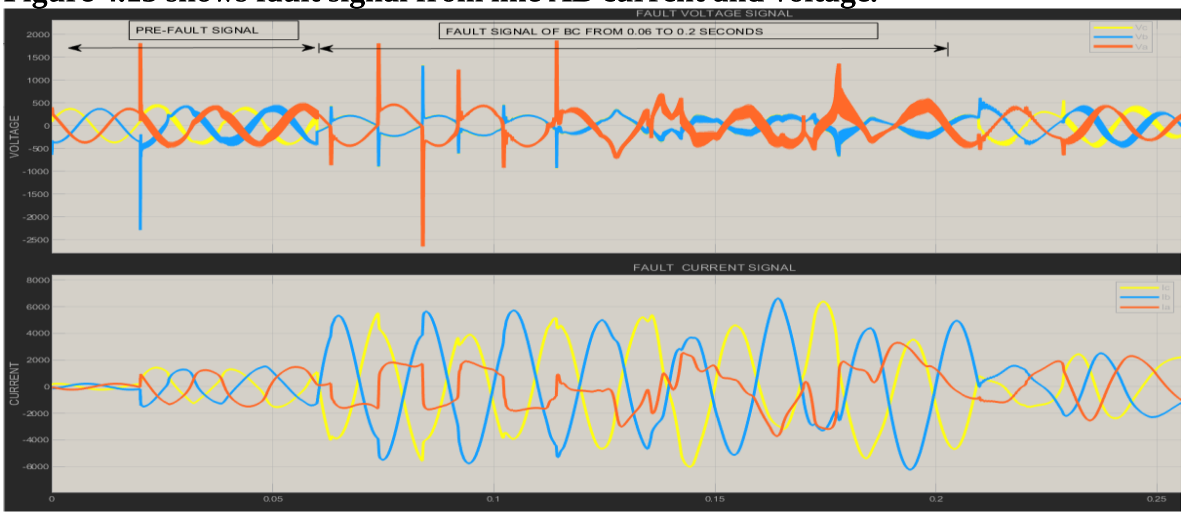


Figure 4.14 Shows fault signal for line BC from pre-fault condition to fault.

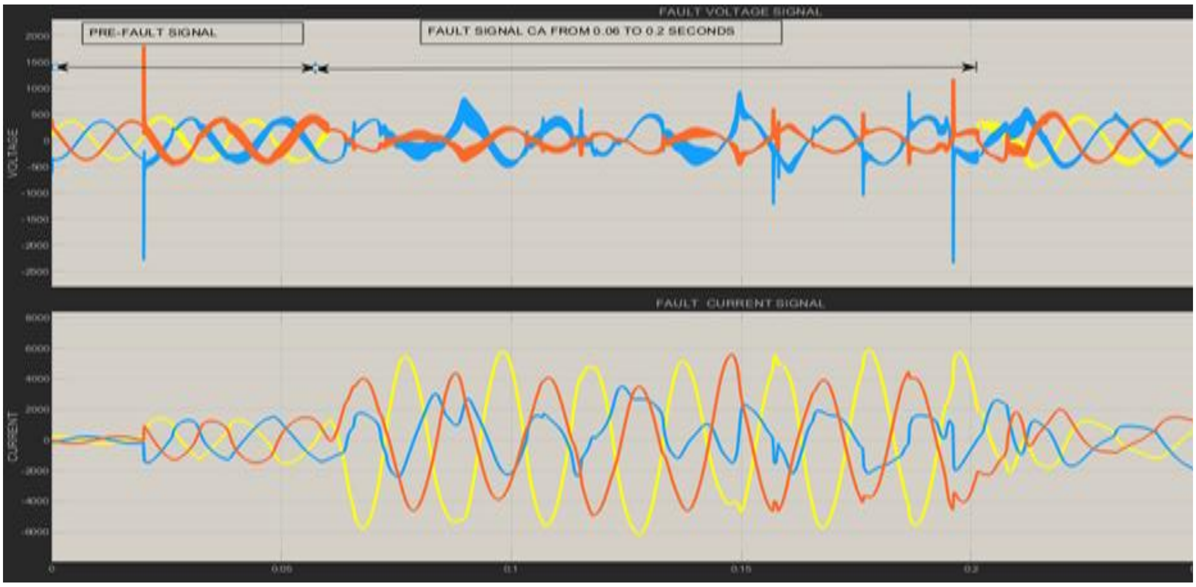


Figure 4.15 shows fault signal of current and voltage from fault line CA

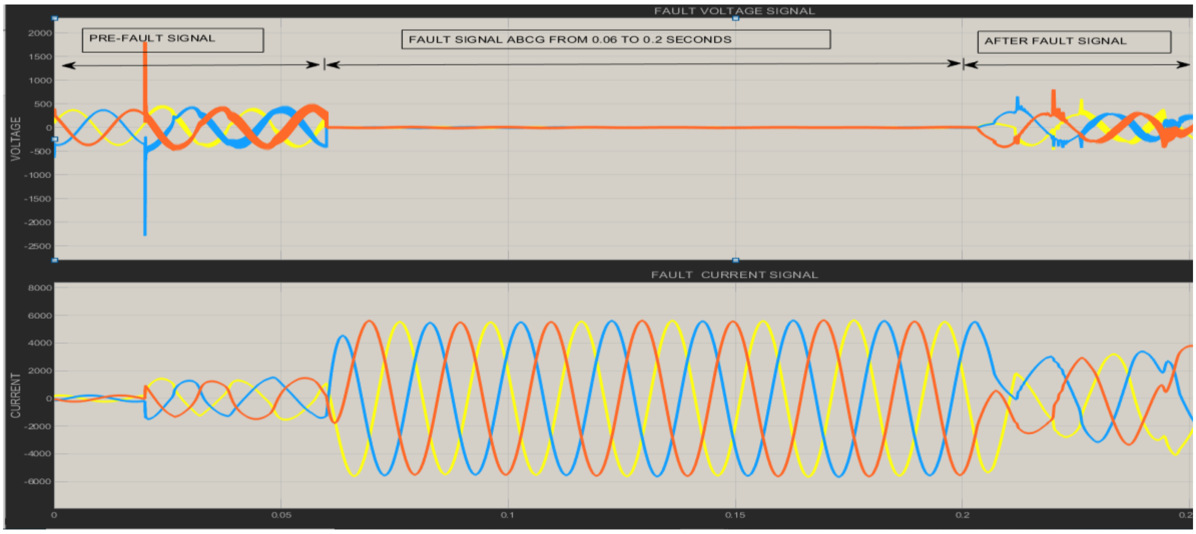


Figure 4.16 Shows resultant of fault current and voltage signal from fault line ABCG

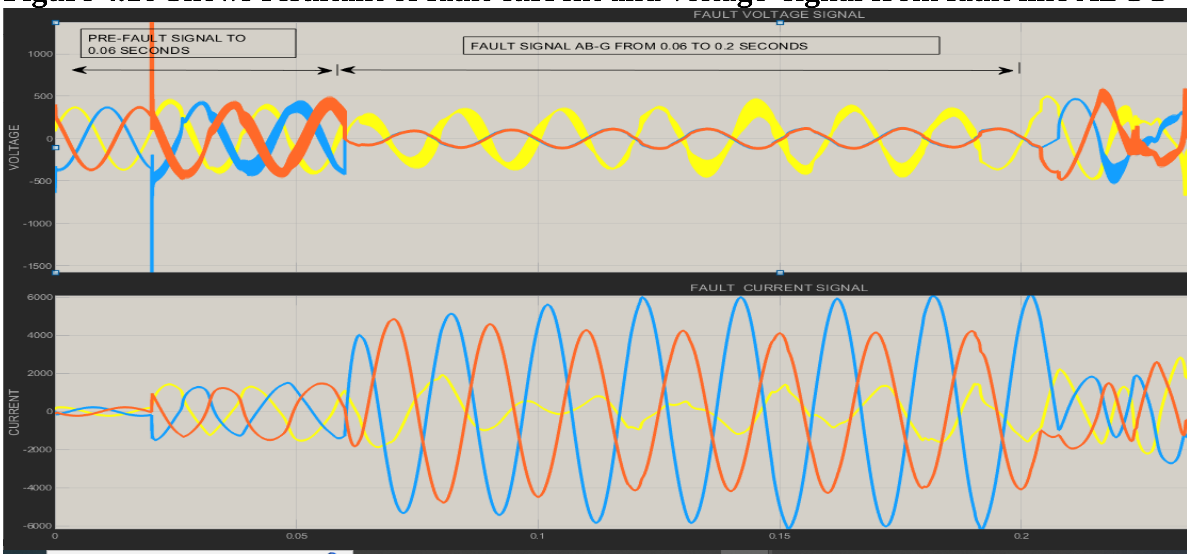


Figure 4.17 Shows resultant voltage and current fault signal when fault is induced in line AB-G

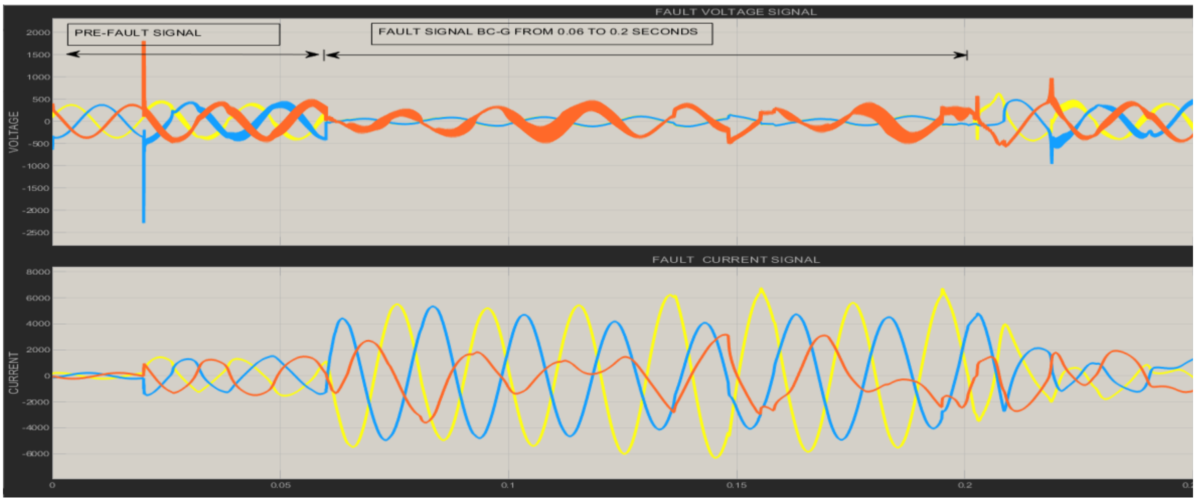


Figure 4.18 Shows resultant of voltage and current fault signal from fault induced in line BC-G

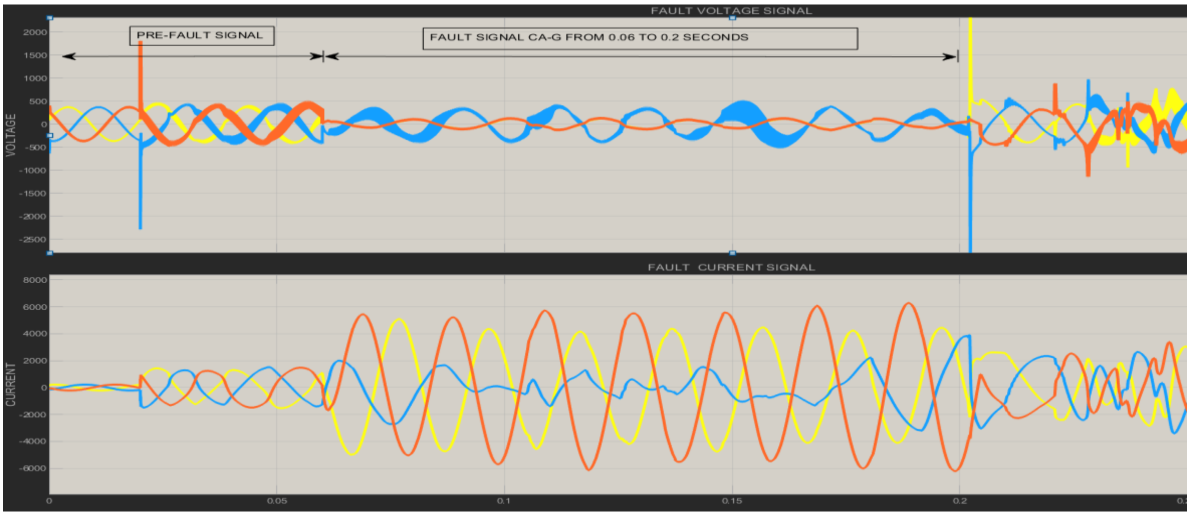


Figure 4.19 Shows resultant of voltage and current fault signal from fault induced in line CA-G

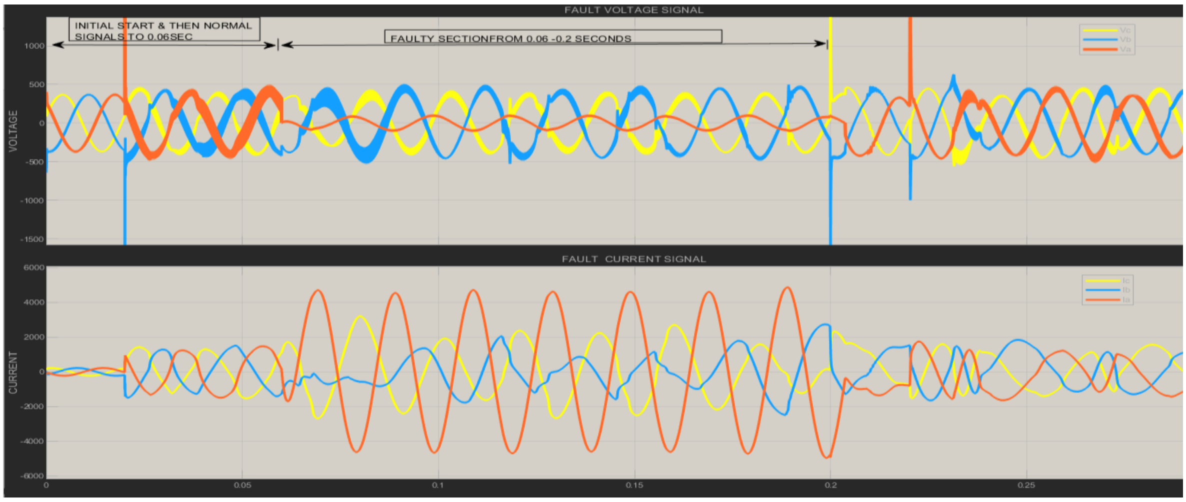


Figure 4.20 Shows resultant of voltage and current fault signal from fault induced in line A-G

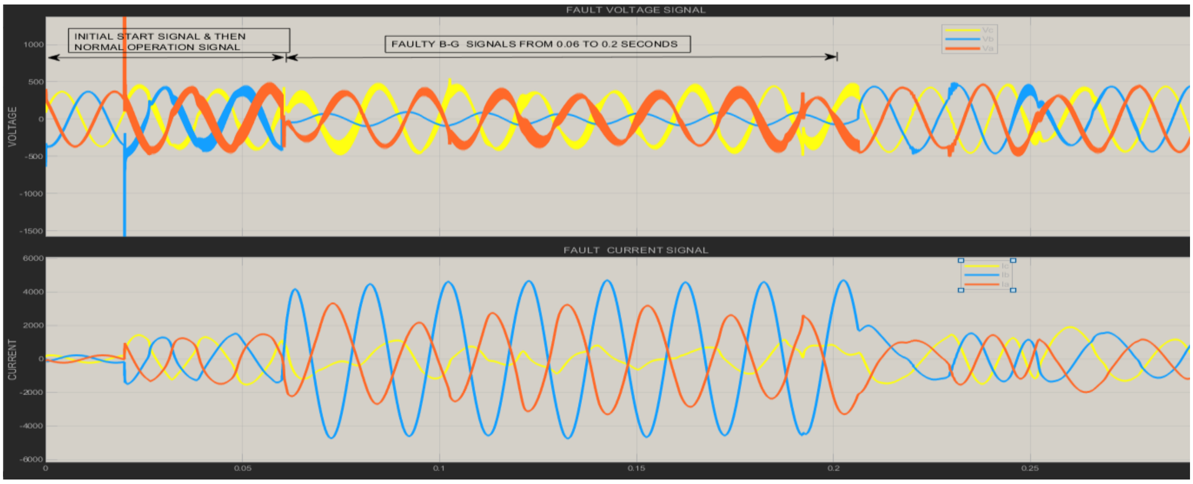


Figure 4.21 Shows resultant of voltage and current fault signal from fault induced in line B-G

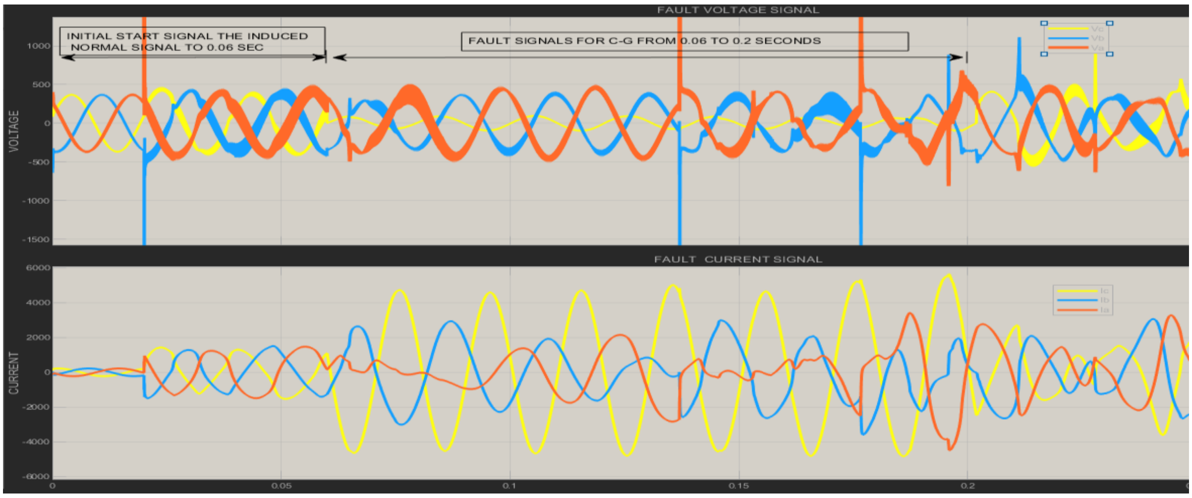


Figure 4.22 Shows resultant of voltage and current fault signal from fault induced in line C-G

The provided information pertains to the voltage and current signal waveforms observed during the pre-fault and fault sessions. These waveforms need to be transformed into numerical data in order to undergo data pre-processing for our machine learning algorithm, which aims to detect faults and classify them accordingly.

4.7 Fault signal Parameter Results summary

The fault signals and normal conditions signals have been simulated for the different symmetrical fault conditions. The symmetrical faults were induced at time from 0.06 seconds to 0.2 seconds. During pre-fault at around 0.02 there is a spike on the voltage signal on all the scenarios which can be attributed to integration of the DGs during that period which causes a small disturbance of overshooting of voltage signal. It can be observed from the overshoot the increase in system current, which is still within the safe limit of the network and is not quantified as fault current. These results now require be conversion from graphical description into numeric value and then pre-processing for machine learning model generation and coding for multiclass classification. The resultant pre-processed different faults and normal data-set has been converted and recorded numerically in the following tabular way, shown in Table 4.4 for which the machine model and the number of data sets are recorded within the fault period per signal. The data of current and voltage signals can be dataset containing numerical data and fault codes, "G-C-B-A fault lines as category dataset and categorical data for use of classification algorithms.

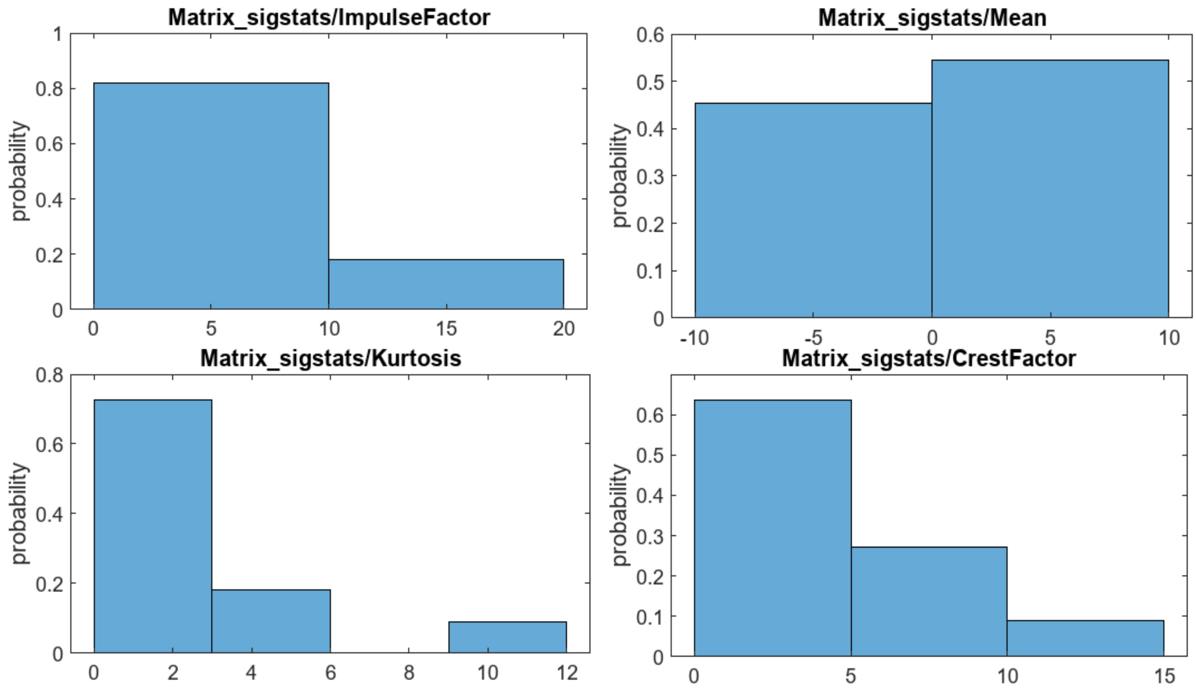
Table 4.4 Shows data amount of pre-processed data of current and voltage signals with corresponding fault codes.

G	C	B	A	Ia	Ib	Ic	Va	Vb	Vc	Fault code	Fault Line
✓			✓	1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	1001	AG
✓		✓	✓	1001-3001	1001-3001	1001-3001	1001-3001	1001-3001	1001-3001	1101	ABG
	✓	✓		1001-3001	1001-3001	1001-3003	1001-3001	1001-3002	1001-3003	0110	BC
✓	✓	✓		1001-3001	1001-3001	1001-3001	1001-3001	1001-3001	1001-3001	0111	BCG
	✓	✓	✓	1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	1110	ABC
✓	✓	✓	✓	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1111	ABCG
		✓	✓	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1100	AB
	✓		✓	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1003-3005	1010	CA
✓	✓		✓	1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	1011	CAG
✓		✓		1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	1002-3003	0101	BG
✓	✓			1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	0011	CG
0	0	0	0	1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	1001-3002	NO FAULT	NORMAL

The dataset focuses on applying Machine Learning to detect and classify defects in micro-grid to create a Machine Learning algorithm to classify the type of fault in power lines in order to avoid incidents.

4.8 Machine Learning Feature extraction results

The machine-learning model has been developed utilizing the parameters of the pre-processed data to deduce the resultant outcome from the feature extraction technique. Through the modelling of various failure scenarios, the collecting of crucial data regarding the system's overall health and operational effectiveness have been accomplished, due to its intrinsic qualities. The implementation of the feature extraction procedure involved the application of signal processing feature-engineering techniques. After the simulation is finished, the characteristics of the simulation are extracted and the data is stored using a simulation ensemble data storage system. The graphical representation of the best feature for the model can analysed using the Figure 4.26 which shows the data probability of the different features that will have the best performance in classifying the fault. This system can be thought of as a database that is specifically designed to meet the needs of engineers working on the development of condition monitoring and predictive maintenance software. The data stored in this system is derived from raw sensor data that has not been processed.



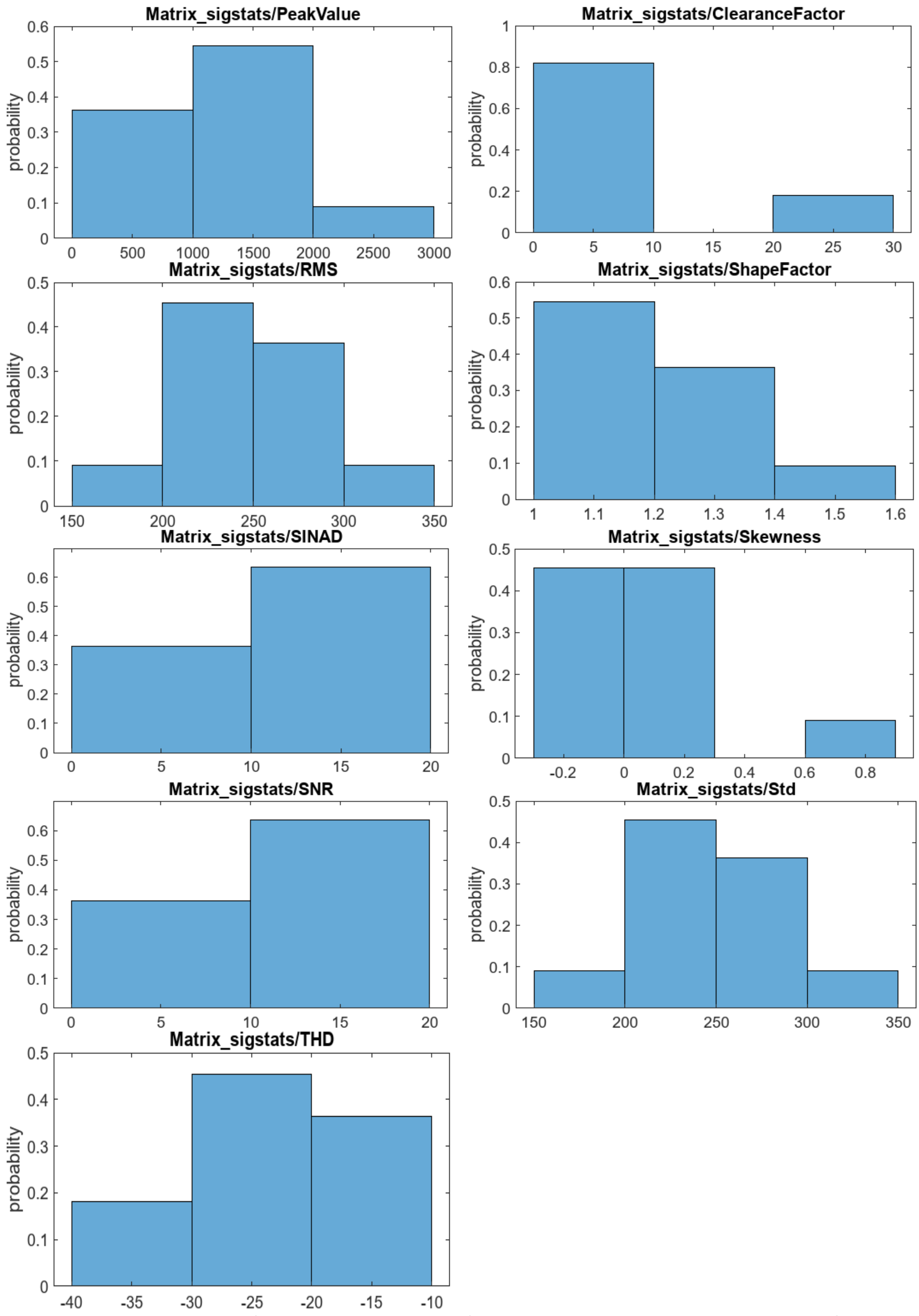


Figure 4.23, shows the feature extraction of the voltage and current signals for data ensemble.

4.9 Histogram Feature extraction

A histogram is a visual depiction that illustrates the frequency distribution of a given dataset. This visual representation illustrates the relative occurrence of different values within a given dataset. In order to generate a histogram pertaining to a particular feature, it is necessary to follow a specified procedure. The provided dataset exhibits values for both current and voltage. In order to generate a histogram for a particular feature. This study explores the application of the Diagnostic Feature Designer software (DFDS) for the extraction of time-domain and spectral characteristics from data, with the aim of designing a fault detection algorithm which is a component of machine condition monitoring,

With the collection of measurements from the micro-grid the location of the “Point of Common Coupling (PCC)” as measurement devices are located at the bus-bars under various fault scenarios. The Diagnostic Feature Designer can greatly improve the capacity of predicting micro-grid electrical defects. Histogram characteristics are valuable for capturing the distribution of data values, which might provide insights into specific circumstances or faults in the micro-grid. The DFDS was utilized in the dissertation to extract certain features from the histogram, including statistical measurements such as mean, median, mode, skewness, and kurtosis etc...Of the distribution. These features were then subjected to further study. The analysis is conducted using DFDS feature ranking techniques to assess the significance of these features. This stage aids in determining which features are the most indicative of the health or faults of a micro-grid. After extracting and selecting the features from the accumulated histogram features, they are exported for usage in the workflow of machine learning models.

In results of the Histogram feature extraction the pre-processed current and voltage data being the inputs of the model are then subsequently visualized interactively, with the readings already categorized based on various fault circumstances. Once the signal and spectral features have been retrieved from the data, the efficacy of these recovered features can be assessed through the utilization of histograms. Additionally, it is possible to assign a number ranking to these qualities in order to ascertain their effectiveness in distinguishing between healthy and flawed behaviour as shown in Figure 4.26 and Figure 4.27 with fault codes included. Ultimately, the most efficient characteristics are transferred to the Classification Learner application for additional assessment of their efficacy and for training machine-learning models.

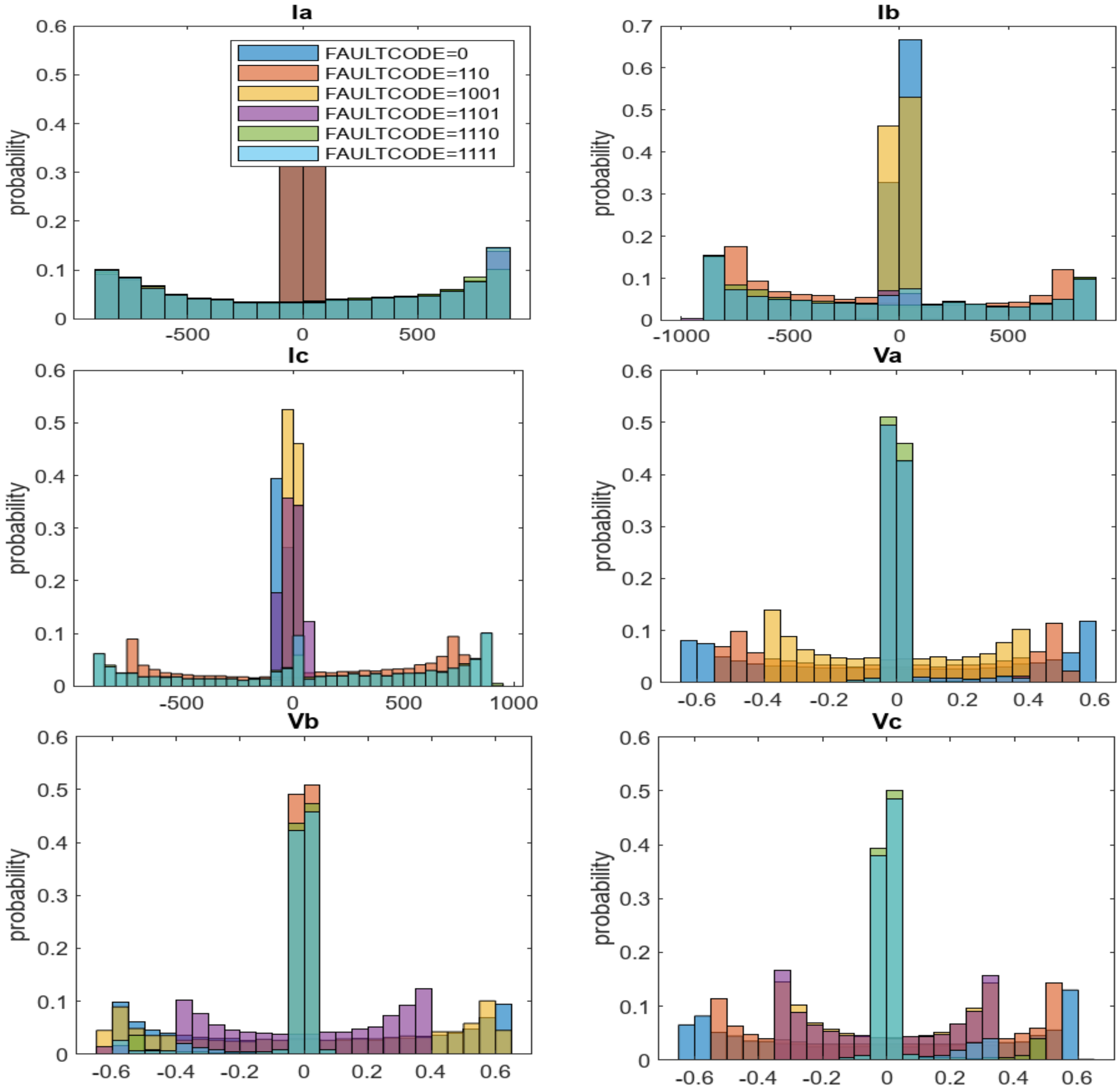


Figure 4.26-b, Histogram probability factor analysis of the feature extraction.

4.10 Classification Learner results

Upon the initial implementation of the classification learner app, in which is implemented in model for training, validating, and comparing different classification models and is efficiently also utilized for fault categorization in the Micro-grid, aiding in the identification and classification of various fault types using both historical and real-time data. There are various classification models available, including Decision Trees, Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), and Ensemble techniques. A comparison will be made based on the performance of several different models.

The primary objective is to partition the data into distinct training and validation sets. Subsequently, the chosen algorithm shall be trained exclusively on the training data. Following the completion of the training process, the application offers a diverse range of performance metrics and visual representations to assess the efficacy of the model's performance. This encompasses metrics such as accuracy, the confusion matrix, ROC curves, and other relevant parameters. From this point onwards, one can utilize the trained model for making predictions on novel data by generating MATLAB code or employing the application's interface. The following are the results from the training of the model of various performing classifiers:

4.10.1 Scatterplot results

A scatterplot, within the framework of a classifier learner, serves as a visual depiction of the data points, wherein each point on the plot corresponds to an instance from the dataset, as a visual representation to effectively depict the dispersion and interdependencies among various attributes or data instances. By implementation the scatter plot in the Classification Learner app incorporates features collected from the Diagnostic Feature Designer Software (DFDS) in MATLAB. The features extracted from the Diagnostic Feature Designer Software (DFDS) are included into the Classification Learner application and subsequently shown visually through a scatter plot, as depicted in Figure 4.27-4.29. These feature extracted are saved into tabular form which are labelled as categorical data (Predictors) and numerical data (Features).

The scatter plot illustration facilitated the comprehension of the associations between various attributes and their efficacy in categorizing fault types. The scatter plot was utilized to efficiently analyse the model, enabling visualization of feature data and enhancing fault classification in the micro-grid. Figure 4.27 to 4.29 shows the scatter plot that identify predictors which are the current and voltage data that separate classes by plotting different pairs of predictors which enable visualization of the training data for the purpose of misclassified point on the scatterplot. The below scatterplot shows model prediction results.

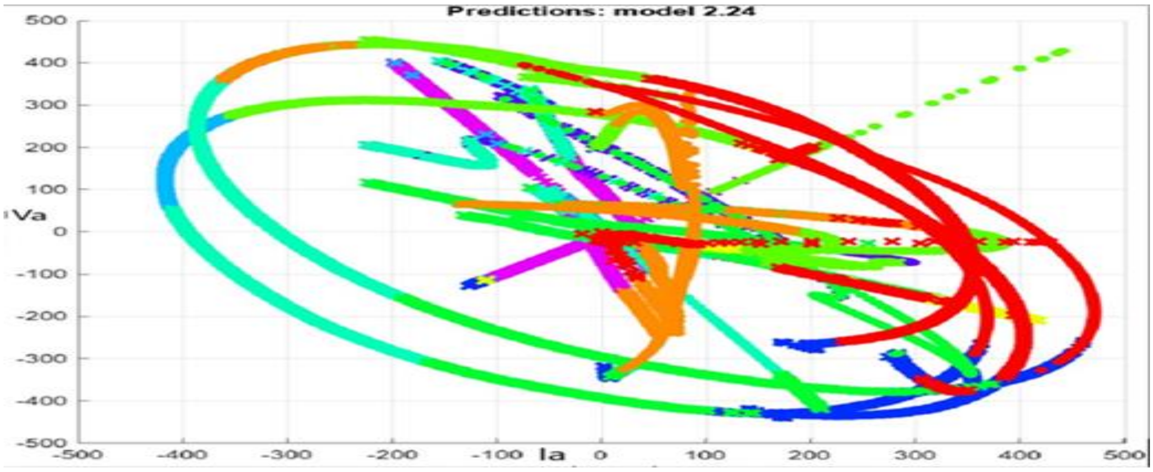


Figure 4.27 X-axis Ia and Va –Yaxis Scatterplot for the classification model.

The scatterplot displays the outcome of the model's prediction for the fault predictor’s data of classes Ia and Va, which are distinct predictor variables. The data representing the faults in the "classes" exhibit an elliptical shape, resembling a two-dimensional figure defined by its axes. The presence of the Red colour, representing fault code-0 denoted as "no fault", along with the other classes and their respective colours such as 110(BC)-Green, 1010(CA)-Cyan, 1011(ACG)-Blue, and 11(CG)-Orange, collectively form an elliptical shape within the predictor framework. The given data set of class 110 (BC)-Green exhibits values that surpass or extend the specified boundary. The remaining data and its corresponding classes are contained within the data "boundary" of the elliptical shape that has been formed. The objective is to identify predictors that exhibit effective class separation. Based on the displayed results, there is clear evidence of this, as the class data is not randomly distributed across the predictor plane. Instead, the data exhibits a distinct elliptical shape, indicating a well-defined boundary within which the data points are scattered.

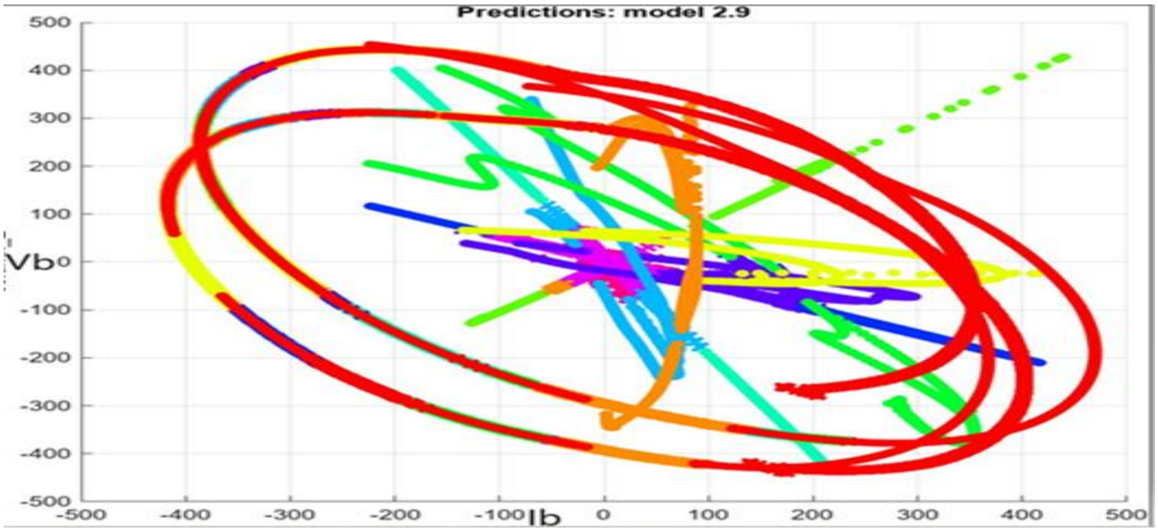


Figure 4.24 X-axis Ia and Va –Yaxis Scatterplot for the classification model.

The scatterplot exhibits the model's predicted outcome for the fault predictor's data of classes Ib and Vb, which are characterized by separate predictor variables. The data pertaining to the flaws observed in the "classes" reveal an elliptical form, reminiscent of a two-dimensional structure characterized by its axes. The inclusion of the Red colour, which signifies fault code-0 or "no fault," provides additional data compared to the scatterplot depicted in figure 4.1. This scatterplot also includes other classes, each represented by a distinct colour: 0101 (BG) in Yellow, 1011 (ACG) in Blue and 11 (CG) in Orange. Together, these classes form an elliptical shape within the predictor framework. The provided dataset for class 110 (BC)-Green displays values that exceed the defined boundary in a linear manner, intersecting the half-plane of the ellipse. The data that remains, together with its matching classes, is encompassed within the data "boundary" of the elliptical shape that has been generated. The aim of this study is to uncover predictors that demonstrate a significant ability to distinguish across different classes. The presented results provide compelling evidence to support the claim that the distribution of class data over the predictor plane is not random. However, the data demonstrates a clear elliptical form, suggesting the presence of a well-defined boundary within which the data points are dispersed.

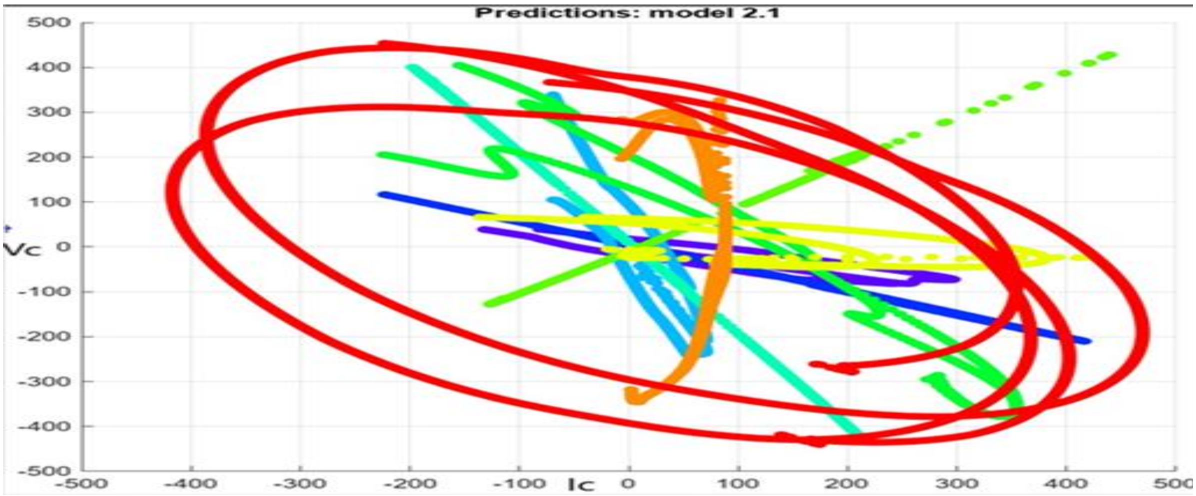


Figure 4.25 X-axis Ia and Va –Yaxis Scatterplot for the classification model.

The scatterplot illustrates the results of the model's predictions for the fault predictor’s data of classes Ic and Vc, which are two separate predictor variables. The data pertaining to the fault observed in the "classes" reveal an elliptical form, reminiscent of a two-dimensional shape defined by its axes. Within the predictor framework, an oval shape is formed by the collective presence of the Red color, which symbolizes fault code-0, indicating the absence of any faults. The provided dataset for class 110 (BC)-Green displays values that exceed or surpass the designated threshold. The data that remains, together with its matching classes, is encompassed within the data "boundary" of the elliptical shape that has been generated. The aim of this study is to uncover predictors that demonstrate a significant ability to distinguish across different classes. The observed results provide compelling evidence to support the notion that the distribution of class data over the predictor plane is not random. On the contrary, the data demonstrates a discernible circular configuration, suggesting the presence of a clearly delineated boundary within which the data points are dispersed.

4.11 Model analysis - Performance per class in the confusion matrix

For the model performance analysis the Confusion Matrix is used as the evaluation tool criteria of the machine learning model due its simple and efficient performance measurement capability. The results of the confusion matrix pertains to the count of instances in which the classifier accurately identifies the positive class as positive. The analysis of the result will be determined. The aforementioned counts are of utmost importance in the calculation of various performance metrics, encompassing but not limited to accuracy and precision. The performance of the confusion matrix will be explained in summary per classifier grouping according to the classification learner app in Tables 4.5. The True Positive and False Negative rate will be used to accomplish in detail the performance of each class of fault in predicting and validation of accuracy. The formula will be used to find these rating using the confusion metrics results

$$TPR = \frac{TP}{TP + FN} \qquad 4.1$$

Where True Positive rate(TP) = correctly classified

False Negative(FN) = Miss-classification

4.11.1 Decision Tree classifiers

In the present investigation, we want to compare classification of different classifiers using the utilization of the Classification Learner tool. Simultaneously training and later testing features of varying granularity, by using Decision tree classifier including fine, medium, and coarse trees. Each model possesses a validation accuracy score, which serves as an indicator of the percentage of accurately anticipated responses. To evaluate the best performing model, a comparison of the confusion matrices for each model is conducted.

The Fine tree classifier

A confusion matrix is a technique used to measure the performance of classification models in machine learning. A confusion matrix is utilized to assess the performance of a fine tree classifier in fault classification within a micro-grid. It provides a comparison between what is true and predicted classifications.

A 57.5% accuracy in the confusion matrix for fault classification in a micro-grid indicates that the model accurately detected a fault type 57.5% of the time. This potential causes for the decreased accuracy may be related to imbalanced data, a phenomenon that occurs If some fault categories are not adequately represented in the training data, the model may fail to accurately categorize these errors. Even a highly detailed model, such as a fine tree classifier, may still be inadequate for capturing the complexity of fault patterns in a micro-grid. While feature selection might improve accuracy, in the case of the Fine tree classifier, the selected features may not encompass all the pertinent information regarding the problems.

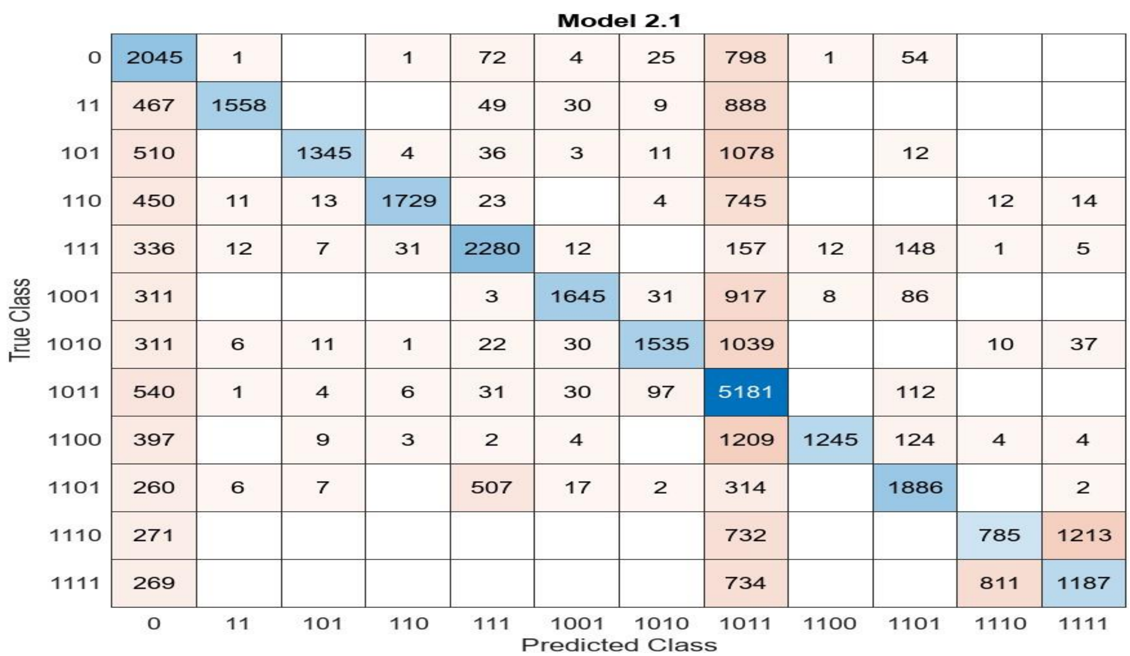


Figure 4.26 Shows the Fine tree classifier confusion matrix results accuracy result 59.9%

Table 4.5 Analysis of confusion metrics results and true Positive rate of the performance of the model.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	2045	956	68.1	31.9
11	1558	2205	51.9	48.1
101	1345	2202	44.8	55.2
110	1729	2338	57.6	42.4
111	2280	2121	76	24
1001	1645	2224	54.8	45.2
1010	1535	2272	51.1	48.9
1011	5181	2082	86.3	13.7
1100	1245	2299	41.5	58.5
1101	1886	2136	62.8	37.2
1110	785	2335	26.2	73.8
1111	1187	2835	39.6	60.4

Table 4.5 shows the summary of the confusion matrix results of the Miss classified and correctly classified data with the True positive rate vs False Negative rate to rate the overall model performance for each fault code . Fault code 1111(Line A-B-C-G) Triple line to ground faults has the highest misclassification data of 2835(60.4%) showing the modelling worst performing fault classification fault code and No fault being the most correctly classified category of fault-code which is the normal operating current and voltage signal.

Medium tree classifier

A medium tree classifier is a type of decision tree classifier, which is a widely used machine learning technique for both classification and regression tasks. A "medium" tree, as the name suggests, is characterized by a balance between shallowness, which may result in under fitting the data, and depth, which may lead to overfitting the data. The workflow has an intermediate level of complexity.

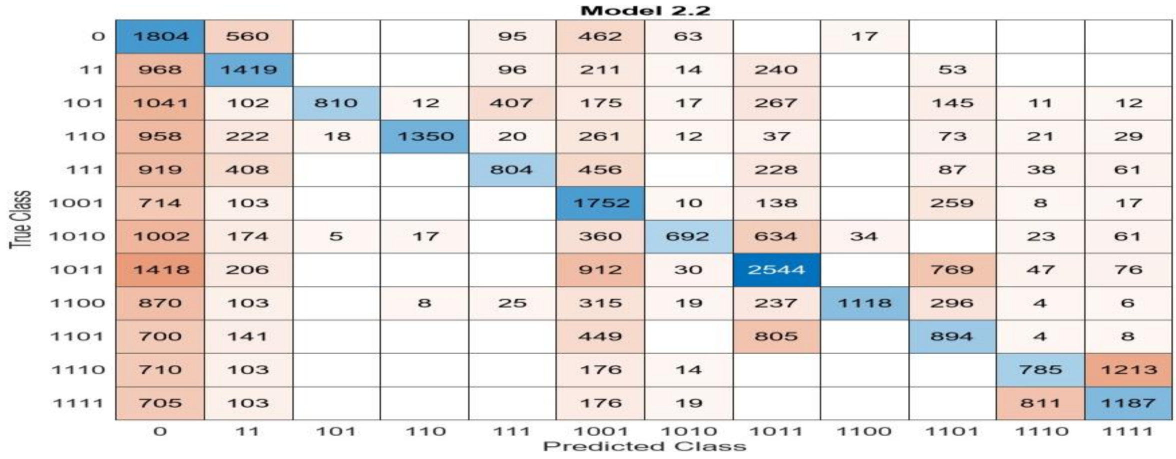


Figure 4.27 Shows the Meduim tree classifier confusion metrics model with total accuracy of 54.5%.

The confusion matrix of this model offers comprehensive understanding of its performance across several classes. Figure 4.31 Showing the Medium tree classifier confusion metrics results, which has a total accuracy of 54.5%. With Accuracy of the model being 54.5%, the suggestion is that it is slightly more than half of the predictions made by the model were true. By analysis this equates that the model may encounter difficulties due to class imbalance if one class is disproportionately represented. Which refers to a classification data set where certain classes have significantly more occurrences than others. Standard classifiers sometimes struggle to handle huge classes and tend to disregard smaller ones in such situations of imbalance.

Table 4.6 Analysis of confusion metrics results and True Positive, Negative rate of the performance of the model.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1804	1197	60.1	39.9
11	1419	1582	47.3	52.7
101	810	2189	27	73
110	1350	1651	45	55
111	804	2197	26.8	73.2
1001	1752	1249	58.4	41.6
1010	692	2310	23.1	76.9
1011	2544	3458	42.4	57.6
1100	1118	1883	37.3	62.7
1101	894	2107	29.8	70.2
1110	785	2216	25.2	73.8
1111	1187	1814	39.9	60.4

In the confusion, metrics we can deduce a high classification number in the main diagonal for the no fault code but for other fault codes a very low number of classified data vs misclassification. The table 4.6 has been used to ensure that the True positive and Negative classification data can be analysed to see which fault could performs in percentage terms.

To improve the model performance, there are a number of strategies that can be developed, including addressing class imbalance, refining feature selection, and even selecting a more sophisticated model in order to increase performance.

Coarse tree classifier

During the process of evaluation, the Coarse Tree classifier model was shown to have an accuracy rate of 41.6% overall. Consequently, this indicates that 41.6% of the total predictions that the model made were correctly predicted. In this particular model, the confusion matrix provides a comprehensive analysis of the model's performance across a variety of classroom settings. The accuracy of this result is 41.6%, which is lower than the accuracy of the decision tree classifier, which is the lowest among the decision tree classifiers. Figure 4.32 shows the Coarse tree classifier which is the worse in performance if you compare the decision tree classifiers. The main diagonal which is the true classification data is shown showing inconsistency in correct classification, with fault codes such as 110(BC) and 1101(ABG) having no data to show correct classification of data.

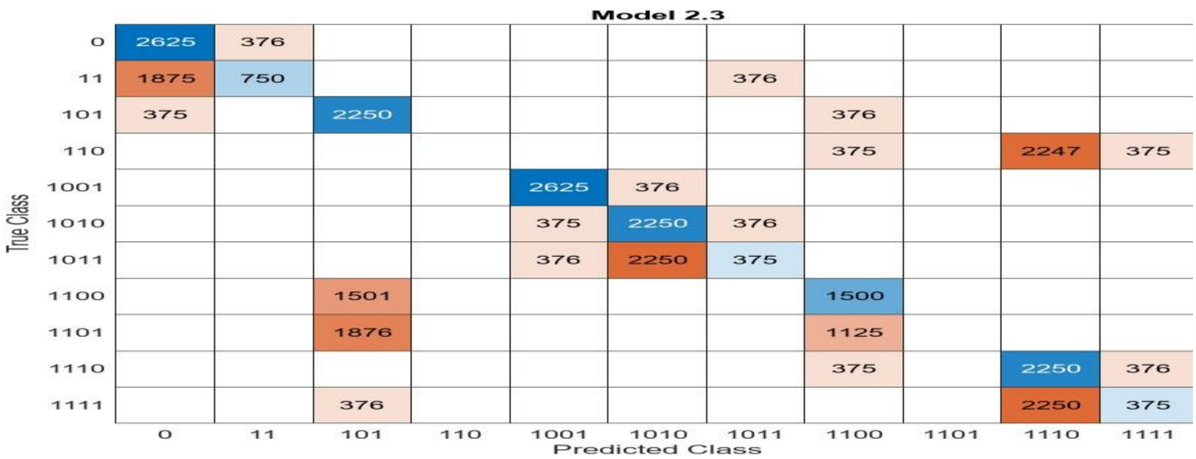


Figure 4.28 Coarse tree classifier confusion matrix results and performance accuracy of 41.6%

When the confusion metrics results were analysed, the result that stood out the most was the Low Accuracy in the model's overall accuracy, which was 41.6%. This is a pretty low percentage, which indicates that the model properly classifies fewer than half of the cases.

In the confusion metrics, the occurrence of high False Positive and False Negative values implies that the model frequently misclassifies instances. This is indicated by the high error rates, which are displayed in orange inside the confusion metrics. For instance, if the false negative value is high, it indicates that the model frequently fails to recognize different instances.

Table 4.7 Table of results for the Coarse tree classifier from confusion metrics performance.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	2625	376	87.5	12.5
11	750	2251	25.0	75.0
101	2250	751	75.0	25.0
110	-	2997	-	100
1001	2625	376	87.5	12.5
1010	2250	751	75	25
1011	375	2626	12.5	87.5
1100	1500	1500	50.0	50.0
1101	-	3001	-	100
1110	751	2250	75.0	25.0
1111	375	2626	12.5	87.5

4.11.1.1 Performance summary by Decision tree classifiers

The confusion matrix for the “Decision tree classifier” is presented which shows an average performance as whole in terms of accuracy in observation of the confusion matrix performance across all the trained models. Although the values on the main diagonal are slightly higher, Figure 4.30 shows the confusion matrix of the Fine Tree classifier and Table 4.5 showing the analysis of the True Positive rate ratio to False Negative for model 2.1 classifier. This is multiclass class analysis with the fault codes in the X-axis and Y-axis as the predicted class and true class respectively. The main diagonal data, which is depicted in light blue colour in the cell, shows the correctly classified data with respective to the axis and the light orange cells shows the misclassified data of the observation and the numbers showing the number of misclassified data with respective to that axis. The blank cell can be interpreted as the non-misclassified data.

4.11.2 Logistic regression classifiers

Logistic regression is a widely used classification algorithm due to its simplicity in interpretation. The classifiers represent the probability of different classes by using a mathematical function that combines the predictors in a linear way. Therefore, because of their straightforwardness, efficiency, and comprehensibility, they are extensively employed in the identification and categorization of faults in micro-grid protection, primarily for tasks involving binary classification. For multi-class classification, extensions of logistic regression like One-vs-Rest (OvR) or Softmax regression can be used [123]

Following the completion of data pre-processing and collection, the raw data was processed to extract pertinent features, such as statistical metrics, harmonic components, or specific indicators of fault states, as described in chapter 3's feature extraction approach. The logistic regression model was then trained using labelled data. The training and validation results are displayed in the figure below of the confusion metrics.

4.11.2.1 Efficient logistic regression

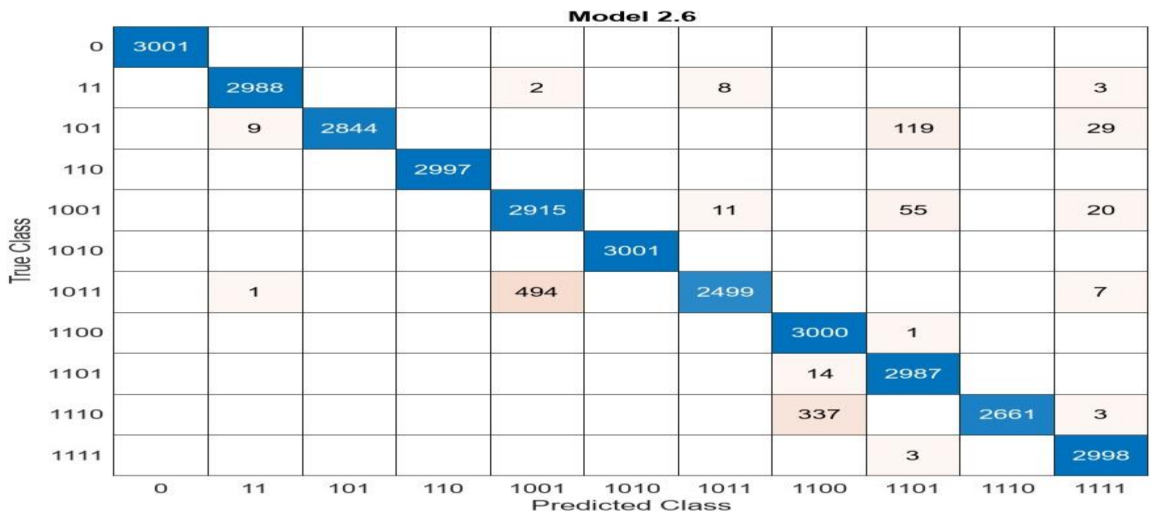


Figure 4.29 Efficient logistic regression classifier confusion metrics

When it comes to the application of micro-grids, the effectiveness of fault detection and classification is of the utmost importance in order to guarantee a continuous and dependable supply of electricity. The utilization of machine learning, in particular logistic regression classifiers, has been demonstrated to be an effective method for accomplishing this objective.

A strong performance has been proved by the implementation of an effective logistic regression classifier, which has achieved an outstanding 96.1% accuracy in defect detection.

Table 4.8 Shows overall summary of Efficient logistic classifies confusion metrics performance.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	2988	13	99.6	0.4
101	2844	157	94.8	5.2
110	2997	0	100	0
1001	2915	86	97.1	2.9
1010	3001	0	100	5.6
1011	2499	502	83.3	16.7
1100	3000	1	100	0
1101	2987	14	99.5	0.5
1110	2661	340	88.7	11.3
1111	2998	3	99.9	0.1

The logistic regression classifier displayed exceptional effectiveness in fault detection within the micro-grid, which was defined by several characteristics as shown in Table 4.8 shows the correctly classification and miss-classification data-set , from data it is observed the in all faults classifications there no fault type that has zero misclassification which is an indication by key metrics of True Positives (TP): Correctly identified fault instances, and False Negatives (FN): which are correctly identified non-fault instances.

4.11.2.2 Efficient linear SVM

Logistic regression is one of the numerous machine learning models and algorithms that are utilized for linear classification. There are many more models and techniques. A linear support vector machine (SVM) is a type of support vector machine (SVM) in which the decision boundary is a linear hyperplane. In particular, it is well-suited for situations in which classes may be split linearly, which means that a straight line (or hyperplane in higher dimensions) can divide the data points of distinct classes.

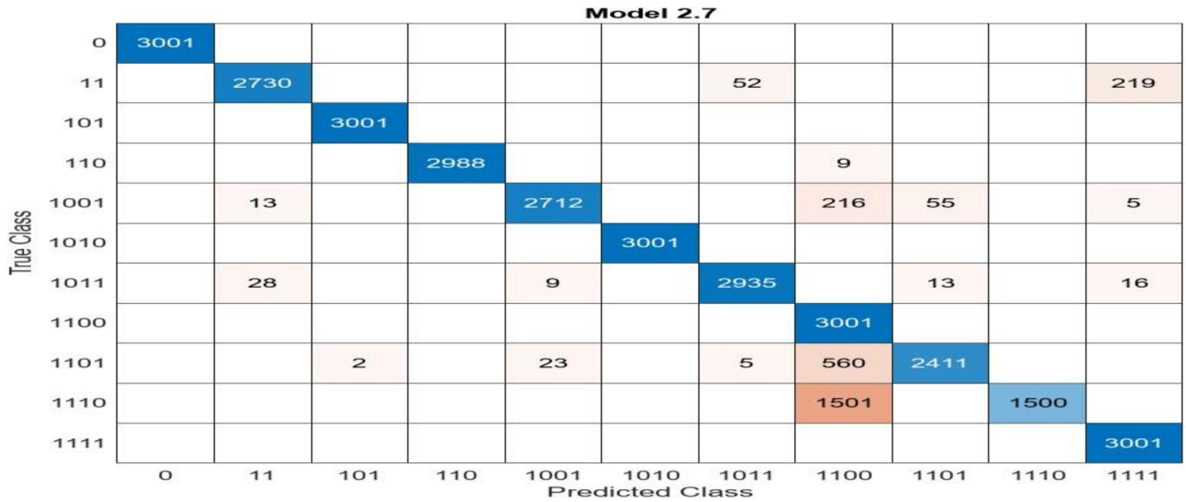


Figure 4.30 Efficient linear SVM confusion metrics performance results

The linear Support Vector Machine (SVM) is highly suitable for real-time fault detection applications in modern micro-grid systems due to its simplicity and computational efficiency. Its usage considerably enhances the dependability and maintenance of these systems.

A highly efficient Linear Support Vector Machine (SVM) has demonstrated its effectiveness in this field, obtaining an impressive accuracy of 90.4% in defect identification. Figure 4.34 presents a comprehensive analysis of the model's performance, specifically in relation to true positives, true negatives, false positives, and false negatives. The model demonstrated a 90.4% accuracy, suggesting a significant proportion of accurately detected occurrences of both faults and non-faults. When compared to Linear Regression, which belongs to the same classifier model as regression classifier, the Efficient linear Support Vector Machine (SVM) has shown worse performance.

Table 4.9 The breakdown of confusion metrics performance and TPR and TNR.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	2730	271	91	9
101	3001	0	100	0
110	2988	9	99.7	0.3
1001	2712	289	90.4	9.6
1010	3001	0	100	0
1011	2935	66	97.8	2.2
1100	3001	0	99.7	0.3
1101	2411	590	94.7	5.3
1110	1500	1501	85.3	14.7
1111	3001	0	91.7	8.3

4.11.2.3 Performance summary by logistic regression classifiers

The multiclass functional logistic regression model utilizes pre-processed and labelled data. The dataset is subsequently divided into training and validation subsets, maintaining consistent proportions. The confusion matrices were constructed using various training and validation subsets derived from the same datasets, focusing on distinct classes that were selected. Figures 4.33 and 4.34 display the matrix outcomes for different subsets derived from the dataset using various fault codes. Following the completion of the fitting process, the model achieves an accuracy of around 96.1% for efficient logistic regression and 90.4% for Efficient Linear SVM when evaluated on the validation subset. Tables 4.8 and 4.9 are utilized to present a concise overview of the primary data analysis pertaining to the true positive rate and false positive rate of the model's performance. This analysis is conducted by considering both the accurately classified data and the misclassified data. The performance of the Logistic Regression classifiers can be deemed excellent, as seen by its high accuracy validation performance of over 90%.

4.11.3 Naive Bayes Classifiers

Naive Bayes classifiers encompass a set of classification methods that derive their foundation from Bayes' Theorem. The classification process is not governed by a singular algorithm, but rather by a collection of algorithms that adhere to a shared principle. This principle dictates that the classification of any given pair of features is entirely independent from the classification of any other pair of features.

The results of this research examines a dataset consisting of extracted features related to electrical voltage and currents in a micro-grid. The dataset aims to characterize various fault conditions in the micro-grid, with the ultimate goal of classifying these faults. Based on the presence of fault conditions, each tuple is assigned to a specific class that categorizes the circumstances in a multiclass fashion, encompassing a total of 12 distinct fault conditions.

4.11.3.1 Kernel Naives Bayes

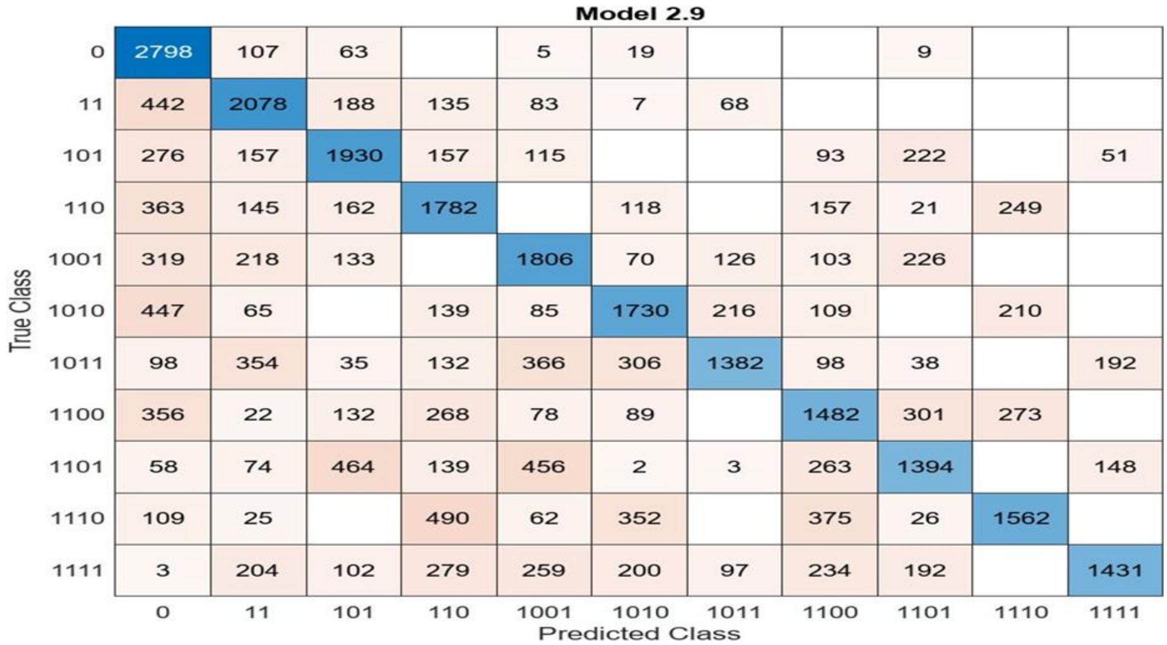


Figure 4.31 Kernel Naives Bayes shows confusion metrics graphical results, with an overall 58.7% performance in it accuracy analysis.

Kernel Naive Bayes is a modified version of the Naive Bayes classifier that use kernel density estimation to estimate the data distribution. The performance of a Kernel Naive Bayes model was evaluated using a confusion matrix. The analysis reveals the utilization of a confusion matrix, which is a commonly employed for evaluating the effectiveness of a classification model. Each row in Figure 4.35 represents the instances in a predicted class, while each column represents the occurrences in an actual class (or vice versa). The main diagonal of the matrix represents the cases that were classified correctly, while the off-diagonal elements reflect the instances that were misclassified. In the case of multiclass faults, the fault code numbers are used to label the misclassifications. With an overall accuracy of 58.7%, it can be concluded that a significant portion of the instances were correctly classified by your Kernel Naive Bayes model. To gain insight into the model's errors, it is important to closely analyse the elements in the confusion matrix that are not on the main diagonal. Significant values in the off-diagonal elements suggest that there may be some confusion between different classes.
Possible enhancements could include gathering additional data, refining the features, or experimenting with alternative modelling approaches.

Table 4.10 The breakdown of confusion metrics performance of Kernal Naïve Bayes classifier with TPR and TNR analysis.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	2798	203	93.2	6.8
11	2078	923	69.2	30.8
101	1930	1071	64.3	35.7
110	1782	1215	59.5	40.5
1001	1806	1195	60.2	39.8
1010	1730	1271	57.7	42.3
1011	1382	1619	46.1	53.9
1100	1482	1519	49.4	50.6
1101	1394	1607	47.2	52.8
1110	1562	1439	52.0	48
1111	1431	1570	47.7	52.3

4.11.3.2 Performance summary of Naives Bayes classifiers

The confusion matrix for the "Naïve Bayes classifier" is presented which shows an excellent performance as whole in terms of accuracy in observation of the confusion matrix across all the trained models. Bayes' Theorem is a mathematical formula used to determine the chance of an event happening, based on the probability of another event that has already taken place, Figure 4.35 shows the confusion matrix of the Kernel Naives Bayes classifier and Table 4.10 showing the analysis of the True Positive rate ratio to False Negative for model 2.9 classifier. This is multiclass class analysis with the fault codes in the X-axis and Y-axis as the predicted class and true class respectively. The main diagonal data, which is depicted in light blue colour in the cell, shows the correctly classified data with respective to the axis and the light orange cells shows the misclassified data of the observation and the numbers showing the number of misclassified data with respective to that axis. The blank cell can be interpreted as the non-misclassified data.

4.11.4 The support vector machine (SVM)

The support vector machine (SVM) has enhanced generalization capabilities in the classification of unseen data. The aforementioned characteristics of Support Vector Machines (SVM) render it an optimal classifier for effectively addressing non-linear problems, such as the one examined in this particular work. Support Vector Machine (SVM) is a type of machine learning technique that is used to classify data in supervised learning tasks. SVM classifiers have been utilized in fault detection to categorize the system as either normal or faulty, as well as to determine the specific type of fault. This classification is based on a range of input features obtained from the voltage and current data of the Micro-grid. Therefore, it is referred to as multiclass fault classification. The process of implementing SVM classifiers for fault detection in micro-grids with solar and BESS using MATLAB requires numerous essential processes, including data collecting, pre-processing, and real-time deployment. The performance of the SVM model is greatly influenced by the quality and representativeness of the training data. Therefore, it is critical to select the appropriate features in order to achieve accurate fault detection. Consequently, it becomes essential to extract relevant features from the collected data in order to obtain the best possible results.

4.11.4.1 Linear SVM

An effective classifier for defect detection in a micro-grid can be a Linear SVM, provided that the data is well pre-processed and the features are precisely engineered. In order to guarantee strong performance, it is essential to assess the model using relevant metrics and rectify any class imbalances.

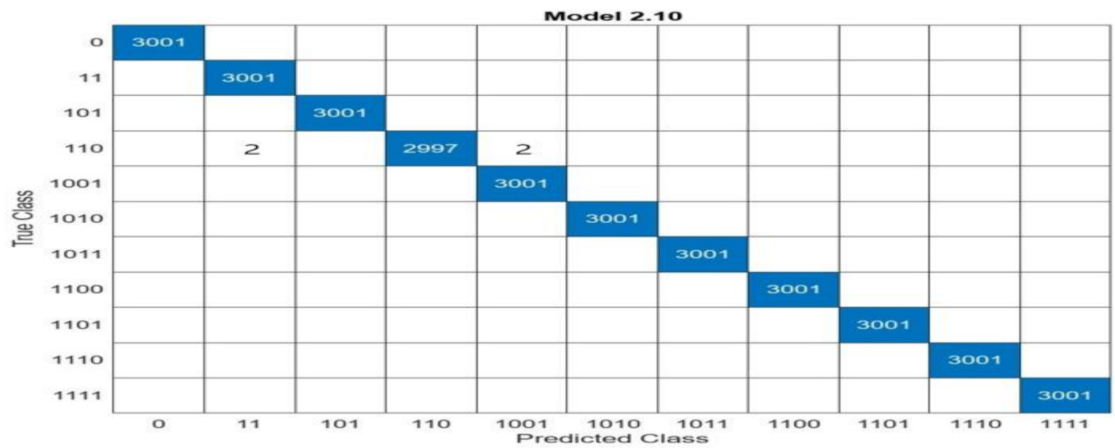


Figure 4.32 Linear SVM confusion metric performance output

The Linear SVM metrics indicate multiclass fault classification accuracy. Using confusion metrics, the classifier has 99.9% accuracy. With Figure 4.41 shows the Main Diagonal in blue of fault true class vs anticipated class. All fault types perform well in accuracy, with only fault class 110 misclassified of 2997 data instead of 3001 as shown.

Table 4.11 Shows the Linear SVM confusion metrics summary for TPR and FNR results.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

Table 4.11 summarizes confusion metrics performance by number of datasets, class datasets, categorized and miss-classified data, and True Positive Vs False Negative proportion.

4.11.4.2 Quadratic SVM

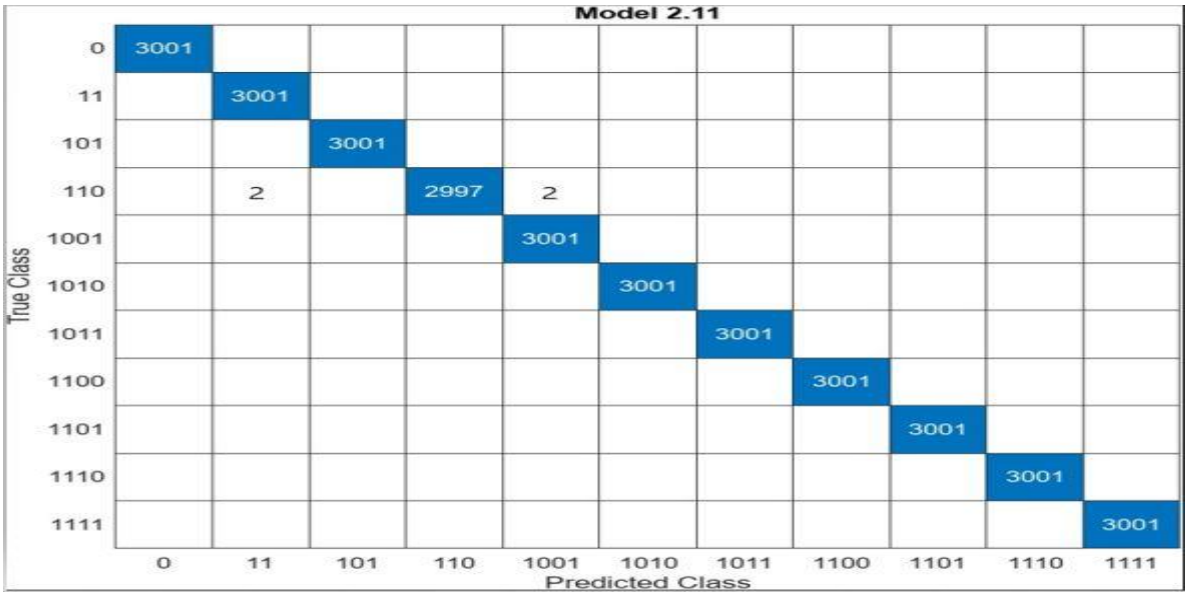


Figure 4.33 shows the confusion metrics of the Quadratic SVM for all fault class.

The Quadratic SVM metrics indicate multiclass fault classification accuracy. Using confusion metrics, the classifier has 99.9% accuracy. With Figure 4.41 shows the Main Diagonal in blue of fault true class vs anticipated class. All fault types perform well in accuracy, with only fault class 110 misclassified OF 2997 data instead of 3001 as shown.

Table 4.12 , Shows the different fault class performance summary from confusion metrics result.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

Table 4.16 summarizes confusion metrics performance by number of datasets, class datasets, categorized and miss-classified data, and True Positive Vs False Negative proportion. Fault class 110 shows misclassification of 4 dataset making the True Positive rate being 99.9% and 0.1 % for False Negative Rates.

4.11.4.2 Cubic SVM

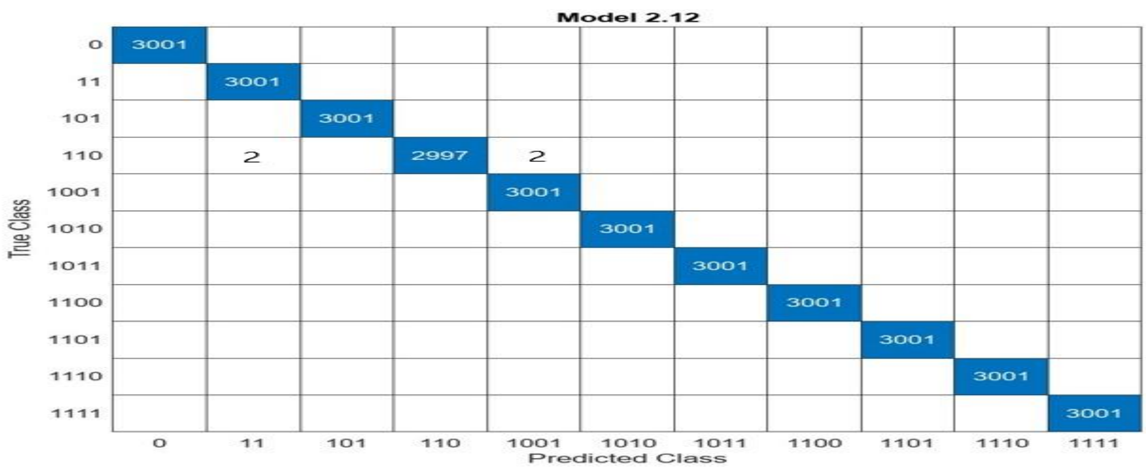


Figure 4.34 Cubic SVM confusion metrics results for model performance analysis.

The Cubic Support Vector Machine metrics display the accuracy of the multiclass fault classification using the results. The confusion metrics can be used to conduct an analysis of the accuracy, and the classifier that was used demonstrated an accuracy of 99.9%. In Figure 4.41, the major diagnostic is depicted in blue, and it is possible to conduct an analysis of the fault true class in comparison to the predicted class. It is observed that all fault types exhibit an exceptional performance in terms of accuracy, with the exception of fault class 110.

Table 4.13 Cubic SVM results summary from confusion metrics model performance.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

Using the amount of datasets and breaking down each class dataset for classified and miss-classified data, Table 4.16 provides an overall review of the performance of the confusion metrics. Additionally, the percentage of True Positives vs False Negatives is also summarized in this table.

4.11.4.3 Fine Gaussian SVM

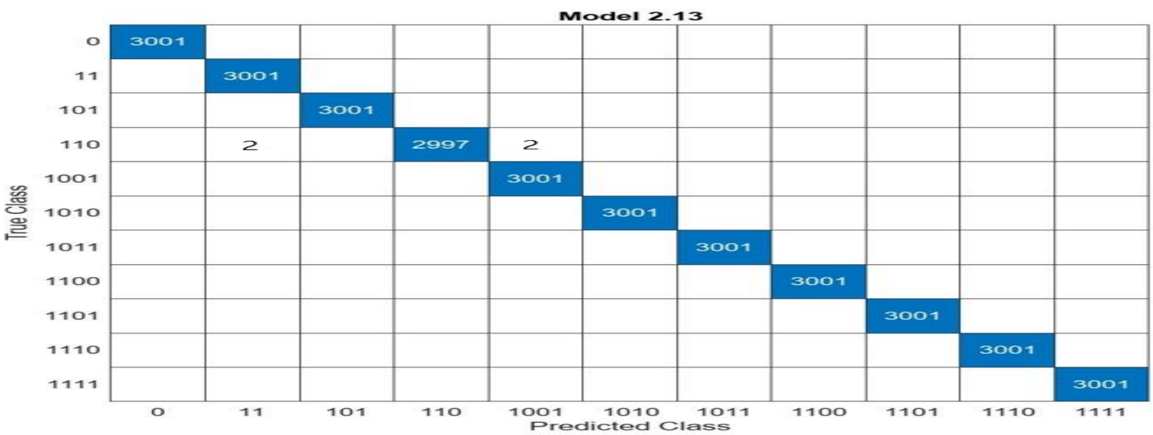


Figure 4.35 Fine Gaussian SVM confusion metrics results.

A Fine Gaussian Support Vector Machine (SVM) classifier that achieves 99% accuracy in testing for fault detection in a micro-grid is a sign of strong model performance within the micro-grid. This can be seen clearly in Figure 4.39, which displays the results of the confusion matrix Main Diagonal, which are shown in blue for all multiclass fault code classifications. It is also possible to carry out an examination of the fault True Class in contrast to the Predicted Class. With the exception of fault class 110, which exhibits a slight case of misclassification of 4 misclassification, it has been noted that other fault classes demonstrate an extraordinary performance in terms of accuracy by 100% in each class except 110.

Table 4.14 Shows the different fault class performance summary from Fine Gaussian SVM confusion metrics result.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

The following Table 4.14 displays the various fault class performance summaries derived from the results of the Fine Gaussian SVM confusion metrics examination. With regard to the true positive rate (TPR), we find that there are four instances of misclassification for a total of 110 fault type classes on the true class.

4.11.4.4 Fine Gaussian SVM

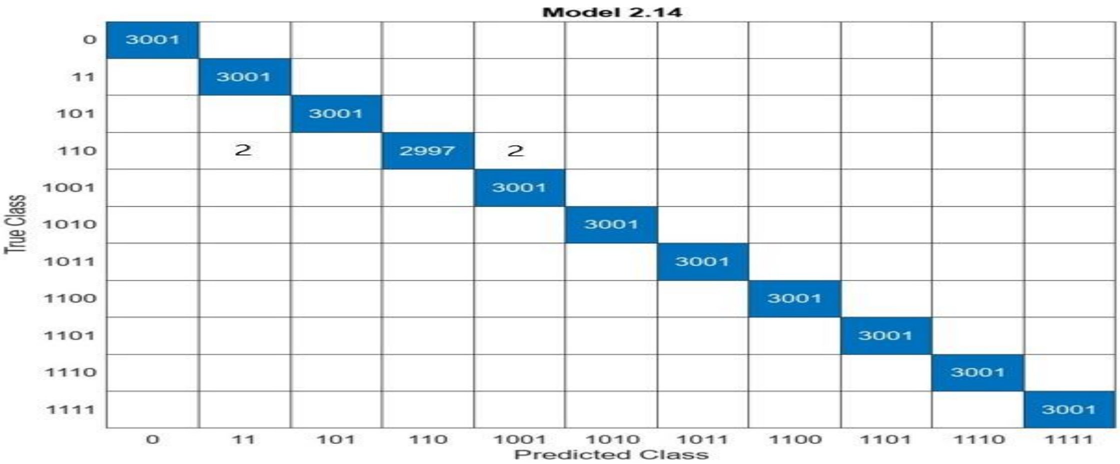


Figure 4.36 Medium Gaussian (SVM) confusion metrics for the training, validation and testing phases.

A Medium Gaussian Support Vector Machine (SVM) classifier that achieves 99.9% accuracy in testing for fault detection in a micro-grid is a sign of strong model performance within the micro-grid. This can be seen clearly in Figure 4.39, which displays the results of the confusion matrix main diagonal, which are shown in blue for all multiclass fault code classifications. It is also possible to carry out an examination of the fault True Class in contrast to the Predicted Class. With the exception of fault class 110, which exhibits a slight case of misclassification of 4 misclassification

Table 4.15 Shows the different fault class performance summary from confusion metrics result from Medium Gaussian model.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

4.11.4.5 Coarse Gaussian SVM

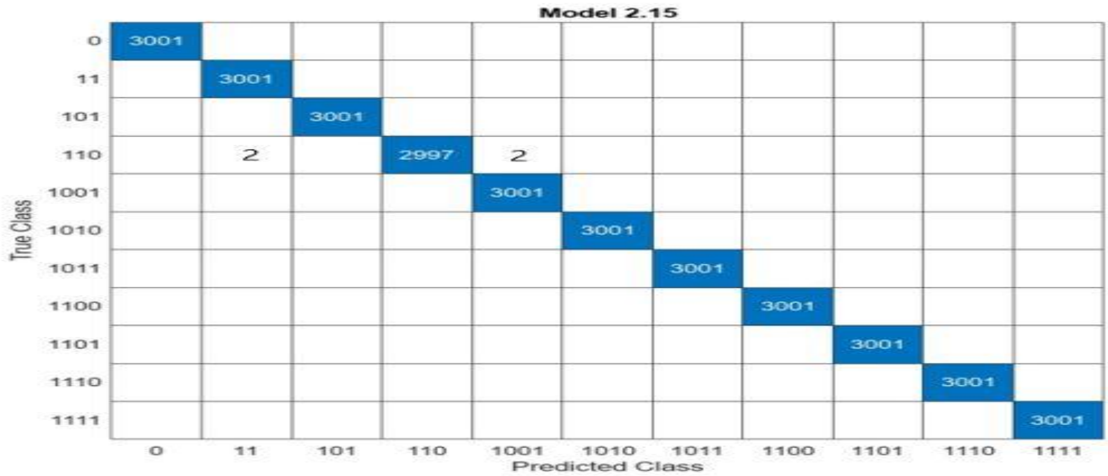


Figure 4.37 Coarse Gaussian SVM metrics for the training, validation and testing phases.

The Coarse Gaussian SVM metrics shows the result of accuracy for multiclass fault classification. The Accuracy can be analysed using the confusion metrics with the classifier showing accuracy of 99.9%. From Figure 4.41 the main diagonal is shown in blue colour can be analysed of the fault true class vs predicted class and all fault types show an outstanding performance in accuracy with only fault class 110 showing slight misclassification.

Table 4.16 is the overall summary of the confusion metrics performance by using the number of dataset and breaking down each class datasets for classified and miss-classified data with the True Positive Vs False Negative percentage also summarised.

Table 4.16 Shows the different fault class performance summary from confusion metrics result from Coarse Gaussian model for TPR and FNR calculation.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

4.11.4.1 Performance summary of Support Vector Machine classifier

Support Vector Machines (SVMs) are characterized by their ability to create nonlinear division hyperplanes, resulting in a high level of discrimination. Furthermore, Support Vector Machines (SVM) exhibit superior generalization capabilities when it comes to classifying unseen data. The aforementioned characteristics of Support Vector Machines (SVM) are employed to effectively address non-linear problems. In addition to SVM performance, it is important to consider accuracy, true positive rate, false positive rate, precision, and false negative rate as performance indicators. The aforementioned factors are utilized in the construction of the confusion matrix for the SVM as shown in Figure 4.36 which is the classifier named Linear SVM, Figure 4.37 Quadratic SVM, Figure 4.38 is the Cubic SVM, Figure 4.39 being the Fine Gaussian SVM and Figure 4.40 Meduim Gaussian . The results indicate that the SVM classifier, when combined with classifier subset evaluation feature selection, achieves the highest level of accuracy compared to not employing classifier subset evaluation. This finding holds true across all model classifier groups, including Decision Tree, Logistic Regression Classifier, Support Vector Machine, Ensemble Classifiers, and Neural Network Classifier classifiers. The aforementioned comparison has been conducted on the data sets Ia, Ib, Ic, Va, Vb, and Vc. Classification accuracy is calculated by considering the proportion of correct predictions considering the positive and negative inputs which is highly dependent on the data set .The results from true positive rate and confusion metrics indicate that the SVM classifier performance in accuracy which enables us to find each fault class percentage of performance in true positive rating or true negative rating.

4.11.5 Ensemble classifier

The present investigation presents an ensemble classification framework that utilizes confusion matrices and analyses the true positive rate and true negative rate. A classification ensemble refers to a predictive model that is constructed by combining numerous classification models with varying weights, resulting in an enhanced prediction performance. The Feature extraction and classification learner tools provides support for a range of ensemble learning methods, such as bagging, random space, and several boosting algorithms.

4.11.5.1 Boosted tree classifier

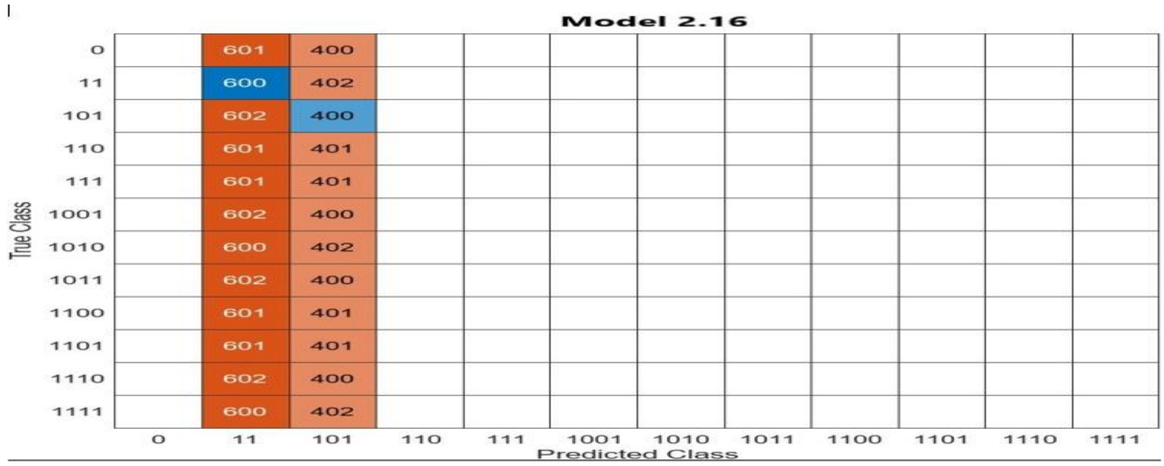


Figure 4.38 , Boosted tree classifier confusion metrics classification observation results.

A Boosted trees are an ensemble method that combines numerous weak learners, usually decision trees, to create a powerful predictive model. The classifier model has a poor accuracy rate of 9.1% during testing for fault detection in a micro-grid, which serves as an indication of subpar performance. The evidence can be seen in Figure 4.42 of the confusion matrix. The primary diagonal of the matrix, represented in blue, shows the results for the multiclass fault code classification. Specifically, it pertains to fault class 11 and 101. Additionally, it is feasible to perform an examination of the accurate fault category in relation to the anticipated category. All fault types are shown to have exceptionally low accuracy, with a misclassification rate of 0%.

Table 4.17 Shows the different fault class performance summary from confusion metrics result from Boosted tree classifier model for TPR and FNR calculation.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	-	1001	93.6	6.4
11	600	402	88.5	11.5
101	400	602	87.4	12.6
110	-	219	-	-
111	-	149	-	-
1001	-	1002	-	-
1010	-	1002	-	-
1011	-	1002	-	-
1100	-	1002	-	-
1101	-	1002	-	-
1110	-	1002	-	-
1111	-	1002	-	-

Table 4.15 displays the performance overview of distinct fault classes based on the confusion metrics result of the Boosted tree. We have observed four instances of misclassification in the true positive rate (TPR) for all fault types in the true class.

4.5.11.2 RUS Boosted Tree

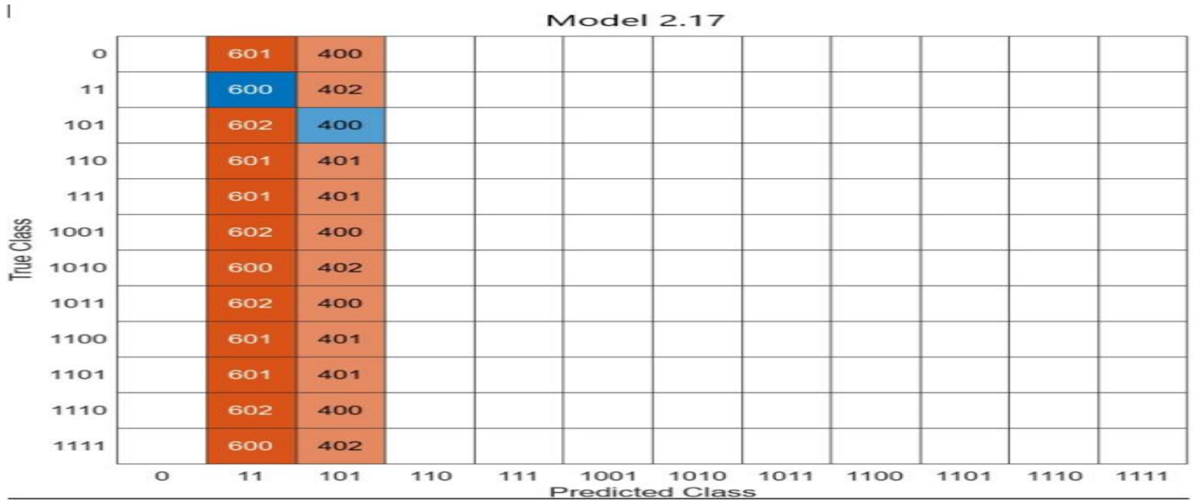


Figure 4.39, RUS-Boosted Trees confusion metrics results.

The confusion measures are analysed according to the representation in Figure 4.43.The confusion matrix of the classification model reveals that the RUS-Boosted Trees classifier achieves an overall accuracy of 8.3% in detecting faults within a micro-grid system. This low accuracy suggests that the model is performing significantly below expectations. There are several possible reasons for this poor performance, which can be investigated to identify the underlying issue. The confusion matrix analysis tool is used to generate and analyse a classification model in order to identify which classes are being misclassified.

Table 4.18 is provided to offer concise information on whether the model exhibits bias towards specific classes or if it struggles to detect certain problems. The available training data may be inadequate to accurately represent the underlying patterns. However, this performance specifically focused on summarizing data from both classified and misclassified datasets. Only the faults with code 11 and 101 were successfully classified, while the remaining faults did not yield any results. Obtaining the True Positive Rate (TPR) and False Negative Rate (FNR) proved to be difficult, as indicated in the table.

Table 4.18 Shows the different fault class performance summary from confusion metrics result from RUS-Boosted tree classifier model for TPR and FNR calculation.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	-	1001	93.6	6.4
11	600	402	88.5	11.5
101	400	602	87.4	12.6
110	-	219	-	-
111	-	149	-	-
1001	-	1002	-	-
1010	-	1002	-	-
1011	-	1002	-	-
1100	-	1002	-	-
1101	-	1002	-	-
1110	-	1002	-	-
1111	-	1002	-	-

4.5.11.3 Bagged Tree

Figure 4.44 depicts the Bagged Tree classifier, in which the confusion matrix results indicate that the fault detection system in the micro-grid has achieved a testing accuracy of 99%, demonstrating a highly successful performance. Bagging, also known as bootstrap aggregating, is a technique used to enhance the stability and accuracy of machine learning systems.

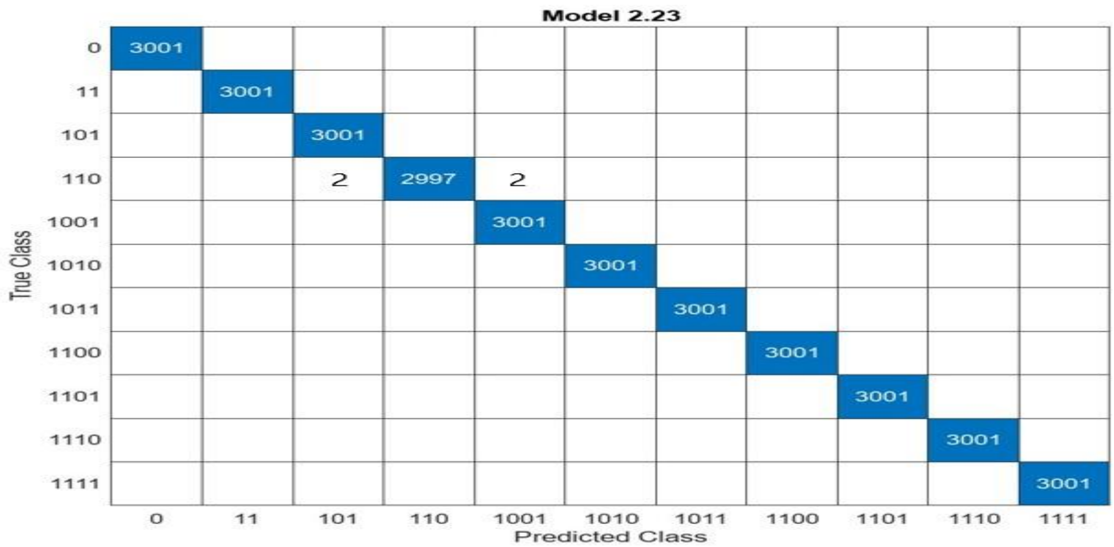


Figure 4.40, Bagged tree confusion metrics training validation accuracy performance.

Attaining a 99% accuracy rate with a Bagged Tree classifier indicates numerous favourable aspects. The findings can be attributed to the ensemble technique, which mitigates overfitting and enhances generalization to unseen data, together with the incorporation of feature relevance.

The remarkable performance of the system can be attributed to the high quality of the training data, which exhibits unambiguous patterns that differentiate between different fault circumstances.

Table 4.19 summarizes the performance by using the percentage overall of the True Positive rating Vs False Negative rating of the correctly classified and misclassified dataset as shown. Fault code 110 is the only fault type which doesn't have all the data correctly classified by Confusion matrix main diagonal and also by using the table data.

Table 4.19 Shows the different fault class performance summary from confusion metrics result.

Fault code	Correctly classified	Miss-classification	TPR%	FNR%
0	3001	0	100	0
11	3001	0	100	0
101	3001	0	100	0
110	2997	4	99.9	0.1
111	3001	0	100	0
1001	3001	0	100	0
1010	3001	0	100	0
1011	3001	0	100	0
1100	3001	0	100	0
1101	3001	0	100	0
1110	3001	0	100	0
1111	3001	0	100	0

4.11.5.1 Performance of the Ensemble classifier

The Ensemble Classifiers are in addition to their performance referred to as a prediction model that is constructed by combining many classification models, with each model being assigned a specific weight. It is important to consider accuracy, true positive rate, false positive rate, precision, and false negative rate as performance indicators to enhance the prediction and true performance.

The data set of current and voltage signals was subjected to the use of ensemble classifiers such as Bagging Ensemble, RUS-Boosted Tree, Bagged Tree, and Ensemble. Subsequently, the performance metrics were assessed. The aforementioned factors are utilized in the construction of the confusion matrix for the Ensemble as shown in Figure 4.42 which is the classifier named Boosted Tree, Figure 4.43 RUS-Boosted Trees, and Figure 4.44 is the Bagged Tree. When an ensemble classifier exhibits inadequate performance, as evidenced by a confusion matrix, it implies that the model is not accurately forecasting the target classes. In fact the misclassification and classification data has no result depicted in the other fault codes of the confusion metrics except for the Bagged Tree classifier. Classification accuracy is calculated by considering the proportion of correct predictions considering the positive and negative inputs. It is highly dependent on the data set.

4.11.6 Neural Network Performance

Neural Network models are organized in a hierarchical manner, consisting of multiple layers that mimic the information processing mechanism of the human brain. This study introduces a novel ensemble classification paradigm that incorporates confusion matrices and examines the true positive rate and true negative rate.

4.11.6.1 Narrow Neutral Network

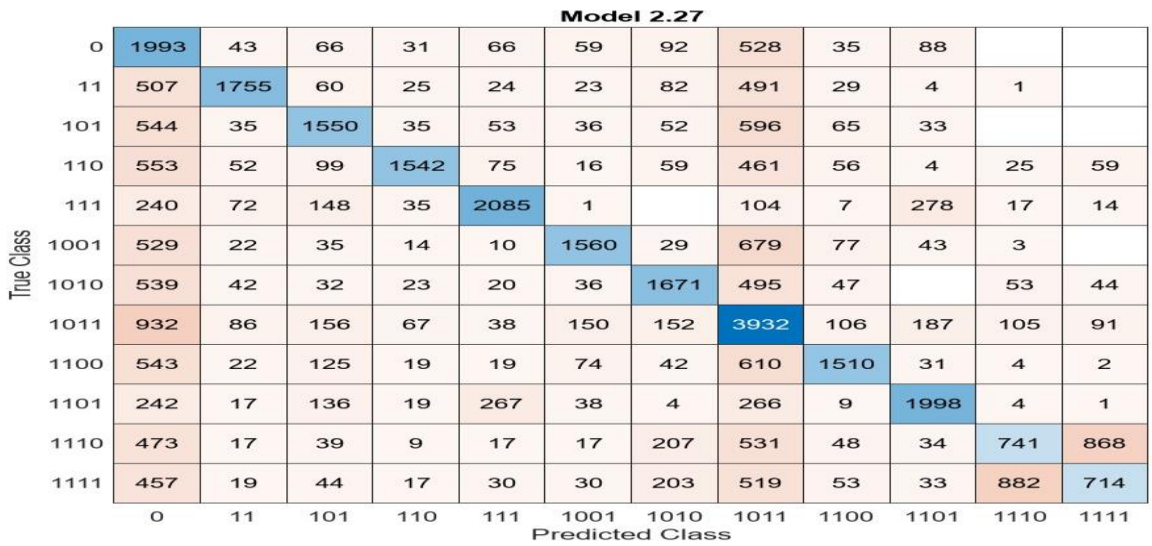


Figure 4.41, Confusion metrics of a Narrow neutral Network with overall training accuracy of 59.7%.

Table 4.20 Shows the different fault class performance summary from confusion metrics result for Narrow neutral network classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1993	1008	66.4	33.6
11	1755	1246	58.5	41.5
101	1550	1449	51.7	48.3
110	1542	1459	51.4	48.6
111	2085	916	69.5	30.5
1001	1560	1441	52	48
1010	1671	1331	55.7	44.3
1011	3932	2070	65.5	34.5
1100	1510	1491	50.3	49.7
1101	1998	1003	66.6	33.4
1110	741	2257	24.7	75.3
1111	714	2287	23.8	76.2

4.11.6.2 Narrow Neutral Network

Figure 4.42, shows the Medium Neural Network with accuracy of 61.7% in it fault classification performance.

Table 4.21 Shows the different fault class performance summary from confusion metrics result for Medium neural Network classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1896	1105	63.2	36.8
11	1978	1023	65.9	34.1
101	1966	1033	65.6	34.4
110	1986	1015	66.2	33.8
111	2418	583	80.6	19.4
1001	1986	1015	66.2	33.8
1010	1997	261	66.5	33.5
1011	5397	605	89.9	10.1
1100	1971	547	65.7	34.3
1101	2454	547	81.8	67.0
1110	891	2110	29.7	70.3
1111	906	2095	30.2	69.8

4.11.6.3 Narrow Neutral Network

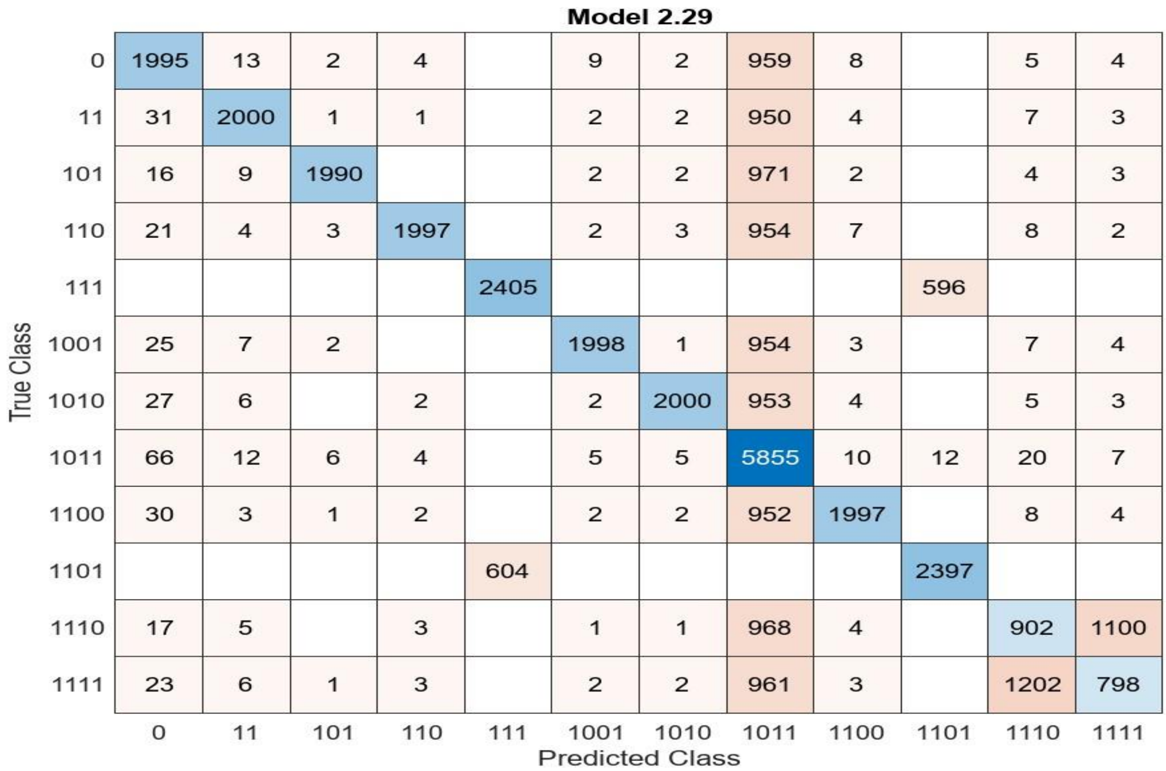


Figure 4.43, shows the classifier of Wide Neural Network with overall performance accuracy of 62.2%.

Table 4.22 Shows the different fault class performance summary from confusion metrics result for Wide Neural Network classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1995	1006	66.5	33.5
11	2000	1001	66.6	33.4
101	1990	1009	66.4	33.6
110	1997	1004	66.5	33.5
111	2405	596	80.1	19.9
1001	1998	1003	66.6	33.4
1010	2000	1002	66.6	33.4
1011	5855	147	97.6	2.4
1100	1997	1004	66.5	33.5
1101	2397	604	79.9	20.1
1110	902	2099	30.1	69.9
1111	798	2203	26.6	73.4

4.11.6.4 Narrow Neutral Network

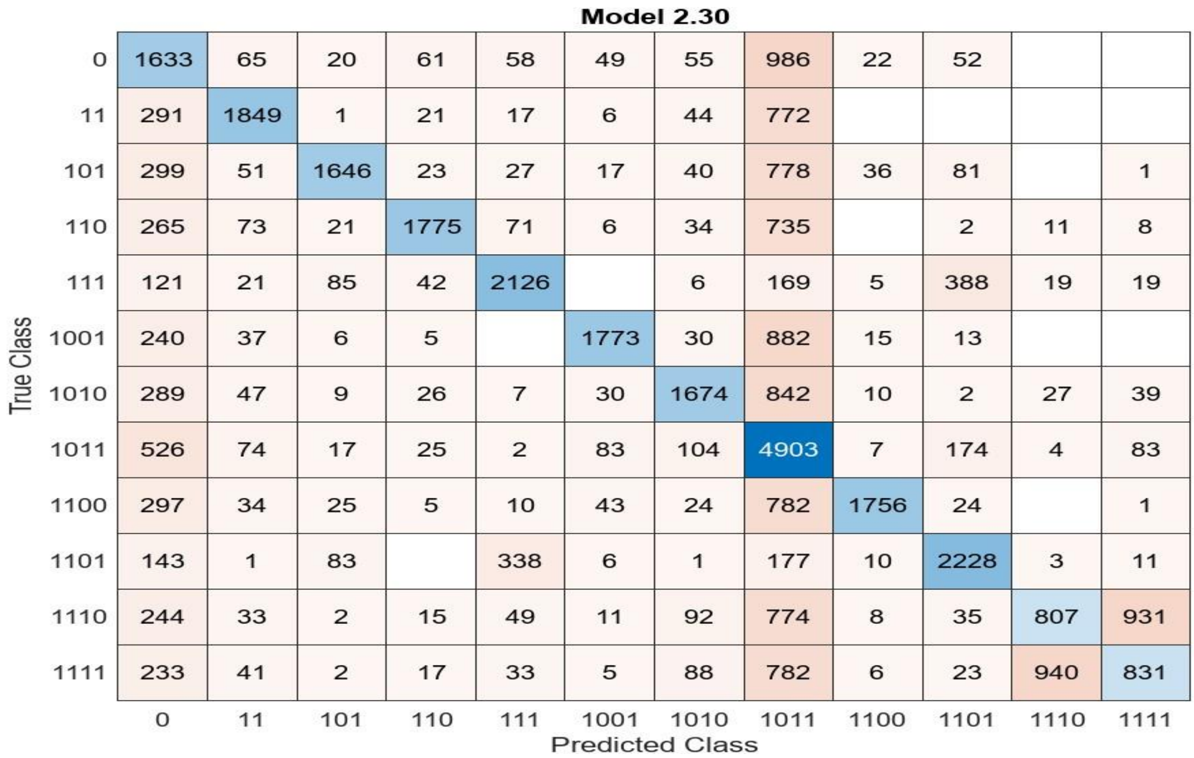


Figure 4.44, shows the confusion metrics of the Bi-layered Neural Network classifier with overall accuracy percentage of 58.7%.

Table 4.23 Shows the different fault class performance summary from confusion metrics result for Bi-layered Neural Network classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1633	1368	54.4	45.6
11	1849	1152	61.6	38.4
101	1646	1353	54.9	45.1
110	1775	1226	59.1	40.9
111	2126	875	70.8	29.2
1001	1773	1228	59.1	40.9
1010	1674	1328	55.8	44.2
1011	4903	1099	81.7	18.2
1100	1756	1245	58.5	41.5
1101	2228	773	74.2	25.8
1110	807	2194	26.9	73.1
1111	831	2170	27.7	72.3

4.11.6.5 Narrow Neutral Network

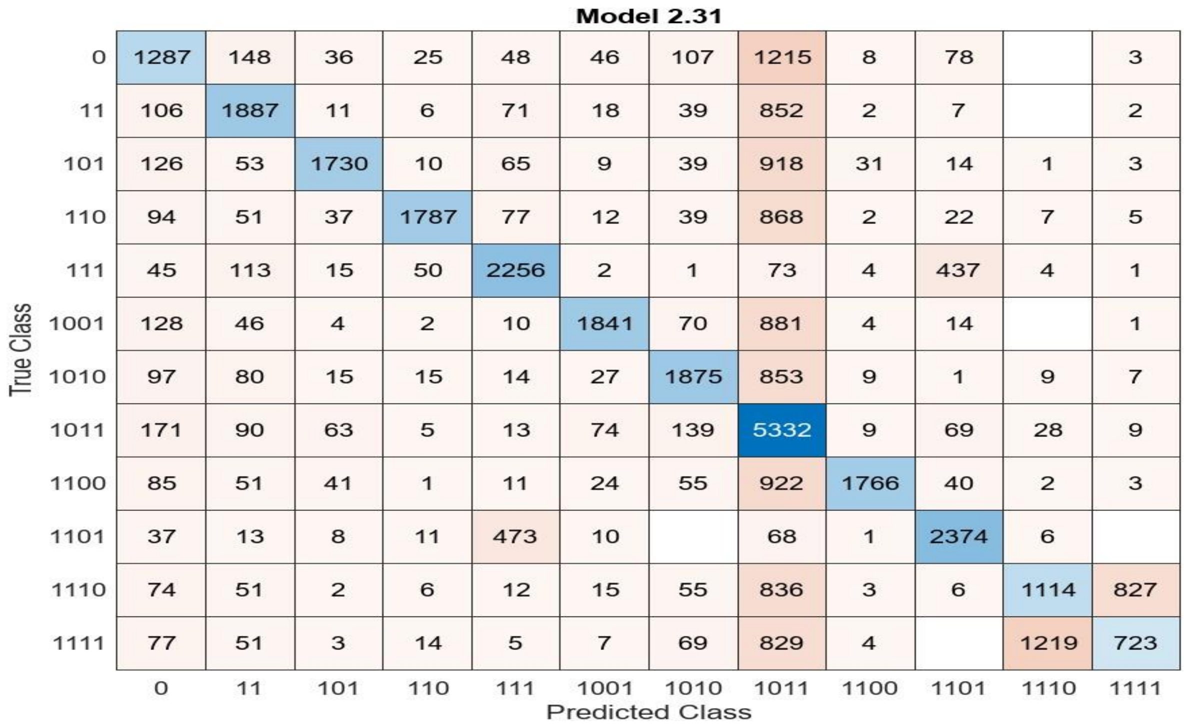


Figure 4.45, shows the Tri-layered Neural Network classifies confusion metrics overall accuracy of 60.1%.

Table 4.24 Shows the different fault class performance summary from confusion metrics result for Tri-layered Neural Network classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1287	1714	42.9	57.1
11	1887	1114	62.9	37.1
101	1730	1165	57.7	42.3
110	1787	1214	59.5	40.5
111	2256	745	75.2	24.8
1001	1841	1160	61.3	38.7
1010	1875	1127	62.5	37.5
1011	5332	670	88.8	11.2
1100	1766	1235	58.8	41.2
1101	2374	627	79.1	20.9
1110	1114	1887	37.1	62.9
1111	723	2278	24.1	75.9

4.11.7. Performance of the Neural Network classifier

The fault classification method required a neural network that allows it to determine the type of fault from the patterns of pre fault and post fault voltages and currents, which are formed from the values measured from a three phase line of an electrical DN with DN. Therefore, the numerous potential combinations can accurately depict each of the diverse flaws. The Neural Network classifier demonstrates the ability to effectively differentiate between the 11 potential fault types, as well as plus one being the absence of any defect. The evaluation of the trained and tested Neural Network is conducted using confusion matrices for six different neural network classifiers. The five implemented classifiers are the Narrow Neural Network, Medium Neural Network, Wide Neural Network, Bi-layered Neural Network, and Tri-layered Neural Network.

The outcomes obtained from the six classifiers indicate that the performance of the Neural Network classifier is generally mediocre to unsatisfactory, as evidenced by the analysis of the confusion matrices presented in figures 4.45, 4.46, 4.47, 4.48, and figure 4.49. The primary diagonal of the metrics reveals that the data in each class is lower than the misclassified data in each fault category, resulting in a low percentage in the "True Positive Rate" (TPR) from Table 4.20 to Table 4.24. This consistency is observed.

4.11.7 Approximate classifier

Kernel Approximation classifiers can be employed to carry out nonlinear classification of datasets including a large number of observations. In the context of handling huge in-memory datasets, it has been observed that Kernel classifiers generally exhibit quicker training and prediction speeds compared to SVM classifiers utilizing Gaussian Kernels.

4.11.7.1 SVM Kernal

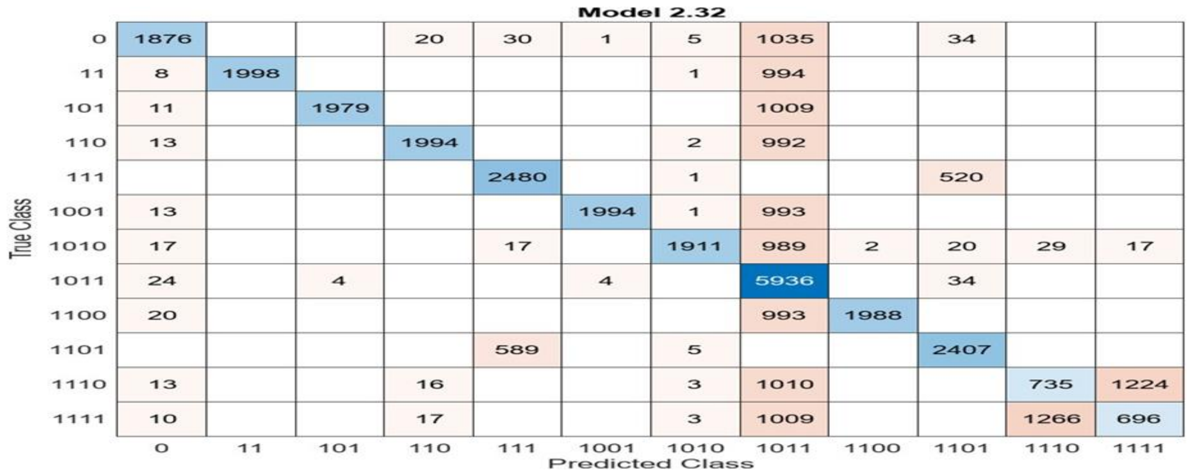


Figure 4.46, shows the results of the SVM Kernal classifies confusion metrics.

The confusion matrix in fault detection for a micro-grid offers a comprehensive analysis of the classification model's performance. Figure 4.50 displays the SVM Kernel confusion matrix for a classifier in a multi-class fault detection scenario, with an accuracy of 66.6%. The dataset consists of 3001 examples, which have been intentionally kept at a manageable size for the sake of simplicity. The distribution of instances across the different classes is reasonably balanced, as depicted in the figure. Table 4.25 below presents the obtained metrics from confusion metrics that offer a comprehensive assessment of the SVM classifier's effectiveness in detecting faults in a multi-class scenario.

Table 4.25 Shows the different fault class performance summary from confusion metrics result for SVM Kernel classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1876	1125	62.5	37.5
11	1998	1003	66.6	33.4
101	1979	1020	66	34
110	1994	1007	66.4	33.6
111	2480	521	82.6	17.4
1001	1994	1007	66.4	33.6
1010	1911	102	63.7	36.3
1011	5936	66	98.9	1.1
1100	1988	1013	66.2	33.8
1101	2407	594	80.2	19.8
1110	735	2266	24.5	75.5
1111	696	2305	23.2	76.8

The table provides a concise overview and visual representation of the classifier's performance in various classes. It emphasizes the classifier's strengths and areas that could be improved. Specifically, it presents a clear assessment of the classifier's effectiveness in detecting faults in a micro-grid, assuming a reasonable distribution of classes. The fault type 1011 shows the best perform class with 98.9% accuracy in True Positive rate and 1111 fault type the worst performing in reference to True Positive Rate

4.11.7.2 Logistic Regression Kernel

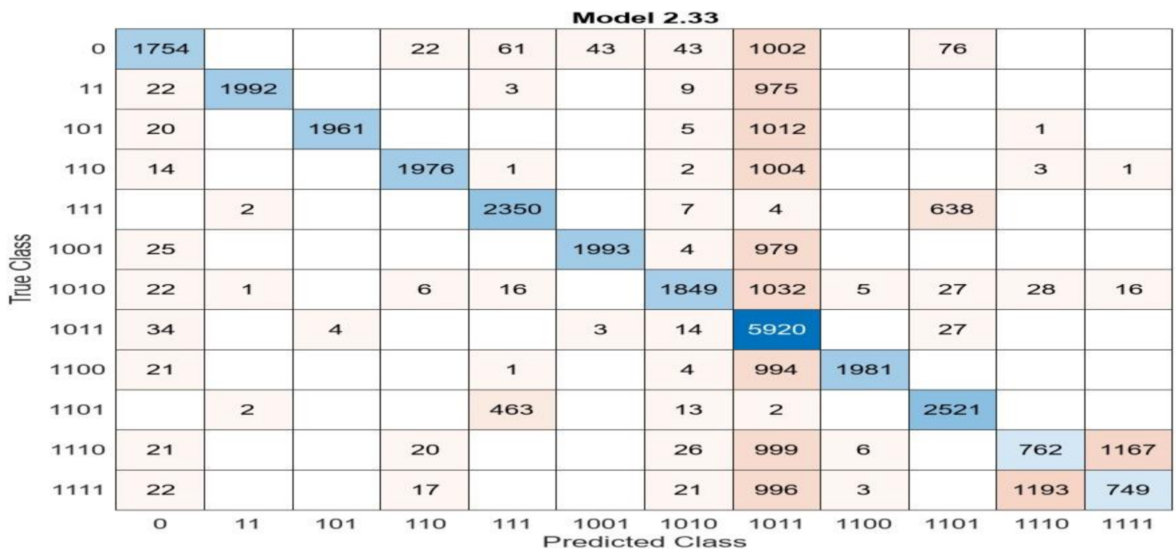


Figure 4.47, shows the results of the Logistic Regression Kernel classifier confusion metrics.

Table 4.26 Shows the different fault class performance summary from confusion metrics result for Logistic Regression Kernel classifier.

Fault code	Correctly classified	Miss-classification	TPR	FNR
0	1754	1247	58.4	41.6
11	1992	1009	66.4	33.6
101	1961	1038	65.4	34.6
110	1976	1025	65.8	34.2
111	2350	651	78.3	21.7
1001	1993	1008	66.4	33.6
1010	1849	1153	61.6	38.4
1011	5920	82	98.6	1.4
1100	1981	1020	66	34
1101	2521	480	84	16
1110	762	2239	25.4	74.6
1111	749	2252	25	75

4.11.7.3 Performance of the approximation classifier

The evaluation of the approximation classifiers, specifically the SVM Kernel and the Logistic Regression Kernel, has been conducted. The accuracy of the predictions in each class has been assessed by utilizing a validation confusion matrix.

The comparison of confusion matrices for the two Kernel models can be shown in Figure 4.50, Figure 4.51, Table 4.25, and Table 4.26. The interpretation of the results is as follows. The SVM Kernel exhibits superior accuracy in classifying the data as compared to its subset classifier, Logistic Regression. In fact, the number of miss-classified predictions is higher than the number of correctly classified predictions. The classification system ranges from the top category, labelled as "0-No-Fault," to the lowest category denoted by the fault code "1110-ABC," indicating poor performance. Through this analysis, it is evident that there are misclassifications in all categories when comparing the predicted and true classes. This further reinforces the inadequate performance of the model in fault classification. There is a lack of gain in performance observed even when utilizing the Logistic Regression Kernel, instead resulting in a fall in accuracy. Notably, the class labeled as "1111-LLLG" has the lowest accuracy. Another notable observation is that the classes between the predicted class and the true class exhibit misclassification for all fault codes, which serves as an additional indicator of subpar performance.

4.12 Multiclass fault classifiers accuracy results

Multiclass fault classifiers refer to methods or models that are specifically developed to classify examples into numerous groups or categories. These classifiers are commonly employed in fault diagnosis or pattern recognition applications. The measure of accuracy for a multiclass fault classifier pertains to its ability to accurately forecast the classes of instances within a given dataset.

Table 4.27 Shows the overall accuracy of the Machine learning classifiers performance of training and testing.

Classifiers	Model number	Testing Accuracy%
Decision Tree		
Fine Tree	2.1	59.9
Medium Tree	2.2	54.5
Coarse Tree	3.3	41.6
Logististic regression classifier		
Efficient Logistic Regression	2.4	96.1
Efficient Linear SVM	2.5	90.4
Naive Bayes		
Gaussian Naive Bayes	2.9	Failed
Kernel Naive Bayes	2.7	58.7
Support Vector Machine		
Linear SVM	2.8	99.9
Quadratic SVM	2.9	99.9
Cubic SVM	2.10	99.9
Fine Gaussian SVM	2.11	99.9
Medium Gaussian SVM	2.12	99.8
Coarse Gaussian SVM	2.13	99.9
Ensemble Classifiers		
Boosted Trees	2.14	8.3
Bagged Trees	2.15	99.9
Rusboosted Trees	2.16	8.3
Neural Network Classifier		
Narrow Neural Network	2.27	54.0
Meduim Neural Network	2.28	61.7
Wide Neural Network	2.29	62.2
Bilayered Neural Network	2.30	59.0
Trilayered Neural Network	2.31	61.4
Kernel Approximation classifier		
SVM Kernel	2.32	66.6
Logistic Regression Kernel	2.33	66.2

4.12.1 Overall classifier accuracy in training and testing performance summary

The outcomes of the classifier for the supervised machine learning techniques employed in the research article are presented in Table 4.27. The objective of this study was to evaluate the performance of 35 available classifiers by running the data signals through them. The focus was on identifying the best-performing classifier for the micro-grid dataset, which had been gathered, pre-processed, and subjected to feature extraction. Out of the total number of classifiers tested and trained, only 23 demonstrated the ability to effectively interact with the provided data signals. The observation of the Totals indicates that the SVM classifier group underwent a larger number of tests and shown superior performance, as evidenced by the greatest accuracy percentage per classifier when compared to the other examined classifiers. The SVM classifiers category demonstrated an accuracy range of 99.9% to 99.8%. These results indicate that the SVM classifier is the recommended choice for implementing the model due to its high level of accuracy.

Table 4.28 Categorization of classifiers performance.

RATING%	CATEGORY	DESCRIPTION
1-20	Unacceptable Performance	Performance does not meet the standard expected. The review/ assessment indicates that the classifier has achieved less than fully effective results against almost all of the performance criteria and indicators as specified in the Performance indicator metrics.
21-40	Performance Not Fully Effective	Performance meets some of the standards expected for the job. The assessment indicates that the classifier has achieved less than fully effective results against more than half of the performance criteria and indicators as specified in the Performance indicator metrics
41-60	Performance Fully Effective	Performance fully meets the standard expected. The assessment indicates that the classifier has achieved as a minimum effective results against all of the performance criteria and Indicators as specified in the Performance indicator metrics.
61-90	Performance Significantly Above Expectations	Performance is significantly higher than the standard expected. The assessment indicates that the classifier has achieved better than fully effective results against more than half of the performance criteria and indicators as specified in the Performance indicators metrics.
91-100	Outstanding Performance	Performance far exceeds the standard expected of a ML classifier at this level. The assessment indicates the classifier has achieved better than fully effective results against all of the performance and indicator criteria metrics

If we consider the classification performance grade presented in Table 4.28, it becomes evident that RUS-Boost and Boosted Tree exhibit the lowest diagnostic accuracy and efficiency, with an accuracy rate of 8.3% in their classification data train testing. Thus, both algorithms can be identified as the poorest performers in this regard.

The Decision Tree classifier, Naïve Bayes and Neural Network classification, which is consisting of Fine, Medium, and Coarse tree models, with, is used for classification based on the categorization presented in Table 4.28 , which ranks the performance. In conclusion, the decision tree classifier determines the appropriate classification model based on the given categorization. Based on the assessment, it can be concluded that the classifier has successfully met the minimum requirements for effectiveness across all performance criteria and indicators during both training and testing phases. This includes accuracy and efficiency in fault detection.

5. Chapter five: Summary, Conclusion and Suggestion for Future Work

This concluding chapter serves to emphasize and summarize the significance, objectives, and results of the present study. The research is examined in relation to its objectives and potential contributions to the electrical protection systems of the power industry's requirements. Additionally, this study delves into the potential avenues for further enhancement and expansion of the research, as well as proposes strategies for achieving these improvements. This encompasses potential avenues for comprehending the behaviour of ML in forthcoming research endeavours. There is an expectation that through the effective utilization of the vast amount of data obtainable from MGs and many dynamics of situations, machine learning (ML) has the potential to play a substantial role in enhancing the philosophy of protection systems for micro grids (MGs) in the power-generating domain.

For example, in the event of a failure occurring in the main generator, it is necessary to classify the problem according to the specific line to which it pertains. The understanding of machine learning behaviour within protective systems is still limited. The objective of this thesis is to construct models that are appropriate for comprehensive analysis and design. The objective of this study was to develop a multi-class fault classification model for analysing both the transient and steady-state characteristics of the micro-grid system. The ultimate objective was to provide a foundation that would enable the examination of fault classification in a more comprehensive dynamic power system model, with the aim of enhancing accuracy in future development. This study has effectively achieved its intended goal. The development of various models will enable the research of machine learning classifiers, hence facilitating the advancement of a more sophisticated model through enhanced comprehension.

5.2 Summary

The proliferation of dispersed generating units in contemporary power systems necessitates the prompt identification of faults. The increasing complexity of electricity networks poses a challenge for conventional methods in detecting symmetrical faults. This paper presents a potential detection approach that utilizes Artificial Intelligence. The cost efficiency of the technology is attributed to the fact that measurements are conducted at the point of common coupling, eliminating the need for extra communication with each grid component. The processed signals refer to the voltage and current measurements of the power network. The research paper presents an implementation of artificial intelligence (AI) techniques for the purpose of detecting and categorizing problems within a micro grid system. The Micro-grid was represented by designing a solar photovoltaic system and a battery energy system using MATLAB/Simulink as shown in Annexure1. The simulation consisted of five distinct fault types: phase to ground, phase-phase, phase-phase-to-ground, three phase faults, and three phase to ground faults occurring along the distribution line. In addition to normal operation, the AI model was trained and evaluated using voltage and current values.

All the available classifier models were regarded as supervised learning networks for the purpose of detecting and categorizing non-fault conditions.

In order to anticipate the evaluation of classification error statistical parameters, an assessment is conducted using the performance of confusion matrices and the rates of True and False Positive outcomes. Therefore, the utilization of Artificial Intelligence techniques for the aforementioned defects is effectively categorized.

5.3 Conclusion and future direction

Uncertainty is one of the most difficult challenges for academics studying ML protection solutions in a micro grid systems. Distribution generation sources integrated into the grid and occupancy behaviour are inherently variable, making effective task and resource scheduling more complex.

There are numerous prospective avenues for future research in the realm of building-integrated micro-grid protection systems. The forthcoming research directions within this domain are anticipated to concentrate on the resolution of challenges, the enhancement of accuracy, and the augmentation of adaptability for machine learning algorithms in the context of micro-grid protection. The next phase of our project involves the enhancement of machine learning models to achieve higher accuracy and faster fault detection and classification in micro-grids. Furthermore, the incorporation of novel sensing paradigms is imperative in the realm of data acquisition and scientific endeavours, particularly in the field of machine learning. It is of utmost importance to delve deeper into the exploration of amalgamating diverse data sources, including synchronized phasor measurements data, smart meter data, and weather data, with the primary objective of enhancing the dependability of fault detection mechanisms. With the increasing complexity of micro-grid models, there arises a heightened demand for optimization strategies that are both flexible and reliable. Consequently, Exploration Machine Learning (ML) algorithms capable of adapting to the dynamic nature of micro-grids are essentially required, Investigate methods for improving classifier learners' adaptability to the dynamic nature of micro-grids. This entails creating algorithms that can adapt to changes in micro-grid architecture, load circumstances, and generation patterns automatically. This is because the topology and operating conditions of micro-grids may undergo frequent changes, necessitating the need for adaptive algorithms to accommodate these dynamic variations. Finally, through the ongoing integration of investigating the utilization of ensemble learning methodologies, such as random forests or gradient boosting, we aim to augment the overall precision and dependability of micro-grid protection classifiers. Ensemble methods have the capacity to combine multiple classification learners to create a stronger and more robust model.

In closing we have to keep in mind that the application of machine learning in the protection of micro-grids is an area that spans multiple disciplines, and that in order to be successful, it is necessary to have collaboration between experts in power systems, Machine Learning, and

Control systems. In addition, keeping up with the most recent developments in machine learning as well as power systems is another important step that will contribute to the creation of creative solutions for the protection of micro-grid.

6. Reference

[1] A. E. Tio, I. B. N. C. Cruz, B. M. Malquisto, and R. D. Mundo, *A binary programming model for reliability optimization considering fuse-blow and fuse-save schemes*. 2012, pp. 1-6.

[2] M. Jamil, R. Singh, and S. K. Sharma, "Fault identification in electrical power distribution system using combined discrete wavelet transform and fuzzy logic," *Journal of Electrical Systems and Information Technology,* vol. 2, no. 2, pp. 257-267, 2015/09/01/ 2015, doi: https://doi.org/10.1016/j.jesit.2015.03.015.

[3] *The South African Grid Code*

The Network Code, N. E. R. of and S. A. (NERSA), 2019.

[4] Z. T. Hassan, "Development of differential protection scheme using positive sequence current for a microgrid," Universiti Tun Hussein Onn Malaysia, 2022.

[5] S. S. Gururajapathy, H. Mokhlis, and H. A. Illias, "Fault location and detection techniques in power distribution systems with distributed generation: A review," *Renewable and Sustainable Energy Reviews,* vol. 74, pp. 949-958, 2017/07/01/ 2017, doi: https://doi.org/10.1016/j.rser.2017.03.021.

[6] S. Simani and S. Farsoni, "Chapter 1 - Introduction," in *Fault Diagnosis and Sustainable Control of Wind Turbines*, S. Simani and S. Farsoni Eds.: Butterworth-Heinemann, 2018, pp. 1-12.

[7] M. A. Hanif, F. Nadeem, R. Tariq, and U. Rashid, "Chapter 6 - Hydropower energy generating systems," in *Renewable and Alternative Energy Resources*, M. A. Hanif, F. Nadeem, R. Tariq, and U. Rashid Eds.: Academic Press, 2022, pp. 325-357.

[8] Z. Yu, "Induction Generator in Wind Power Systems," in *Induction Motors*, R. Raul Igmar Gregor Ed. Rijeka: IntechOpen, 2015, p. Ch. 1.

[9] https://www.thesouthafrican.com/carbon-tax-sa-budget-speech-2018/ (accessed.

[10] F. Katiraei and J. R. Agüero, "Solar PV Integration Challenges," *IEEE Power and Energy Magazine,* vol. 9, no. 3, pp. 62-71, 2011, doi: 10.1109/MPE.2011.940579.

[11] J. J. Grainger, *Power system analysis*. McGraw-Hill, 1999.

[12] J. Driesen, P. Vermeyen, and R. Belmans, "Protection issues in microgrids with multiple distributed generation units," in *2007 Power Conversion Conference-Nagoya*, 2007: IEEE, pp. 646-653.

[13] M. R. Aghaebrahimi and A. Hosseinabadi, "THE EFFECTS OF DIFFERENT DG GENERATORS ON DISTRIBUTION NETWORK'S PROTECTION AND EXTERNAL FAULT DETECTION."

[14] W. Chen and A. M. Bazzi, "Logic-Based Methods for Intelligent Fault Diagnosis and Recovery in Power Electronics," *IEEE Transactions on Power Electronics,* vol. 32, no. 7, pp. 5573-5589, 2017, doi: 10.1109/TPEL.2016.2606435.

[15] S. Ntalampiras, "Fault Diagnosis for Smart Grids in Pragmatic Conditions," *IEEE Transactions on Smart Grid,* vol. 9, no. 3, pp. 1964-1971, 2018, doi: 10.1109/TSG.2016.2604120.

[16] M. Kezunovic and B. Perunicic, "Automated transmission line fault analysis using synchronized sampling at two ends," *IEEE Transactions on Power Systems,* vol. 11, no. 1, pp. 441-447, 1996, doi: 10.1109/59.486131.

[17] J. Gao, X. Wang, X. Wang, A. Yang, H. Yuan, and X. Wei, "A High-Impedance Fault Detection Method for Distribution Systems Based on Empirical Wavelet Transform and Differential Faulty Energy," *IEEE Transactions on Smart Grid,* vol. 13, no. 2, pp. 900-912, 2022, doi: 10.1109/TSG.2021.3129315.

[18] B. D. Russell and C. L. Benner, "Intelligent Systems for Improved Reliability and Failure Diagnosis in Distribution Systems," *IEEE Transactions on Smart Grid,* vol. 1, no. 1, pp. 48-56, 2010, doi: 10.1109/TSG.2010.2044898.

[19] V. C. Gungor, B. Lu, and G. P. Hancke, "Opportunities and Challenges of Wireless Sensor Networks in Smart Grid," *IEEE Transactions on Industrial Electronics,* vol. 57, no. 10, pp. 3557-3564, 2010, doi: 10.1109/TIE.2009.2039455.

[20] V. K. Sood, D. Fischer, J. M. Eklund, and T. Brown, "Developing a communication infrastructure for the Smart Grid," in *2009 IEEE Electrical Power & Energy Conference (EPEC)*, 22-23 Oct. 2009 2009, pp. 1-7, doi: 10.1109/EPEC.2009.5420809.

[21] A. Kurniawan, A. Afandi, S. Sendari, A. Wibawa, and I. Fadlika, "Application of Fuzzy Logic to Electrical Protection Devices," in *Journal of Physics: Conference Series*, 2020, vol. 1501, no. 1: IOP Publishing, p. 012013.

[22] S. Samonto, S. Kar, S. Pal, and A. A. Sk, "Fuzzy Based Intelligent Fault Protection Schemes: A Critical Review to Explore the Future," *SSRN Electronic Journal,* 01/01 2020, doi: 10.2139/ssrn.3811213.

[23] S. Heo and J. H. Lee, "Fault detection and classification using artificial neural networks," *IFAC-PapersOnLine,* vol. 51, no. 18, pp. 470-475, 2018.

[24] L. C. Acácio, P. A. Guaracy, T. O. Diniz, D. R. R. P. Araujo, and L. R. Araujo, "Evaluation of the impact of different neural network structure and data input on fault detection," in *2017 IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America)*, 20-22 Sept. 2017 2017, pp. 1-5, doi: 10.1109/ISGT-LA.2017.8126699.

[25] V. Jakkula, "Tutorial on support vector machine (svm)," *School of EECS, Washington State University,* vol. 37, no. 2.5, p. 3, 2006.

[26] "<A NEW POWER FACTOR COMPENSATION STRATEGY FOR.pdf>."

[27] H. M. M. Maruf, F. Müller, M. S. Hassan, and B. Chowdhury, "Locating Faults in Distribution Systems in the Presence of Distributed Generation using Machine Learning Techniques," in *2018 9th IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG)*, 25-28 June 2018 2018, pp. 1-6, doi: 10.1109/PEDG.2018.8447728.

[28] M. A. Awadallah and M. M. Morcos, "Application of AI tools in fault diagnosis of electrical machines and drives-an overview," *IEEE Transactions on Energy Conversion,* vol. 18, no. 2, pp. 245-251, 2003, doi: 10.1109/TEC.2003.811739.

[29] M. Sanaye-Pasand and H. Khorashadi-Zadeh, "Transmission line fault detection & phase selection using ANN," in *International Conference on Power Systems Transients*, 2003, pp. 1-6.

[30] K. M. Silva, B. A. Souza, and N. S. D. Brito, "Fault detection and classification in transmission lines based on wavelet transform and ANN," *IEEE Transactions on Power Delivery,* vol. 21, no. 4, pp. 2058-2063, 2006, doi: 10.1109/TPWRD.2006.876659.

[31] R. Mahanty and P. D. Gupta, "Application of RBF neural network to fault classification and location in transmission lines," *IEE Proceedings-Generation, Transmission and Distribution,* vol. 151, no. 2, pp. 201-212, 2004.

[32] Y. Zhu, H. Limin, and J. Lu, "Bayesian networks-based approach for power systems fault diagnosis," *IEEE Transactions on Power Delivery,* vol. 21, no. 2, pp. 634-639, 2006, doi: 10.1109/TPWRD.2005.858774.

[33] X. Liu, H. Li, and L. Li, "Building method of diagnostic model of Bayesian networks based on fault tree," in *Seventh International Symposium on Instrumentation and Control Technology: Sensors and Instruments, Computer Simulation, and Artificial Intelligence*, 2008, vol. 7127: SPIE, pp. 529-534.

[34] L. E. G. Castañón, J. P. Nieto González, M. A. G. Castañon, and R. Morales-Menéndez, "Fault diagnosis of industrial systems with Bayesian networks and neural networks," in *MICAI 2008: Advances in Artificial Intelligence: 7th Mexican International Conference on Artificial Intelligence, Atizapán de Zaragoza, Mexico, October 27-31, 2008 Proceedings 7*, 2008: Springer, pp. 998-1008.

[35] A. S. Bretas and N. Hadjsaid, "Fault diagnosis in deregulated distribution systems using an artificial neural network," in *2001 IEEE Power Engineering Society Winter Meeting. Conference Proceedings (Cat. No.01CH37194)*, 28 Jan.-1 Feb. 2001 2001, vol. 2, pp. 821-823 vol.2, doi: 10.1109/PESW.2001.916970.

[36] Syafaruddin, E. Karatepe, and T. Hiyama, "Controlling of artificial neural network for fault diagnosis of photovoltaic array," in *2011 16th International Conference on Intelligent System Applications to Power Systems*, 25-28 Sept. 2011 2011, pp. 1-6, doi: 10.1109/ISAP.2011.6082219.

[37] E. B. M. Tayeb, "Faults detection in power systems using artificial neural network," *American Journal of Engineering Research,* vol. 2, no. 6, pp. 69-75, 2013.

[38] A. C. Adewole, R. Tzoneva, and S. Behardien, "Distribution network fault section identification and fault location using wavelet entropy and neural networks," *Applied Soft Computing,* vol. 46, pp. 296-306, 2016/09/01/ 2016, doi: https://doi.org/10.1016/j.asoc.2016.05.013.

[39] Q. Yang, J. Li, S. Le Blond, and C. Wang, "Artificial Neural Network Based Fault Detection and Fault Location in the DC Microgrid," *Energy Procedia,* vol. 103, pp. 129-134, 2016/12/01/ 2016, doi: https://doi.org/10.1016/j.egypro.2016.11.261.

[40] "Harmonic Sources from Industrial Loads I." www.idc-online.com/technical.../Harmonic_Sources_from_Industrial_Loads_I.pdf (accessed.

[41] M. R. Zaidan, "Power System Fault Detection, Classification And Clearance By Artificial Neural Network Controller," in *2019 Global Conference for Advancement in Technology (GCAT)*, 18-20 Oct. 2019 2019, pp. 1-5, doi: 10.1109/GCAT47503.2019.8978400.

[42] S. R. Fahim, S. K. Sarker, S. M. Muyeen, S. K. Das, and I. Kamwa, "A deep learning based intelligent approach in detection and classification of transmission line faults," *International Journal of Electrical Power & Energy Systems,* vol. 133, p. 107102, 2021/12/01/ 2021, doi: https://doi.org/10.1016/j.ijepes.2021.107102.

[43] A. R. Aqamohammadi, T. Niknam, S. Shojaeiyan, P. Siano, and M. Dehghani, "Deep Neural Network with Hilbert–Huang Transform for Smart Fault Detection in Microgrid," *Electronics,* vol. 12, no. 3, p. 499, 2023.

[44] P. Singh, N. Singh, and N. K. Choudhary, "Fault Detection and Classification in Microgrid Using Wavelet Transform and Artificial Neural Network," in *Advances in VLSI, Communication, and Signal Processing*, Singapore, D. Harvey, H. Kar, S. Verma, and V. Bhadauria, Eds., 2021// 2021: Springer Singapore, pp. 17-31.

[45] D. Thukaram, H. P. Khincha, and H. P. Vijaynarasimha, "Artificial neural network and support vector Machine approach for locating faults in radial distribution systems," *IEEE Transactions on Power Delivery,* vol. 20, no. 2, pp. 710-721, 2005, doi: 10.1109/TPWRD.2005.844307.

[46] B. Ravikumar, D. Thukaram, and H. Khincha, "Application of support vector machines for fault diagnosis in power transmission system," *IET generation, transmission & distribution,* vol. 2, no. 1, pp. 119-130, 2008.

[47] A. Rahmani, J. Haddadnia, and O. Seryasat, "Intelligent fault detection of electrical equipment in ground substations using thermo vision technique," in *2010 2nd International Conference on Mechanical and Electronics Engineering*, 1-3 Aug. 2010 2010, vol. 2, pp. V2-150-V2-154, doi: 10.1109/ICMEE.2010.5558469.

[48] D. Thukaram, H. P. Khincha, and B. Ravikumar, "An Intelligent Approach Using Support Vector Machines for Monitoring and Identification of Faults on Transmission Systems," in *2006 IEEE Power India Conference*, 10-12 April 2006 2006, pp. 184-190, doi: 10.1109/POWERI.2006.1632511.

[49] D. Alahakoon and X. Yu, "Smart electricity meter data intelligence for future energy systems: A survey," *IEEE transactions on industrial informatics,* vol. 12, no. 1, pp. 425-436, 2015.

[50] R. Agrawal and D. Thukaram, "Identification of fault location in power distribution system with distributed generation using support vector machines," in *2013 IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, 24-27 Feb. 2013 2013, pp. 1-6, doi: 10.1109/ISGT.2013.6497853.

[51] N. Shahid, S. A. Aleem, I. H. Naqvi, and N. Zaffar, "Support Vector Machine based fault detection & classification in smart grids," in *2012 IEEE Globecom Workshops*, 3-7 Dec. 2012 2012, pp. 1526-1531, doi: 10.1109/GLOCOMW.2012.6477812.

[52] Y. Ping and L. Sui Sheng, "Fault diagnosis for boilers in thermal power plant by data mining," in *ICARCV 2004 8th Control, Automation, Robotics and Vision Conference, 2004.*, 6-9 Dec. 2004 2004, vol. 3, pp. 2176-2180 Vol. 3, doi: 10.1109/ICARCV.2004.1469502.

[53] M. Singh, B. K. Panigrahi, and R. P. Maheshwari, "Transmission line fault detection and classification," in *2011 International Conference on Emerging Trends in Electrical and Computer Technology*, 23-24 March 2011 2011, pp. 15-22, doi: 10.1109/ICETECT.2011.5760084.

[54] H. Livani and C. Y. Evrenosoğlu, "A fault classification method in power systems using DWT and SVM classifier," in *PES T&D 2012*, 7-10 May 2012 2012, pp. 1-5, doi: 10.1109/TDC.2012.6281686.

[55] M. R. Alam, K. M. Muttaqi, and A. Bouzerdoum, "An Approach for Assessing the Effectiveness of Multiple-Feature-Based SVM Method for Islanding Detection of Distributed Generation," *IEEE Transactions on Industry Applications,* vol. 50, no. 4, pp. 2844-2852, 2014, doi: 10.1109/TIA.2014.2300135.

[56] R. Pérez and C. Vásquez, "Fault Location in distribution systems with distributed generation using Support Vector Machines and smart meters," in *2016 IEEE Ecuador Technical Chapters Meeting (ETCM)*, 12-14 Oct. 2016 2016, pp. 1-6, doi: 10.1109/ETCM.2016.7750857.

[57] M. Sarlak and S. Shahrtash, "SVM-based method for high-impedance faults detection in distribution networks," *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering,* vol. 30, no. 2, pp. 431-450, 2011.

[58] B. M. Aucoin and B. D. Russell, "Distribution high impedance fault detection utilizing high frequency current components," *IEEE Transactions on Power Apparatus and Systems,* no. 6, pp. 1596-1606, 1982.

[59] S. Samantaray, B. Panigrahi, and P. Dash, "High impedance fault detection in power distribution networks using time–frequency transform and probabilistic neural network," *IET generation, transmission & distribution,* vol. 2, no. 2, pp. 261-270, 2008.

[60] H. R. Baghaee, D. Mlakić, S. Nikolovski, and T. Dragicević, "Support Vector Machine-Based Islanding and Grid Fault Detection in Active Distribution Networks," *IEEE Journal of Emerging and Selected Topics in Power Electronics,* vol. 8, no. 3, pp. 2385-2403, 2020, doi: 10.1109/JESTPE.2019.2916621.

[61] F. Mohammadi, G. A. Nazri, and M. Saif, "A Fast Fault Detection and Identification Approach in Power Distribution Systems," in *2019 International Conference on Power Generation Systems and Renewable Energy Technologies (PGSRET)*, 26-27 Aug. 2019 2019, pp. 1-4, doi: 10.1109/PGSRET.2019.8882676.

[62] M.-Y. Cho and T. T. Hoang, "Feature Selection and Parameters Optimization of SVM Using Particle Swarm Optimization for Fault Classification in Power Distribution Systems," *Computational Intelligence and Neuroscience,* vol. 2017, p. 4135465, 2017/07/11 2017, doi: 10.1155/2017/4135465.

[63] A. Ahmadi, E. Aghajari, and M. Zangeneh, "High-impedance fault detection in power distribution grid systems based on support vector machine approach," *Electrical Engineering,* vol. 104, no. 5, pp. 3659-3672, 2022/10/01 2022, doi: 10.1007/s00202-022-01544-1.

[64] C. G. Wester, "High impedance fault detection on distribution systems," in *1998 Rural Electric Power Conference Presented at 42nd Annual Conference*, 26-28 April 1998 1998, pp. c5-1, doi: 10.1109/REPCON.1998.666955.

[65] T. Edwards, "Discrete wavelet transforms: Theory and implementation," *Universidad de,* pp. 28-35, 1991.

[66] K. L. Butler-Purry and M. Bagriyanik, "Characterization of transients in transformers using discrete wavelet transforms," *IEEE Transactions on Power Systems,* vol. 18, no. 2, pp. 648-656, 2003, doi: 10.1109/TPWRS.2003.810979.

[67] R. H. Salim, K. R. C. d. Oliveira, A. D. Filomena, M. Resener, and A. S. Bretas, "Hybrid Fault Diagnosis Scheme Implementation for Power Distribution Systems Automation," *IEEE Transactions on Power Delivery,* vol. 23, no. 4, pp. 1846-1856, 2008, doi: 10.1109/TPWRD.2008.917919.

[68] J. Joe-Air, F. Ping-Lin, C. Ching-Shan, Y. Chi-Shan, and S. Jin-Yi, "A fault detection and faulted-phase selection approach for transmission lines with Haar wavelet transform," in *2003 IEEE PES Transmission and Distribution Conference and Exposition (IEEE Cat. No.03CH37495)*, 7-12 Sept. 2003 2003, vol. 1, pp. 285-289 Vol.1, doi: 10.1109/TDC.2003.1335236.

[69] K. Silva, K. Dantas, B. Souza, N. Brito, F. Costa, and J. Silva, "Haar wavelet-based method for fast fault classification in transmission lines," in *2006 IEEE/PES Transmission & Distribution Conference and Exposition: Latin America*, 2006: IEEE, pp. 1-5.

[70] R. M. de Castro Fernandez and H. N. D. Rojas, "An overview of wavelet transforms application in power systems," *Proc. 14th PSCC, Sevilla, Session,* vol. 1, pp. 1-8, 2002.

[71] N. Perera and A. D. Rajapakse, "Agent-based protection scheme for distribution networks with distributed generators," in *2006 IEEE Power Engineering Society General Meeting*, 18-22 June 2006 2006, p. 6 pp., doi: 10.1109/PES.2006.1709528.

[72] X. Zeng, K. K. Li, W. L. Chan, and S. Su, "Multi-agents based protection for distributed generation systems," in *2004 IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies. Proceedings*, 5-8 April 2004 2004, vol. 1, pp. 393-397 Vol.1, doi: 10.1109/DRPT.2004.1338528.

[73] A. Ngaopitakkul, C. Apisit, C. Pothisarn, C. Jettanasen, and S. Jaikhan, "Identification of fault locations in underground distribution system using discrete wavelet transform," in *Proceedings of the International MultiConference of Engineers and Computer Scientists*, 2010, vol. 2.

[74] R. H. Salim, K. R. C. De Oliveira, and A. S. Bretas, "Fault detection in primary distribution systems using wavelets," in *International Conference on Power Systems Transients (IPST)*, 2007.

[75] J. G. Moreno, F. E. Pérez, and E. A. Orduña, "Protection functions for distribution networks with distributed generation applying wavelet transform," in *2012 Sixth IEEE/PES Transmission and Distribution: Latin America Conference and Exposition (T&D-LA)*, 3-5 Sept. 2012 2012, pp. 1-5, doi: 10.1109/TDC-LA.2012.6319073.

[76] M. Shaik, M. N. Rao, and S. A. Gafoor, "A wavelet based protection scheme for distribution networks with Distributed Generation," in *2012 International Conference on Emerging Trends in Electrical Engineering and Energy Management (ICETEEEM)*, 13-15 Dec. 2012 2012, pp. 33-37, doi: 10.1109/ICETEEEM.2012.6494440.

[77] K. Saravanababu, P. Balakrishnan, and K. Sathiyasekar, "Transmission line faults detection, classification, and location using Discrete Wavelet Transform," in *2013 International Conference on Power, Energy and Control (ICPEC)*, 6-8 Feb. 2013 2013, pp. 233-238, doi: 10.1109/ICPEC.2013.6527657.

[78] A. C. Adewole and R. Tzoneva, "Fault detection and classification in a distribution network integrated with distributed generators," in *IEEE Power and Energy Society Conference and Exposition in Africa: Intelligent Grid Integration of Renewable Energy Resources (PowerAfrica)*, 9-13 July 2012 2012, pp. 1-8, doi: 10.1109/PowerAfrica.2012.6498611.

[79] K. De Kerf *et al.*, "Wavelet-based protection strategy for DC faults in multi-terminal VSC HVDC systems," *IET Generation, Transmission & Distribution,* vol. 5, no. 4, pp. 496-503, 2011.

[80] B. Moeil, M. Gandomkar, M. Gooran, and S. Mirsaeidi, "Distinction of permanent and transient faults in microgrids using wavelet transform," *Journal of Applied Environmental and Biological Sciences (JAEBS),* vol. 3, no. 10, pp. 41-51, 2013.

[81] P. K. Ray, S. R. Mohanty, and N. Kishor, "Disturbance detection in grid-connected distributed generation system using wavelet and S-transform," *Electric Power Systems Research,* vol. 81, no. 3, pp. 805-819, 2011/03/01/ 2011, doi: https://doi.org/10.1016/j.epsr.2010.11.011.

[82] F. B. Costa, A. Monti, and S. C. Paiva, "Overcurrent protection in distribution systems with distributed generation based on the real-time boundary wavelet transform," *IEEE Transactions on Power delivery,* vol. 32, no. 1, pp. 462-473, 2015.

[83] X. G. Magagula, Y. Hamam, J. A. Jordaan, and A. A. Yusuff, "Fault detection and classification method using DWT and SVM in a power distribution network," in *2017 IEEE PES PowerAfrica*, 27-30 June 2017 2017, pp. 1-6, doi: 10.1109/PowerAfrica.2017.7991190.

[84] Y. Mohammadnian, T. Amraee, and A. Soroudi, "Fault detection in distribution networks in presence of distributed generations using a data mining–driven wavelet transform," *IET Smart Grid,* vol. 2, no. 2, pp. 163-171, 2019.

[85] P. Ray, D. P. Mishra, K. Dey, and P. Mishra, "Fault Detection and Classification of a Transmission Line Using Discrete Wavelet Transform & Artificial Neural Network," in *2017 International Conference on Information Technology (ICIT)*, 21-23 Dec. 2017 2017, pp. 178-183, doi: 10.1109/ICIT.2017.24.

[86] M. Shafiullah, K. A. AlShumayri, and M. S. Alam, "Machine learning tools for active distribution grid fault diagnosis," *Advances in Engineering Software,* vol. 173, p. 103279, 2022/11/01/ 2022, doi: https://doi.org/10.1016/j.advengsoft.2022.103279.

[87] Y. D. Mamuya, Y.-D. Lee, J.-W. Shen, M. Shafiullah, and C.-C. Kuo, "Application of Machine Learning for Fault Classification and Location in a Radial Distribution Grid," *Applied Sciences,* vol. 10, no. 14, p. 4965, 2020. [Online]. Available: https://www.mdpi.com/2076-3417/10/14/4965.

[88] N. K. Sah and P. K. Pankaj, "Distribution System Fault Detection and Classification using Wavelet Transform and Artificial Neural Networks," 2022.

[89] M. Y. Suliman and M. T. Alkhayyat, "High impedance fault detection in radial distribution network using discrete wavelet transform technique," *Archives of Electrical Engineering,* pp. 873-886-873-886, 2021.

[90] T. Biswal and S. Parida, "High impedance fault detection in Microgrid system using Discrete wavelet transform and Support vector machine," in *2022 Second International Conference on Power, Control and Computing Technologies (ICPC2T)*, 2022: IEEE, pp. 1-6.

[91] F. Lucas, P. Costa, R. Batalha, D. Leite, and I. Škrjanc, "Fault detection in smart grids with time-varying distributed generation using wavelet energy and evolving neural networks," *Evolving systems,* vol. 11, pp. 165-180, 2020.

[92] N. Vineeth and P. Sreejaya, "High impedance fault detection in low voltage distribution systems using wavelet and harmonic fault indices," in *2020 IEEE International Conference on Power Electronics, Smart Grid and Renewable Energy (PESGRE2020)*, 2020: IEEE, pp. 1-6.

[93] W. Rebizant and K. Feser, "Fuzzy logic application to out-of-step protection of generators," in *2001 Power Engineering Society Summer Meeting. Conference Proceedings (Cat. No.01CH37262)*, 15-19 July 2001 2001, vol. 2, pp. 927-932 vol.2, doi: 10.1109/PESS.2001.970179.

[94] C. Hong-Chan, "Fault section diagnosis of power system using fuzzy logic," *IEEE Transactions on Power Systems,* vol. 18, no. 1, pp. 245-250, 2003, doi: 10.1109/TPWRS.2002.807095.

[95] H.-J. Cho and J.-K. Park, "An expert system for fault section diagnosis of power systems using fuzzy relations," *IEEE Transactions on power systems,* vol. 12, no. 1, pp. 342-348, 1997.

[96] S. Min, S. Lee, and J. Park, "An Expert System for Fault Section Diagnosis in Power Systems using the information including operating times of actuated relays and tripped circuit breakers," in *Proceedings of the KIEE Conference*, 2000: The Korean Institute of Electrical Engineers, pp. 125-127.

[97] O. A. S. Youssef, "Combined fuzzy-logic wavelet-based fault classification technique for power system relaying," *IEEE Transactions on Power Delivery,* vol. 19, no. 2, pp. 582-589, 2004, doi: 10.1109/TPWRD.2004.826386.

[98] B. Das and J. V. Reddy, "Fuzzy-logic-based fault classification scheme for digital distance protection," *IEEE Transactions on Power Delivery,* vol. 20, no. 2, pp. 609-616, 2005, doi: 10.1109/TPWRD.2004.834294.

[99] D. Srinivasan, R. L. Cheu, Y. P. Poh, and A. K. C. Ng, "Automated fault detection in power distribution networks using a hybrid fuzzy–genetic algorithm approach," *Engineering Applications of Artificial Intelligence,* vol. 13, no. 4, pp. 407-418, 2000/08/01/ 2000, doi: https://doi.org/10.1016/S0952-1976(00)00012-9.

[100] S. Vasilic and M. Kezunovic, "Fuzzy ART neural network algorithm for classifying the power system faults," *IEEE Transactions on Power Delivery,* vol. 20, no. 2, pp. 1306-1314, 2005, doi: 10.1109/TPWRD.2004.834676.

[101] P. K. Dash, M. Padhee, and T. K. Panigrahi, "A hybrid time–frequency approach based fuzzy logic system for power island detection in grid connected distributed generation," *International Journal of Electrical Power & Energy Systems,* vol. 42, no. 1, pp. 453-464, 2012/11/01/ 2012, doi: https://doi.org/10.1016/j.ijepes.2012.04.003.

[102] P. Mahat, Z. Chen, and B. Bak-Jensen, "Review on islanding operation of distribution system with distributed generation," in *2011 IEEE Power and Energy Society General Meeting*, 2011: IEEE, pp. 1-8.

[103] V. Calderaro, C. N. Hadjicostis, A. Piccolo, and P. Siano, "Failure Identification in Smart Grids Based on Petri Net Modeling," *IEEE Transactions on Industrial Electronics,* vol. 58, no. 10, pp. 4613-4623, 2011, doi: 10.1109/TIE.2011.2109335.

[104] A. Mouzakitis, "Classification of fault diagnosis methods for control systems," *Measurement and Control,* vol. 46, no. 10, pp. 303-308, 2013.

[105] P. Bunnoon, "Fault detection approaches to power system: state-of-the-art article reviews for searching a new approach in the future," *International Journal of Electrical and Computer Engineering,* vol. 3, no. 4, p. 553, 2013.

[106] M. Dehghani, M. H. Khooban, and T. Niknam, "Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations," *International Journal of Electrical Power & Energy Systems,* vol. 78, pp. 455-462, 2016/06/01/ 2016, doi: https://doi.org/10.1016/j.ijepes.2015.11.048.

[107] A. Zidan *et al.*, "Fault Detection, Isolation, and Service Restoration in Distribution Systems: State-of-the-Art and Future Trends," *IEEE Transactions on Smart Grid,* vol. 8, no. 5, pp. 2170-2185, 2017, doi: 10.1109/TSG.2016.2517620.

[108] S. Adhikari, N. Sinha, and T. Dorendrajit, "Fuzzy logic based on-line fault detection and classification in transmission line," *SpringerPlus,* vol. 5, no. 1, p. 1002, 2016/07/07 2016, doi: 10.1186/s40064-016-2669-4.

[109] R. Perez *et al.*, "Fault Diagnosis on Electrical Distribution Systems Based on Fuzzy Logic," in *Advances in Swarm Intelligence*, Cham, Y. Tan, Y. Shi, and Q. Tang, Eds., 2018// 2018: Springer International Publishing, pp. 174-185.

[110] F. G. Y. Souhe, A. T. Boum, P. Ele, C. F. Mbey, and V. J. F. Kakeu, "Fault Detection, Classification And Location In Power Distribution Smart Grid Using Smart Meters Data," *Journal of Applied Science and Engineering,* vol. 26, no. 1, pp. 23-34, 2022/03/18 2022, doi: 10.6180/jase.202301_26(1).0003.

[111] R. Verma and M. A. Ansari, "Fault Detection and Classification Using Fuzzy Logic Controller and ANN," in *Micro-Electronics and Telecommunication Engineering*, Singapore, D. K. Sharma, S.-L. Peng, R. Sharma, and D. A. Zaitsev, Eds., 2022// 2022: Springer Nature Singapore, pp. 25-35.

[112] S. Roy and S. Debnath, "PSD based high impedance fault detection and classification in distribution system," *Measurement,* vol. 169, p. 108366, 2021/02/01/ 2021, doi: https://doi.org/10.1016/j.measurement.2020.108366.

[113] B. K. Chaitanya and A. Yadav, "An intelligent fault detection and classification scheme for distribution lines integrated with distributed generators," *Computers & Electrical Engineering,* vol. 69, pp. 28-40, 2018/07/01/ 2018, doi: https://doi.org/10.1016/j.compeleceng.2018.05.025.

[114] K. K. Shukla and A. K. Tiwari, *Efficient algorithms for discrete wavelet transform: with applications to denoising and fuzzy inference systems*. Springer Science & Business Media, 2013.

[115] P. Sharma, D. Saini, and A. Saxena, "Fault detection and classification in transmission line using wavelet transform and ANN," *Bulletin of Electrical Engineering and Informatics,* vol. 5, no. 3, pp. 284-295, 2016.

[116] Y. Ahuja, "Multiclass classification and support vector machine," *Global Journal of Computer Science and Technology,* vol. 12, no. G11, pp. 15-19, 2012.

[117] P. Dondi, D. Bayoumi, C. Haederli, D. Julian, and M. Suter, "Network integration of distributed power generation," *Journal of Power Sources,* vol. 106, no. 1, pp. 1-9, 2002/04/01/ 2002, doi: https://doi.org/10.1016/S0378-7753(01)01031-X.

[118] J. Driesen and R. Belmans, "Distributed generation: challenges and possible solutions," in *2006 IEEE Power Engineering Society General Meeting*.

[119] A. D. Udgave and H. T. Jadhav, "A review on Distribution Network protection with penetration of Distributed Generation," in *2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO)*, 9-10 Jan. 2015 2015, pp. 1-4, doi: 10.1109/ISCO.2015.7282387.

[120] Y. Ma, G. Li, and R. Tang, "Optical performance of vertical axis three azimuth angles tracked solar panels," *Applied Energy,* vol. 88, no. 5, pp. 1784-1791, 2011.

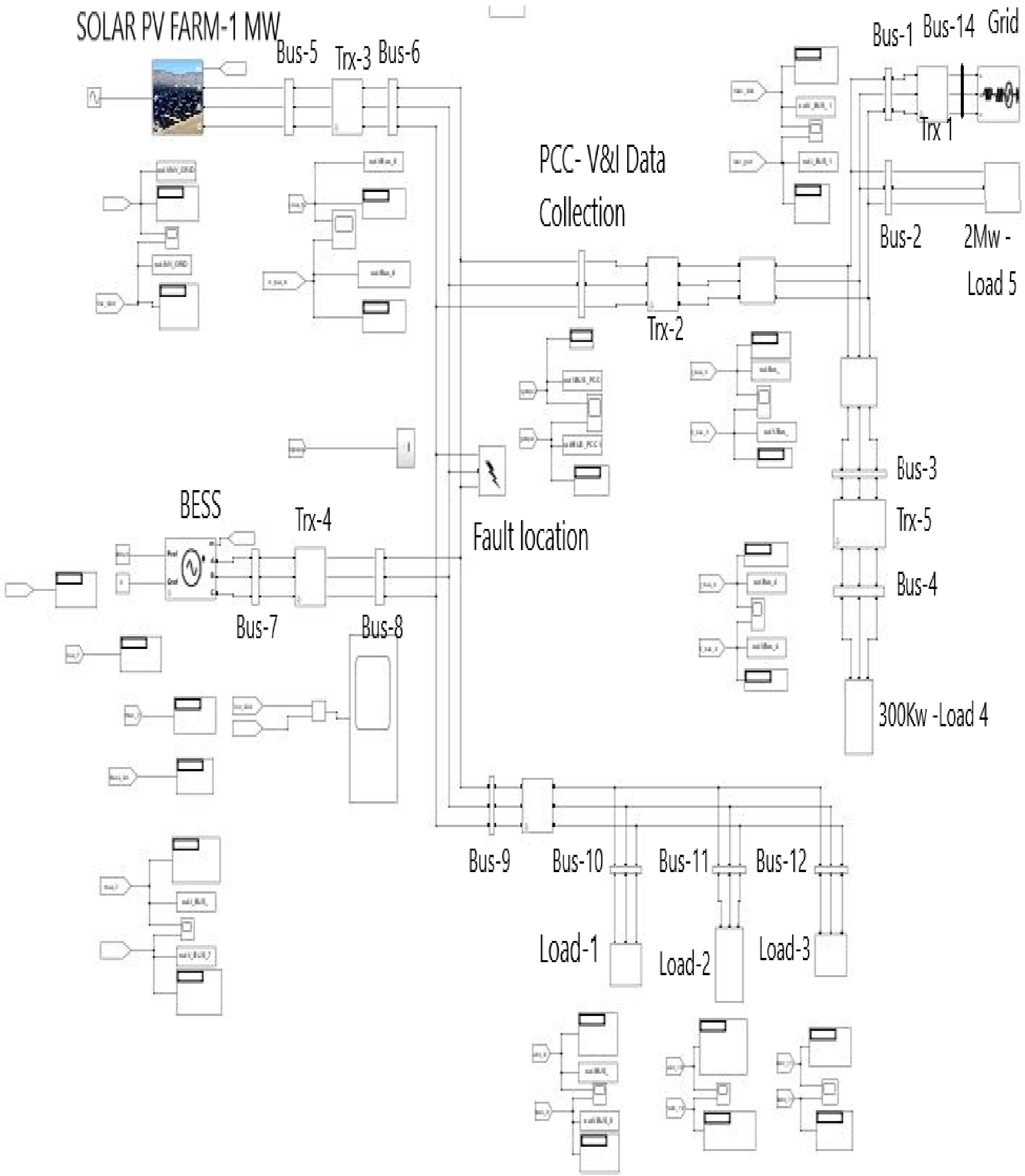
[121] J. Yang, J. E. Fletcher, and J. O'Reilly, "Short-circuit and ground fault analyses and location in VSC-based DC network cables," *IEEE transactions on Industrial Electronics,* vol. 59, no. 10, pp. 3827-3837, 2011.

[122] I. Kasikci, *Short circuits in power systems: A practical guide to IEC 60909-0*. John Wiley & Sons, 2018.

[123] J.-H. Hong and S.-B. Cho, "A probabilistic multi-class strategy of one-vs.-rest support vector machines for cancer classification," *Neurocomputing,* vol. 71, pp. 3275-3281, 10/01 2008, doi: 10.1016/j.neucom.2008.04.033.

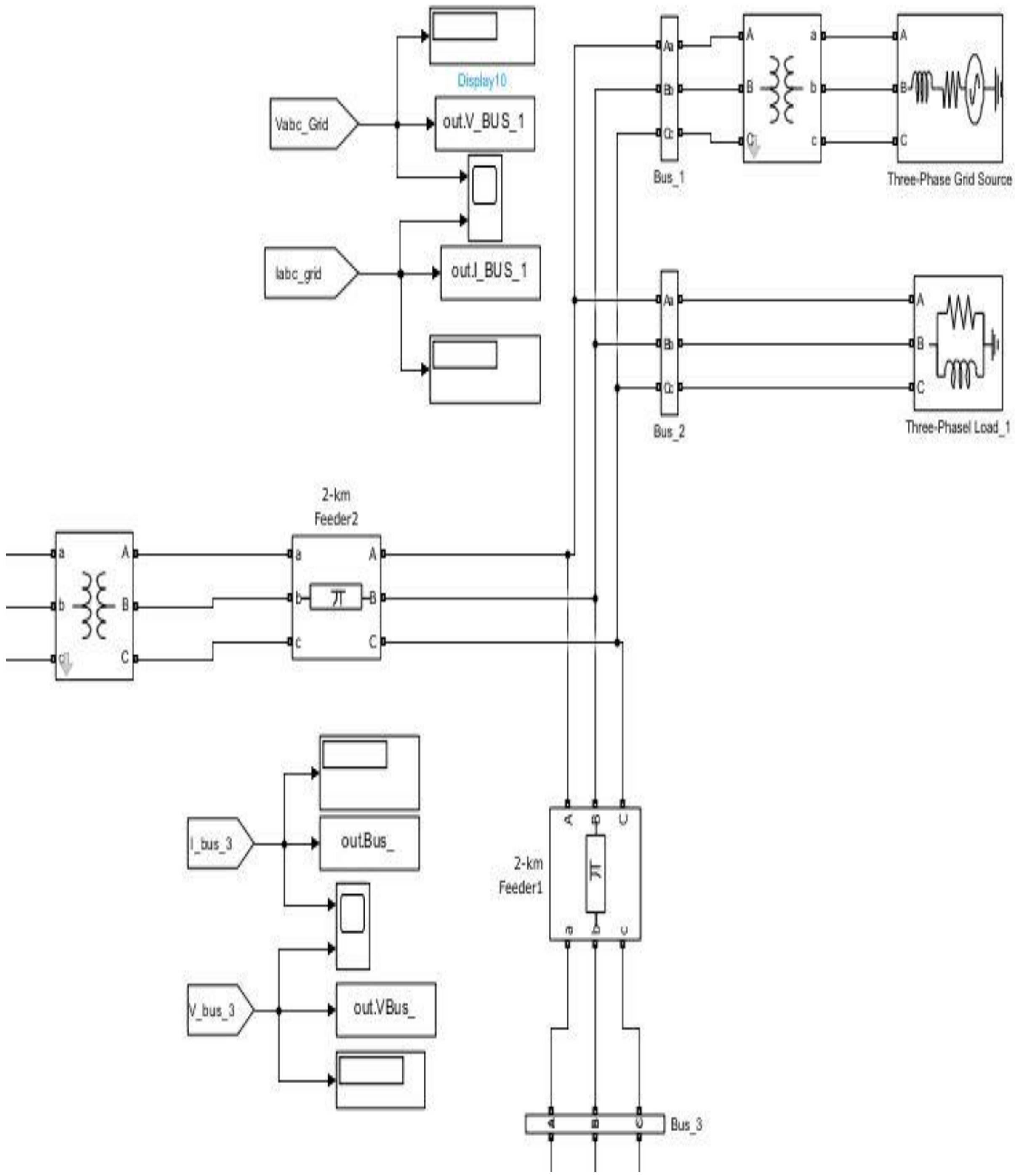
7. Appendix A

7.1 Complete micro-grid model



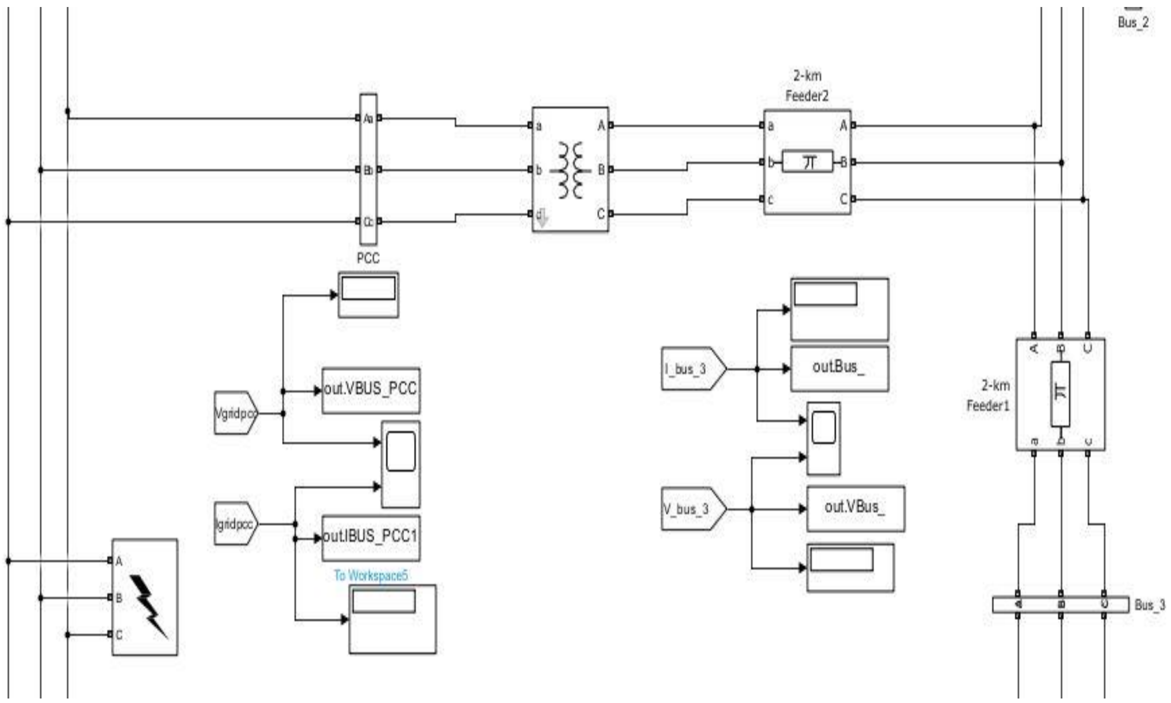
7.2 Appendix B

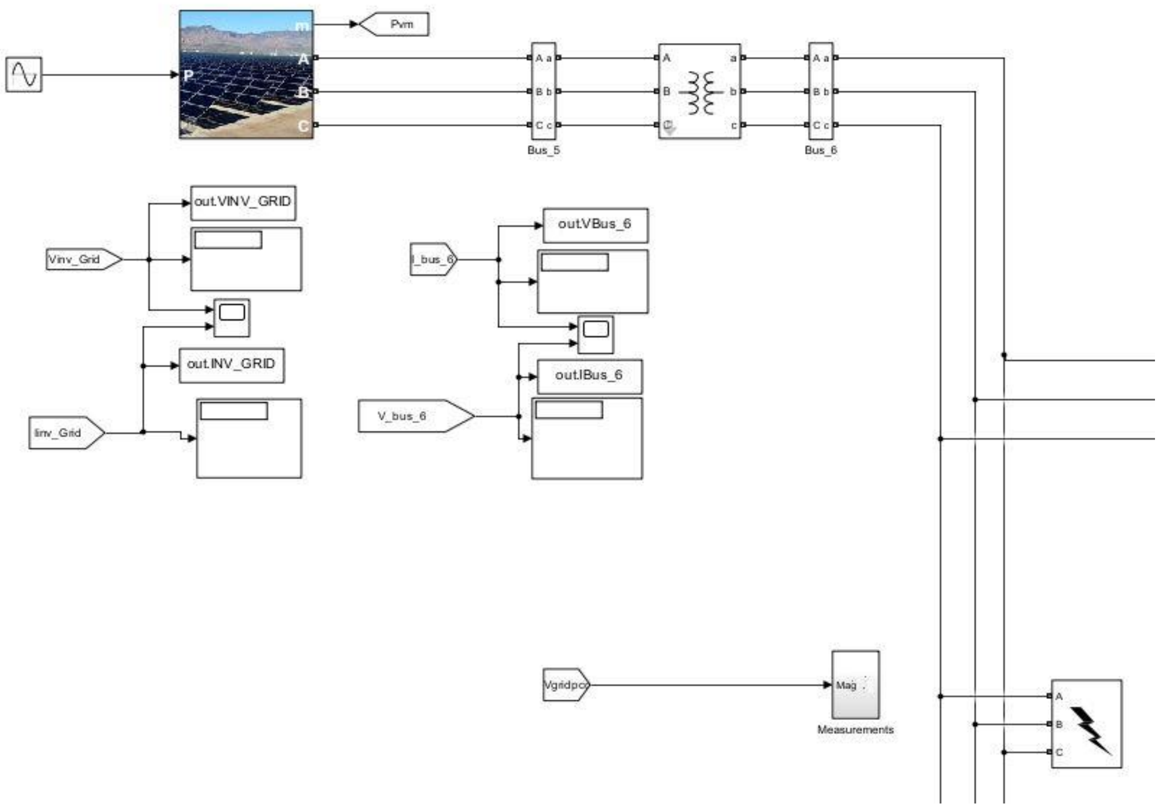
Grid side of the Micro-grid model



7.3 Appendix C

From the PCC and fault initiation location.





7.4 Appendix D

Battery side of the model and different types of loads