
Automatic Identification of Fault Types in the Distribution Network using Supervised Learning

by

Archana Ranganathan

to obtain the degree of Master of Science
at the faculty of Electrical Engineering, Mathematics, and Computer Science
at Delft University of Technology,
to be defended publicly on Tuesday August 31, 2021 at 16:00

Student Number:	5043727		
Project Duration:	04.01.2021 - 31.08.2021		
Thesis Committee	Dr. Simon Tindemans	Supervisor	IEPG, TU Delft
	Dr. ir. Frans Provoost	Supervisor	Qirion, Alliander
	Prof. dr. Peter Palensky	Chair	IEPG, TU Delft
	Dr. ir. Jianning Dong	Committee Member	DCE&S, TU Delft



Abstract

Electrical faults in the distribution system can lead to interruptions in customer power supply resulting in penalties that are borne by the distribution system operator. Accurate fault classification is an important step in locating the fault to achieve faster network restoration times. This reduces the operational costs of the system operator by reducing the duration of interruptions in the power supply. Furthermore, it has been found from practical experience that distortions or instabilities in fault waveforms can result in their classification and subsequent localisation being delayed, causing the loss of valuable customer minutes. The objective of this thesis is to investigate the potential of modern supervised learning techniques for the classification of faults in the distribution network. The problem is split into two parts: one aspect is to study the ability of a supervised learning classifier to differentiate between types of stable faults, and in this pursuit, to also identify better criteria for stability.

First, after a review of pertinent literature, it is found that discrete wavelet transforms and support vector machines (SVM) are suitable for fault signal processing and classification, respectively. Next, identifying characteristics, or features, from the three-phase fault current and voltage waveforms are engineered with the help of the db4 wavelet. The features are used as inputs to an SVM classifier model, and the model is tuned and validated to ensure optimal classification performance. Results showed that the developed SVM can differentiate between real-world instances of single-phase, two-phase and three-phase stable faults with a classification accuracy of 95%. A set of business rules are also developed to characterise instability by performing a windowed Fourier analysis and studying the strength of the fundamental frequency component of fault waveforms. The rules are tested on a set of fault data whose stability is uncertain, and it is found that the developed rules are able to improve upon the older method of stability analysis by increasing the rate of stable fault identification.

Acknowledgements

This thesis is my final work as part of my Master's degree in electrical engineering, and looking back, I am extremely thankful for my experience at TU Delft. It was a difficult yet rewarding journey, to work on a research project amidst the pandemic, but I am grateful for everything I learnt, and more importantly, the people I learnt from. I would like to use this space to thank the people who helped make this thesis possible.

Firstly, I would like to thank my supervisor at TU Delft, Dr. Simon Tindemans. His guidance and experience were instrumental to my research and his feedback always pushed me to do my best. I thank my supervisor at Qirion, Dr.ir. Frans Provoost, for his invaluable expert knowledge in electrical power engineering and his advice that was integral to my efforts during my thesis. Together, my supervisors ensured my research went in the right direction and I have learnt much about data analytics and power engineering from them both. I am grateful to Ir. Tongyou Gu from Alliander for her support and input during vital stages of my project. I would also like to extend my thanks to Prof.dr. Peter Palensky and Dr.ir. Jianning Dong for agreeing to be part of my thesis committee and evaluating my work.

For the good ideas, company, and dinners I'd like to thank my friends. The long hours spent studying would have been significantly harder without the prospect of meeting them at the next possible break.

Finally, I thank my parents for their ability to make me feel supported and motivated through the large physical distance between us. Thank you for your constant patience, words of advice, and encouragement. This thesis is dedicated to you.

*Archana Ranganathan
Delft, August 2021*

Contents

Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 Background	1
1.2 Motivation	2
1.3 Thesis Objectives and Scope	3
1.4 Research Questions	3
1.5 Thesis Structure	4
2 Literature Review	5
2.1 Introduction to Faults in the Distribution Network	5
2.2 Overview of the Present Classification Model at Alliander	6
2.3 Review of Classification Methods for Electrical Faults	8
2.4 Discussion	13
2.5 Conclusion	14
3 Theoretical Background	16
3.1 Signal Decomposition using Discrete Wavelet Transforms	16
3.2 Supervised Learning for Classification	20
3.2.1 Overview of the Classification Process	20
3.2.2 Support Vector Machines	21
3.3 Statistical Tests for Determining Feature Importance	23
3.4 Metrics for Evaluating Classifier Model Performance	24
4 Development of Features for Stable and Unstable Faults	27
4.1 Description of the Sample Space	27
4.1.1 Steps for Data Preprocessing	29
4.1.2 Fault Segmentation Algorithm	29
4.2 Development of Features for Stable Faults	33
4.3 Development of Features for Unstable Faults	37
4.3.1 Introduction to Unstable Faults	37
4.3.2 Windowed Fourier Analysis for Detecting Stability	39
4.4 Conclusion	44
5 Classification of Unstable Faults	45
5.1 Introduction	45

5.2	Thesholds for the Stability Analysis of Unstable Faults	45
5.3	Considerations of the Classifier Algorithm	46
5.4	Development of the Thresholding Algorithm	46
5.4.1	Feature Importance	46
5.4.2	Cross-Validation for Determining Thresholds	48
5.4.3	Validation of the Thresholds	48
5.5	Business Rules for Unstable Fault Classification	49
5.6	Application of the Rules to Unstable Fault Data	50
6	Classification Model for Stable Faults	53
6.1	Introduction	53
6.2	Overview of the Model Development Process	54
6.2.1	Feature Selection	54
6.2.2	Hyper-Parameter Tuning	55
6.2.3	Cross-Validation	55
6.2.4	Unifying the Model Selection Processes	56
6.3	Selection of the Mother Wavelet	57
6.4	Results of the Model Selection	57
6.4.1	Results of ANOVA-F Scoring Feature Selection	58
6.4.2	Results of Mutual Information Scoring Feature Selection	60
6.4.3	Comparison of ANOVA and MI Feature Selection Methods	62
6.4.4	Results of Sequential Feature Selection	62
6.5	Selection of the Model	63
6.6	Conclusion	64
7	Evaluation of the SVM Classifier	65
7.1	Introduction	65
7.2	Evaluation of the Support Vector Machine Performance	65
7.2.1	Defining the Test Cases for Model Evaluation	66
7.2.2	Results of the Test Cases	68
7.3	Selection of the Mother Wavelet	72
7.4	Evaluation of the Data Pre-processing Techniques	73
7.5	Conclusion	74
8	Conclusion and Outlook	75
8.1	Research Questions	75
8.2	Avenues for Future Research	77
A	Supplementary Figures	83
B	Supplementary Results	86

List of Figures

1.1	Fault prediction system in the 10kV MV cable networks as described in [1]	2
1.2	Outline of the thesis.	4
2.1	Current (left) and voltage (right) waveforms for a single phase extinguishing fault.	6
2.2	Comparison of the frequency and time resolution for different signal processing methods.	9
2.3	A simple decision tree classifier.	11
2.4	A comparison between a real-world two-phase fault (left) and a synthetic two-phase fault (right).	13
3.1	Different families of discrete (top) and continuous wavelets (bottom)	17
3.2	Plot depicting the low-pass scaling and high-pass wavelet filter for the daubechies 4 wavelet.	19
3.3	Diagram of wavelet decomposition using cascading filter banks.	19
3.4	Diagram depicting the separating hyperplanes in a two-class classification problem using support vectors [38].	22
3.5	Error matrix that enables the visualisation of the performance of a supervised learning classification problem.	25
4.1	Current (left) and voltage (right) waveforms for a two-phase-to-ground fault. . . .	28
4.2	Current (left) and voltage (right) waveforms for a three-phase-to-ground fault. . .	28
4.3	Current (left) and voltage (right) waveforms for a single-phase fault that develops into a three-phase fault.	28
4.4	Current (left) and voltage (right) waveforms for a two-phase fault that develops into a three-phase fault.	28
4.5	Plot comparing the different states in a permanent and self-extinguishing fault. . .	30
4.6	Change points for the phase currents of the three-phase fault in Figure 4.3	32
4.7	Segmented section of the three-phase fault in Figure 4.3	32
4.8	Current and voltage waveforms for a two-phase synthetically created fault.	33
4.9	Level 1 detail coefficients for the current and voltage of a two-phase fault (in the 1-2 kHz bandwidth).	36
4.10	Current (left) and voltage (right) waveforms for a single-phase unstable fault. . . .	38
4.11	Current (left) and voltage (right) waveforms of an single-phase permanent fault that was identified as unstable.	39
4.12	Current (left) and voltage (right) waveforms of a single-phase self-extinguishing fault that was identified as unstable.	39
4.13	Comparison between zero-sequence currents of a single-phase permanent fault (left) and single-phase self-extinguishing fault (right).	40
4.14	Comparison between the normalised zero-sequence current (I^0), its derivative (I_{deriv}^0) and the reference (I_{ref}). Note: This image shows a window of two cycles for representative purposes.	41
4.15	Comparison of the 50Hz Fourier components between the derivative of the zero-sequence current and its reference sinusoidal waveform.	42

4.16	Comparison of the 50 Hz Fourier component of unstable single-phase permanent (left) and single-phase self-extinguishing (right) faults.	42
5.1	Comparison of the distributions of the f_{max} feature of single-phase permanent and self-extinguishing faults.	47
5.2	Depiction of the rules developed for the determination of acceptable unstable behaviour in a fault.	50
5.3	Subset of stable faults that are a result of the Fourier analysis classification process.	51
5.4	Subset of (possibly) unstable faults that are a result of the Fourier analysis classification process.	52
6.1	Pictorial representation of cross-validation. Feature selection and hyper-parameter tuning occurs at each fold.	56
6.2	Comparison of cross-validation scores using ANOVA-F and Mutual Information (MI) feature selection techniques.	58
6.3	Correlation between the top nine features ranked by the ANOVA-F test.	59
6.4	Correlation between the top 12 features ranked by the mutual information test.	61
6.5	Cross-validation scores of the sequential feature selection techniques.	63
7.1	Precision, recall and accuracy scores of the different test cases.	71
7.2	Precision, recall and accuracy scores of the classifier for different mother wavelets.	73
A.1	Signal decomposition results for a waveform that does not possess any transients using the db4 mother wavelet.	84
A.2	The variation of F_{ratio} for the faults classified as stable with the new business rules.	84
A.3	The variation of F_{ratio} for the faults classified as unstable with the new business rules.	85

List of Tables

2.1	Criteria for the identification of short- and long-term extinguishing faults by Gu et al. [10]	7
2.2	Summary of research papers that use DWT for fault feature extraction and the corresponding mother wavelet used.	14
2.3	Summary of research papers that utilise different types of machine learning classification techniques.	14
3.1	Filter coefficients for the Daubechies-4 low-pass and high-pass wavelet function	18
3.2	Bandwidths of a signal sampled at 4kHz after the wavelet transform at each level of the filter functions.	20
4.1	Summary of the number of DFRs for each fault type collected from the SASensor.	29
4.2	Change points for the fault in Figure 4.3	32
4.3	Summary of the complete sample space of the fault considered for classification	34
4.4	Table detailing the features extracted from the fault signals.	37
4.5	Criteria developed by Alliander for identifying a fault as stable ^a [7]	38
4.6	Table summarising the features extracted from the Fourier analysis for the study of unstable faults	43
5.1	List of features for the characterisation of unstable faults in decreasing order of importance.	47
5.2	Cross-validation scores and values of each feature determined by the decision tree algorithm.	49
5.3	Table describing the classification of unstable faults based on their stability.	51
6.1	List of features selected using the ANOVA-F ranking method (ordered in descending order of importance).	59
6.2	Updated of features selected using the ANOVA-F ranking method.	60
6.3	List of features selected using the mutual information ranking method (ordered in descending order of importance)	61
6.4	Updated of features selected using the mutual information ranking method.	62
6.5	List of features selected using the sequential feature selection technique	63
6.6	The selected model for the support vector machine based on the results from the mutual information feature selection method.	64
7.1	The features selected for the optimised support vector machine classifier model.	67
7.2	The selected model for the support vector machine based on the results from the mutual information feature selection method.	67
7.3	Classification error matrix when classifier is trained and tested on synthetic faults.	68
7.4	Performance scores of the SVM trained and tested on synthetic faults.	69
7.5	Classification error matrix when classifier is trained and tested on real faults.	69
7.6	Performance scores of the SVM trained on synthetic faults and tested on real faults.	69
7.7	Classification error matrix when classifier is trained on real faults and tested on synthetic faults.	70

7.8	Performance scores of the SVM trained on real faults and tested on synthetic faults.	70
7.9	Classification error matrix when classifier is trained and tested on real-world faults.	70
7.10	Performance scores of the SVM trained and tested on real faults.	70
7.11	Classification error matrix when classifier is trained on a combination of real-world and synthetic faults and tested on real faults.	71
7.12	Performance scores of the SVM trained on a combination of real and synthetic faults, and tested on real faults.	71
7.13	Summary of precision, recall and accuracy scores of the different test cases.	72
7.14	Comparison of the performance of the classifier with and without the fault segmentation step.	74
B.1	Cross-validation scores and hyper-parameter values for the ANOVA-F feature selection technique.	87
B.2	Cross-validation scores and hyper-parameter values for the mutual information feature selection technique.	88

1 Introduction

1.1 Background

Alliander is a distribution system operator (DSO) in the Netherlands that is responsible for the development and operation of energy networks. Key features of the distribution system in the 10/20 kV medium voltage (MV) grid involve high power quality and reliability of the electricity supply, both in the present and the future [1]. In the context of distribution network operation, the minimisation of power interruption time is vital for the quality of supply for the customers. Hence, a key facet of Alliander's responsibilities is fast location of faulted cables to quickly restore power supply. A fault in the power system is defined as an abnormal state that is characterised by large currents (in the kA range) that are switched off by protective devices, such as circuit breakers, which cause a blackout in a part of the MV network. After the faulted cable section is located, it can be isolated from the rest of the network and the energy supply can be restored.

The principal goal of grid improvement is to reduce outages and the loss of customer minutes. One way to achieve this is to reduce the time in the manual fault localisation process. It is also possible to predict future faults and repair weak cable joints by studying self-extinguishing faults as they are usually precursors to permanent faults, i.e., faults with large currents that need to be switched off by a circuit breaker. Self-extinguishing faults occur for a small duration of between half a cycle to two cycles. The time between the first occurrence of self-extinguishing faults and their evolution into permanent faults can be between several weeks to months. Furthermore, self-extinguishing faults themselves do not require activation of the fault protection system as they do not interrupt the power supply; however, as they can evolve into actionable persistent faults, it is also necessary to study how these faults can be detected. The system for the detection of these faults can be observed in Figure 1.1.

It is also important to note that the classification and localisation of faults in distribution networks are more difficult than in transmission networks. This is due to the complex nature of the distribution network and the limited number of measurements — the MV lines are underground and have taps that connect to the customers, as opposed to the transmission network, where the cables run for long distances between substations without interruptions. Additionally, with the growth of the population and the need for decentralised generation, the load carried by the distribution network has increased considerably [2][3].

Presently, a substation automation tool called the SASensor is used for modernizing the grid with its centralised protection and control architecture. Alliander's fault detection and localisation system compares certain values of the fault current and voltage with preset thresholds that were set by studying past instances of faults. As the complexity of the distribution network increases with the increase in distributed generation and two-way power flows, accurate fault classification is more crucial than ever. This thesis uses modern signal processing techniques and data-driven classification algorithms to study if newer methods can improve the classification of faults in the distribution network. For this purpose, supervised learning techniques are used

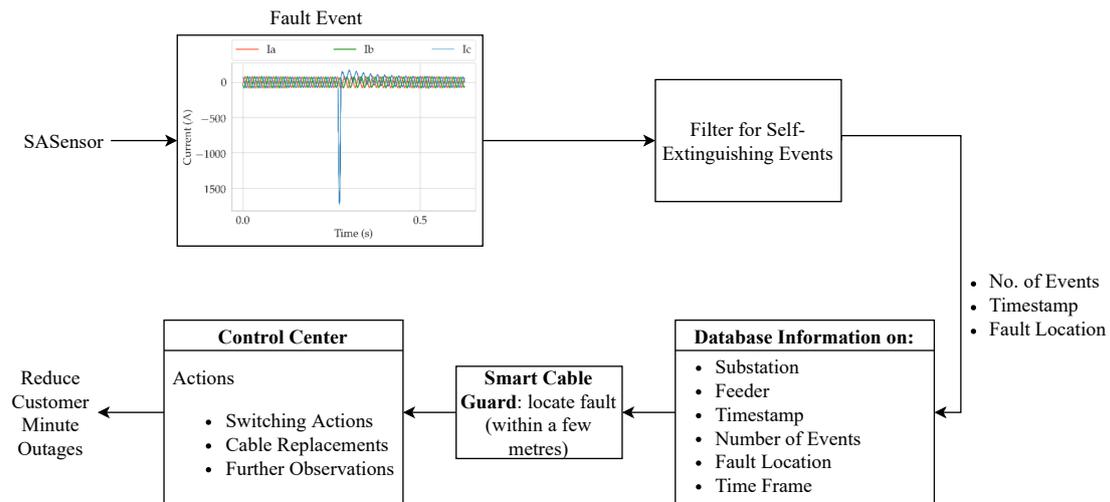


Figure 1.1: Fault prediction system in the 10kV MV cable networks as described in [1]

where a function is inferred between (labelled) data and their classes to make predictions about the class of new/unseen data. In this thesis, the classes are the fault types, and the labelled data are selective properties of the faults. This chapter first provides more context to the motivations for conducting this thesis in Section 1.2. After identifying the incentive for this study, the main goals of the thesis are outlined in Section 1.3 in the form of research objectives. From here, Section 1.4 narrows down on the specific research questions that this study will attempt to answer. Finally, the structure of this thesis is presented in Section 1.5

1.2 Motivation

The metrics Customer Average Interruption Duration Index (CAIDI) and System Average Interruption Duration Index (SAIDI) are used to quantify the reliability of a system in terms of the power that the electric grid is unable to deliver, from the customer and system side, respectively. It is found that the cause of 65% of all interrupted customer minutes originated in the distribution network in the 10kV level[4]. The high voltage (transmission) level has a contribution of 20% to the customer minutes lost due to the redundant lines and superior network automation, however, faults in the transmission system have a wider area of influence than in the distribution system [4]. Hence, priority has to be given to the faster location and isolation of faulted cables in the MV network to reduce the CAIDI and SAIDI. Moreover, Dutch DSOs are incentivised to minimise customer interruption minutes through a penalty and reward system where utilities are mutually benchmarked based on their quality of performance, or q-factor [5]. In the year 2020, Alliander had a SAIDI of 23.2 minutes and a CAIDI of 70.6 minutes [6].

The MV grid is connected to the low voltage (LV) network with MV/LV transformer stations called Ring Main Units (RMU). When a fault occurs in the distribution network, a larger section of the RMU is first disconnected from the supply before the exact location of the faulted cable connecting the the units is identified. A critical aspect of the network protection philosophy in the MV network is based on selectively switching off faulted parts of the network so power can be restored to the rest of the network. The switching times are based on the magnitude of the current and the location of the fault in the network. The identification of the fault type is necessary for locating the fault because this determines the how the fault-loop impedance is calculated. This fault-loop impedance is then used in the location calculations by network analysis tools, such as Vision software. Misclassifications of faults result in the calculation of the wrong loop impedance

which affects the localisation.

The location of faults whose nature are certain are automatically calculated and sent to the control center so the faults can be isolated from the rest of the network. Self-extinguishing faults occur only for short durations and will therefore not be located, or interrupted by a circuit breaker. It is currently a challenge to distinguish between a series of self-extinguishing faults and a permanent fault due to their similarity in waveforms. Additionally, when a fault waveform is not sufficiently sinusoidal, it is classified as an unstable fault. In case there is a possibility that the fault is not classified accurately, such as with an unstable fault, the location of the fault is not sent to the control centre. This is to prevent incorrect information from switching off the wrong section of the network. The possible location of such a fault is later manually calculated by experts. The question therefore arose as to how the criteria for classifying (un)stable faults could be improved.

Unstable faults are currently loosely defined with preset thresholds as described by Gu et al. in [7]; however, faults that could be classified as stable faults occasionally end up classified as unstable as some of the parameters on which they are assessed cross the border that defines stability. Improving the criteria for stability can therefore increase the number of correct fault locations sent to the control center. It can also reduce the time spent on manually identifying these misclassified faults if a better definition of unstable faults can be devised that prevents stable faults from being classified as unstable.

The control center must receive the correct location for isolating the faulted section of the distribution network. If the wrong location is sent to the control center, valuable fault isolation time is wasted. It is therefore important, due to the necessity of sending only correct information to the control center, and correctly classifying stable faults (that are classified as unstable in the current method), to study if there are better ways to study fault waveforms. This culminates in the definition of the research objectives and questions which are elaborated in the following two sections.

1.3 Thesis Objectives and Scope

The overarching objective of the thesis is to study the potential of using machine learning to classify faults in the distribution network. Signal processing techniques for studying fault waveforms and modern classification algorithms are first analysed by performing a literature review. It is the aim of the thesis to develop a classification algorithm that can classify stable faults and to gain insights into what aspects of the faults are vital for their differentiation. The thesis objective can be split broadly into two parts: one is to develop a classifier for accurately labelled data and the other is to classify data whose labels are suspect and bring them under the umbrella of accurately labelled data.

The scope of the study is bounded by the reliable fault data available — the nature of the faults should be exactly known before they can be studied. Insights are then made on the performance of the supervised learning classifier on stable and unstable faults. The thesis objectives are refined in the next section, in terms of research questions that are answered over the course of this study.

1.4 Research Questions

The following are the research questions that are addressed in this thesis. The overarching question is:

In what capacity can modern signal processing and machine learning techniques improve the classification of faults in the distribution network?

The sub-questions that are answered during the course of this thesis are:

1. Which machine learning approaches exist for detection and classification of faults?
2. What are the criteria that must be used to differentiate between stable and unstable fault behaviour?
3. Within the umbrella of stable faults, what are the aspects of stable fault waveforms that lend to the identification of fault types?
4. How can self-extinguishing faults be differentiated from faults that are directly switched-off by a circuit breaker or a fuse?
5. What are the methods to evaluate the performance of the machine learning model and how does the developed classifier perform in its ability to distinguish between faults?

1.5 Thesis Structure

The thesis is broadly structured as shown in the outline in Figure 1.2.

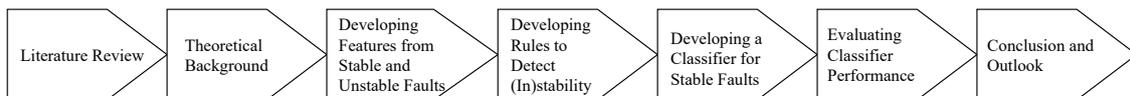


Figure 1.2: Outline of the thesis.

First, Chapter 2 surveys the relevant literature in the field of fault classification and studies techniques that can be applied to the fault classification problem at hand. Gaps in the present research in this field are identified and their implications to the research questions are discussed. Chapter 3 covers the theoretical background required to understand the signal processing techniques and supervised learning classification methodologies that are used in the later chapters. Chapter 4 describes the process of feature engineering where pertinent features from the stable and unstable faults are developed. The features developed for unstable faults in Chapter 4 are used to develop business rules for the classification of unstable faults in Chapter 5. The methodology for the development and optimisation of the supervised learning model for the classification of stable faults is discussed in 6. The performance of the classifier is then assessed using test cases; the results of which are presented in Chapter 7. Finally, in Chapter 8, conclusions are drawn from the study, and the research questions are reiterated with explanations as to how they have been answered and to what extent. Recommendations are then given for how this research can be taken forward for future research.

2 Literature Review

Introduction

Fault analysis is integral to the energy network protection scheme. This chapter reviews the techniques for studying and classifying faults in the power system. First, Section 2.2 will lay the base with the techniques that Alliander uses to implement the present fault location system in the distribution network. This is followed by a review of papers that aims to assess the current methods for fault analysis in Section 2.3. Section 2.4 identifies the gaps and limitations of the reviewed research. The chapter concludes with Section 2.5 that describes how the present research can be taken forward by discussing the thesis objectives in the context of the existing research.

2.1 Introduction to Faults in the Distribution Network

Electric faults in networks are defined as any failure in the circuit that disrupts the normal current flow. All of Alliander's distribution network is underground and one cause of faults can be attributed to short circuits due to insulation failure. Underground faults can also be caused by factors such as damage to cables caused due to digging and interference from tree roots [4]. Incipient faults are temporary over-currents that occur due to defects or voids in the cable insulation and other accessories. These faults also occur due to localised degradation of the insulation due to aging, and will, if no action is taken, finally result in a permanent fault. [8].

Faults can be classed as either symmetrical or unsymmetrical faults. Unsymmetrical faults involve a combination of one or two phases and the neutral/ground. Of these faults, single-phase-to-ground faults are the most commonly occurring faults, with nearly 70-80% of faults being this type of fault [9]. Two-phase faults occur between two phase conductors of the network and involve the ground in two-phase-to-ground faults. Symmetrical faults are three-phase faults that involve all three current-carrying conductors when they come in contact with each other or the ground. They are the rarest faults but also the most severe for the system. Hence, faults can be differentiated based on their waveforms.

Faults can also be differentiated based on how they are extinguished. Permanent faults last for longer than four cycles and must be switched off externally by a circuit-breaker. Self-extinguishing faults are short-term faults that occur from half a cycle up to four cycles and the extinction of the fault occurs at the zero-crossing of the fault current. Figure 2.1 shows the behaviour of a single-phase self-extinguishing fault that lasts for half a cycle at 0.2s at phase C. Finally, faults can also be characterised by their stability. Stable faults are defined as faults whose current and voltage waveforms are "sufficiently close to an ideal sinusoidal waveform [7]." A formal definition for unstable faults has not been established yet, and this thesis studies possible ways to make such a distinction between stable and unstable sinusoidal waveforms.

2.2. Overview of the Present Classification Model at Alliander

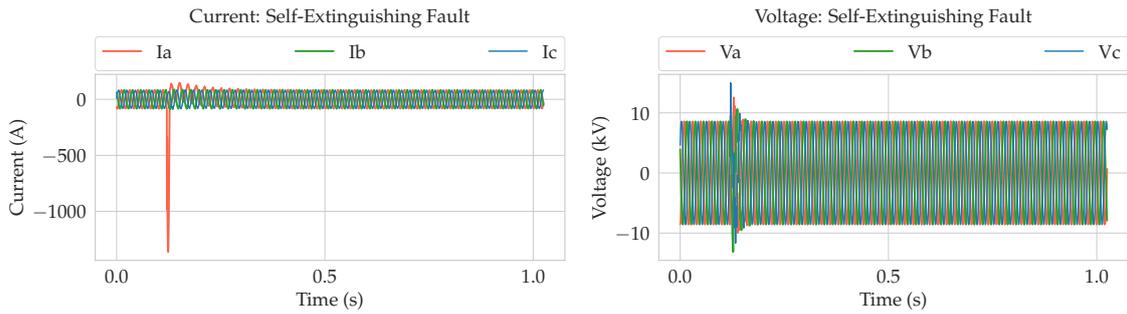


Figure 2.1: Current (left) and voltage (right) waveforms for a single phase extinguishing fault.

Self-extinguishing fault waveforms are characterised by a high-frequency transient that is superposed on the wave shape in steady-state [10]. The discharge of the faulted phase creates transient behaviour. These faults last for short time periods between half a cycle and multiple cycles [8]. Repetitive sequences of self-extinguishing faults can lead to permanent faults and they are hence called incipient faults.

Single-phase faults, single-phase self-extinguishing faults, two-phase faults, three-phase faults, and unstable single-phase faults are in the purview of this thesis as these classes have the highest strength in reliably labelled data. The following section explains how these faults are classified in Alliander's fault location system.

2.2 Overview of the Present Classification Model at Alliander

Operation of Alliander's Network Protection System

Alliander implemented a pilot digital protection system in a substation at Zaltbommel in the year 2002, where data from fault recordings were sent to a control centre for analysis of the faults and their locations. In 2007, a substation automation system called the SASensor was implemented in 10 substations as a protection and control system. The sensor has protection, control, energy metering, and power quality monitoring functionalities. The SASensor samples the current and voltage at a high sampling rate of 28 kHz.

For the case of fault classification, the SASensor creates digital fault recordings (DFR) that store the data, resampled at 4 kHz, for the relevant process variables — which, in this case, are the three-phase voltage and current waveforms. This data then is sent to the control centre for further automatic analysis. First, these recordings of voltage and currents are analysed to find their fault type. After this, the fault-loop impedance between the substation and fault location is calculated. Next, the impedance of the feeder at which the fault occurs is ascertained from simulating the fault on a network model. The exact fault location can be determined by verifying if the measured short-circuit impedance of the fault correspond to the simulated values for a variety of locations. The network model reflects the actual network data such as network nodes, cables, and switch positions in the feeder. This model is stored and regularly updated by a Geographic Information System (GIS). Following the identification of the fault location, the control center is tasked with clearing the faults. Dutch MV networks generally utilise circuit-breakers for clearing faults and these circuit-breakers are activated a few periods after the incidence of a fault [4].

The determination of fault type is vital to the calculation of the fault location. The location of the fault can only be sent to the control center for further isolation from the network if the

location is reliable. Unstable faults have irregular and distorted sinusoidal waveforms. Unstable behaviour in a fault can prevent the fault protection system from identifying the type of the fault, and thereby its location. The location of such faults whose nature is uncertain is not sent to the control unless they are manually inspected to verify their fault type and location. Additionally, due to the broad definition of unstable faults, faults that can otherwise be classified as (stable and) permanent or extinguishing, are labelled unstable due to a certain degree of distortions. Hence, it is important to investigate methods for improving the definition of unstable faults for two reasons — to prevent misclassifications, and to enable efficient network restoration by way of calculating the correct fault location.

Fault Classification Methodology at Alliander

A desired requirement of digital fault location schemes is that they use minimal data for faster calculations to enable faster fault interruption. The fault analysis consists of two stages - the first is the detection and classification using short term Fourier transforms (STFT), and the second stage is locating the fault by performing short-circuit calculations on a model of the actual network. It was found that a drawback of using STFT to study the events meant that single phase extinguishing faults, which consists of repetitive restrike and extinction, have similar frequency registrations as single-phase-to-ground faults [3].

Multi-phase faults are often preceded by a series of single-phase self-extinguishing faults that can last days or weeks before developing into a permanent fault. These self-extinguishing faults should be identified and localised in order to prevent/reduce the impact of a future multi-phase permanent fault. Gu et al. [10], at Alliander, develop practical rules for identifying single phase extinguishing faults using historical data as a reference to create thresholds for the peak current and rate of change of the root mean square (RMS) of the zero-sequence current (I_0) to set benchmarks for what classifies as a self extinguishing fault. This model was implemented in 2012 and has been able to sufficiently identify self-extinguishing faults [10]. Table 2.1 indicates the criteria for the classification of short-term and long-term extinguishing faults. Short-term extinguishing faults have a duration of around $2ms$, and long-term extinguishing faults last for about $10ms$, which is around half a cycle. In order to makes the rules generalised (as opposed to network specific) it is necessary to compare the maximum instantaneous value and the RMS value of the zero sequence current at every cycle (20ms) [10]. This method is indeed successful in identifying self-extinguishing faults, however, it requires frequent recalculations and updating. This is done to manually obtain a global location for the fault as the network model is subject to changes.

Table 2.1: Criteria for the identification of short- and long-term extinguishing faults by Gu et al. [10]

Parameter	Short	Long	Unit
$ I_{0_max} $	>30	>30	A
$ I_{0_max} /I_{0_rms}$	>2	>2	-
$ dI_{0_max} /dI_{0_rms}$	>3	-	-

Gu et al. [7] improve on their fault analysis model from their previous work in [10], by decreasing the number of inputs. The new model takes only the transient currents and voltages as inputs and solves misclassifications that occur due to self-extinguishing faults having similar waveforms as single-phase faults, or inrush currents, for example. This algorithm uses the negative and zero sequence currents, as well as the loop impedance/reactance to perform the classification. The difference between three-phase, two-phase, and single-phase faults are set based on thresholds for the loop impedance/reactance.

Unstable faults are characterised by fault current or voltage waveforms that do not closely

match an ideal sinusoidal waveform. Unstable faults are currently differentiated from stable faults by the maximum value of the waveform, the RMS of the waveform, and the differences between sample data points. The criteria and thresholds for classification are obtained from practical experience and mathematical principles. The algorithm first checks if the fault is stable or not and if stable, then proceeds to classify it as a single-phase, two-phase, or three-phase fault. The fault location is then determined based on the short-circuit impedance, which is calculated from the current and voltage [7].

A desired property of a fault protection system is the minimisation of false alarms; while this is implemented in [7] with a set of business rules, it is worth looking into whether modern methods of data processing and/or machine learning perform better. Hence, a part of the project is to study how the aforementioned method compares to other modern methods of signal processing and classification algorithms. The following section is a review of literature in the field of fault signal analysis and fault classification.

2.3 Review of Classification Methods for Electrical Faults

To locate a fault in power networks, the type of fault should be correctly identified. This location, calculated as a function of the fault type and the fault impedance, is sent to the control center for further isolation from the network for protective purposes. As mentioned previously, much of the existing scientific literature focuses on fault analysis problems in transmission lines as they carry a higher amount of power, and can affect a larger portion of the power network [2]. This section investigates the different signal processing and classification techniques for faults in the distribution network.

Review of Signal Processing Techniques for Studying Faults

During the course of the literature review, it was observed that the discrete wavelet transform (DWT) was a popular method for extracting features from the fault signals. It is useful to understand how wavelet transforms work, briefly, before moving on to classification techniques. It was found that an increasing amount of modern research on fault analysis involves studying faults through the lens of wavelet transforms. Historically, temporal analysis (by analysing the current and voltage waveforms in time) or frequency analysis (by using the Fourier transforms) were utilised for the analysis of electromagnetic transients and fault signals [11] [12]. The following paragraph reviews the reasons why wavelet transforms have gained popularity in the field of fault analysis by comparing it with the conventional signal processing methods — such as fast Fourier transforms (FFT) and short-time Fourier transforms (STFT).

Fault signals consist of different frequency components that can be analysed for further fault classification. In older fault analysis studies, Fourier transforms (FT) were used to analyse these signals. FT converts the signal in the time domain to the frequency domain, and the frequency spectrum of the constituent frequencies in the signal along with their contributions can be studied. Sinusoidal waves are used as the basis functions in FT due to them being precisely located in frequency. The transform details which frequencies constitute the original fault signal; however, this does not exactly suit the purpose of fault analysis because they are not localised in time, and hence, give no information on what time the fault occurs [13]. STFT overcomes the lack of temporal information in the FT by performing the FT in windows of the input signal. It does not, however, allow for multi-resolution analysis in time as the length of the window is fixed. This means that both high-frequency components (such as transients during a fault) and the low-frequency components (such as the signal during normal operating conditions) are treated with a window that is constant in time.

The wavelet transform enables both temporal and frequency analysis at multiple resolutions by the wavelet function. The wavelet can capture the short high-frequency components with a contracted version of a wavelet function called the "mother wavelet", and can capture the longer low-frequency components by using a dilated version of the mother wavelet. The wavelet is therefore translated and scaled to provide both time and frequency localisation [14] [11]. The comparison of resolutions between the Fourier transform, short time Fourier transform, and wavelet transform is shown in Fig. 2.2. The property that allows for the time-frequency analysis at different resolutions is useful in studying transients in the current and voltage signals. Chapter 3 explains the theory behind wavelet transforms by laying out the equations.

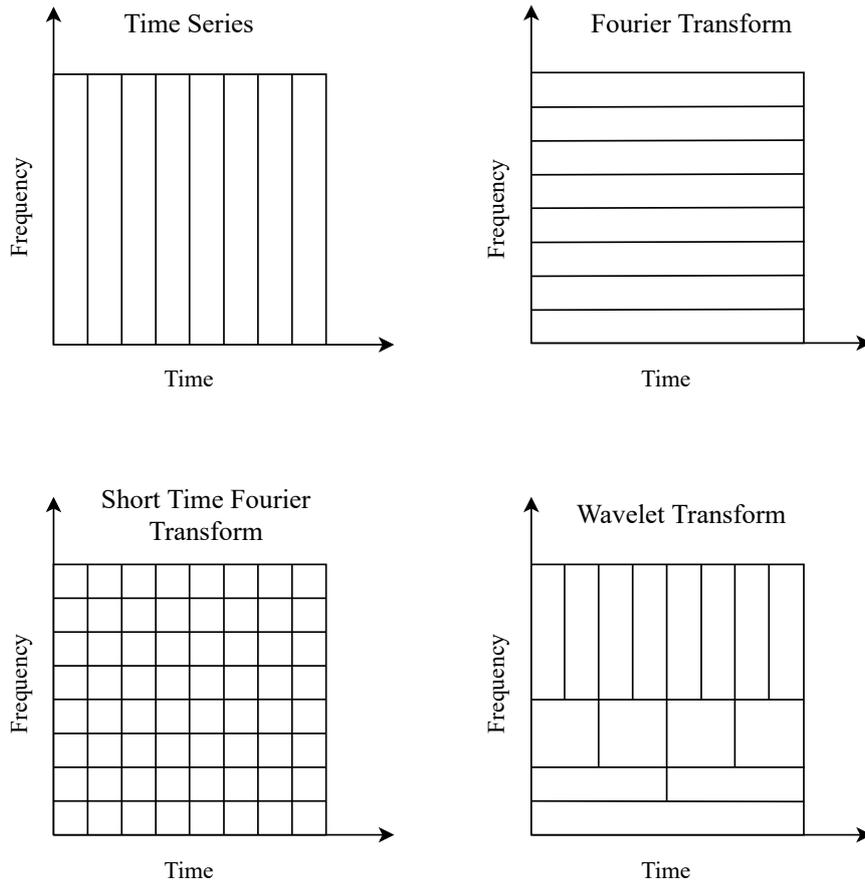


Figure 2.2: Comparison of the frequency and time resolution for different signal processing methods.

The results of performing DWT on a fault signal are coefficients containing information relating to the "shape" and the "details" of the waveform. Approximation coefficients carry information on the general "shape" of the signal and detail coefficients carry information regarding the high-frequency transients. Individual representative characteristics, or features, can then be extracted from these coefficients to assist in distinguishing faults from one another. For example, Mahanty et al. [15] consider the peak and the sum of the detail coefficients of each phase of the three-phase fault currents in the 250 - 500 Hz frequency range as input features to their classifier. Majd et al. [16], on the other hand, consider the phase-wise sum of the detail coefficients in the 625-1250 Hz range. Some other features calculated from the detail coefficients are the energy, used by Livani et al. [17], standard deviation, in the work of Li et al. [18], and RMS, by Tarlochan et al. [8].

It should be noted that the chief factor for the feature extraction from the detail coefficients is the bandwidth at which the features are calculated. To optimally capture the high-frequency

information from the faults, it is necessary to identify the appropriate level of decomposition. It is gathered from the literature review that this selection of bandwidth for feature extraction can be made with the use of expert knowledge of which frequency band carries the transient frequencies. This selection must be made with consideration for the sampling frequency of the signals.

An important facet of transient analysis using wavelets is the selection of the mother wavelet. Results of the wavelet transform differ based on the choice of the mother wavelet; an appropriate choice is especially vital for the study of a certain class of waveforms, such as current/voltage restrike behaviour that can be observed in self-extinguishing faults. A rule of thumb is to select a mother wavelet that closely resembles the signal that is being studied. It has been observed that fast transients are more effectively detected with the use of longer wavelets [19]. Kam et al. [20] analyse restrike from the operation of circuit breakers using the Daubechies 5 mother wavelet. They narrow down on the Daubechies 5 wavelet from a set of other mother wavelets including symlets, coiflets and biorthogonal wavelets by individually testing the performance of each wavelet on a transient signal. To elaborate, this is done by using each mother wavelet to decompose a sample transient signal and by analysing the detail coefficients. The detail coefficients are plotted and studied visually to see if the resulting peaks and energies correspond to the transients present in the original signal. It can be gathered that this is a time-intensive exercise that can be improved upon with modern data analytical techniques.

Review of Fault Classification Techniques

In the previous section, the technique of extracting high-frequency fault information using DWT was explained. With the case of using DWT as a strong fault signal processing tool being made, this section elaborates how features can be used in various modern classification algorithms for distinguishing between faults in the power network.

Das et al. [2] and Hessine et al. [21] use a classification system based on fuzzy logic to differentiate between single-phase, two-phase, and three-phase faults. This algorithm does not require expert knowledge in the domain of fault analysis by way of the status of the circuit breakers/isolators, or the voltage and currents in other parts of the distribution system. Additionally, the only input required is the three-phase fault current. Relationships between the three-phase currents are established with "approximations" for differentiating between the types of faults. This method is superior to one that specifies hard thresholds since the computation of the values of the latter requires a large number of fault samples, and even then, some faults currents may have measurement errors and noise [2] [21]. A drawback of methods with fuzzy logic is that it can only differentiate between stable faults and does not consider unstable or incipient faults in its study. Additionally, events like inrush currents are not in the scope of the two papers.

Classification of fault signals using neural networks are explored by Guo et al. [22] and Mahanty et al. [15]. Guo et al. [22] use Hilbert-Huang Transforms - a type of bandpass filter - for fault feature extraction and use the features as an input to a convolutional neural network (CNN) that performs the classification. On the other hand, Mahanty et al. [15] use DWT to extract fault features and use the features as input to a decision tree classifier. The decision tree algorithm is then used to classify the faults on the basis of the maximum absolute value of the sum of the details of three-phase currents and the ratio of the peak value of the details of phase/delta currents. It is found that the decision tree method performed better than the CNN by possessing a higher classification accuracy [22] [15]. Additionally, neural networks require a large number of samples and an extensive training process and are perhaps too complex to be considered for the case of power system faults. Both Guo et al. [22] and Mahanty et al [15] use PSCAD/EMTC or EMTP software to simulate the large sample space containing different fault types which is not the case for the project at hand that is limited by a relatively smaller space of real-world data.

Decision trees are used in both classification and regression problems. Each node indicates a decision that splits into two and divides the data on whether a condition is met (at the node) or not, as seen in Figure 2.3. Decision trees have good interpretability as the importance of features is clear and relations between target variables and results can be easily observed [23]. Liu et al. [24] catalogue the use of decision trees in the field of power systems for system security assessment, protection and relays, load forecasting, state estimation, and fault diagnosis. Sheng et al. [25] propose a method for fault recognition using decision trees by using only the RMS and harmonic values of the current signals with high accuracy results. Fault signals are simulated and the classification model was trained using EMTP software. The results state that the classifier has 100% accuracy. While these results indicate perfect classification ability, it should be noted that the classifier was not trained (or tested on) on real-world data.

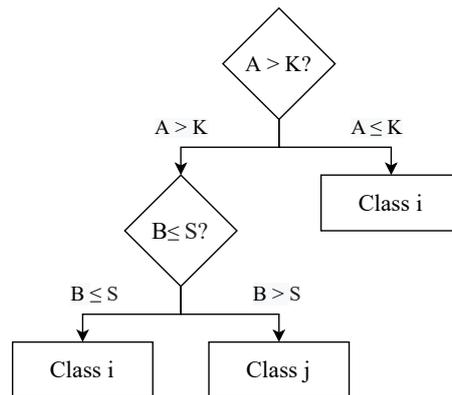


Figure 2.3: A simple decision tree classifier.

Majd et al. [16], propose a method of detection and classification using the k-Nearest Neighbours (k-NN) method. A decision on if a fault has occurred is made based on the distance between two consecutive current signal samples and comparing that with a threshold value. Classification is performed by comparing each sample with its 5th nearest (sample) neighbour and comparing this distance with a threshold. Each sample is a snapshot of the phase current. This classification method uses a decision tree algorithm and is performed for each phase. This method is hence capable of classifying multi-phase faults. The algorithm does not use waveform signatures, but only the distance between samples, which may lead to misleading results.

Thus far, most of the classification techniques in this review use pre-calculated thresholds for the features in the original signal to check if a fault has occurred or not. It can be interesting to see how modern data analytics tools, such as machine learning (ML), can help with the classification problem at hand. Supervised learning (SL) (a type of ML) methods make predictions by learning the relationships between the data points and their features through training a function on labelled data [23] [26]. This literature study proceeds with the critical review of papers that use ML to perform fault recognition.

A procedure to classify faults based on DWT called Wavelet Detection Method (WDM) is described by Guo et al. in [27]. The detail coefficients at the 5th level of decomposition of a phase current signal, i.e., the 125Hz-250Hz frequency band, in this case, are compared with predetermined thresholds for three consecutive sampling points. Livani et al. [17] use DWT to extract features from the fault voltage and current signals. A support vector machine (SVM) is utilised for the classification process. SVMs determine the best boundary between two classes by finding a separating hyper-plane that has the most distance between the support vectors [28]. The theory behind support vector machines is explained in Chapter 3. ATP software was used to simulate the transients, and the classification method was tested using MATLAB.

Magagula et al. in papers [29] and [30] extract the high frequency components (the detail coefficients of the DWT) from the fault voltage and current signals. These features are used as

inputs to a multi-class SVM classifier, and the classifier hyper-parameters are selected by performing sequential minimal optimisation. The detail features are extracted using the Daubechies 7 (db7) mother wavelet filter. The mother wavelet is selected after comparing the performance of the db7 wavelet with the Daubechies 4 (db4) and Daubechies 6 (db6) wavelets. This paper uses a similar SVM classification technique as [17] to detect all the phase and ground faults.

Abdelgayed et al. [31], propose a method of semi-supervised learning to classify faults in power systems. The classification algorithm operates by co-training two types of classifiers - a decision tree and k-NN - on labelled data and testing on a subset (without replacement) of the unlabelled data. This results in probability estimates for each class the data point might belong to. A resulting feature of the DWT on the fault current and voltage waveforms is the energy of the detail coefficients. This feature is used as input to the classifiers. The labelled data set is then updated by adding data points that have the largest difference in probability estimates between classes - with the class having the greater probability being the new label for the previously unlabelled data [26]. In this manner, all the unlabeled data points are added to the labeled set, and the algorithm is cross-validated on these values before testing. The mother wavelet is chosen through the harmony search optimisation technique [31].

Li et al. [18] use feature extraction by way of DWT and an SVM to detect and classify short circuit faults in low voltage AC systems. The hyper-parameters of the classifier are optimised using a black hole particle optimisation technique (PSO) that uses the concept of the physics behind black holes to improve the searching ability of particles. Black hole PSO accelerates the search process by introducing black hole regions where the particle has a certain probability of clustering into or escaping. The classifier is implemented as a multi-level SVM to classify single-phase-to-ground, two-phase-to-ground, three-phase-to-ground, and phase-to-phase faults. This method appears to have good results stating 99.86% classification accuracy.

Classification of Incipient Faults and Detection of Unstable Behaviour

It can be observed from the literature review that most research focuses on detecting and classifying the standard phase-to-phase and phase-to-ground faults and not incipient or unstable faults. The scope of the literature study was expanded to search for papers that deal with these events; while the number of papers (besides the Alliander research that was described earlier in Section 2.2) were not many, the methodologies in the following research provide a summary of the work done in this field and help to identify the scope for implementation and improvement for this thesis.

Tarlochan et al., [8], use two methods to detect and classify faults - DWT and superimposed sequence components. The latter involves calculations based on the negative and zero-sequence currents and comparing with thresholds. The DWT method extracts details from the high-frequency transients that are used for fault detection. It was found that the DWT method showed better detection and classification accuracies. This paper, however, does not consider other regular faults such as phase-to-phase and phase-to-ground faults in its scope and focuses only on incipient faults. The concept of using human-level concept learning (HLCL) for classifying incipient faults is explored by Xiong et al. [32]. In this method, signal analysis is inspired by human perception (that can easily make out fault signals and transients) and hierarchical probabilistic learning. Additionally, as stated by Xiong et al. [32] in their research on incipient faults, they differ from regular faults in the RMS of their approximation coefficients and energy of their detail coefficients. HLCL performs well for small sample problems - such as 50 samples.

A two-step strategy based on incipient fault detection and classification is described by Jan-nati et al. in [33]. First, the algorithm detects the transient occurrence using the CUMulative-SUM (CUSUM) test — which is a technique that can detect abrupt changes. Next, the classification of the type of incipient is implemented using ADaptive LInear NEuron (ADALINE). The

ADALINE classifier, a type of neural network, was trained to classify faults and was found to be able to quickly and precisely extract harmonics in the signal.

The problem of identifying unstable faults is a classification problem in name but is in effect, an operational problem. Unstable faults can be single-,two-, or three-phase faults and have highly distorted waveforms. Faults whose single-,two-, or three-phase nature cannot be identified with reasonable accuracy are classified as unstable for further manual inspection of fault type and location. From the review of literature, it was observed that the faults considered for the study were simulated, and hence did not consist of the distortions that lend to the non-sinusoidal nature of faults that make them unstable, by Alliander’s definition [7]. Consider Figure 2.4, where a real-world and synthetic two-phase fault are compared. It can be seen that there are more distortions present in the initial cycles of the real-world fault waveform.

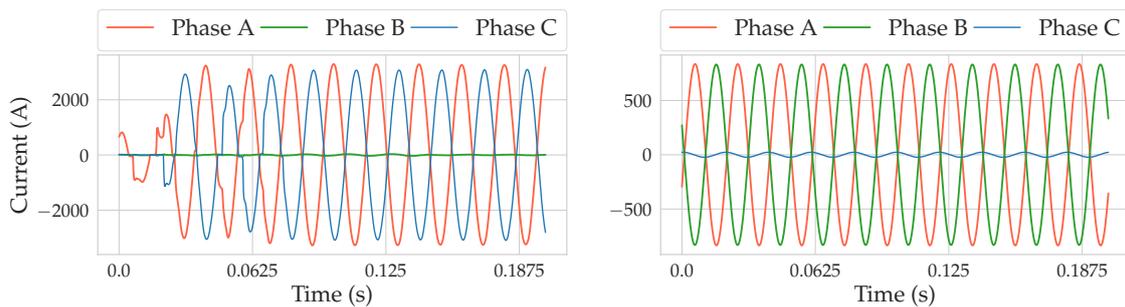


Figure 2.4: A comparison between a real-world two-phase fault (left) and a synthetic two-phase fault (right).

Having summarised the research in the field of fault analysis and classification, the following section discusses the merits, shortcomings, and gaps in the research.

2.4 Discussion

This literature review attempted to cover a considerable number of research papers and studies on fault analysis to assess the state-of-the-art methods and algorithms in the fault classification field. The papers that were studied as part of the literature review were found through online searches on Google Scholar and IEEE Xplore. Papers were chosen on the basis of how recently they were published, how many citations they possessed, and the relevance of their contents to the research questions. The goal was to study the types of fault feature extraction techniques, classification methods, and how faults in the distribution network can be distinguished from each other with modern signal processing.

The selection of the most suitable mother wavelet is essential for extracting useful features from fault waveforms. Table 2.2 summarises the different and most popular mother wavelets that have been used for this purpose. It can be observed that the Daubechies wavelets are the most commonly used, followed by the cubic b-spline and symlet wavelets. It was, however, found that it is difficult to determine the ideal wavelet for a study *a priori*, and this aspect is an interesting avenue for further investigation.

A common drawback of published classification methods, are that the papers use software (such as MATLAB/SIMULINK/ETAP/SPSS) to simulate the electromagnetic transients — for example, papers [17], [16], and [33] use simulated faults. These would be highly “ideal” transient signals that do not accurately represent the unpredictable nature of real-world fault signals. An interesting part of the study would hence also be to see how well the methods prescribed

Table 2.2: Summary of research papers that use DWT for fault feature extraction and the corresponding mother wavelet used.

Mother Wavelet	Papers
Cubic B-Spline	[18][27]
Daubechies-4	[8][32][17][34]
Daubechies-7	[30]
Daubechies-8	[14]
Symlet-4	[29]

in the paper compare with real-world values and signals. It must also be investigated if a suitable algorithm can be developed for the detection of unstable behaviour in distribution network faults.

There are many modern classification techniques, and a critical aspect of this thesis is the selection of an appropriate classification model. Table 2.3 presents a summary of the research papers covered in this literature review and the classification techniques they each use. It can be observed that SVMs are the most popular method of fault classification, followed by ANNs and decision trees. In addition to being the most accepted technique for the fault classification problem, SVMs strike a balance between the high interpretability of decision trees that neural networks lack, and the superior processing power of neural networks that decision trees cannot compare to [28]. A decision is therefore made to use support vector machines for the fault classification problem at hand.

Table 2.3: Summary of research papers that utilise different types of machine learning classification techniques.

Classification Techniques	Papers
Fuzzy Logic	[2][21]
k-Nearest Neighbour	[16][31]
Decision Tree	[31][24][25][15]
Support Vector Machine	[29][30][18][17] [28][32] [12]
Artificial Neural Networks	[22][15][33][28] [13]

2.5 Conclusion

Section 2.2 introduced the current state of the classification system in DSO Alliander. To study if recent developments in data processing can improve fault analysis, research methodologies for fault feature extraction and classification techniques were critically reviewed in Section 2.3. It became clear that the support vector machine classification technique would be a good choice for this thesis' fault study as evidenced in Table 2.3. Furthermore, the DWT was identified as a promising signal processing technique for extracting high-frequency information from fault waveforms. Section 2.4 discussed the main takeaways from the literature review and highlighted the gaps in research that could be studied in this thesis. Finally, the literature review also helps put the research questions from the Introduction in Section 1.4 into context.

Can supervised learning improve the performance of the classifier on real-world fault data? Fault classification algorithms have evolved into using machine learning techniques for the classification of electrical faults. It has been found through the literature survey that these data-driven techniques are increasingly common and perform robustly for simulated faults.

What are existing machine learning approaches for detection and classification of faults? It was found that support vector machines, among decision trees and neural networks are a popular choice for the classification model. A starting point for a for this study can hence be the SVM as discussed in section 2.4.

What are the features that are required to distinguish faults from one another? The db4 wavelet has been identified as a suitable filter for the DWT of fault signals, among other wavelets. It is important to validate that this is the best choice for fault classification in the distribution network. Another interesting aspect of this thesis is to study if DWT can be used to extract differentiating features from real-world faults, or if simpler features are better in capturing the differences. To enable a clear understanding of support vector machines and discrete wavelet transforms, along with supervised learning performance metrics, the theoretical background for this thesis is laid out in the following chapter, Chapter 3.

3 Theoretical Background

This chapter introduces the theoretical background for various concepts used in the thesis. This section is broadly divided into three parts that pertain to different stages of the classification process identified in Chapter 2.2 — feature extraction using wavelets, classification using supervised learning techniques, and classifier performance evaluation. Section 3.1 describes the working of the discrete wavelet transform with a focus on the how the wavelet filters enable signal decomposition for feature extraction. Section 3.2 describes the process behind supervised learning classification techniques and elaborates on the working of support vector machines. Section 3.3 explains the statistical tests used to facilitate the selection of the most appropriate features for a classification model. This chapter concludes with Section 3.4 that presents suitable metrics for evaluating classifier performance.

3.1 Signal Decomposition using Discrete Wavelet Transforms

In this section, the theory behind discrete wavelet transforms (DWT) is discussed. A wavelet is a function that is localised in time and integrates to zero. The wavelet analysis is used for studying the physical state of signals - such as sudden disruptions, sharp spikes, and transients - by splitting the signal into different bandwidths. This enables the study of the different frequency components in the frequency levels at which disturbances occur [8].

Wavelets can be used to decompose a signal into their high-frequency or "detail" components and low-frequency or "approximation" components. Wavelet transforms were originally developed for studying images at different resolutions without losing valuable temporal information. A multi-resolution analysis (MRA) of images allows for easy interpretation of information. Different resolutions depict different aspects of the image. At coarser resolutions, one can observe the "background" of the image and at finer resolutions, one can observe the more intricate details. MRA enables a scale-invariant interpretation of a signal, i.e. the most pertinent components of the signal can be studied on the same scale as the original data. This concept of studying images at different resolutions can be extended to signal processing with the aim of observing transient behaviour. At lower frequencies, the steady-state 50 Hz component of the signal can be observed, and at higher frequencies, the transient behaviour can be observed. This is done by creating a hierarchy of bandwidths where, at each frequency band, the behaviour of the signal in time can be observed and studied. [35]

There are two types of wavelet transforms - the continuous and discrete wavelet transform (abbreviated as CWT and DWT, respectively). It was gathered from the literature review, in Section 2.3, that DWT is more popular for the purpose of transient analysis. DWT is more commonly used for applied signal processing as it is computationally less expensive, more compact, and faster than the CWT. The wavelets are implemented in the DWT process with cascading filter banks. Each low-pass filter, or scaling function, preserves the time domain properties in the form of approximation coefficients. The high-pass filter, or wavelet function, captures the

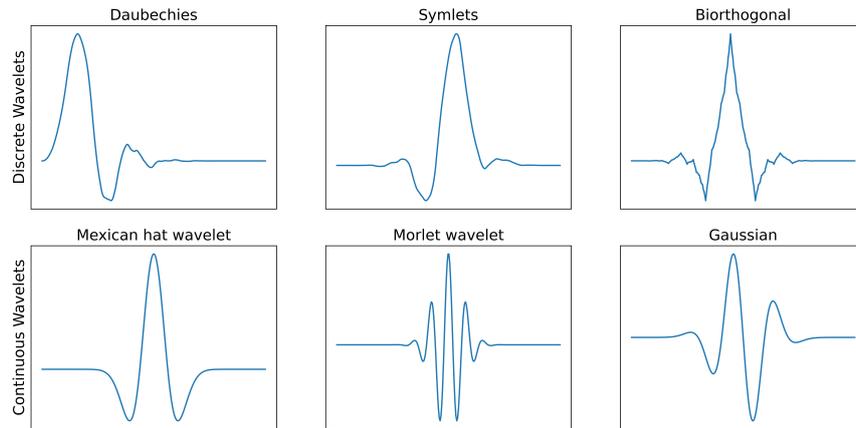


Figure 3.1: Different families of discrete (top) and continuous wavelets (bottom)

high-frequency components such as transients. The scaling function and the wavelet function are orthonormal - i.e. the integral of the product of the two functions equals to zero [35]. The result of each level of decomposition arises from the convolution of the filter with the signal under analysis.

There are different families of wavelets that are used for continuous and discrete wavelet transforms as shown in Figure 3.1. The wavelet function, or mother wavelet, is shifted, scaled, and convolved with different subsequent intervals of the input signal so as to capture the transients in time. The mother wavelet function should be short, oscillatory, and integrate and average to zero [11].

The equation for the continuous wavelet transform is shown in Equation 3.1. The input signal under analysis is $x(t)$. The wavelet function is $g(t)$ is shifted by a factor b and scaled by a factor a . $g\left(\frac{t-b}{a}\right)$ is hence scaled in time, i.e. stretched or contracted by a factor of a . The function is stretched or expanded in time if $a > 1$. This decreases in the frequency of oscillations of the wavelet function. The function is contracted in time if $a < 1$, which increases the frequency of oscillations of the wavelet function [11]. The contracted wavelet can hence capture the high-frequency oscillations of the input signal, and the expanded wavelet can capture the overall "shape" of the input signal. Robertson et al. [11] explain the derivation of the filter bank implementation of the DWT as follows, in Equations 3.1, 3.2, 3.3, and 3.4.

$$WT(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) g\left(\frac{t-b}{a}\right) dt \quad (3.1)$$

The digital realisation of the CWT is the DWT which can be defined by the Equation 3.2. $g[n]$ is the mother wavelet, and the scaling and translation parameters a and b are a function of (integer) parameter m , where $a = a_0^m$ and $b = na_0^m$.

$$DWT(m, k) = \frac{1}{\sqrt{a_0^m}} \sum_n x[n] g\left[\frac{k - na_0^m}{a_0^m}\right] \quad (3.2)$$

On simplification, this results in Equation 3.3,

$$DWT[m, k] = \frac{1}{\sqrt{a_0^m}} \sum_n x[n] g[ka_0^{-m} - n] \quad (3.3)$$

Equation 3.3 can be rewritten as shown in 3.4 by swapping k and x to represent the wavelet transform implemented in the form of a filter bank. It can be observed that this equation holds a similarity to the general equation of finite impulse response (FIR) digital filters shown in Equation. 3.5.

$$DWT[m, n] = \frac{1}{\sqrt{a_0^m}} \sum_k x[k] g[na_0^{-m} - k] \quad (3.4)$$

$$y(n) = \frac{1}{c} \sum_k x(k) y(n - k) \quad (3.5)$$

Through this similarity it can be inferred that $g[na_0^{-m} - k]$ is an FIR of a digital (low-pass) filter [35]. To simplify computations, Crowley et al. [36] have introduced a pyramid type implementation of the calculations at the different levels. This is done by selecting $a_0 = 2$ and $m = 1, 2, 3, \dots$

Similarly, a high-pass implementation of Equation 3.2 is constructed by using a dual, $h(n)$, of the low-pass filter $g(n)$. These two filters are chosen such that they form an orthogonal basis. The coefficient $\sqrt{2^{-m}}$ normalises the two functions. Functions $g(n)$ and $h(n)$ are filters that can be represented in terms of their coefficients. Table 3.1 shows the filter coefficients of the db4 wavelet as an example, with Figure 3.2 depicting the function in its continuous and discrete form. The coefficients of the high-pass filter is a rearranged form of the low-pass filter with negation in every other coefficient [11].

Table 3.1: Filter coefficients for the Daubechies-4 low-pass and high-pass wavelet function

Filter Type	Filter Coefficients							
Low-Pass	-0.011	0.033	0.031	-0.187	-0.028	0.631	0.715	0.23
High-Pass	-0.23	0.715	-0.631	-0.028	0.187	0.031	-0.033	-0.011

As convolution is a commutative operation, the equation can also be interpreted as seen in Equation 3.6 which represents a single pass of the filter. Functions $x(k)$ and $g(a_0^{-m}n - k)$ are interchanged, and $a_0 = 2$. This equation can now be interpreted as: for each value of m in $m = 0, 1, 2, 3, \dots$, the signal bandwidth is halved and passed through a high-pass and low-pass filter. The output of the high-pass filter can be passed again through the same high-pass filter, by repeatedly using Equation 3.6, until the desired bandwidths are obtained.

$$DWT(m, n) = \frac{1}{\sqrt{2^m}} \sum_{k=-\infty}^{+\infty} g(k) x(2^{-m}n - k) \quad (3.6)$$

This can be visualised as the filtering of signal X using filter $g(k)$. Figure 3.3 depicts the implementation of DWT in the form of cascading high-pass and low-pass filters up to the 3rd level of decomposition. The input signal $S_{2^0} f$ is fed into high-pass filter H0 and low-pass filter G0. H0 filters out signal frequencies below one half of the maximum frequency f_{max} of the input signal and the G0 filters out signal frequencies greater than one-half of f_{max} . The outputs of the two filters are subsampled by a factor of 2 — i.e., every other sample is discarded in order

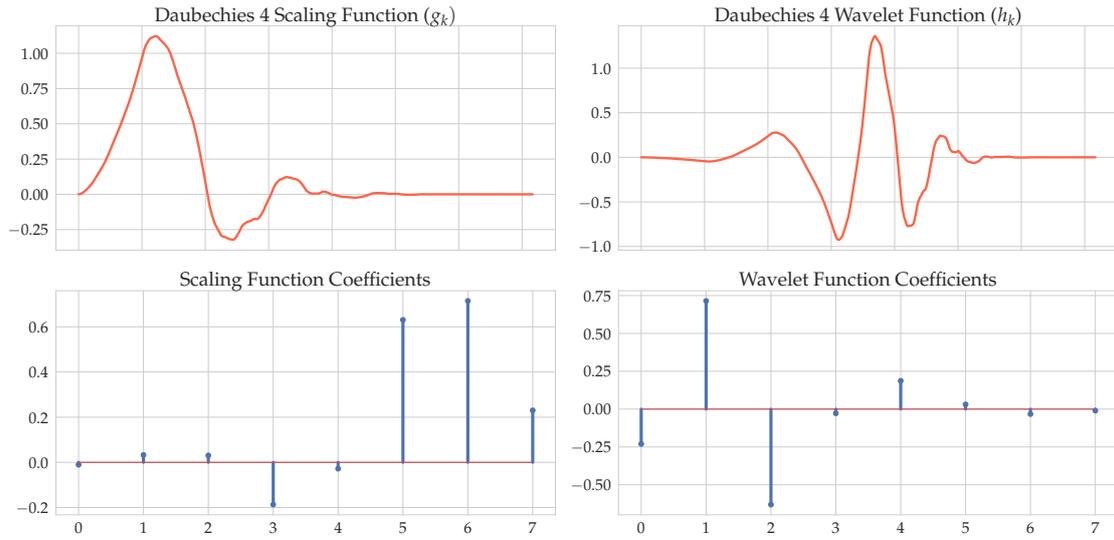


Figure 3.2: Plot depicting the low-pass scaling and high-pass wavelet filter for the daubechies 4 wavelet.

to effectively double the scale of the signal for the subsequent stage. Equation 3.7 describes the high-pass and low-pass filter process.

The coefficients of the high-pass and low-pass filters are h_k and g_k , respectively. S_{2^m} is the low-frequency approximation component, and W_{2^m} is the high-frequency detail component at the decomposition scale m . For the purpose of fault classification, more focus is given to the detail coefficients as they can capture useful features from high-frequency transient behaviour in each sub-band [37] [8].

Figure 3.2 shows the graph of the Daubechies 4 wavelet’s implementation as a high-pass wavelet filter and low-pass scaling filter. Each (low- and high-pass) filter is convolved with the input signal to produce the detail or approximation components. Subsequently, this is subsampled by a factor of 2.

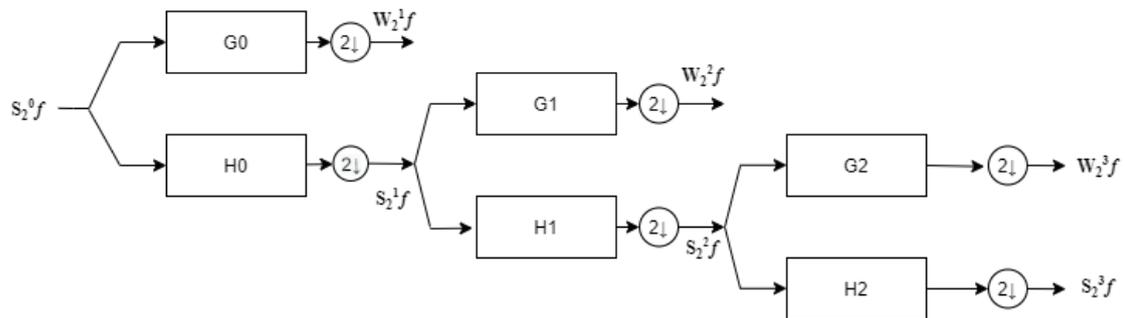


Figure 3.3: Diagram of wavelet decomposition using cascading filter banks.

$$\begin{aligned}
 S_{2^m} f(n) &= \frac{1}{\sqrt{2^m}} \sum_k h_k S_{2^{m-1}} f(2^{-m}n - k) \\
 W_{2^m} f(n) &= \frac{1}{\sqrt{2^m}} \sum_k g_k S_{2^{m-1}} f(2^{-m}n - k)
 \end{aligned}
 , m > 0 \tag{3.7}$$

The signals from the SASensor are sampled at 4 kHz. Since the Nyquist-Shannon sampling theorem determines that the sampling frequency f_s should be at least two times greater than f_{max} , the f_{max} of the signals from the sensor is at most 2 kHz. Table 3.2 lists the bandwidths of the output of the high-pass filters at each level of decomposition. Note that the fundamental 50 Hz frequency is present in the sixth level of decomposition and the transients are present in the 1000 - 2000 Hz frequency range, and the harmonics are present at frequencies above 250 Hz. The applications of these formulas are elaborated further in Chapter 4.

Table 3.2: Bandwidths of a signal sampled at 4kHz after the wavelet transform at each level of the filter functions.

Decomposition Level	Bandwidth of the Detail Coefficients (Hz)
None (Original Signal)	0 - 2000
1	1000 - 2000
2	500 - 1000
3	250 - 500
4	125 - 250
5	62.5 - 125
6	31.25 - 62.5

3.2 Supervised Learning for Classification

The problem of classification is to identify the category that a set of observations belong to. In this thesis, current and voltage measurements of fault signals are classified to predict their fault type. Classification problems can be either supervised or unsupervised. With each observation of the predictor values $x_i, i = 1, \dots, n$, there is a corresponding response y_i . The goal of a classifier is to fit a function or model that can accurately predict the response for future observations. Unsupervised learning, on the other hand, is when each observation $i = 1, \dots, n$ has a measurement of x_i , but no associated response y_i . This is referred to as unsupervised learning as no outcome variable can supervise the analysis. The fault classification problem in this thesis is a supervised learning problem, as, for each measurement of the three-phase currents and voltage, there exists a response variable, i.e., the type of fault.

The classification consists of sets of training observations $(x_i, y_i), \dots, (x_n, y_n)$ that can be used to fit a classification model. The model must not only have a good performance on the training data, but also on test data, i.e. observations that were not used for the classification training. The following section provides an overview of the a popular classifier (SVM) used for the classification of faults in the distribution network [23].

3.2.1 Overview of the Classification Process

For N training samples $(X = x_1, x_2, \dots, x_n)$, where $x_i \in \mathbf{R}^p$ is the i th sample, each sample has p dimensions or features, each training sample belongs to a class with labels $y_i \in 1, 2, \dots, c$. The training data set is then $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, where the labels for x_1, x_2, \dots, x_n are y_1, y_2, \dots, y_n respectively. While training a classifier, the model attempts to learn the relationship between the observations (X) and their labels (Y) [38]. This process is applied to the training set alone. The model obtained from this process is tested on unseen data (that was not a part of the training set) to check how well the fitted model performs on data that it was not trained on — or, how well it is able to generalise with respect to the test data. In the case of classification using supervised learning, the trained model classifies the test data with their labels removed so

the true class information is “hidden” from the model. After the classification, it is possible to compare the actual class and predicted class of each sample. The following section elaborates on the performance metrics that can be derived from correct and incorrect classifications.

The model must also strike a balance between bias and variance. The goal of classification development is to create a generalised model that maximises the accuracy of future classifications of (unseen) data from the same source. A model that fits very closely to the training data has a propensity to generalise poorly on future test data as the model fits noisy or random aspects of the samples in addition to the differentiating features. The more a model learns the smaller irregularities of a sample data set, the less likely it is to catch the larger trends present in the data set. This is called overfitting, and the opposite — generating a model fit that is too general and does not learn the legitimate trends of the data — is called underfitting [39].

An effective classification model must hence be flexible enough that it catches the broader regularities of a data set but must also be capable of learning the smaller distinguishing features without becoming too focused on the noisy aspects of the data, mistaking it for trends. This trade-off between the bias (former) and the variance (latter) is useful to keep in mind while calibrating the classifier model. A sign of overfitting is if the classification scoring metrics, such as accuracy, are high with the training data set, but significantly lower with the testing data set. This is also the case with the selection of the number of features which will be elaborated in Chapter 6.

3.2.2 Support Vector Machines

Support vector machines find the optimal separating hyperplane between two classes of observations. SVMs can be used for linearly and non-linearly separable classes. This section focuses on the use of SVMs for non-linearly separable data.

In a p -dimensional space, a hyperplane is a flat subspace of $p - 1$ dimensions. For example, in a two-dimensional space, a hyperplane is a line, and in a 3-dimensional space, a hyperplane is a plane. This is explained in Figure 3.4 for the case of two linearly separable classes. Often, observations in a classification setting need not always be linearly separable by a flat hyperplane. In this case, fitting a perfect margin between two classes can result in a loss of sensitivity, as a single observation can then change the positioning of the margin [23].

A support vector machine (SVM) uses a hyperplane that does not separate two classes perfectly but permits misclassifications of a few observations from the training data in order to preserve the bias and perform well on the classification of the rest of the observations. Additionally, in the case of classes with non-linear boundaries between them, a linear classifying margin will perform poorly. This problem of non-linear boundaries is overcome in SVMs by expanding the feature space using quadratic, cubic, or high-order polynomial functions of predicting variables [38].

In the case of non-separable data, kernel functions are utilised to transform the data into a higher-dimensional space where the classes then consist of linear boundaries. In order to ensure the trade-off between bias and variance is not skewed, some samples will be misclassified to ensure the overall general applicability of the classifier [38].

Figure 3.4 shows how two (linearly separable) classes can be classified using support vector machines. If two classes are linearly separable, the decision boundary between the classes can be described by the line $w^T x + b = 0$. \mathbf{w} is the weight vector, b is the bias and \mathbf{x} is the training sample data. The label vectors for two classes are $y_i \in -1, +1$. The SVM attempts to construct planes H_1 and H_2 to form a hyper-plane that is as far as possible from the closest samples. In the case of the data being non-separable, misclassifications are unavoidable. The hyperplane that separates the

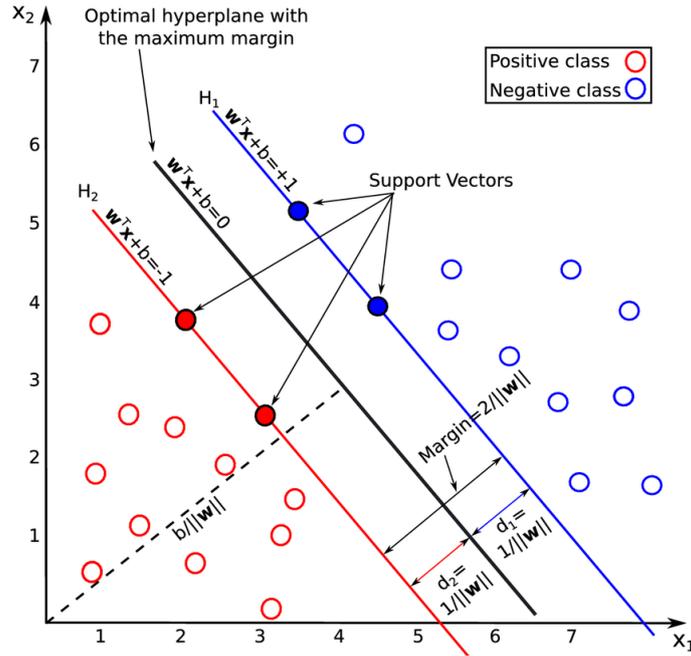


Figure 3.4: Diagram depicting the separating hyperplanes in a two-class classification problem using support vectors [38].

two classes is:

$$w^T x_i + b = 0 \tag{3.8}$$

With the following equations being the two planes on either side of the hyperplane:

$$\begin{aligned} w^T x_i + b &\geq +1 - \epsilon_i \text{ for } y_i = +1 \text{ (hyperplane for the positive class)} \\ w^T x_i + b &\geq -1 + \epsilon_i \text{ for } y_i = -1 \text{ (hyperplane for the negative class)} \end{aligned} \tag{3.9}$$

The combined equation is hence:

$$y_i(w^T x_i + b) - 1 + \epsilon_i \geq \forall i = 0, 1, 2, \dots, n \tag{3.10}$$

A slack variable $\epsilon_i \geq 0$ is introduced to relax the constraints in case the margin between the classes is non-separable. If ϵ is between 0 and 1, the sample is correctly classified between the margin and the hyperplane of the correct side. If $\epsilon_i > 1$, the decision function $w^T x_i + b$ and the label have different values and hence imply that the sample has not been correctly classified. The optimisation problem can be expressed as below [40]:

$$\begin{aligned} \min \frac{1}{2} \|x\|^2 + C \sum_{i=1}^n \epsilon_i \\ \text{such that } y_i(w^T x_i + b) - 1 + \epsilon_i \geq \forall i = 1, 2, \dots, n \end{aligned} \tag{3.11}$$

$\frac{2}{|w|}$ is the margin between the two planes and must be maximised. This is shown in the first

term of equation 3.11. The second term, $C \sum_{i=1}^n \epsilon_i$ is the penalty for classifications where ϵ_i is the error vector and C is the penalty factor for misclassifications [18].

If two classes are non-linearly separable, the data points are mapped onto a higher dimensional space using a kernel function K where the data can be linearly separable wherein ϕ is embedded to implicitly map the vectors. Kernel functions are defined as the scalar product of two non-linear functions as shown in Equation 3.12. For fault classification, the Gaussian Radial Basis (RBF) kernel function is used, where $k(x, y)$ represents the RBF kernel. Equation 3.13 describes the function [17]:

$$K(x, y) = \phi(x)^T \phi(y), \text{ where } x \text{ and } y \text{ are two input vectors} \quad (3.12)$$

$$k(x, y) = \exp\left(-\frac{\|x - y\|^2}{\gamma}\right) \quad (3.13)$$

The parameter C from Equation 3.11 and γ from Equation 3.13 are considered the hyper-parameters of the support vector machine that cannot be determined *a priori* but through a process called hyper-parameter tuning that is discussed in Chapter 6. Parameter C is the trade-off factor of the decision function margin against the correct classification of the training data sample. C is directly proportional to the bias and inversely proportional to the variance of the classifier. C hence acts as a regularisation mechanism for the SVM. Parameter γ determines the range of influence of a single training sample to classify the future samples. The value of γ is directly proportional to the range of influence of the separating hyperplane.

3.3 Statistical Tests for Determining Feature Importance

This section introduces the statistical tests used for ranking features and the theory behind the tests.

ANOVA-F test

The ANOVA (ANalysis Of VAriance) is a procedure that checks if the different classes have means that are statistically different from each other based on the null hypothesis that there are no significant differences between the means. The statistical significance of the ANOVA is tested with the F-statistic. In the context of feature selection, this is used to quantify how much the features are related to the target variable for the classification [41]. The equation for the ANOVA-F test that compares two variances can be seen in Equation 3.14

$$F = \frac{\text{Variance between Sample Means}}{\text{Variance within Samples}} = \frac{SS_{\text{Samples}}/(I - 1)}{SS_{\text{Error}}/(n_S - I)} \quad (3.14)$$

SS_{Samples} is the sum of squares based on the total deviations of the samples from the grand mean. SS_{Error} is the sum of squares based on the deviation of the sample mean from the grand mean. n_S and I represent the total number of cases and the number of samples, respectively.

The numerator hence represents, for a feature, the distance between the classes. The value represents how spread out the samples are within each class. For a feature to be a good discriminator between classes, the value of the numerator, i.e. the distance between the samples *between* classes should be high, and the denominator should be low. Hence, the F-score is directly proportional to how well a feature separates classes for a given sample set. Features can then be ranked by importance on the basis of their F-scores.

Mutual Information Method

Mutual information is a statistical test from the field of information theory for measuring the interdependence of two random variables. Similar to the ANOVA-F test, mutual information can be used to rank the importance of features in terms of how much they contribute to the classification into the target class. The mutual information, or information gain, is a quantification of how much information one variable, or in this case data set, contains about the other [42]. The formula for mutual information is seen in Equation 3.15. $p(x,y)$ is the joint probability mass function of two data sets X and Y , and p_x and p_y are the marginal probability mass functions of X and Y , respectively [43].

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (3.15)$$

In the context of the classification problem, the mutual information I is calculated between the features and the target - i.e. the class labels in accordance with Formula 3.15. The higher the I for the feature, the higher the dependence of the class label on the particular feature. The features can hence be ranked by importance on the basis of the magnitudes of the calculated I .

It is important to mention that these two statistical tests calculate the importance score per feature and relate the feature to its individual discriminating ability. The importance scores can hence determine which features share the strongest relationships with the classes but not the relationships between the features themselves. For this purpose, sequential feature selection, a greedy technique for feature selection is implemented in Chapter 6. Additionally, while these statistical tests are useful for ranking the importance of discriminating features, they do not indicate how many features are required to model a classifier that performs well. Hence, the next step is to identify which subset of features from the ranked feature-set yields the optimal performance - this feature selection process is elaborated in Section 6.2.1.

3.4 Metrics for Evaluating Classifier Model Performance

The metrics used to evaluate the performance of the classifiers are introduced and defined in this section. This thesis uses the metrics precision, recall, and the f1-score for classifier performance analysis. Figure 3.5 represents an error matrix that facilitates the understanding of these metrics. Each column represents the samples in a predicted class and each row represents the samples in the actual class. This error matrix represents a two-class classification problem where "positive" and "negative" indicate the properties of two classes; i.e., a sample can be classified as either positive or negative. This section concludes with how the metrics from this error matrix can be extended to a multi-class problem, such as the classification problem in this thesis.

True positives indicate samples that are actually from the positive class and have been classified correctly as being from the positive class. True negatives indicate samples from the negative

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positives	False Positives
	Negative	False Negatives	True Negatives

Figure 3.5: Error matrix that enables the visualisation of the performance of a supervised learning classification problem.

class that have been classified as being from the same. True positives and negatives hence represent the samples that have been classified correctly. False positives represent samples from the positive class that have been misclassified in the negative class, with false negatives representing the opposite. From here, definitions for the performance metrics can be drawn.

Recall, or sensitivity, is the ratio of actual positive samples that have been correctly classified as being positive. This metric calculates the fraction of the relevant samples that were successfully classified. It should be noted that the recall alone is not a meaningful metric to measure classifier performance as it only pertains to the information retrieval of one (the positive) class [44]. The equation for the recall metric can be seen in Equation 3.16.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3.16)$$

Precision, or confidence, conversely represents the ratio of the positively classified instances that are actually from the positive class. The equation for precision can be seen in Equation 3.17.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (3.17)$$

Precision and recall each do not represent the performance of a classifier as a whole as they are only focused on the "positive class", and do not capture how the classifier treats the other, in this case equally important, (negative) class. The **f1-score** is a metric that combines the two metrics by calculating the weighted average of precision and recall as shown in Equation 3.18. The f1-score is a robust metric for the evaluation of a classifier that deals with imbalanced classes.

$$\text{F1-Score} = 2 \cdot \frac{\text{Recall} \cdot \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3.18)$$

The above metrics can be extended for a multi-class problem in a One-vs-Rest evaluation approach where the classification performance for each class is evaluated against the remaining classes. The multi-class problem is converted to a combination of binary classification problems.

3.4. Metrics for Evaluating Classifier Model Performance

The classification performance of each class is evaluated by treating the target class as the positive class and the other classes together as the negative class. These smaller tasks result in a precision, recall, and f1-score for each classification. The weighted average is further calculated (as a function of the samples in each class) for each metric so it is representative of the performance for all the classes together. Unless it is mentioned otherwise, it is the weighted average that is being referred to by the terms precision, recall, and f1-score in the subsequent chapters.

4 Development of Features for Stable and Unstable Faults

To develop the classifier algorithm it is important to first develop an input feature set that represents different aspects of the data that is being studied. Features are identifiable packets of data functionality that are useful to characterize the data from the user's perspective. This section introduces the description of the fault voltage and current signal data with examples for each type of fault. Section 4.1 describes the motivation and steps for processing the input data before suitable features can be extracted. This step is predominantly facilitated by a binary segmentation algorithm. Section 4.2 describes the features that are developed from the fault waveforms for stable faults. Additionally, the reasoning is provided as to why each feature is selected, and what aspect of the data is expected to be represented by the feature. Section 4.3 introduces unstable faults and the strategy for developing features for such faults.

4.1 Description of the Sample Space

The input for the classifier is the three-phase voltage and current waveforms. These signals are sampled by the SASensor at a rate of 4 kHz and are recorded from 1s prior to the instant an event occurs. An event is assumed to have occurred if the currents cross a preset threshold that is calculated based on the thermal restrictions of the components [4]. The frequency of the steady-state current and voltage signals is 50 Hz. As the sampling rate is 4 kHz, the number of samples per cycle is $(4000/50) = 80$ samples/cycle. These sampled waveforms are then recorded and saved as Digital Fault Recordings (DFR) files.

Fig. 4.1 shows an example of how faults are recorded by the sensor. The recorded waveform can be observed in three sections. The pre-fault steady-state exists in the time period from $t = 0$ s to the beginning of the event at $t = 1$ s. The two-phase-to-ground fault starts at 1s and extinguishes at 1.55s. The post-fault state occurs after the fault extinguishes at 1.55s up to 1s post extinction at $t = 2.5$ s. It is important to note that while the pre-fault state always lasts 1s, the post fault state duration that is recorded is not strictly 1s. For instance, consider Figure 4.3 where the fault is recorded for around 1.5s post fault extinction. Furthermore, in some cases, faults can evolve from one type of fault into a different fault. For example, Figure 4.3 depicts a single-phase fault that progresses into a two-phase fault, and finally, a three-phase fault.

Another example is a two-phase fault that evolves into a three-phase fault as shown in Figure 4.4. In these cases, the fault is labelled with the type of fault in its final stage before extinction. This is because the calculation of the fault location is more accurate for two-phase or three-phase faults within Alliander's fault localisation framework. Hence, in the cases of Figure 4.3 and Figure 4.4, the waveforms should be labelled as three-phase faults for accurate fault localisation.

4.1. Description of the Sample Space

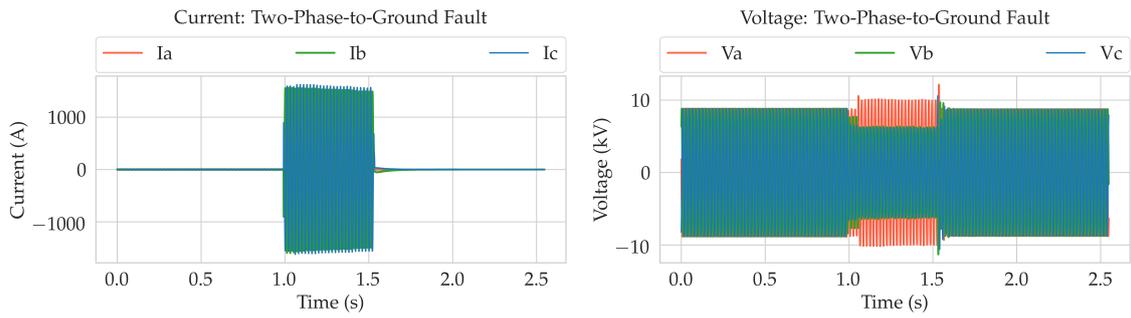


Figure 4.1: Current (left) and voltage (right) waveforms for a two-phase-to-ground fault.

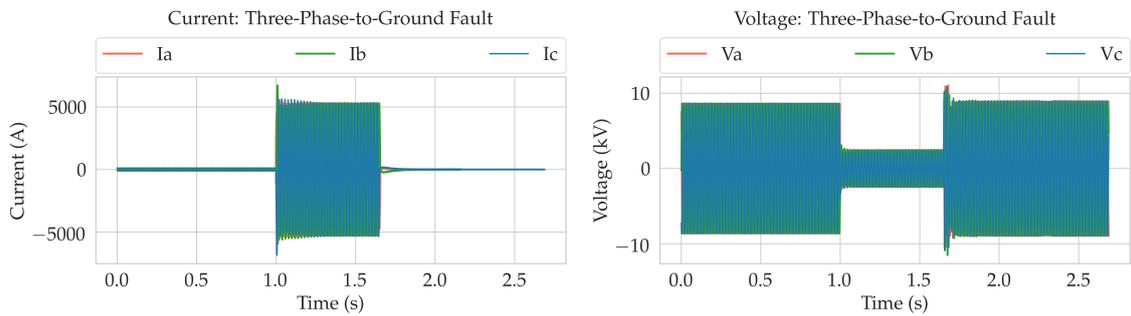


Figure 4.2: Current (left) and voltage (right) waveforms for a three-phase-to-ground fault.

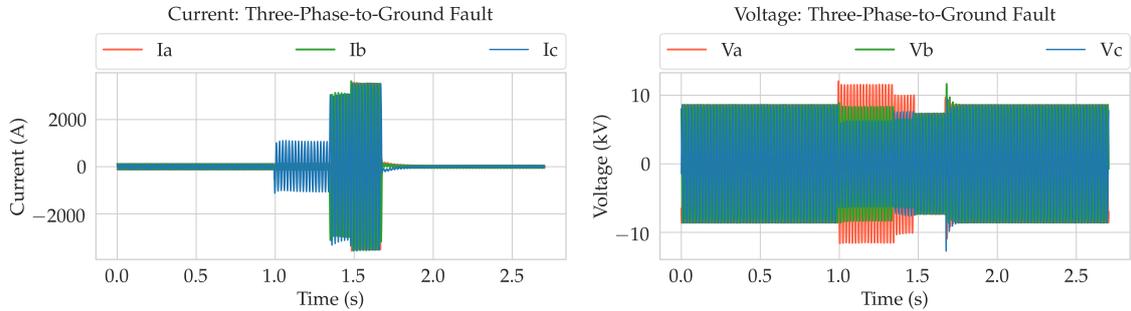


Figure 4.3: Current (left) and voltage (right) waveforms for a single-phase fault that develops into a three-phase fault.

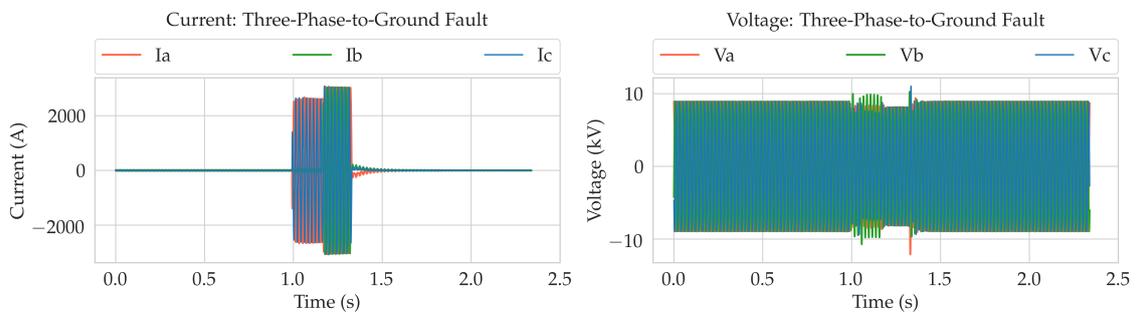


Figure 4.4: Current (left) and voltage (right) waveforms for a two-phase fault that develops into a three-phase fault.

4.1.1 Steps for Data Preprocessing

Introduction

Each digital fault recording (DFR) consists of three-phase voltage and current signals that are recorded from $1s$ pre-fault to around $1s$ post fault. The signal can be divided into three parts. In the pre-fault state, the three-phase current and voltage are in a sinusoidal steady-state condition; i.e., the signals have an amplitude and frequency (of 50 Hz) that is constant in time [45]. The fault state directly succeeds the pre-fault state and is characterised by low voltage magnitudes and large current magnitudes of a few thousand amperes that exceed the current carrying capacity of the cables.

The currents that flow immediately after the occurrence of a fault are different from the currents that are interrupted by a fuse or circuit breaker. For instance, first few cycles of single-phase faults exhibit unstable behaviour such as distortions at the zero-crossing due to the extinguishing and restrike process [46]. The third and final section of the recording is the post-fault state. In the case of permanent faults, the isolating device switches the fault signal off, thereby interrupting the sustained flow of large currents that could damage the network [9]. Hence, the post-fault currents are close to 0 if they are switched off at the substation, and lower than than pre-fault currents otherwise. In the case of self-extinguishing faults, the faults extinguish at the zero crossing of the signals [10]. The post-fault state, in this case, is the same as the pre-fault steady-state conditions for self-extinguishing faults. These states can be observed more clearly in Figure 4.5

Overview of the Sample Space

The classes of faults that are considered for this thesis are single-phase-to-ground permanent faults, single-phase self-extinguishing faults, two-phase-to-ground permanent faults, and three-phase faults. The distribution of samples among the four classes can be observed in Table 4.1. It is clear that the data set is imbalanced. This is a function of the frequency of occurrence of each type of fault and the reliability of their data labels, as only faults that have already been correctly classified can be used for accurately determining the performance of a new classification model.

Table 4.1: Summary of the number of DFRs for each fault type collected from the SASensor.

Fault Type	Number of Samples
Single-Phase Permanent	118
Single-Phase Self-Extinguishing	81
Single-Phase Unstable	51
Two-Phase Permanent	100
Three-Phase Permanent	96

4.1.2 Fault Segmentation Algorithm

For the purpose of identifying the fault type, the most pertinent portion of the fault signal recorded by the SASensor is between the trigger of the fault and the extinction/return to steady-state, as it is this section that consists of the information about the faults and their type. Hence, it can be useful to section out only this part of the signal for further fault analysis. The start of the fault can be fixed at $1s$ (or 4000 samples) into the signal, by virtue of the recording starting exactly $1s$ pre-fault; however, the terminating point of the fault is not as certain. A segmentation algorithm

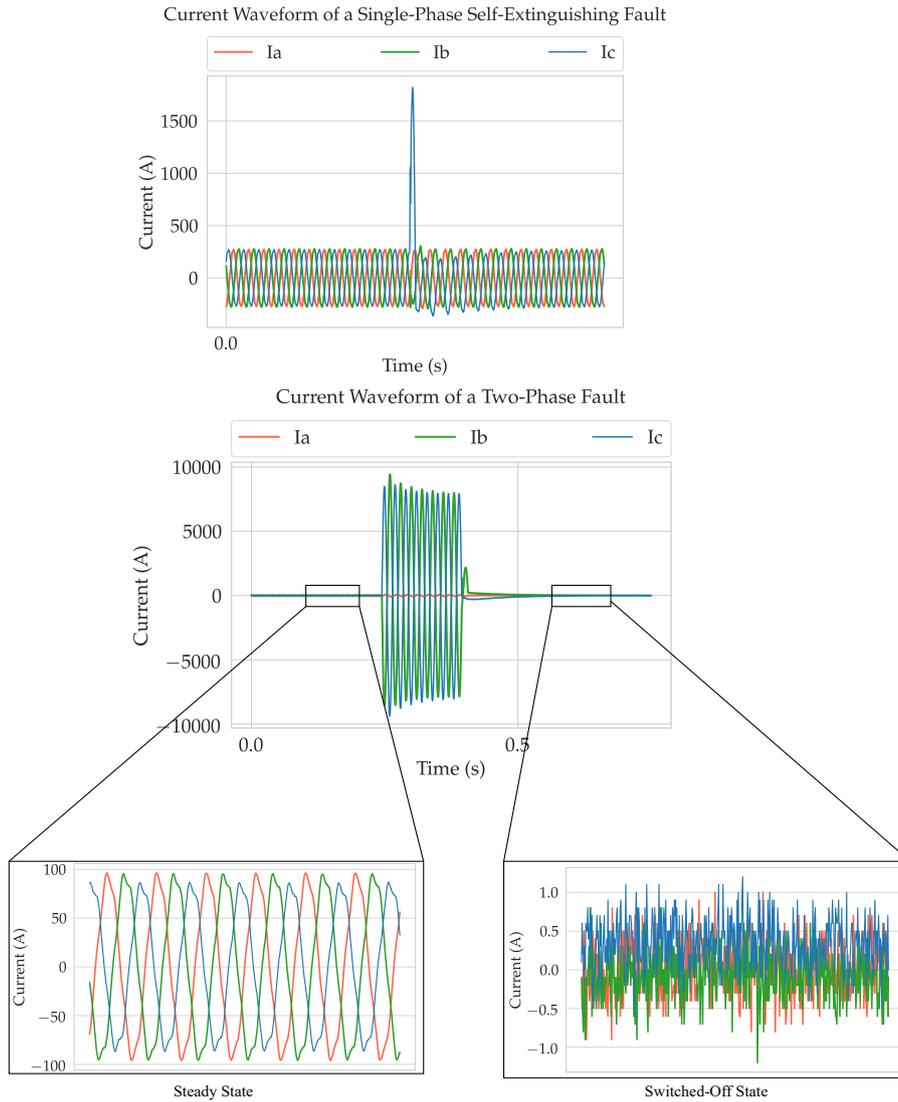


Figure 4.5: Plot comparing the different states in a permanent and self-extinguishing fault.

can be used to discern the second state change point. Additionally, in the case of faults that evolve from one type to another, it can be beneficial to further partition these faults into the final stage of the fault. This is because, from practical experience, it is easier to calculate the accurate location for multi-phase faults. For instance, in the case of the three-phase fault shown in Figure 4.3, the fault starts as a single-phase fault at $1s$, progresses into two-phase-to-ground fault at $1.3s$ and finally evolves into a three-phase fault at $1.5s$. For the location of the fault to be accurately calculated, it is easier to only use the final stage of the fault, which, in this case, is the three-phase fault portion. In such faults, it can be investigated if the performance of the classifier can improve with a segmentation algorithm that sections out the fault stage before it is switched off.

This partitioning of the fault signal into its final fault section is performed with a binary segmentation algorithm. Binary segmentation is a technique to detect change points such as transients and sudden disruptions in a time series.

A change point occurs in time series $y_t, t \in 1, \dots, n$, if, for time τ , there exists a difference in mean, variance or another such criterion between y_1, \dots, y_τ and $y_{\tau+1}, \dots, y_n$. If, for the time series, the number of change points is m , with the (ordered) change points being τ_1, \dots, τ_m . The

last changepoint at the end of the timeseries is $\tau_{m+1} = n$. The data is hence segmented into $m + 1$ sections where the i^{th} segment contains $y_{(\tau_{i-1}+1):\tau_i}$. The cost function C is present for the segment and overfitting is prevented by introducing a penalty term βm , that is linear in m . The equation for the cost function utilised by the binary segmentation algorithm is presented in Equation 4.1.

$$\sum_{i=1}^{m+1} [C(y_{(\tau_{i-1}+1):\tau_i})] + \beta m \quad (4.1)$$

First, for a detected changepoint τ applied to time series y_1, \dots, y_n , it is checked if the cost function of the two resulting subsections in addition to the penalty is smaller than the cost function over the complete series. That is, it is checked if the expression 4.2 is satisfied for $\tau \in \{1, \dots, n-1\}$.

$$C(y_{1:\tau}) + C(y_{(\tau+1):n}) + \beta < C(y_{1:n}) \quad (4.2)$$

If this condition is satisfied, τ is considered a change point, and the algorithm checks within the segments for further change points until the expression does not hold true anymore. This method of segmentation has a time complexity of $O(n \log n)$ and is not computationally expensive [47].

Implementation of the segmentation algorithm on the fault signal

The implementation of the segmentation algorithm is discussed in this section. The desired outcome of the segmentation step is the final fault state before extinction. This is implemented using a combination of segmentation and peak detection techniques. To make the change points more explicit, the signal is squared to eliminate the zero mean of the sinusoidal signal. Additionally, since it is known that the faults start at 1s, the segmentation algorithm is only applied to the portion of the signal after 1s. The binary segmentation technique iteratively detects change points per segment. For each (current and voltage) phase, the most prominent change points are identified for further analysis.

The results of the change point detection algorithm are three sets of change points for each phase. Each set has two values that indicate the first and second change point respectively and are values in time. Figure 4.6 shows an example of the calculated change points for each phase of the currents. To ensure only the final stage of the fault is captured for faults that evolve from one type to another, the largest value of the first and second change points from the three sets are selected. For example, in the case of the fault in Figure 4.6, the resultant change points shown in table Table 4.2. To segment the fault into just its final fault component, the largest values of the first and second change points each, i.e., 1.47s and 1.67s are selected to represent the fault. These values from the segmentation process are extended to the voltages as well — a separate segmentation process would be computationally expensive, and as the current and voltage follow the same pattern of change points, this decision is reasonable. The result of the segmentation can be seen in Figure 4.7. This process also works well in cases with only a single fault (for example only a single-phase or two-phase fault) as all three sets will have similar change point values.

Self-extinguishing faults are too short to be detected accurately by the change point algorithm. As self-extinguishing faults typically last around half a cycle to two cycles, if the difference between two change points is smaller than two cycles (0.04s), the entire signal is selected instead of a section. This choice is justified by the fact that further feature engineering is done to ensure relevant (and distinguishing) aspects of the faults are identified for the purpose of classification in Section 4.2.

4.1. Description of the Sample Space

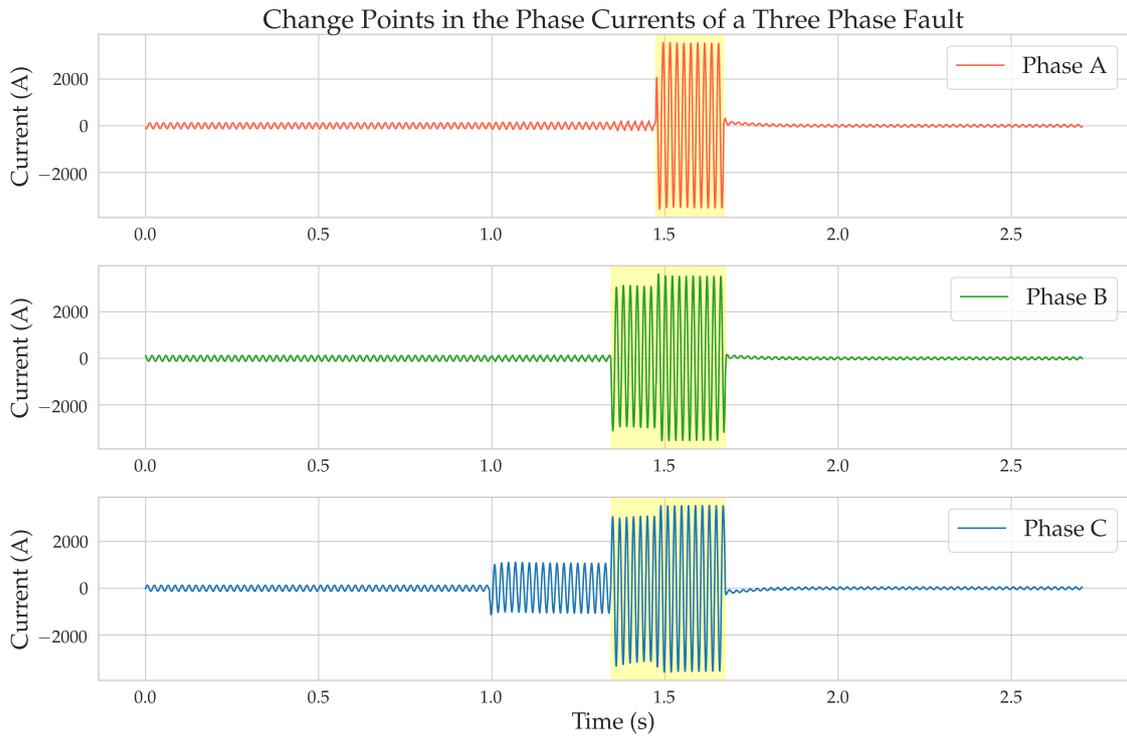


Figure 4.6: Change points for the phase currents of the three-phase fault in Figure 4.3

Table 4.2: Change points for the fault in Figure 4.3

Phase	First Change Point (in seconds)	Second Change Point (in seconds)
A	1.47	1.67
B	1.35	1.67
C	1.35	1.67

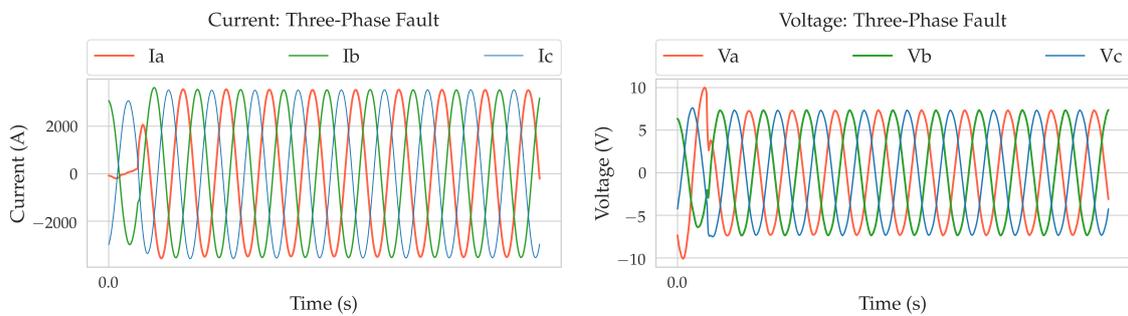


Figure 4.7: Segmented section of the three-phase fault in Figure 4.3

4.2 Development of Features for Stable Faults

A feature is a characteristic or property of an event. It is important to choose features that are informative and discriminating between classes [48]. The features identified for the faults are numeric and are explained in this section. First, a raw set of features are developed based on the properties of fault waveforms. Following this, feature selection is performed to develop a more general classification model in Chapter 6.

An important aspect that lent to the selection of features is that the fault file does not explicitly indicate in which phase the fault occurs. In literature, such as in [18], [29] and [27], each phase is treated individually and checked for a fault, and features like the magnitude or RMS can be applied independently without it being phase-dependant. In the case at hand, since fault phase information is not present in the DFRs, when voltage and current features must be calculated for each phase and used in the feature set, the phase of each feature becomes implicit in the decisions made for classification. As the type of fault is not dependant on the phases that are affected, it is important to decouple the phase information from the features.

One way to do this is by using features that contain information on all the three phases for the voltage and current each, such as the zero-sequence values. As zero-sequence currents and voltages are not sufficient to completely represent a fault, the choice was made to consider the signal behaviour of each phase current and voltage, and order them in (decreasing order of) magnitude. It was considered reasonable to assume that the magnitude of the features has a larger bearing on the type of fault than the phase, particularly since faults are characterised by large/small values of fault currents/voltages relative to the currents/voltages in steady-state. This section explains the features selected, with the motivation for their choosing.

Inclusion of Synthetic Faults

From the images shown of single-phase, two-phase, and three-phase fault data collected from occurrences in the distribution network, it is clear that the waveforms consist of distortions that could skew the feature set in a way that could prevent future faults from being classified correctly. To mitigate this, a set of synthetic faults are added to the permanent faults to improve the feature set with reference values that accurately characterise textbook single-/two-/three-phase permanent faults. The fault data for the synthetic faults are created in the same format as the segmented faults from the DFRs. They consist of three-phase current and voltage waveforms, and they do not include the initial pre-fault or final switched-off periods. An example of a synthetic two-phase fault is shown in Figure 4.8.

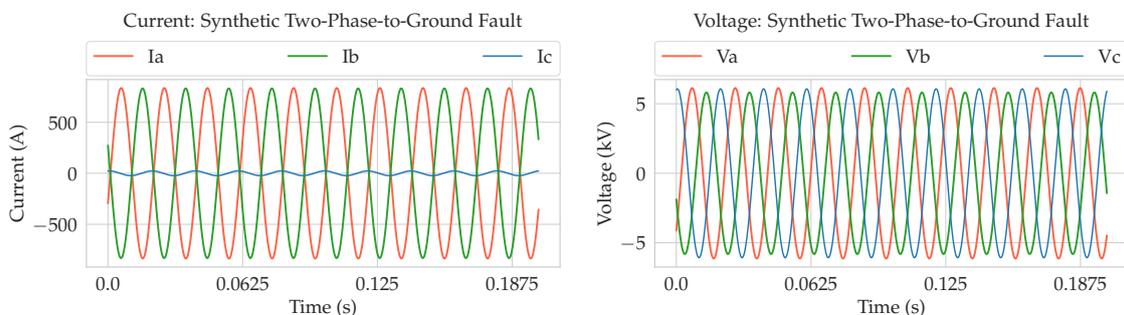


Figure 4.8: Current and voltage waveforms for a two-phase synthetically created fault.

Once these faults are created, the data set is added to the samples of the permanent faults. A synthetic fault is effectively the same as a permanent fault sample from a DFR after the segment-

Table 4.3: Summary of the complete sample space of the fault considered for classification

Fault Class	Fault Type	Fault Data Origin	Number of Samples
Single-Phase	Permanent	DFR	118
		Synthetic	12
	Self-Extinguishing Unstable	DFR	81
		DFR	51
Two-Phase	Permanent	DFR	100
		Synthetic	12
Three-Phase	Permanent	DFR	96
		Synthetic	12
Total			482

ation processing. The updated description of the sample space can be observed in Table 4.3. The labels of the twelve faults for each class of the synthetic faults are added to the complete sample set. At this point, it is important to note, that in order to assess the performance of the classifier on faults in the distribution network, the synthetic faults are used only to train the classifier model and are not included in the test set. The labels of the DFR fault files are determined automatically from the fault identification process explained in Section 1.2, and have been cross-checked by an expert to ensure that only correctly labelled faults are used in this thesis.

Features from the Current

The first set of features are taken from the three-phase fault current. As seen in Figures 4.1 and 4.2, faults are characterized by large currents and lower voltages at the phases in which the faults occur. The RMS values (or the DC component) of the fault phases are proportionally large or small for a current and voltage signal, respectively. The RMS of each phase current and voltage, as calculated using Equation 4.3 is hence selected as a feature. The RMS is calculated as a function of the sampled signal values. The number of samples that make up the signal under consideration is represented by n , and x_i is the value of the signal at each sample step i .

$$RMS = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2} \quad (4.3)$$

Only the RMS of the currents in each phase is selected as they show larger deviations between the fault- and steady-state. This is also done in the interest of keeping the number of features to a minimum. This results in three features that each represent the RMS for the three-phase fault current. As mentioned previously in Section 4.2, using only these features directly as inputs to the classifier can result in misleading results by creating an implication that the phase is a determining factor for fault classification. The RMS values are further processed by being sorted in descending order, in order to dissociate the value from the phase.

A high-level observation of single-phase, two-phase, and three-phase faults is that they each have one, two, and three currents that have a value significantly higher than the steady-state values. A metric that compares the magnitudes of the RMS currents is calculated using Equation 4.4. It is in the interest of classifier simplicity to have features that combine aspects of all three phases into a single metric and can possibly be useful to reduce the dimensions of the feature space. As the RMS values of the current phases are of different magnitudes for different faults,

this metric helps capture the relational aspect between the three phases, which is vital to the identification of the fault type.

$$irms_comparison = 2 * \frac{irms_1}{irms_2} - \frac{irms_2}{irms_3} - \frac{irms_3}{irms_1}, \quad (4.4)$$

where $irms_1, 2, 3$ are ordered in decreasing order of magnitude

For a single-phase fault, $irms_1 \gg irms_2, irms_3$, so it is expected that $irms_comparison$ will be a positive value. For a two-phase fault, $irms_1, irms_2 \gg irms_3$, hence, $irms_comparison$ will be negative. Finally, for a three-phase fault, where all the RMS values of the currents are approximately equal, $irms_comparison$ will be close to zero.

An important assumption in the usage of this formula is that none of the RMS values are zero as this causes $irms_comparison$ to become undefined. A recommendation for future use of this formula is to create a test case for zero RMS currents — the presence of zero current in a phase also indicates that the phase does not contain a fault. This thesis does not include this test case as an occurrence of such a fault is rare from practical experience, and the data set at hand does not consist of such outliers.

Features from the Results of the Discrete Wavelet Transform

Discrete wavelet transforms (DWT) can be used to decompose a signal into smaller frequency bands as explained in Chapter 3. The detail coefficients from levels 1 and 2 capture the high-frequency disturbances in the ranges, 1 kHz - 2 kHz, and 500 Hz - 1 kHz, respectively. The frequency bands below these levels capture the harmonics and the fundamental frequency. For this study, the detail coefficients of the current and voltage at the first and second levels are considered as they contain the majority of the transient and harmonic frequencies.

The detail coefficients for a two-phase-to-ground fault are shown in Figure 4.9. It can be observed that the detail coefficients capture the transients in the fault signals - i.e. when the signals are nearly a 50 Hz sinusoidal wave, the coefficients are close to zero and when there are transients in the signal, the coefficients take higher values. It can be observed that the detail coefficients in Figure 4.9 show significant deviations from zero at the start of the fault and return to values close to the zero-mean as the fault progresses. The detail coefficients not only capture the transient at the moment of fault occurrence but also the distortions at the zero-crossing as the fault attempts to extinguish during the initial cycles. As the fault progresses, the extinguishing-restrike cycle is less significant and the detail coefficients are nearly zero again.

In a perfectly sinusoidal signal with a frequency of 50 Hz, with no disturbance or transient, the detail coefficients are zero in the first two levels of decomposition. This has been checked for all wavelets and an example of the results of the decomposition of a signal with the db4 wavelet can be observed in Figure A.1 in Appendix A. As the sinusoid does not contain any high-frequency components present in the 0.5 - 2 kHz range, the detail coefficients are correspondingly zero in this bandwidth. Hence, it can be gathered that the magnitude of the detail coefficients is proportional to the degree of disturbance in the sinusoidal signal. Additionally, the results of the DWT on the fault signal are similar in all three phases, only differing in magnitude depending on the degree of signal distortion, and if the phase in question contains the fault or not.

The features selected as a function of the detail coefficients of the voltage and current are the energy and standard deviation of both the levels using Equations 4.5 and 4.6. The standard deviation carries information on how dispersed the detail coefficients are from the zero-mean.

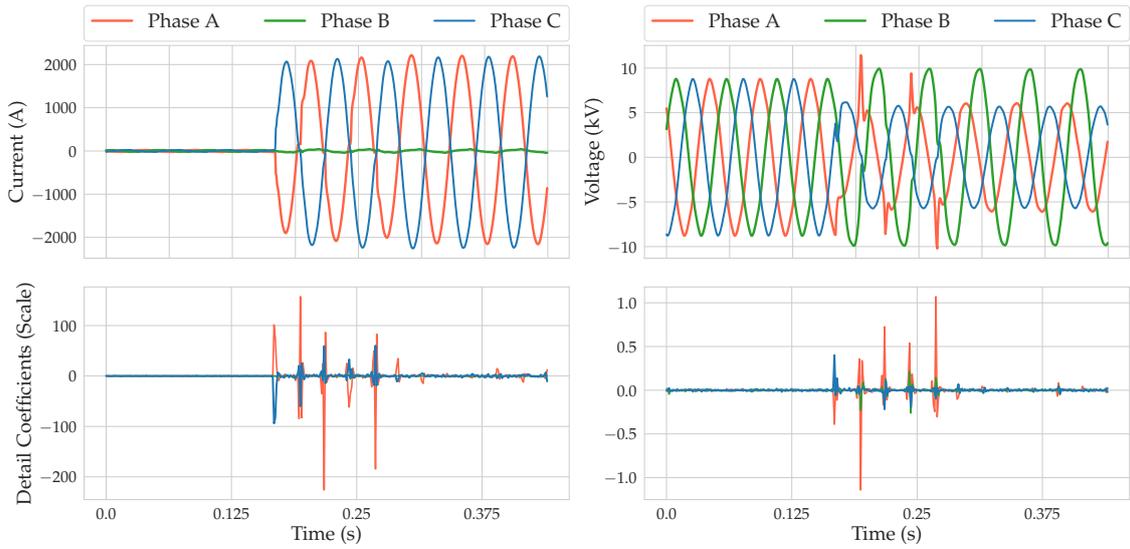


Figure 4.9: Level 1 detail coefficients for the current and voltage of a two-phase fault (in the 1-2 kHz bandwidth).

The energy carries the information on the “strength” of the detail coefficients. The higher the transients in a signal in the time-domain, the higher the number of non-zero detail coefficients and the subsequent energy. The standard deviation and energy are calculated for the first and second levels of decomposition.

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}, \quad (4.5)$$

where x_i is each sample, \bar{x} is the sample mean, and N is the number of sample observations.

$$E = \sum_{i=1}^N x(i)^2, \quad (4.6)$$

where x_i is each sample between 1 and N

Features from the Zero-Sequence Voltage and Current

The zero-sequence component is a vector sum of the phase current (or voltage). This is the current that flows in the cable sheath or ground. Equation 4.7 represents the formula for the calculation of the zero-sequence current and voltage, I^0 and V^0 respectively. A balanced, or symmetrical system, is one where all phases of currents (and voltages) have equal amplitudes and frequency, and are each shifted by 120° . The zero-sequence components for such a system, i.e., the vector sum of the balanced phases, are zero. In the case of an unsymmetrical fault, the zero-sequence component will have a large value of current and voltage when the earth is involved.

Particularly, in the case of a single-phase fault, the zero-sequence component will follow the same pattern as the original signal in the phase containing the fault. This is beneficial for characterizing restrike and extinguishing fault behaviour; this concept is further explained in detail in Section 4.3. In the case of a symmetrical three-phase fault, however, the zero-sequence components are zero due to the balanced nature of the fault currents.

$$\begin{aligned}
 I^0 &= \frac{1}{3}(I_a + I_b + I_c) \\
 V^0 &= \frac{1}{3}(V_a + V_b + V_c)
 \end{aligned} \tag{4.7}$$

The features from the zero-sequence components are the RMS and standard deviation calculated from the signal in time, and the detail coefficients. The features in the time-domain signal carry information on the magnitude of the transient signal in the context of the sinusoidal current and voltage. The features from the DWT carry the time-frequency details about the transients alone.

Table 4.4 consists of a summary of the features determined important for the identification of the different types of faults. The features in the three-phase signals are ordered in descending order of magnitude to eliminate the dependency on the phase while creating the features. For example, the features v_{l1_std1} , v_{l1_std2} , v_{l1_std3} are the standard deviation of the first level of detail coefficients for the three voltage phases listed in descending order of magnitude. The total number of features amounts to 36. These 36 features do not all characterise unique aspects of the faults and can be narrowed down to a smaller subset to improve performance, increase model training times, and develop a more broadly applicable model [23].

Table 4.4: Table detailing the features extracted from the fault signals.

			Voltage Features	Current Features	
Features from the Signal in Time	3-Phase Signal	RMS	-	irms_1	
			-	irms_2	
			-	irms_3	
		RMS Comparison	-	irms_comparison	
Zero-Sequence Signal		σ	V0_regular_std	I0_regular_std	
		RMS	V0_regular_rms	I0_regular_rms	
Features from the Detail Coefficients	3-Phase Signal	Level 1	σ	v_{l1_std1}	i_{l1_std1}
				v_{l1_std2}	i_{l1_std2}
				v_{l1_std3}	i_{l1_std3}
		Energy	v_{l1_e1}	i_{l1_e1}	
			v_{l1_e2}	i_{l1_e2}	
			v_{l1_e3}	i_{l1_e3}	
	Level 2	σ	v_{l2_std1}	i_{l2_std1}	
			v_{l2_std2}	i_{l2_std2}	
			v_{l2_std3}	i_{l2_std3}	
		Energy	v_{l2_e1}	i_{l2_e1}	
			v_{l2_e2}	i_{l2_e2}	
			v_{l2_e3}	i_{l2_e3}	
Zero-Sequence Signal	Level 1	σ	V0_wavelet_std	I0_wavelet_std	
		Energy	V0_wavelet_E	I0_wavelet_E	

4.3 Development of Features for Unstable Faults

4.3.1 Introduction to Unstable Faults

An unstable fault, as defined by Gu et al. in [7], is a fault whose current/voltage is not sufficiently close to a sinusoidal waveform. The present criteria at Allander for classifying a fault as

stable is presented in Table 4.5. MaxDelta is the maximum numerical differential between data points between two consecutive samples, and max is the maximum value of the waveform. Two consecutive periods of a fault must be identified as a stable event to be classified as a stable fault. It has been observed that the values for the thresholds developed are too severe as, from practical experience, faults that can otherwise be classified as stable permanent or self-extinguishing (by an expert), are sometimes identified as unstable. An example of a fault that has a high degree of unstable behaviour can be seen in Figure 4.10. It can be observed that the amplitudes of the fault voltage and current are not constant, and the waveform is characterised by a sequence of restrike and extinction at the zero-crossing.

 Table 4.5: Criteria developed by Alliander for identifying a fault as stable^a [7]

Type	Criteria	Single-Phase	Two-Phase	Three-Phase
Phase Voltage	max > 0	x	x	x
	RMS/max > 0.6	x	x	x
	maxDelta/max < $8\pi/N_s$	x	x	x
Phase Current	max > 0	x	x	x
	RMS/max > 0.6	x	x	x
	maxDelta/max < $4\pi/N_s$	x	x	x
Zero-Sequence Voltage	max > 0	x		
	RMS/max > 0.6	x		
	maxDelta/max < $8\pi/N_s$	x		
Zero-Sequence Current	max > 0	x		
	RMS/max > 0.6	x		
	maxDelta/max < $4\pi/N_s$	x		

^a N_s is the number of samples per period.

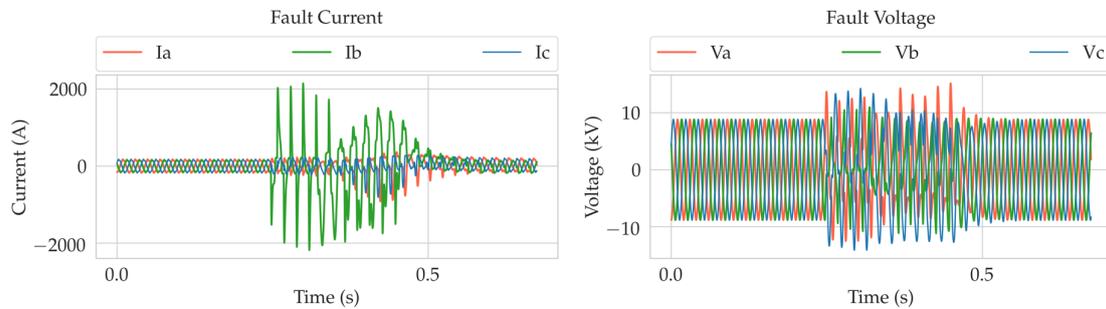


Figure 4.10: Current (left) and voltage (right) waveforms for a single-phase unstable fault.

A reason for stable permanent faults being misclassified as unstable is due to certain sections of the fault waveform being unstable even though the fault develops into a stable fault. An example of a stable fault classified as unstable can be observed in Figure 4.11 where the fault in phase A starts with significant distortions but develops into a visually stable waveform a few cycles before it is extinguished. Faults are labelled as being unstable in an automatic process through the rules in Table 4.5. It is possible that this fault was classified as unstable due to the distortions in the first half of the fault. Self-extinguishing faults can also be classified as unstable faults when they occur in a series as shown in Figure 4.12. It can be seen that the fault in phase C is a single-phase fault that displays restrike and extinguishing behaviour close to every zero-crossing. This sequence of self-extinguishing behaviour is difficult to capture with the rules that are presented in Table 4.5. Additional features are hence developed to better identify the differences between what makes a fault unstable and stable.

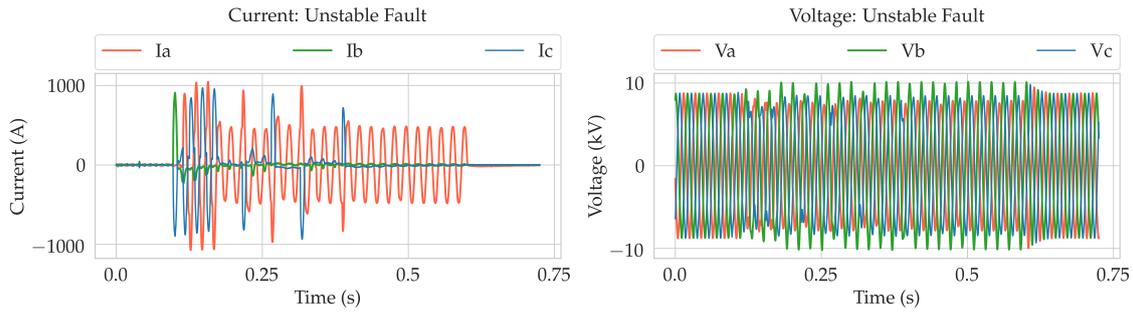


Figure 4.11: Current (left) and voltage (right) waveforms of an single-phase permanent fault that was identified as unstable.

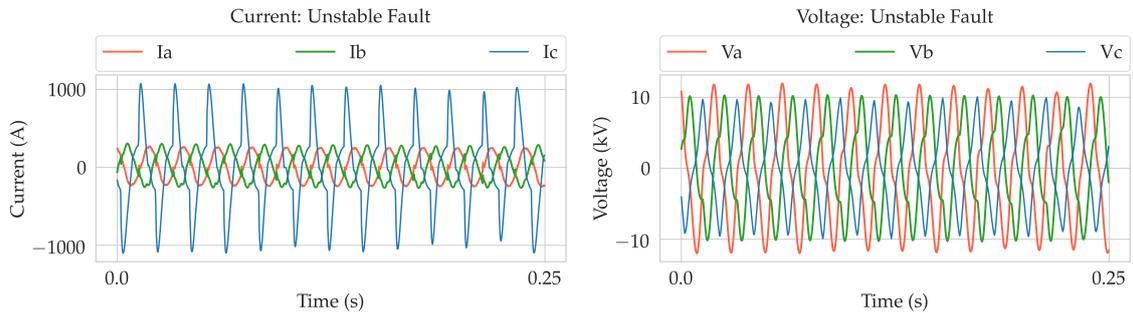


Figure 4.12: Current (left) and voltage (right) waveforms of a single-phase self-extinguishing fault that was identified as unstable.

4.3.2 Windowed Fourier Analysis for Detecting Stability

The purpose of this analysis is to develop an improved system for classifying unstable faults. Due to the current definition of unstable faults being less concrete than desired, some faults classified as unstable were found to be stable permanent or self-extinguishing faults under further manual inspection by experts. The objective is hence to prevent these misclassifications of single-phase permanent and self-extinguishing faults and to add more certainty to what makes a fault (un)stable.

When faults are classified as unstable, their locations are not automatically calculated to avoid the risk of sending a wrong location to the control centre for isolation. Further inspection must be manually performed by experts to calculate the location of unstable faults. In the case of an incorrectly classified stable fault, however, the fault location information is still sent to the control center for further isolation. If the calculated location is greater than the tolerated margin of 0.5 - 1 km, the incorrect section is isolated by the control center and the actual fault is not mitigated. While it is important to ensure stable faults are not classified as unstable, it is also vital to ensure that the definition does not leave room for additional misclassifications. The need for this balance is taken into account while determining the features for unstable fault classification.

As Gu et al. [7] state that the identification of unstable faults is influenced by the degree of sinusoidal nature that the waveform displays, further studies were performed using Fourier transforms. The Fourier transform was used to analyse and select features in the frequency spectrum of the faults that were classified as unstable. This method attempts to connect the notion of "insufficiently" sinusoidal behaviour with identifying unstable faults using a windowed analysis of the fault waveforms. This section focuses on single-phase permanent and self-extinguishing faults due to there being an abundance of accurately labelled data. The following sections describe the new methodology devised to develop better features for single-phase unstable faults.

4.3. Development of Features for Unstable Faults

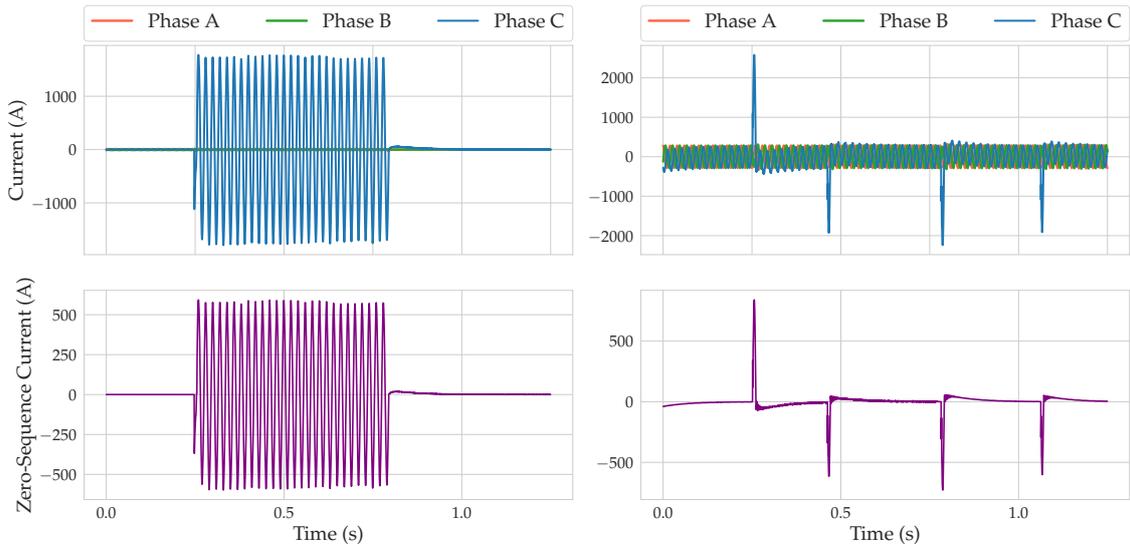


Figure 4.13: Comparison between zero-sequence currents of a single-phase permanent fault (left) and single-phase self-extinguishing fault (right).

The Importance of Zero-Sequence Signals

The first step in the analysis is to calculate the zero-sequence current/voltage from fault waveforms using Equation 4.7. As the fault occurs only in a single phase, the zero-sequence component reflects the fault waveform alone, during the fault period, and values close to zero when the three-phase current/voltage is balanced (theoretically, balanced currents/voltages sum to zero). Examples of the zero-sequence current for a single-phase permanent fault and single-phase self-extinguishing fault are shown in Figure 4.13. As the point of interest is the fault portion of the waveform, studying the zero-sequence current simplifies the process by representing this section more prominently for the study.

Sliding Fourier Analysis

The stability of a fault waveform is studied by analysing which periods are stable and which periods are unstable. For a signal to be considered stable, and for accurate fault localisation, practical experience suggests that at least two consecutive stable periods must be present [7]. In this thesis, the stability is hence studied period-by-period with a moving window of one cycle (80 samples). The calculation of the derivative of the zero-sequence current is useful for highlighting the restrike behaviour in a single-phase fault as explained below.

The derivative of a sinusoidal wave remains sinusoidal, but has accentuated distortions and transients as the instantaneous rate of change of the wave is high for this portion. An example of this effect can be seen in Figure 4.14. I^0 is the zero-sequence current calculated as per Equation 4.7, normalised to have unit energy. The derivative is calculated by taking the difference between every two consecutive samples to obtain the zero-sequence-derivative current, I_{deriv}^0 . The window is set at one period of the fault waveform, which consists of 80 samples. In this window, a comparison is drawn between I_{deriv}^0 and a reference ideal sinusoidal waveform I_{ref} that contains no transients. Waveform I_{ref} is developed to have (i) a fundamental frequency of 50 Hz, (ii) unit energy, and (iii) no harmonics. I_{deriv}^0 is first normalised to unit energy using Equation 4.8:

$$\text{Normalised } I_{deriv}^0 = \frac{I_{deriv}^0}{\sqrt{E_{I_{deriv}^0}}}, \quad (4.8)$$

where $E_{I_{deriv}^0}$ is the energy of I_{deriv}^0 calculated for one window as per Equation 4.6. The frequency spectrum of the normalised I_{deriv}^0 is then compared with that of I_{ref} in the next step.

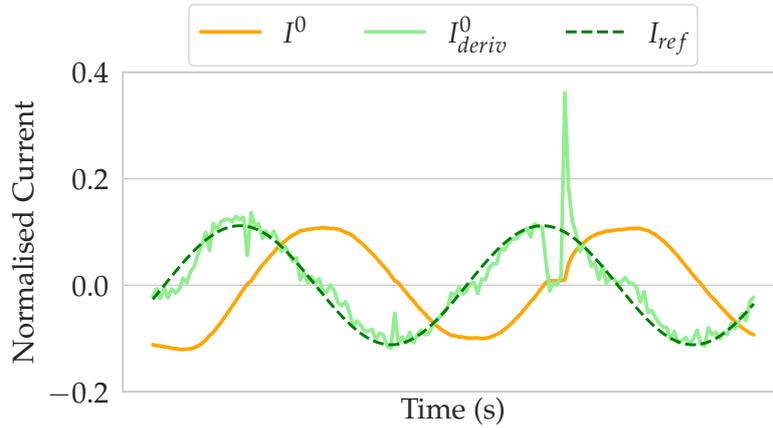


Figure 4.14: Comparison between the normalised zero-sequence current (I^0), its derivative (I_{deriv}^0) and the reference (I_{ref}). **Note:** This image shows a window of two cycles for representative purposes.

From Figure 4.14, it can be stated that, in terms of distortions and harmonics present, I_{deriv}^0 is less stable than the reference waveform I_{ref} (that can be considered to be perfectly stable). A metric F_{ratio} is developed to quantify this degree of instability. F_{ratio} compares the power of the fundamental frequency (50 Hz) of I_{deriv}^0 with that of I_{ref} . This power of the fundamental frequency is calculated as a result of the Fourier transform of the two signals.

For each window, F_{ratio} can be calculated as per Equation 4.9:

$$F_{ratio} = \frac{a_{0_{fault}}}{a_{0_{ref}}}, \quad (4.9)$$

where, $a_{0_{fault}}$ is the power of the fundamental frequency of I_{deriv}^0 , and $a_{0_{ref}}$ is the power of the fundamental frequency of I_{ref} . As example of this can be observed in Figure 4.15, which represents the comparison of the intensities of the constituent frequencies present in I_{deriv}^0 and I_{ref}^0 . As the DFR signals are sampled at 4 kHz, the frequencies between 0 - 2 kHz are captured. Both signals have a dominant 50 Hz component, and it can be seen that $a_{0_{fault}} < a_{0_{ref}}$. For this transform, $a_{0_{fault}} = 3.63$ and $a_{0_{ref}} = 4.47$. F_{ratio} for this window is hence $3.63/4.47 = 0.812$.

For two equal energy waveforms, the amplitude of the 50 Hz Fourier component of the fault waveform will always have a value that is smaller than or equal to that of the reference sinusoidal waveform due to the presence of distortions. That is, the relation $a_{0_{fault}} \leq a_{0_{ref}}$ always holds true. F_{ratio} will equal 1 for a fault waveform that is exactly sinusoidal and will have lower values for a signal with a high degree of distortions. As seen in Figure 4.15, the Fourier component of the fundamental frequency of the reference sinusoidal waveform is higher than that for the waveform that consists of distortions. It is clear, that while the values may be close, the ratio $a_{0_{fault}}/a_{0_{ref}}$ cannot be greater than one. Hence, this developed metric F_{ratio} can be used for assessing the relative sinusoidal nature of the fault waveform.

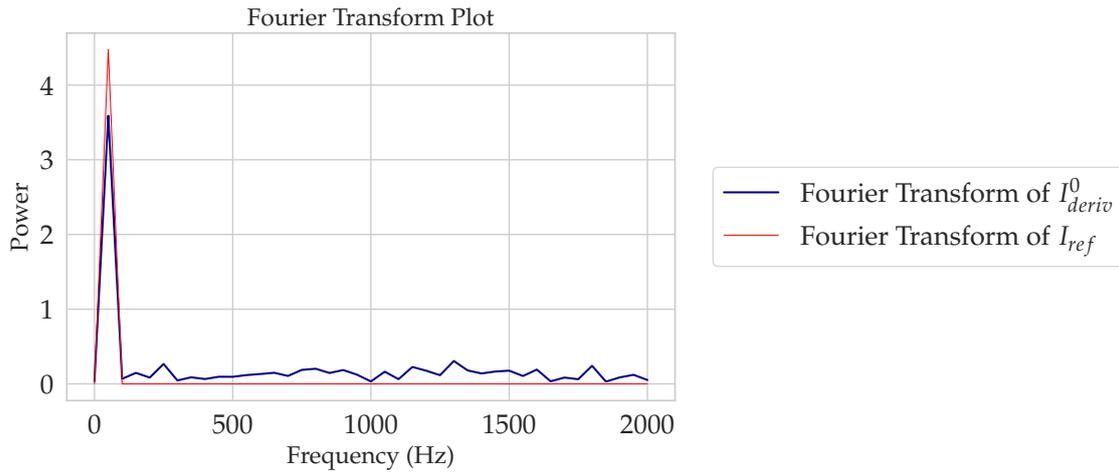


Figure 4.15: Comparison of the 50Hz Fourier components between the derivative of the zero-sequence current and its reference sinusoidal waveform.

Metric F_{ratio} is calculated for each period of the fault current waveform, enabling a period-wise analysis of stability. The results of this moving window analysis of the 50 Hz component for a sample permanent and extinguishing fault can be seen in Figure 4.16. In the case of the permanent fault (to the left of Figure 4.16), it can be observed that, at the initiation of the fault, the distortions in the waveform reflect in the magnitude F_{ratio} . An assumption is made that a sinusoidal wave of 50 Hz frequency with no distortions defines the upper limit of stability; therefore, comparisons to this reference wave can be made to evaluate stability. As the fault becomes more stable, this ratio reaches magnitudes close to one as well. This fault can be considered a stable fault on this basis. In the case of the extinguishing fault the right of Figure 4.16, it can be observed that the values of F_{ratio} are significantly lower than that of the permanent fault for all periods of the fault.

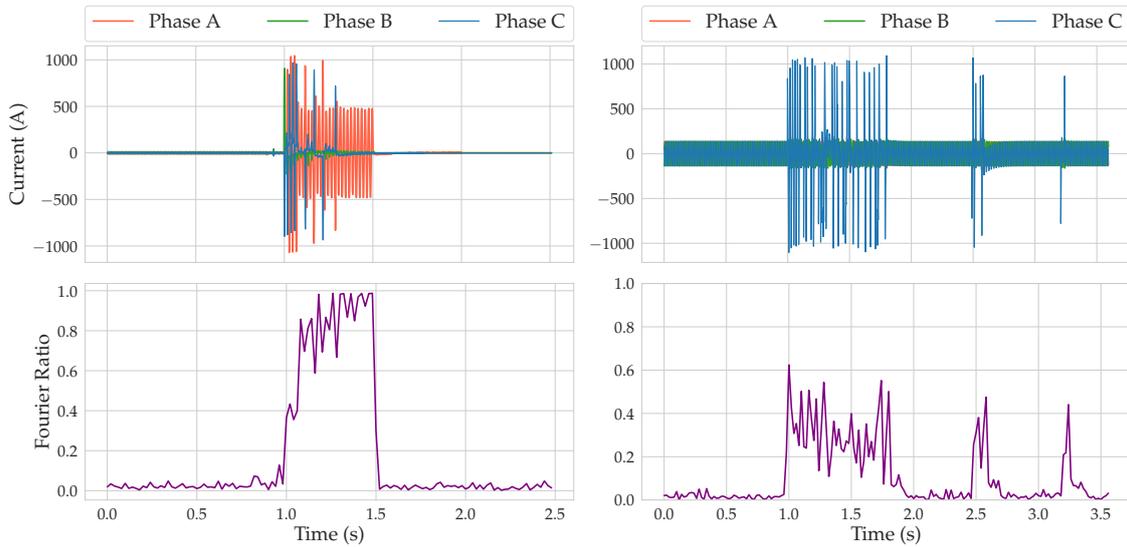


Figure 4.16: Comparison of the 50 Hz Fourier component of unstable single-phase permanent (left) and single-phase self-extinguishing (right) faults.

A key observation of the windowed analysis of F_{ratio} is the difference in the magnitudes of the metric between the permanent and self-extinguishing fault (as seen in Figure 4.16). It can be noted that F_{ratio} takes values close to 1 for the seemingly stable section around 1.5 s for the

permanent fault, with lower values being observed for the self-extinguishing fault. This justifies the process of dividing $a_{0_{fault}}$ with respect to a reference $a_{0_{ref}}$ to calculate F_{ratio} for each window instead of normalising the entire series of $a_{0_{fault}}$ post-calculation for all windows. The former ensures that there is a difference between the most stable periods of a permanent fault and self-extinguishing fault. Within the same waveform, this process also ensures the relative magnitudes between the section of the waveform containing the fault and the balanced section are preserved.

Features from the Fourier Ratio

This section discusses the development of features from the Fourier analysis for the classification of unstable faults. The first step in this process is to identify typical values of F_{ratio} that stable permanent and self-extinguishing faults possess. Three different data sets are used - stable single-phase permanent and self-extinguishing faults and unstable faults. Currently, stable permanent and self-extinguishing faults that exhibit atypical behaviour are classified as unstable, even if they are stable under closer inspection. Therefore, features from the stable single-phase permanent and self-extinguishing faults are first studied, and based on this data, a decision is made on what thresholds can be used for classifying unstable faults.

To better classify single-phase permanent and self-extinguishing faults on the basis of their stability, the first feature is the maximum value taken from the result of the Fourier comparison over the entire period of the fault. As the analysis is done using a moving window of 1 period, this feature carries information on how close each cycle is to reaching the devised definition of stability. Another important aspect is the number of periods the fault is greater than a threshold that divides single-phase extinguishing and permanent faults. It is important that this value is at least greater than two consecutive periods to ensure stability [7]. Furthermore, a feature is added to carry the information on the fraction of the total time the Fourier ratio is greater than the set threshold. This threshold is set and validated during cross-validation of the model and is explained in Chapter 5.

Additionally, two more features are added to signify the mean and standard deviation of the values of F_{ratio} . With these two features, it is possible to ascertain around what value the Fourier ratios lie and how much deviation exists between them. A larger mean and lower standard deviation indicate that a fault is more stable. These features are summarised in Table 4.6.

Table 4.6: Table summarising the features extracted from the Fourier analysis for the study of unstable faults

Features	Description
f_max	Maximum value of the normalised 50 Hz Fourier Component
f_periods	Number of periods the Fourier analysis is greater than a preset threshold ^a
f_time	Fraction of time the Fourier comparison waveform is greater than a threshold
f_mean	Mean of the normalised 50 Hz Fourier components greater than a threshold
f_stdev	Standard deviation of the normalised 50 Hz Fourier components

^a The threshold is determined during the cross-validation process.

4.4 Conclusion

The introduction to the different classes of faults, preliminary data preprocessing steps and the identification of important features for stable and unstable faults were covered in this section. The next chapter, Chapter 5, discusses the development of the business rules for the classification of unstable faults.

5 Classification of Unstable Faults

5.1 Introduction

In this chapter, the considerations for the classification of unstable faults, as well as the development of the rules for their classification are discussed. In Chapter 4 the feature engineering process for unstable faults is explained. The stability of faults is analysed using the zero-sequence currents from single-phase self-extinguishing and permanent faults. A windowed Fourier analysis of the derivative of the zero-sequence currents was performed for a period-wise analysis of stability. The 50 Hz Fourier component of the waveform was then compared with that of a 50 Hz ideal sinusoidal reference wave.

This chapter outlines the determination of the feature values that differentiate between stable and unstable faults. This setting of boundaries is performed with the use of decision tree classifiers, in their simplest forms, which offer high interpretability [25]. The motivation for choosing decision trees are further explained in Section 5.2, with Section 5.3 introducing the considerations for unstable fault classification. The practicalities for the development of the algorithm with regards to the importance of features and cross-validation results are discussed in Section 5.4 along with the results. Section 5.5 expounds the business rules for unstable fault determination that can be established from the results. Finally, Section 5.6 discusses some of the results and implications of applying these rules to classify unstable faults.

It is important to underline that, while decision trees were used to determine the thresholds for instability, the trees do not make more than a single decision and they are used as a simple thresholding operation. Hereafter, when decision trees are mentioned, unless stated otherwise, they are for comparison with a threshold value, and not a more complex structure.

5.2 Thesholds for the Stability Analysis of Unstable Faults

The principal reason for selecting decision trees for the analysis of unstable faults is their ability to break a complex decision-making problem into smaller and easily interpretable decisions [49]. The problem of classifying unstable faults is different from the classification of stable faults for several reasons. For one, the task of classifying stable faults involves distinguishing between single-phase, two-phase, and three-phase permanent and self-extinguishing faults, while instability is a property of these faults that is characterised by the lack of stability in their waveforms. Additionally, as explained in Section 4.3, the existing definition of instability does not yet contain the nuance to prevent misclassifications of stable faults as unstable faults. This thesis attempts to add certainty to what defines an unstable fault. Furthermore, as unstable faults do not yet have a concrete definition, using a decision tree can help with understanding what characteristics lend to its instability.

5.3 Considerations of the Classifier Algorithm

For this study, single-phase permanent, self-extinguishing, and unstable faults were considered. A limitation of the data is that, while there are sufficient samples of the permanent and single-phase faults, there were only 51 samples of unstable faults that were available for the study. To better define unstable faults, the characteristics of the stable faults were first studied by analysing the spectral properties of the waveforms using sliding Fourier transforms in Section 4.3.2. From this, features from the Fourier analysis were developed for the differentiation between permanent and self-extinguishing faults.

Stability is first established through the Fourier analysis of single-phase permanent faults, and instability is considered as a state that does not meet the constraints that characterise permanent faults. Permanent and self-extinguishing fault waveforms represent two extremes of the nature of a fault. A self-extinguishing fault lasts for half a cycle to two cycles and returns to the steady-state on its own. Permanent faults persist for a longer duration and need to be externally extinguished by a circuit breaker. With the assumption that the data provided for the single-phase permanent and extinguishing faults are correctly labelled, the permanent faults can be considered as stable, and the self-extinguishing faults can be considered as instances of instability. It is also understood that a stable fault will have higher values of the metric F_{ratio} (defined in Section 4.3.2) across more periods due to their longer duration.

A thresholding algorithm is useful to determine the values of the features in Table 5.1 that differentiate the permanent faults from self-extinguishing faults. It should be noted that each "decision tree" consists of only one decision node and two resultant leaves, effectively rendering it a simpler thresholding mechanism. The threshold for the features `f_periods`, `f_time`, and `f_mean` are determined during the cross-validation process to ensure that the optimal value is selected for the discriminating thresholds.

5.4 Development of the Thresholding Algorithm

In this section, the development of the decision tree algorithm for the determination of unstable faults is discussed. The first subsection discusses the contribution of each feature to the target variable, followed by the next subsection that elaborates on the validation of the thresholds for each feature.

5.4.1 Feature Importance

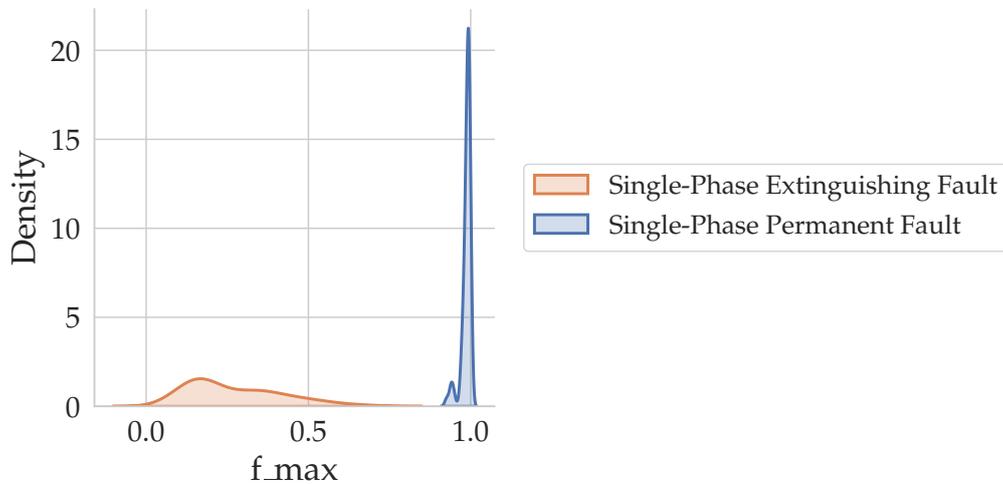
It can be useful to first observe the importance of the features; i.e., how relevant each feature is to make a classification decision between the target classes. As the mutual information technique was selected for the ranking of the stable fault features, the same is used to determine the order of importance for the unstable fault features. The order of importance can be observed in Table 5.1.

It is found that `f_max`, the highest value of F_{ratio} for a fault waveform, is the most important feature that differentiates extinguishing faults from permanent faults. A look into the distribution of the feature `f_max` for the two types of faults in Figure 5.1 corroborates the result of the mutual information test as the feature takes values from two very different distributions. An F_{ratio} of value 1 indicates a perfectly sinusoidal waveform. Feature `f_max` represents the F_{ratio} of the single most sinusoidal period in the fault waveform. The low values of this ratio for self-

Table 5.1: List of features for the characterisation of unstable faults in decreasing order of importance.

Unstable Fault Features (in Descending Order of Importance)	Description
f_max	Maximum value of the normalised 50 Hz Fourier component
f_stdev	Standard deviation of the normalised 50 Hz Fourier components
f_time	Fraction of time the Fourier comparison waveform is greater than a threshold
f_periods	Number of periods the Fourier analysis is greater than a preset threshold
f_mean	Mean of the normalised 50 Hz Fourier components greater than a threshold

extinguishing faults, and high values for permanent faults indicate that these two classes differ widely in their most sinusoidal period. The second most important feature from the mutual information test is f_stdev , another feature that is highly discriminating as a result of the different distributions of F_{ratio} .

Figure 5.1: Comparison of the distributions of the f_max feature of single-phase permanent and self-extinguishing faults.

Features f_mean , $f_periods$, and f_time are each a function of the Fourier ratio within a certain percentage of f_max . For a stable fault, the values of f_period will be high (or more than two, in accordance with [7]) as there are a larger number of periods that are similar in magnitude to the most stable period. A self-extinguishing fault consists of a transient for half a cycle up to two cycles, and hence has a lower number of periods that are as stable; hence, $f_periods$ will also be correspondingly smaller. Similarly, f_mean , and f_time , each represents the average value of metric F_{ratio} above a certain threshold, and the fraction of the total time for which the fault has higher F_{ratio} values than a threshold respectively.

As stable faults will have a higher number of periods that are stable (more than two periods), the thresholds for each feature must be set to a high value so it can help with discerning stable

from unstable states. Cross-validation using the decision tree algorithm is hence performed in the following section to arrive at the optimal value for the thresholds.

5.4.2 Cross-Validation for Determining Thresholds

To ensure that the values of the features f_{\max} , f_{stdev} , f_{time} , f_{periods} , and f_{mean} are optimised for enabling the identification of unstable behaviour, five independent thresholding models are developed. Each "tree" for thresholding consists of two leaves and a decision node that indicates the feature value that differentiates self-extinguishing faults from permanent faults.

The sample space of 160 fault samples consists of 80 single-phase permanent and self-extinguishing faults each. The data is shuffled and split in a ratio of 85:15 between training and testing data, respectively. The cross-validation procedure is performed on the training data alone, and the resultant decision tree model is tested on the remaining 15% of the data that was not used for training. From here, the performance of the model can be evaluated. For each of the five decision trees, the feature values are optimised using cross-validation to ensure each feature possesses optimal discriminating ability.

A result of this particular implementation of a decision tree model is the value of the feature or threshold, that yields the best split between classes. The decision tree is first trained on the training data to learn the value of the feature that is most discriminating between the classes and tests the true differentiating ability of the feature value on the remaining test data. In an ideal case, the value of the feature will result in an accuracy of 100%. This implies that, while testing, none of the faults were misclassified on the basis of that particular feature (and its value). Features f_{\max} and f_{stdev} are a function of F_{ratio} for the entire period of the fault waveform. For each of these two features, a two-leaf decision tree is trained on the value that optimally separates the two classes. These values can be observed in Table 5.2. It should be noted that the cross-validation score for f_{\max} is 100% indicating that this feature is perfectly capable of distinguishing between self-extinguishing and permanent faults in the training set. The cross-validation score for f_{stdev} is also 98% indicating its high discriminating ability.

5.4.3 Validation of the Thresholds

The features f_{time} , f_{periods} , and f_{mean} are handled in a different manner than the first two features as they are also a function of f_{\max} . First, the f_{\max} threshold percentage that is considered for each f_{time} , f_{periods} , and f_{mean} must be determined. The performance that is judged using cross-validation is the f1-score. For each feature, the 5-fold cross-validation is performed for different fractions of f_{\max} and the value that results in the highest cross-validation score is selected as the threshold for the feature. Subsequently, the value of the feature at the percentage of f_{\max} is taken to be the optimal discriminating value between the two classes under consideration. These values, along with their cross-validation scores, and the thresholds can be observed in Table 5.2.

From Table 5.2, it can be observed that the feature f_{time} has significantly lower cross-validation scores than the other four features. This is not in accordance with the order of important features determined through the mutual information test in Table 5.1. This could be due to the fact that the feature importances were determined using a decision tree model that uses the five features together, but the cross-validation is performed using five decision tree models that are each optimised for each feature. The choice was made to optimise five different decision tree models for each feature, as using one decision tree with all five features resulted in feature f_{\max} alone being responsible for a 100% accuracy in the classification of the test cases. It would be remiss to

Table 5.2: Cross-validation scores and values of each feature determined by the decision tree algorithm.

Criteria for Stability	Threshold (% of f_{\max})	Cross-Validation F1-Score (in %)	Test F1-Score (in %)
$f_{\max} > 0.8$	-	100	100
$f_{\text{stdev}} > 0.123$	-	98	99.2
$f_{\text{time}} > 0.015$	90	93.7	91
$f_{\text{periods}} > 2.5$	93	95.3	97
$f_{\text{mean}} > 0.011$	94	96	97

use only one feature to determine stability as this perfect score might be specific to the data set at hand. For this reason, the other four features are used to determine stability as well.

5.5 Business Rules for Unstable Fault Classification

The five individually optimised decision trees are then tested based on the feature values and thresholds that were calculated from the cross-validation process. The test f1-scores can be seen in Table 5.2. Feature f_{\max} has a test f1-score of 100% that confirms the hypothesis that the maximum value is capable of completely separating single-phase permanent and self-extinguishing faults. This high accuracy can also be a function of the relatively smaller sample space for the two classes. It is therefore recommended to not utilise just one feature to distinguish between the two classes, but all five features.

A scoring system can be developed to assess how many counts the fault features comply with the criteria for stability. As stated in Section 5.3, single-phase permanent faults are considered to be stable faults and single-phase extinguishing faults are considered to be an extreme case of unstable faults. Therefore, a single-phase permanent fault would have a score of 5, and a single-phase extinguishing fault would have a score of 0. The scores for each count of features can then be used as new criteria to re-categorise the faults that were initially classified as unstable. It can be stated with certainty, that unstable faults with a score of 5 are actually stable and permanent. Furthermore, it can be reasoned that unstable faults with a score of 0 are self-extinguishing faults.

A fundamental aspect of faults classified as unstable is the distortions that prevent the original classifier from making correct decisions on the type of fault. For example, a single-phase fault that was classified as unstable due to imperfectly sinusoidal waveforms can be found to be stable after manual inspection. To make the distinction less strict, i.e. to prevent misclassifications of this nature, it can be useful to allow a slightly broader margin for what classifies as a stable fault. For this purpose, it was determined that unstable faults that have a score of 4 can be classified as stable, with the underlying condition that the two most important features f_{\max} and f_{stdev} meet the criteria for stability. These two features were chosen as they have the highest importance in determining the target class (from the mutual information test), as well as high (and similar) cross-validation and test f1-scores, as shown in Table 5.2. If an unstable fault possesses a score of 4 but does not meet the criteria for the values of f_{\max} and f_{stdev} — further manual inspection is deemed necessary.

In a similar manner, unstable faults which have a score of 1 can be classified as self-extinguishing faults — with the prerequisite that f_{\max} and f_{stdev} both have lower values than the criteria for stability. Similarly, if one (or both) of the two features does meet the stability criteria, then further manual inspection is recommended.

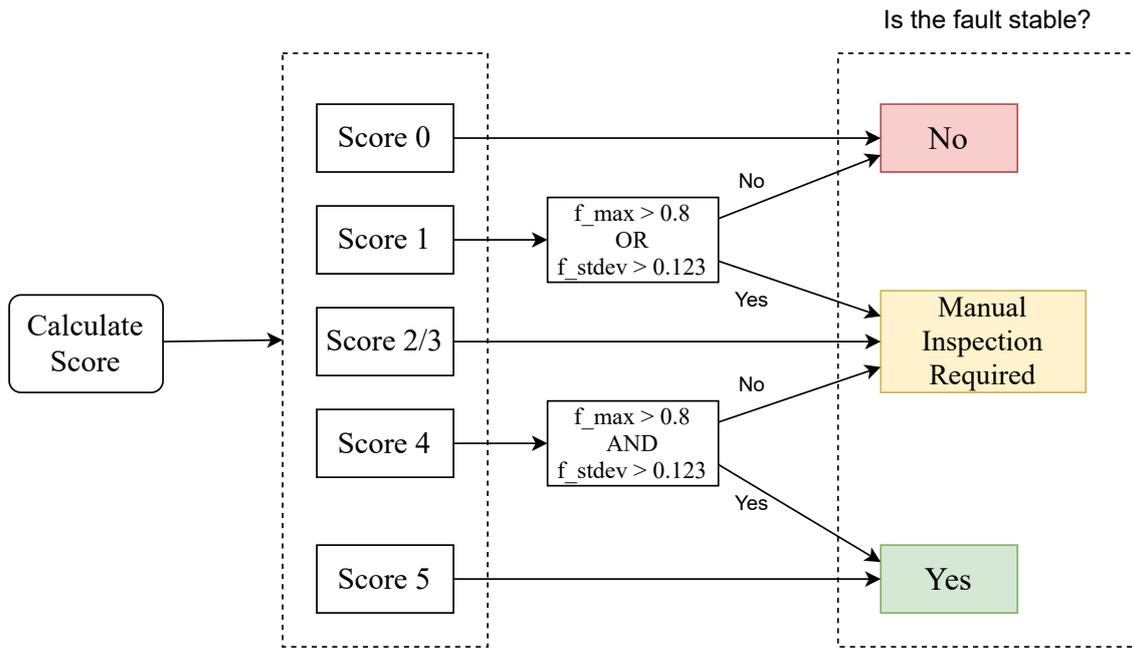


Figure 5.2: Depiction of the rules developed for the determination of acceptable unstable behaviour in a fault.

It is not attempted to classify unstable faults that have scores 2 or 3 as either stable or self-extinguishing. This decision is motivated by the fact that it is vital to only send the information of faults whose nature is certain to the control center for further network restoration. It is hence recommended that unstable faults that score a 2 or 3 are to be manually studied for their fault type. These business rules are summarised in Figure 5.2.

5.6 Application of the Rules to Unstable Fault Data

This section discusses the results of applying the developed rules to the test set. The test set consists of 51 faults that were labelled as unstable faults by Alliander's stability rules. Since the older method of fault detection sometimes classified stable faults as unstable, the business rules defined in this chapter attempt to improve the definition of (in)stability. In this section, it is studied if the rules developed in the thesis can automatically detect faults that exhibit stable behaviour that is recognisable to an expert but not by the older classification rules.

The objective of this exercise was to increase the rate of identification of faults that are stable, i.e., to increase the number of "true positives" of stable faults so they can be correctly isolated from the rest of the network. With the rules, it is important to catch more stable faults without causing an increase of false positives. False positives result in truly unstable faults being mistakenly classified as stable.

The rules classify faults based on their stability as being either definitely stable, definitely unstable, or "requiring manual inspection". This final class consists of faults that have the possibility of being stable through expert inspection. The rules were applied to classify the test set at hand and the results are presented in Table 5.3. It can be observed that 10 faults have been classified as being definitely stable. This implies that within the subset of faults classified as unstable, there are characteristics from the moving-window Fourier analysis that can indicate that these faults are actually stable. It was corroborated with expert opinion if the faults con-

sidered stable with the analysis developed in this thesis can be rightfully considered so.

Table 5.3: Table describing the classification of unstable faults based on their stability.

Number of Faults	Stability Decision		
	Stable	Manual Inspection Needed	Unstable
	10	23	18

Some examples of stable faults can be observed in Figure 5.3. Figures (1), (2), and (3) are a subset from the 10 stable faults from from Table 5.3. These were faults that were initially identified as unstable by Alliander’s algorithm but found to be stable by the new rules. Image (1) is the fault in Figure 4.11 from Section 4.3. It can be observed that the fault has been found to be stable with the new rules, validating the fact that the fault does indeed stabilise during the later cycles before extinction. Images (2) and (3) in Figure 5.3 are also examples of single-phase permanent faults that are evidently stable and have been identified as being stable by the new rules.

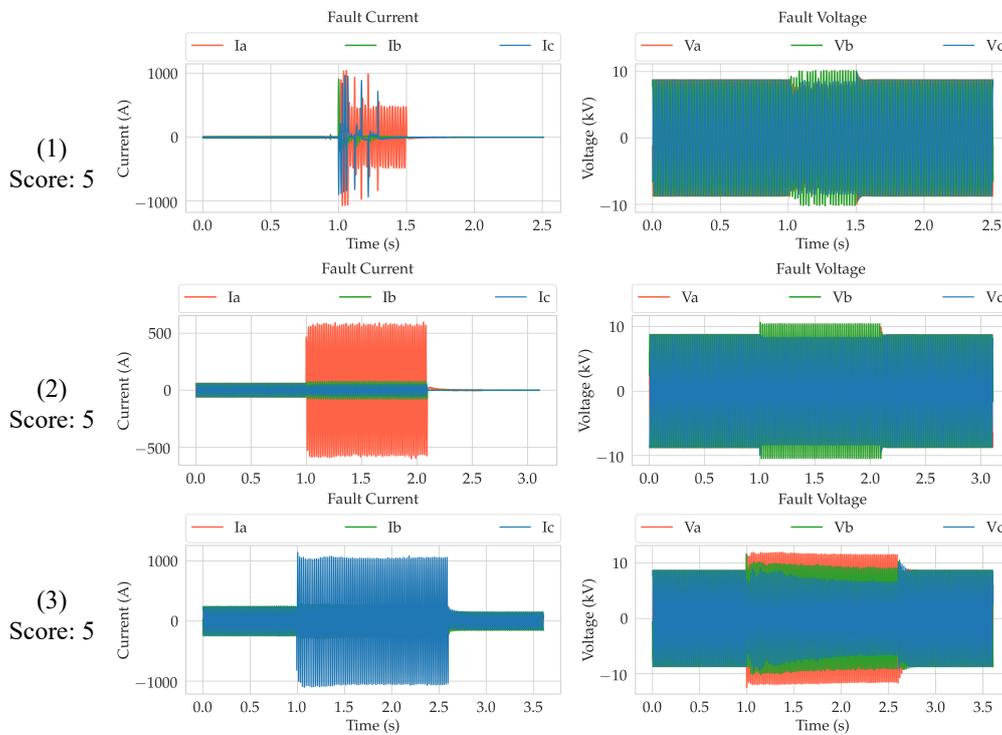


Figure 5.3: Subset of stable faults that are a result of the Fourier analysis classification process.

Figure 5.4 represents faults that have not been considered stable from the Fourier analysis. In image (4), it is evident that the unstable classification arises from the distortions that occur during the zero-crossing, particularly in phase B. In image (5), it can be observed that the fault waveform consists of a series of self-extinguishing faults. The continuous extinguishing and restrike behaviour of the fault results in the unstable classification as self-extinguishing faults were considered to be an extreme case of instability in order to develop the rules and train the classifier.

Image (6) shows an example of a fault that scored a 3, indicating that it requires manual inspection to determine stability, thereby leaving room for the possibility that the fault is stable. Correspondingly, between 2.5s and 3s, it can be observed that the current and voltage at phase B exhibit the behaviour of a single-phase fault, due to the larger currents sustained for longer than 2-4 cycles. However, it is also seen that there exist shorter periods of current restrike and

extinction after around 3.5s. While the fault does not develop into an actual permanent fault, the waveforms indicate longer periods of large currents and hence could warrant manual inspection into the nature of the fault.

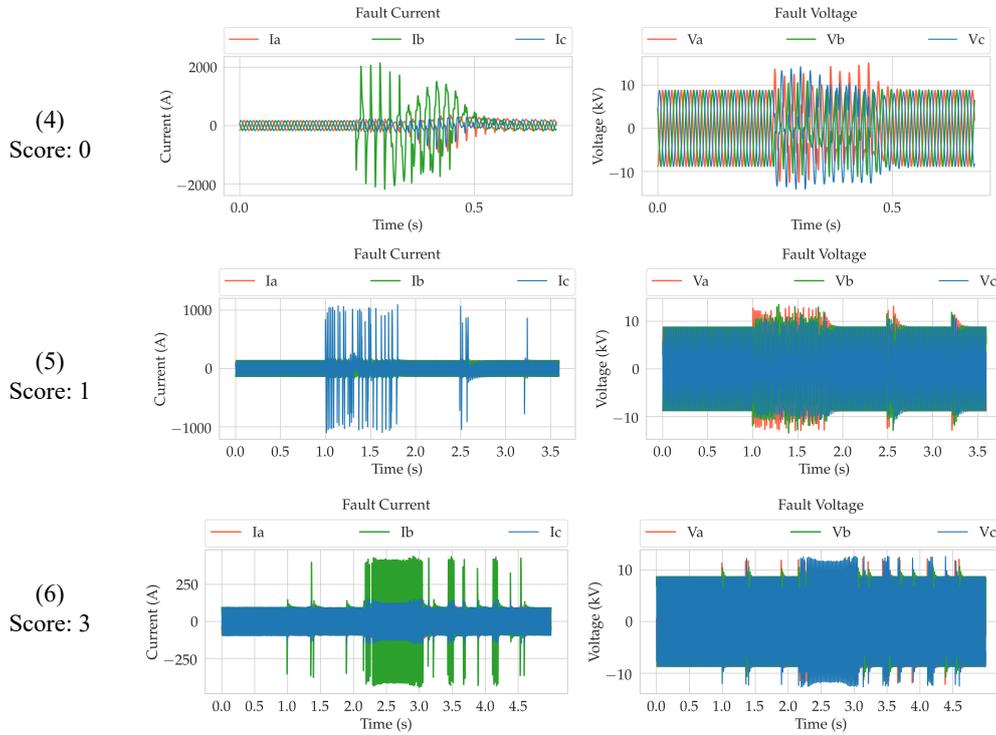


Figure 5.4: Subset of (possibly) unstable faults that are a result of the Fourier analysis classification process.

It can be interesting to observe the variation of the metric F_{ratio} between the faults that were identified as stable and unstable. This can be observed in Figures A.2 and A.3 in Appendix A.

To summarise, rules were developed by performing a windowed Fourier analysis of faults to develop an improved method to characterise stability. It has been seen from Figure 5.3, that these rules are effective in detecting stable and permanent faults when they were previously misclassified as being unstable. This helps to increase the number of faults whose correct locations are sent to the control center for fault mitigation in the distribution network. Furthermore, some examples of the rules "confirming" the unstable nature of faults are presented with Figure 5.3.

6 Classification Model for Stable Faults

6.1 Introduction

This chapter discusses the development of the classifier model for stable faults. Broadly, this consists of three steps - feature selection, hyper-parameter tuning, and cross-validation.

In Chapter 4, several features were engineered and considered important for differentiating between faults. Feature selection is the process of narrowing down the number of features used to develop a classification model. The different methods of feature evaluation techniques used in this thesis are — ANOVA-F, mutual information, and the sequential feature selection method. The ANOVA-F and mutual information (MI) techniques are filter methods of feature selection. Filter methods make use of statistical tests to evaluate the relationship between a feature and the class by giving each feature a score. These scores can then be used to select the most important features for the model. Another test must be done to find the minimum number of the ranked features that result in the optimal model. On the other hand, wrapper techniques like sequential feature selection create multiple classifier models with different subsets of the features and provide information on which feature subset results in the best performance. By adding or removing features, they find the optimal feature combination that maximises the accuracy of the model [50].

As mentioned in Chapter 3, SVMs separate non-linear data by making use of kernel functions to map the data to higher dimensional spaces where they can be separated. The user needs to set the parameters for the kernel function (known as hyper-parameters). These values of the hyper-parameters can affect the performance of the machine learning model significantly. Knowledge of the model and trial-and-error is required to set the hyper-parameters on a case-by-case basis [51]. In this thesis, a study is done to select the optimal hyper-parameters, C and γ , for the support vector machine. To ensure that the hyper-parameters selected are indeed the ones that yield the best classifier performance, their values are tuned over a search space that checks the performance of the model over different combinations of C and γ .

Cross-validation is performed to assess the performance of the model before testing it on new and unseen data. It can increase the classifier model's generalisation ability. It uses data resampling to create smaller train and test sets, and the performance in these smaller sets can give an idea as to how the model will perform on data that was not used to develop the model [23][52].

These three processes, feature selection, hyper-parameter tuning, and cross-validation, are vital to improving the performance of the classifier model. If they are performed in discrete steps, it is possible that the output from one step influences the performance of the next step in a way that increases the bias of the model. For example, if feature selection is performed first, and the hyper-parameter tuning next, it is possible that the tuned hyper-parameters are optimal for that particular feature set only (and vice versa if performed in the other order). To reduce

the bias, a function pipeline is created where, for each cross-validation test, feature selection and hyper-parameter tuning are performed simultaneously.

This chapter first discusses the application of the feature selection techniques in the context of this thesis in Section 6.2.1. This is followed by a detailed explanation of the hyper-parameter tuning and cross-validation process in Section 6.2.2 and Section 6.2.3. Section 6.2.4 elaborates on how these processes are implemented simultaneously. Section 6.3 introduces the motivation and methodology for selecting the optimal mother wavelet for this thesis. Lastly, the results of the model parameters are presented and discussed in Section 6.4. The section concludes with the recommendations for the parameters of the support vector machine classifier.

6.2 Overview of the Model Development Process

6.2.1 Feature Selection

Feature selection is the process of reducing the set of identifying features into a smaller set that remains representative of the data that is being studied. In the case of this fault classification problem, 36 features have been identified that carry identifying information about the fault signals. In order to transform the classification algorithm into a more general concept — i.e. one applicable to a larger data set that was not used to model it — it can be useful to reduce the number of features used to develop the model. A redundant feature can, at best, add no value to the classifier in terms of classification performance, and at worst, increase the program run times and result in an overfit model. The goal of feature selection is to realise a smaller set of features from the original list that does not reduce the accuracy of the classifier and have a distribution that is similar to the original distribution of the classes [53]. It has been studied that the accuracy of SVM classifiers, among others, is adversely affected with the inclusion of more features. This holds true, particularly for smaller data sets [54].

First, the feature selection methodology for the filter methods of feature selection (ANOVA-F and mutual information) is discussed. These two techniques involve using statistical tests to rank the importance of the features based on how much they contribute to the target class. They are useful for evaluating the relative importance between features, but not the number of features that are optimal for the classifier performance. To facilitate this, the feature selection is performed simultaneously during the hyper-parameter tuning and cross-validation processes.

Further, from the results of the selected features from the ANOVA-F and mutual information tests, an analysis of the correlation between the features is performed. Correlation is the strength of the linear relationships of variables, and in this case, features. It is desirable that the classifier model is as simple as possible, or in other words, the classifier should be able to classify data with the smallest number of features as possible. To reiterate the merits of a “simple” classifier, the fewer the features that can perform the classification, the more general the classifier will be for future and unseen data. For this reason, the correlation between the selected features of each case is analysed and depending on the degree of positive correlation, removed. This is done to further narrow down on a feature set that can characterise the differences between fault types.

Sequential Feature Selection

The sequential feature selection technique must be integrated with the tuning and validation process differently as it is a self-contained method that performs validation on its own. This method iteratively searches through combinations of feature subsets to find the optimal features.

For each subset of features, the performance of the subset is validated using an instance of the classifier model. The parameters for this classifier must hence be set *a priori*, without tuning. For the case at hand, including hyper-parameter tuning with this particular wrapper technique was found to be computationally expensive. Hence, the hyper-parameters are set preemptively using values from experience, and the feature-selection and cross-validation processes are performed simultaneously instead.

Sequential feature selection is a wrapper, or self-contained feature selection process that learns which features contain the most information at each timestep where the selection of the next feature is contingent on the previously picked features and the performance of the classifier using the subset of features alone [55]. The feature selection algorithm consists of a search algorithm that searches the feature subset space, an evaluation function that the search algorithm maximises, and a performance function which, in this case, is the classification task.

There are two types of sequential feature selection techniques — forward feature selection and backward feature selection. The forward feature selection process starts with zero features and consecutively adds a feature that results in the best classifier performance. This continues until there is no improvement in the performance of the classifier with the addition of the next best feature. The backward feature selection process starts with all the features and removes one feature at a time until the classifier improvement is maximised. The results of both the forward and backward feature selection will vary, but a high degree of difference between the two can imply the features do not sufficiently represent the classes [56]. For this thesis, forward feature selection was the technique that was considered.

6.2.2 Hyper-Parameter Tuning

Hyper-parameter tuning is performed during the model development task to select the optimal values of the SVM radial basis kernel parameters, C and γ . Kernels in SVMs map linearly inseparable data to a higher dimension where they are separable by a boundary. C controls the error margins of the SVM. If the value of C is too high, the model will face a high penalty for misclassifications. The class boundaries will then be highly dependant on the outliers of each class, and not the more representative data points in the classes. γ is a kernel parameter that sets the range of influence of each data point. Smaller values of γ indicate that data points have a larger range of influence, i.e. more data can be grouped as being in the same class. Large values of γ only group data points that are very close to each other. If γ is too low, the rate of misclassifications can increase due to an overly general boundary; if γ is too high, the model could fail to learn from outlier data. The values of C and γ are hence very specific to the data that the SVM is trained on, and a balance must be made between bias and variance [23].

For each hyper-parameter, a search space is set in the form of a grid, where combinations of C and γ are applied to the SVM for assessing the performance. A logarithmic grid between 10^{-3} and 10^3 for each kernel parameter is deemed sufficient. If the best parameters lie on the grid's boundaries, a future search can be extended in that direction [57].

6.2.3 Cross-Validation

The dataset is split into two sets — 85% of the total is used to train the data, and 20% is held out to be used for model testing. It is important to validate the trained model before testing. To maximise the use of the limited data, instead of splitting the train set further for a validation data set, nested cross-validation techniques like k -fold cross-validation can be performed on the entire data set used for training the model.

With k -fold cross-validation, the training set is split into k equally sized “folds” which are sampled randomly from the entire data set. From here, for each of the k folds, $k - 1$ folds are used as the training data. The resulting model is validated on the remaining data which is used as a test set. The results of the cross-validation process are the model parameters that lead to the best average of the training performance results for each loop [57]. A pictorial representation of k -fold cross-validation is shown in 6.1.

6.2.4 Unifying the Model Selection Processes

To accurately assess the performance of a model, i.e. not the result of the performance of the classifier itself, but its validity in being a model that can apply to future unseen data, it is important that information about the test set does not leak into the model training process [57]. Information from the test data affecting the model parameters can decrease the generalisation ability of the classifier. Therefore, to have an optimally performing model, it should be ensured that the performance is not at the cost of an overfit model.

If feature selection, hyper-parameter tuning, and cross-validation are performed one after the other, it is possible that the features selected are optimal for the test set at hand alone. This could create a cascading leak of test set information into the following steps. For this reason, these three processes are performed simultaneously, as shown in Figure 6.1. For each combination of hyper-parameters and features, the 5-fold cross-validation, feature selection, and hyper-parameter tuning processes are executed. The optimal solution is the one with the best average of the cross-validation f1-score of each fold. Furthermore, as stated in Section 6.2.1, the filter feature selection techniques do not provide information on how many features should be selected. Therefore, for every run of the loop, an increasing number of features are selected. With the results of this cross-validation method, it can be determined how many features result in the optimal model. The features are automatically ranked in terms of importance during the cross-validation process.

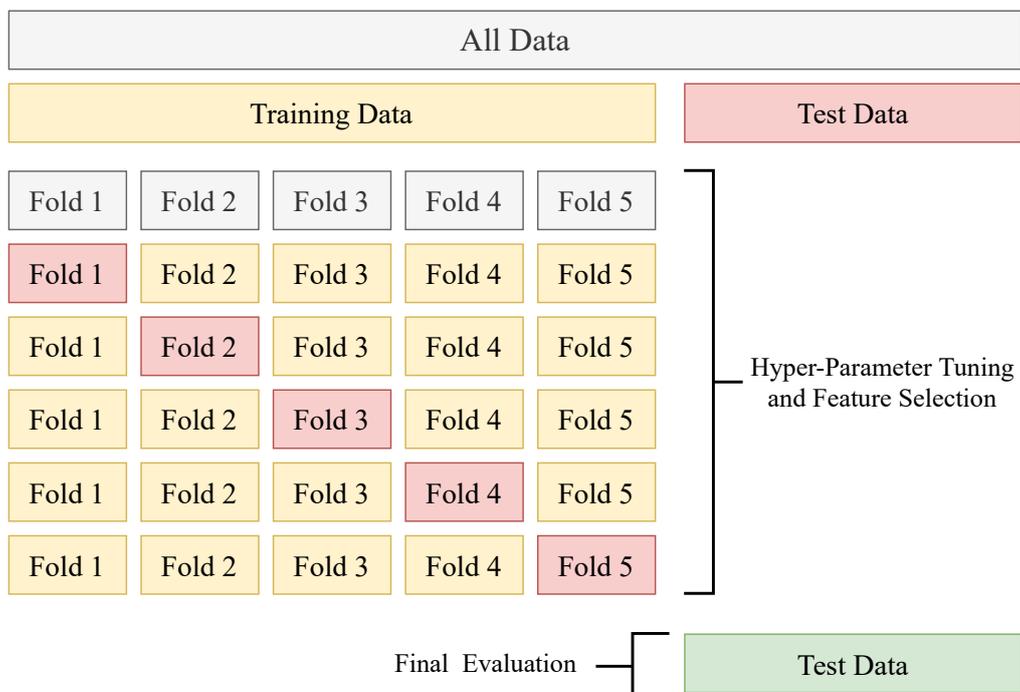


Figure 6.1: Pictorial representation of cross-validation. Feature selection and hyper-parameter tuning occurs at each fold.

To optimise the search process over every set of hyper-parameter combinations, Python's `Optuna` package was used. This is an improvement to the simpler grid search algorithm that checks over every single combination. `Optuna` enables parallel processing and works by maximising an objective function (the SVM performance), subject to constraints, which are the hyper-parameters [58].

In the case of the wrapper feature selection method, sequential feature selection, this process is slightly altered. As this technique already requires a set of hyper-parameters to instantiate an SVM model for each iteration, the tuning cannot be performed in the same way as with the filter methods. Therefore, the results of this method are only used to corroborate the results of the filter methods and to see if the results vary drastically with the different techniques.

6.3 Selection of the Mother Wavelet

An integral aspect of feature engineering using DWT is the selection of the mother wavelet, or wavelet filter. Each wavelet captures transients in a different manner, due to their coefficient values, and their filter length. In Section 2.4, a set of five mother wavelets were identified as being the most popular choices for the study of electrical fault waveforms. Out of the five mother wavelets — db4, db7, db8, sym5, and cubic b-spline, it was found that the db4 was the most frequently chosen mother wavelet for its ability to capture transients in electromagnetic signals [8] [34].

In accordance with the findings of the literature survey, the db4 wavelet was primarily used for fault feature extraction. It is, however, useful to identify if the results of the literature survey indeed hold true, by comparing the performance of the classifier using the db4 wavelet with the performance of the classifier when the other four wavelets are used separately. As this process is too computationally expensive to be executed for the purpose of determining the right wavelet alone, this section describes a method for retroactively determining if the db4 wavelet is indeed the right choice for this classifier.

The choice was made to perform a comparison of the performance scores of the classifier using the different wavelets. Therefore, for each wavelet, features are engineered and selected, SVM hyper-parameters are optimised, and cross-validation is performed. The performance of the different wavelets was studied as a comparative analysis, and not for actually selecting the optimal wavelet for the classifier model. To preemptively make the best choice of mother wavelet for the classifier, a more rigorous analysis of the different types of wavelets, and their properties must be performed, which is considered to be beyond the scope of this study. The results of this comparative analysis are hence elaborated in Section 7.3 of the Results chapter.

6.4 Results of the Model Selection

This section presents and discusses the results of the model development process. The hyper-parameter tuning and cross-validation steps remain the same, with different feature selection techniques being compared. First, a comparison between the ANOVA-F, and mutual information statistical tests are made using the performance of the classifier during the cross-validation process.

The cross-validation scores versus the number of features for the filter methods, ANOVA-F and mutual information, are plotted in Figure 6.2. It can be observed that, initially, with the increase in the number of features, the f1-score increases, and then the score remains more or

less constant even as new features are added. The first highest f1-score occurs at 9 features for the ANOVA-F technique and 12 features for the mutual information technique. The features that must be chosen are based on the feature importance scores calculated by each method. Including more than this subset of features can result in a marginal improvement in test scores at the cost of increased computation times and overfitting. The scores for each subset of features along with the SVM hyper-parameters for both the methods can be referred in Tables B.1 and B.2 in Appendix B.

Cross Validation Scores: F1-Score vs. Number of Features

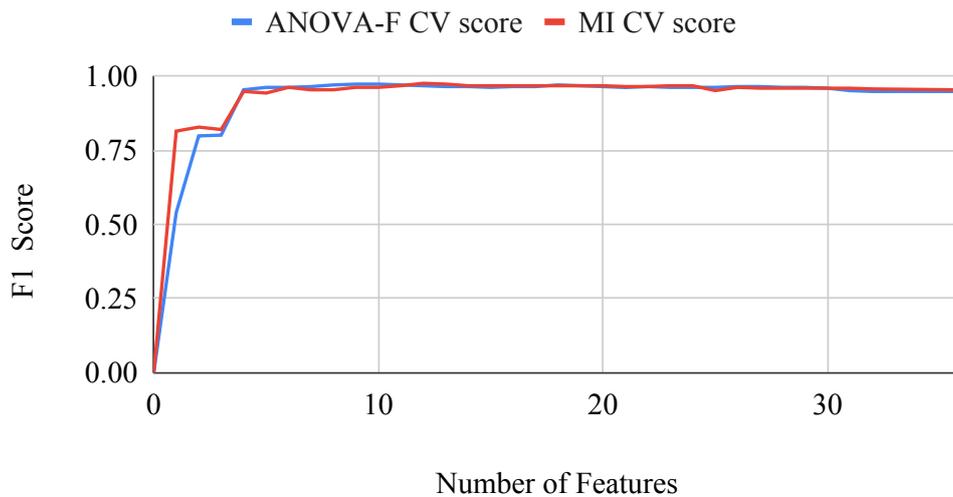


Figure 6.2: Comparison of cross-validation scores using ANOVA-F and Mutual Information (MI) feature selection techniques.

6.4.1 Results of ANOVA-F Scoring Feature Selection

This feature selection method uses the ANOVA-F test to statistically determine (and rank) the importance of each feature after which the performance of the classifier (on the training) is assessed for the incremental addition of each feature during the cross-validation process.

The nine features were selected based on the features that had the highest scores from the ANOVA-F ranking technique. These features can be observed in Table 6.1. It can be observed that the most important feature for the fault classification according to the ANOVA-F test is `irms_3`. As mentioned in Section 4.2, the RMSs are decoupled from the actual phases (a,b,c) and are instead ordered in terms of their magnitudes. The feature `irms_3` represents the smallest RMS value of the three phases. It can also be seen that four features related to the standard deviation and RMS of the zero-sequence current and voltage — `V0_regular_rms`, `V0_regular_std`, `I0_regular_std`, `I0_regular_std` — were selected. This corroborates with the idea that zero-sequence values are vital to the identification of single-phase faults [10]. It should be noted that this could also be due to the fact that single-phase faults (permanent and self-extinguishing faults combined) form the largest group in the sample space. The results of the feature selection may hence be skewed to represent single-phase fault features. It can, however, be stated that this is an unavoidable result of the amount of reliable data set available for the study.

Table 6.1: List of features selected using the ANOVA-F ranking method (ordered in descending order of importance).

ANOVA-F Method Features
irms_3
irms_2
I0_regular_std
I0_regular_rms
V0_regular_std
V0_regular_rms
irms_1
i_l2_std3
i_l1_std3

ANOVA-F: Analysis of the Features Selected

The correlation of the features, i.e. how similar the selected features are to each other, is presented in Figure 6.3. The higher the positive correlation between features, the stronger is the linear relationship between them; therefore, highly correlated features will have a similar effect on the dependant variable/class.



Figure 6.3: Correlation between the top nine features ranked by the ANOVA-F test.

From Figure 6.3, it can be observed that some features are highly correlated with each other, notably feature pairs (irms_1, irms_2), (I0_regular_std, I0_regular_rms), (V0_regular_std, V0_regular_rms), and (I_l2_std3, I_l1_std3). In each pair of features mentioned, the feature that is ranked lower than the other is removed from the feature set. This is done with the motivation of reducing the number of redundant features thereby decreasing the computation time (in calculating these features) and increasing the generalisation ability of the classifier.

In the case of the features that represent the RMS and standard deviation of certain aspects of the fault waveforms (such as the zero-sequence currents, or detail coefficients), it is apparent that both values do not independently represent unique aspects of the signals as they are both a function of the sum of squares of the signals. It is also reasonable that i_rms1 (largest RMS of the current) is highly correlated with i_rms2 (the second-largest RMS of the current), as two-phase

and three-phase faults both have similar values their largest and second-largest RMS values of their phase currents.

For the above reasons, features `i_rms1`, `V0.regular_rms`, `I.l1.std3` and `I0.regular_rms` are removed from the feature set, with Table 6.2 representing the updated feature set. It can also be observed from Figure 6.3 that features `I0.regular_std` and `V0.regular_std` are highly correlated, but the choice is made to keep them in the final feature set as it is considered useful to have both voltage and current information.

Table 6.2: Updated of features selected using the ANOVA-F ranking method.

ANOVA-F Method Features (Updated)
<code>irms_3</code>
<code>irms_2</code>
<code>I0.regular_std</code>
<code>V0.regular_std</code>
<code>i.l2.std3</code>

ANOVA-F: Analysis of the Hyper-Parameters

The result of the tuning of the hyper-parameter values is $C = 1000$ and $\gamma = 0.1$, where C represents the penalty for misclassified faults, or the bias, and γ is the hyper-parameter controlling the variance of the SVM. These hyper-parameters are selected by the optimisation algorithm `Optuna` during the cross-validation process. It can be noted that the selected value of C is the highest in the provided search space, indicating that the SVM model has a high penalty for misclassified faults. The implications of this value are evaluated together with the results of the mutual information feature selection method in Section 6.4.3.

6.4.2 Results of Mutual Information Scoring Feature Selection

This section discusses the results of the feature selection and cross-validation process when mutual information (MI) is used as the method for ranking features based on their importance. Based on the performance of the classifier during the cross-validation process, the highest performance on the training set was at 12 features as shown in Figure 6.3. Mutual information ranks features based on quantitatively determining how much information (derived from entropy) each feature contains about the classes.

The highest cross validation f1-score, which is at 12 features, is 97.54% with hyper parameters $C = 100$ and $\gamma = 1$.

The cross-validation scores indicate that twelve features result in the highest performance when the f1-score is used as the scoring metric. These twelve features are shown in Table 6.3. The following subsection discusses the features that have been selected.

MI: Analysis of the Selected Features

It can be observed that the most important feature when features are ranked using mutual information, is `irms_comparison`. Feature `irms_comparison` was developed to capture, in one

Table 6.3: List of features selected using the mutual information ranking method (ordered in descending order of importance)

MI Method Features
irms_comparison
irms_2
irms_3
irms_1
V0_regular_std
V0_regular_rms
I0_regular_std
I0_regular_rms
i_l1_std2
i_l1_e3
i_l2_e3
i_l2_e2

metric, the relationships between the three current phases. It can also be noted that features $i_{rms1/2/3}$ are also among the top features in the mutual information ranking. This is because, while the mutual information technique is able to capture non-linear dependencies between the variables and classes, it does not capture the redundancy of information between the features themselves. This is why a study of the correlation between the features is performed.

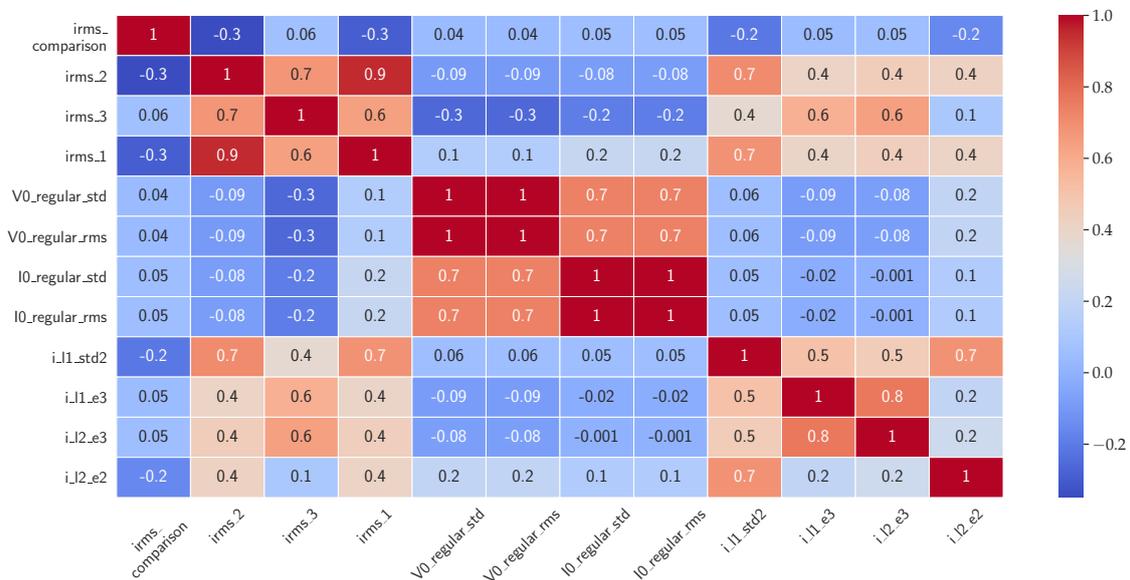


Figure 6.4: Correlation between the top 12 features ranked by the mutual information test.

The correlation matrix for the comparison of the linear relationships between the selected features can be seen in Figure 6.4. With the same reasons mentioned in previous subsection 6.4.1, the highly correlated features i_{rms1} , $V0_regular_rms$, $I0_regular_rms$, i_{l2_e3} are removed from the feature set, resulting a final feature set as shown in Table 6.4.

Table 6.4: Updated of features selected using the mutual information ranking method.

MI Method Features (Updated)
irms_comparison
irms_2
irms_3
V0_regular_std
I0_regular_std
i_l1_std2
i_l1_e3
i_l2_e2

MI: Analysis of the Hyper-Parameters

The hyper-parameters that were selected as part of this cross-validation were $C = 100$ and $\gamma = 1$. These values lie comfortably between the boundaries of the search space provided for their optimisation, and it can hence be stated that these values pose less of a danger of overfitting than the hyper-parameters in the case with the ANOVA-F method. A deeper discussion on the comparison of the mutual information and ANOVA-F techniques is performed in the following section.

6.4.3 Comparison of ANOVA and MI Feature Selection Methods

From the performance of the classifier on the training set as seen in Figure 6.2, it can be observed that the performance increases with an increase in the addition of the features. It can be concluded that the "step"-like nature of the performance chart for both the methods can be attributed to correlated features being added one after the other, resulting in a local plateau of training performance.

With regards to the actual features that have been ranked, it can be observed that the mutual information technique is superior in identifying non-linear relationships between features and their classes, as evidenced by the top ranking of feature `irms_comparison` by the mutual information technique.

Finally, from the results of the hyper-parameter tuning, it was observed that the value of C for the ANOVA-F test was the largest in the search space. Upon increasing the upper limit of the search space, it was still found that C was chosen at the boundary, up to values of 10^6 . At this point, this investigation into the value of C for the ANOVA-F test was halted. As high values of C result in a classifier that works very well on the given data set, but possibly not as well on unseen data, it can be concluded that the method involving the mutual information technique can be used for the model selection.

6.4.4 Results of Sequential Feature Selection

Sequential feature selection requires the initialisation of the hyper-parameters before the feature selection and cross-validation processes. The hyper-parameters are hence loosely set based on the results of the cross-validation with the filter methods. As lower values of hyper-parameters result in a more general model [57], the lower values of C and γ are chosen between the two

previous results from the filter methods. The value of C is set as 100 and γ as 10^{-1} . With the filter feature selection techniques, there is an outer loop that validates the performance of the model with the number of features, and an inner loop that performs the k -fold cross-validation. To similarly ensure that different combinations of the training data set are considered for the cross-validation, the sequential feature selection is performed in 10 trials. The results of each trial are the cross-validation f1-score and the optimal (number of) features. A pictorial representation of the cross-validation scores for the 10 trials can be observed in Figure 6.5.

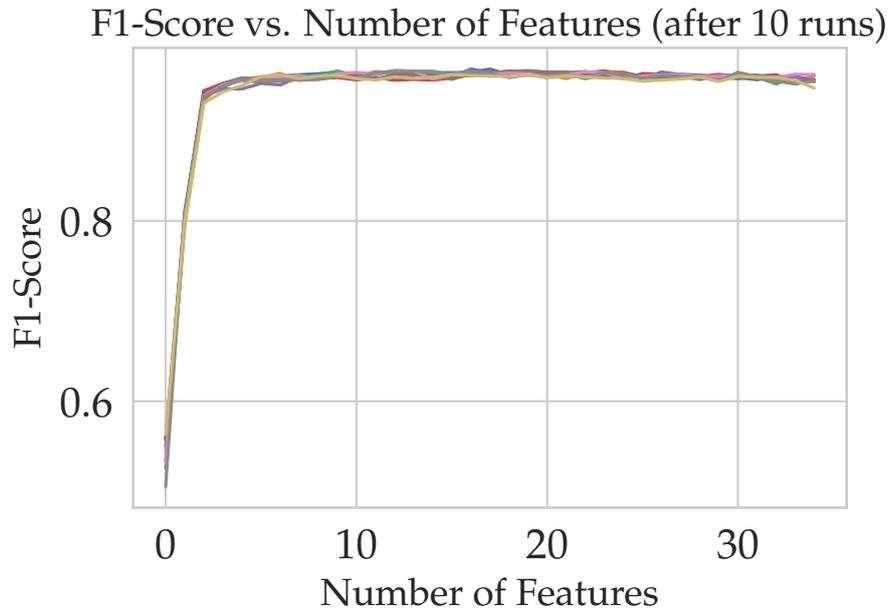


Figure 6.5: Cross-validation scores of the sequential feature selection techniques.

Table 6.5: List of features selected using the sequential feature selection technique

Selected Features
i_l1_std1
v_l1_std3
irms.1
irms.2
v_l2_e1
V0_wavelet_rms
I0_wavelet_std
V0_regular_std

The sequential feature selection technique has a cross-validation f1-score of 96.16% with 8 features. These features are listed in Table 6.5. It can be observed that, as a method that does not deterministically establish feature importance before ranking, the RMS values are no longer the top features.

6.5 Selection of the Model

Between the three feature selection techniques, mutual information is chosen for feature selection to integrate with the hyper-parameter tuning and cross-validation of the classifier model. First,

sequential feature selection is ruled out, as it requires prior knowledge of the hyper-parameters. Between the ANOVA-F and mutual information filter techniques, it was found that mutual information performed better with its ability to measure non-linear relationships and select suitably low values of hyper-parameters. The ANOVA-F technique was also ruled out as the choice of hyper-parameter C was too large.

If the ANOVA-F technique did not have obvious disadvantages, as with this case, that lead to its disqualification, the feature selection technique would be chosen by comparing the cross-validation scores between the ANOVA-F and mutual information methods. The final set of parameters for the SVM can be seen in Table 6.6.

Table 6.6: The selected model for the support vector machine based on the results from the mutual information feature selection method.

Parameters of the Support Vector Machine Classifier	
Number of Features	8
Selected Features	irms_comparison, irms_2, irms_3, V0_regular_std, I0_regular_std, i_l1_std2, i_l1_e3, i_l2_e2
Hyper-Parameters	$C = 100, \gamma = 1$

6.6 Conclusion

In this section, the development of the classification model was discussed. The principal motivation behind performing the above analyses is to develop the simplest, robust SVM classifier possible without the risk of it being only applicable to the fault data at hand.

The different feature selection techniques were explored and it was concluded that the mutual information statistical method for feature importance ranking in combination with hyper-parameter tuning is the most suitable for the problem at hand. In the next chapter, the different test cases for analysing the performance of the support vector machine on the test data are introduced.

7 Evaluation of the SVM Classifier

7.1 Introduction

The SVM model for the classification of unstable faults was developed in Chapter 6, where the model-specific features and parameters were finalised. This chapter describes the steps taken to evaluate the performance of the developed classifier. Subsection 7.2 introduces the test cases that are utilised to judge the behaviour of the classifier under different scenarios. The results of the SVM performance under these test cases are presented in subsection 7.2.1. Section 7.3 provides reasons for the selection of the mother wavelet by comparing the performance of the db4 wavelet with other wavelets mentioned in literature. Finally, this chapter concludes with Section 7.4 where the performance of the classifier with and without the data-preprocessing techniques are compared.

7.2 Evaluation of the Support Vector Machine Performance

The determination of relevant test cases is a vital prerequisite for a thorough test of the model [59]. Test cases involve setting up different subsets for training and testing data. From the results of each test case, it is possible to assess how well the classifier learns from different data sets. In this chapter, the description of the test cases developed for the evaluation of the classification model are discussed. In Chapter 6, the parameters of the SVM classifier were decided using feature selection, hyper-parameter tuning and cross-validation.

The performance of the classifier can be evaluated using different data sets for training and testing. The input data for the model consists of "real" faults that were recorded in the distribution network, and synthetic faults that were generated to represent ideal faults. The synthetic faults were created for single-phase-to-ground, two-phase-to-ground and three-phase-to-ground permanent faults as described in Section 4.2.

By training and testing on different subsets of the real and synthetic faults, it is possible to determine the extent to which the selected features are capable of identifying different fault types. Section 7.2.1 explains each test case considered for the SVM evaluation and what can be learnt from each scenario. Next, Section 7.2.1 presents the results for each test case.

7.2.1 Defining the Test Cases for Model Evaluation

Test Case I: Train and Test on Synthetic Faults

As mentioned in Section 4.1.1, synthetic fault voltage and current waveforms were included to the input data set to provide reference waveforms that precisely model single-phase, two-phase and three-phase faults. The first test case is to check whether the selected features are capable of differentiating between the synthetically created faults themselves. If the features selected accurately represent discriminating features between the faults, the results of the classification for this case should have high values of the performance metrics — precision, recall and f1-score. The data set of the synthetic faults are randomised, split into a training and testing data set in a ratio of 85:15 for training and testing. It is important to note that since only 35 fault instances are available between the training and testing data sets, it may be the case that high performance scores can be a function of the limited data, thereby indicating overfitting or random chance [60]. Larger data sets lead to better performance of classifiers, however, collecting large samples of data is not always possible in a practical setting. To further study the generalisation ability of the classifier, test cases including the real-world data set for single-phase, two-phase and three-phase grounded faults are discussed in the following sections.

Test Case II: Train on Synthetic Faults and Test on Real Faults

In this test case, the SVM classifier is trained on all the synthetically created faults and tested on the real-world single-phase, two-phase and three-phase faults. The real-world self-extinguishing faults are excluded from this test case as there are no synthetic self-extinguishing faults that can be used for training. The performance of the classifier in this test case would indicate the degree of similarity between practical fault instances in the distribution network and the synthetic faults. The magnitude of the classifier performance metrics are directly proportional to the similarity between the real-world and synthetic faults. This test case is useful for understanding how much can be learnt by the classifier from the synthetic faults, and how relevant these features are to actual faults in the distribution network.

Test Case III: Train on Real-World Faults and Test on Synthetic Faults

As an inverse of the previous test case, in this test case the SVM classifier is trained on all the real-world single-phase, two-phase, and three-phase faults and tested on the synthetic faults. This test case is useful for determining if the classifier learns the important behaviour of the real faults and if it is able to accurately classify the simpler synthetic faults — which is expected. Higher performance metrics indicate that the classifier is functioning with its expected capabilities.

Test Case IV: Train and Test on Real Faults

In this test case, the classifier is trained and tested on the real-world faults. This consists of the single-phase, two-phase, three-phase and single-phase self-extinguishing faults. The data set is split in a ratio of 85:15 for training and testing respectively. The results of the performance of the classifier in this test case will provide information on whether the model learns discriminating information from the current set of real faults to accurately classify future real faults.

Test Case V: Train on a Combination of Real and Synthetic Faults, and Test on Real Faults

In this final test case, the synthetic single-phase, two-phase and three-phase faults are added to the training data. This is done to add stronger reference waveforms for the aforementioned practical faults to make the features more robust. The performance scores of this test case are the most important to study the behaviour of the classifier.

Results of the Classification Development Process

It is useful to reiterate the feature set that is used as an input to the classifier model. From the 36 features engineered on the scientific understanding of faults, 8 features were narrowed down on to be the most relevant to the target classes. These features, and their description can be observed in Table 7.1. Using these features calculated for every instance of the four classes as input, and the performance metrics as output, the performance of the classifier for each test case can be evaluated. The optimised classifier model consists of the parameters as shown in Table 7.2.

Table 7.1: The features selected for the optimised support vector machine classifier model.

Feature	Description
irms_comparison	Metric for the relationship between the RMS of the phase currents
irms_2, irms_3	The smallest and second-smallest RMS values of the phase currents
V0_regular_std, I0_regular_std	Standard deviation of the zero-sequence voltage and current of the fault waveform
i_l1_std2	Central value of the standard deviation of the detail coefficients of the fault current waveform in the 1-2 kHz bandwidth
i_l1_e3	Smallest value of the energy of the detail coefficients of the fault current waveform in the 1-2 kHz bandwidth
i_l2_e2	Central value of the energy of the detail coefficients of the fault current waveform in the 0.5-1 kHz bandwidth

Table 7.2: The selected model for the support vector machine based on the results from the mutual information feature selection method.

Parameters of the Support Vector Machine Classifier	
Number of Features	8
Selected Features	irms_comparison, irms_2, irms_3, V0_regular_std, I0_regular_std, i_l1_std2, i_l1_e3, i_l2_e2
Hyper-Parameters	$C = 100, \gamma = 1$

7.2.2 Results of the Test Cases

Interpreting the Classification Error Matrices

This section evaluates the performance of the classifier by using classification error matrices. An example of such a matrix can be seen in Table 7.3. The theory behind these matrices and the metrics that are derived from them are explained in Section 3.4. The metrics used to compare the classifier performances are the precision, recall and f1-score. The following paragraph provides a short guide to interpret the error matrices that are presented further in this work.

The error matrix is drawn up for the classifier performance on the test set alone. Each diagonal element indicates the number of elements that have been correctly identified, and each off-diagonal element indicates the number of misclassified samples. For example, with Table 7.3 (Case I), values are present only in the diagonals. This means that the predicted class of every sample matches the true class of each sample resulting in a value of 100% in all of the performance metrics. On the other hand, with Table 7.5 (Case II), values in the off diagonals of the matrix indicate that misclassifications have occurred. For example, with the two-phase fault class, 27 faults were mistakenly identified to be single-phase faults and 10 faults were mistakenly identified to be three-phase faults. In this way, it is easy to compare the performance of the classifier for each class.

Result of Case I: Train and Test on Synthetic Faults

In this test case, the classifier is trained and tested on the synthetic faults. The rate of (in)correct classifications in the test set can be observed in the error matrix in Table 7.3. From the absence of values in the off-diagonals of the matrix it can be observed that there are no misclassifications. The performance scores for this test case are presented in Table 7.4. It can be gathered that the classifier has scores of 100% for the precision, recall and accuracy. The first inference from the scores is that, as there are no misclassifications, the selected features and SVM model is sufficient to differentiate between the synthetically developed faults.

The objective of adding the synthetic faults to the test cases is to provide ideal reference values for single-phase, two-phase and three-phase faults. Real faults in the distribution network consist of additional disturbances or distortions, at the zero crossing, for example. Adding synthetically created faults to the DFRs of actual faults in the distribution of the training data can help the ML classifier model anchor class labels around certain values of the features. It was also important to ensure that adding the synthetic faults to the classifier does not adversely affect the performance of the classifier. For this purpose, the classifier is first trained and tested on the synthetic faults alone. From the performance scores shown in Table 7.4, it can be concluded that the classifier is capable of perfectly differentiating between these faults, and that their addition to the training data will not harm the classifier performance.

Table 7.3: Classification error matrix when classifier is trained and tested on synthetic faults.

		Predicted Class			
		1	2	3	
True Class	1	3	0	0	1: 1-Phase Permanent Fault
	2	0	2	0	2: 2-Phase Permanent Fault
	3	0	0	3	3: 3-Phase Permanent Fault

Table 7.4: Performance scores of the SVM trained and tested on synthetic faults.

Test-Case	Precision (%)	Recall (%)	Accuracy (%)
Train Data: Synthetic Faults	100	100	100
Test Data: Synthetic Faults			

Result of Case II: Train on Synthetic Faults and Test on Real Faults

This section analyses the performance of the SVM classifier that is trained on real faults and tested on synthetic faults. The objective of this test was to study if the features trained on the synthetic faults are capable of generalising to real faults in the distribution network. The error matrix for this test case can be seen in Table 7.5. From the second row, it can be observed that two-phase faults are commonly misclassified as single-phase faults and three-phase faults.

The scores of the precision, recall and accuracy in Table 7.6 present less than ideal values. The scores are, however, higher than if the classifier were to classify at random. It can therefore be confirmed that the developed features and classifier are capable of extrapolating to more complex fault waveforms to some degree. One reason for the low scores could be due to the fact that the number of training data is limited to 35 points. This small sample space of training data combined with the inability of the synthetic data to capture the uncertainties in fault identification that arise from distortions in the fault waveforms can lead to lower performance scores. Hence, it can be concluded that the a classifier model trained on synthetically developed faults is not sufficient to characterise the faults in the distribution network.

Table 7.5: Classification error matrix when classifier is trained and tested on real faults.

		Predicted Class			
		1	2	3	
True Class	1	116	1	1	1: 1-Phase Permanent Fault
	2	27	63	10	2: 2-Phase Permanent Fault
	3	7	3	86	3: 3-Phase Permanent Fault

Table 7.6: Performance scores of the SVM trained on synthetic faults and tested on real faults.

Test-Case	Precision (%)	Recall (%)	Accuracy (%)
Train Data: Synthetic Faults			
Test Data: Real Faults	86	84	84

Result of Case III: Train on Real Faults and Test on Synthetic Faults

The values of the performance metrics for the SVM classifier trained on the real faults and tested on the 35 synthetic faults are presented in Table 7.8. Table 7.7 presents the error matrix for this case, and similar to Case I, it can be observed that the classifier does not misclassify any of the faults. The precision, recall and accuracy have scores of 100% each. It is expected that the classifier performs well in this scenario as the synthetic faults are variations of the real-world faults, with no distortions in the fault waveforms. As an inversion of the Case II, it can be useful to note that the features from real-world faults are indeed capable of perfectly classifying the synthetic faults.

Table 7.7: Classification error matrix when classifier is trained on real faults and tested on synthetic faults.

		Predicted Class			
		1	2	3	
True Class	1	12	0	0	1: 1-Phase Permanent Fault
	2	0	12	0	2: 2-Phase Permanent Fault
	3	0	0	12	3: 3-Phase Permanent Fault

Table 7.8: Performance scores of the SVM trained on real faults and tested on synthetic faults.

Test-Case	Precision (%)	Recall (%)	Accuracy (%)
Train Data: Real Faults			
Test Data: Synthetic Faults	100	100	100

Result of Case IV: Train and Test on Real Faults

In this is test case, the real-world faults are split in a ratio of 85:15 for training and testing respectively. It should be noted that in this case there are four classes — single-phase, two-phase, three-phase, and single-phase self-extinguishing faults. The error matrix for this test case can be studied Table 7.9. It can be seen that the three-phase faults (row 3) are not misclassified, with the brunt of the misclassifications occurring with single-phase and two-phase faults.

The values of the performance metrics of this test case are presented in Table 7.10. It can be observed that the classifier has an f1-score of 90%. The performance of the classifier for this test case indicates that the selection of the features and the model parameters results in a classifier model that is indeed capable of learning from the real-world fault data to classify future faults that were not used to train the model.

Table 7.9: Classification error matrix when classifier is trained and tested on real-world faults.

		Predicted Class				
		1	2	3	4	
True Class	1	15	3	0	0	1: 1-Phase Permanent Fault
	2	0	14	2	0	2: 2-Phase Permanent Fault
	3	0	0	12	0	3: 3-Phase Permanent Fault
	4	0	1	0	13	4: 1-Phase Extinguishing Fault

Table 7.10: Performance scores of the SVM trained and tested on real faults.

Test-Case	Precision (%)	Recall (%)	F1-Score (%)
Train Data: Real Faults			
Test Data: Real Faults	91	90	90

Result of Case V: Train on a Combination of Real and Synthetic Faults, and Test on Real Faults

In the final test case, the classifier is trained on a combination of the real-world and synthetic faults, and tested on a different subset of the real-world faults. The real faults are split in a ratio of 85:15 for the train and test set respectively. The synthetic faults are appended to the train set to

attempt to improve the classifier's performance by providing reference values for properties that are intrinsic to the faults without consideration for the distortions. The synthetic faults hence provide reliable data for the actual nature of single-phase, two-phase, and three-phase faults and the real-world faults provide samples of the possible distortions. It can be observed that the classifier has higher scores for the precision, recall, and accuracy compared to the previous case.

Table 7.11 shows the error matrix for the classifier that is trained on a combination of the real and synthetic data and tested on a (different) subset of the real-world data. It can be observed that there are only three misclassifications that occur.

Table 7.11: Classification error matrix when classifier is trained on a combination of real-world and synthetic faults and tested on real faults.

		Predicted Class				
		1	2	3	4	
True Class	1	15	1	0	0	1: 1-Phase Permanent Fault
	2	0	16	1	0	2: 2-Phase Permanent Fault
	3	0	0	13	0	3: 3-Phase Permanent Fault
	4	0	1	0	13	4: 1-Phase Extinguishing Fault

Table 7.12: Performance scores of the SVM trained on a combination of real and synthetic faults, and tested on real faults.

Test-Case	Precision (%)	Recall (%)	F1-Score (%)
Train Data: Real and Synthetic Faults			
Test Data: Real Faults	95	95	95

There are two conclusions that can be made from the results of this test case. The first is that the addition of the synthetic faults to the test data enables better performance of the classifier. This can be gathered from Figure 7.1, where the classifier is trained on both real and synthetic faults (test case V), has higher performance scores compared to when the classifier is trained on the synthetic or real faults alone. A summary of the results of all five test cases can be seen in Table 7.13.

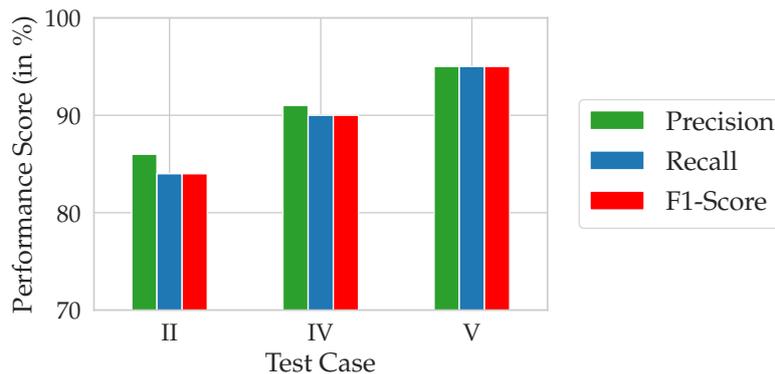


Figure 7.1: Precision, recall and accuracy scores of the different test cases.

Second, it can be concluded that the overarching aim of studying the performance of supervised learning techniques for the purpose of fault classification has been realised. The classifier has an accuracy of 95% - with three misclassifications in the test data set of strength 60 faults.

It should be noted that a classification accuracy of 100% is not the goal of this study. A perfect classification accuracy in the testing data leaves room for the possibility that the classifier will

Table 7.13: Summary of precision, recall and accuracy scores of the different test cases.

Number	Test-Case		Precision (%)	Recall (%)	F1-Score (%)
	Training Fault Data	Testing Fault Data			
I	Synthetic	Synthetic	100	100	100
II	Synthetic	Real-World	86	84	84
III	Real-World	Synthetic	100	100	100
IV	Real-World	Real-World	91	90	90
V	Real-World + Synthetic	Real-World	95	95	95

not perform as well on future unseen data. This means that, while the model would be perfectly capable of differentiating between the characteristics that lend to their classification for the data set at hand, it would be optimistic to assume that this would extend to new/future data. In an exploratory study as in this case, it is more important to identify which features and parameters contribute to the most generally applicable model as possible. The more general a model is, the more closely the classification on future data matches the classifier performance presented in this thesis.

7.3 Selection of the Mother Wavelet

The validation of the selection of the mother wavelet is discussed in this section. In the discrete wavelet transform, signal decomposition is realised through a cascading filter bank implementation [35]. The output of each filter is a convolution of the mother wavelet with the input signal. Hence, the mother wavelet is implicit in the filter banks in the form of the filter bank coefficients. The selected mother wavelet should possess the ability to capture the transients in the fault signal [15]. There exist different families of wavelets, such as Daubechies, Symlets, Biorthogonal, and Haar families. Each family of wavelets have subdivisions of mother wavelets that differ from each other in the length and value of their filter coefficients. The filter coefficients are ordered in two different combinations — one that performs a smoothing function on the input waveform, and one that works to extract the details of the input data [61]. The results of the smoothing filter and the detail-extracting filter are the approximation and detail coefficients, respectively. The detail coefficients are further analysed to study the behaviour of the transients in faults.

For this study, it was identified from the review of literature that the Daubechies family of wavelets was a popular choice for the analysis of fault transients as seen in Section 2.4. It was found that the Daubechies-4, or db4, wavelet is adept at capturing rapid changes in signals caused by power transients [34]. In Chapter 4, the discrete wavelet transform features were extracted from the fault waveforms using the db4 wavelet. This choice was made on the assumption that the literature citing the db4 wavelet to be reliable in extracting transients was applicable to the faults studied in this thesis. From the results in Section 7.2.1, it can be gathered that the results validate this assumption, as the accuracy for the test of the performance of the classifier is 95%. It is, however, worth confirming if this wavelet is indeed the superior choice for this classification problem by comparing the performance of the classifier with models that extract transient features using other mother wavelets.

Four other wavelets were considered for the purpose of comparison with the db4 wavelet — Daubechies-7 (db7), Daubechies-8 (db8), Symlet-4 (sym4), and the derivative of the cubic b-spline wavelet (b-spline). The selection of the wavelets are not integrated in the cross-validation process due to the computational expense of extracting the features, performing the feature selection, hyper-parameter tuning, and cross-validation for each wavelet type. Instead, this work is done under the assumption that the db4 wavelet is superior to the other wavelets cited in literature;

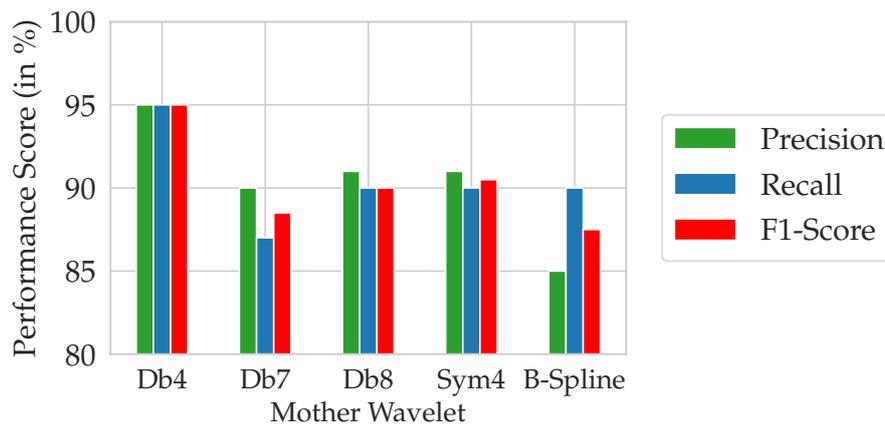


Figure 7.2: Precision, recall and accuracy scores of the classifier for different mother wavelets.

however, the validity of the assumption is cross-checked by comparing classifier performance with each type of wavelet.

To make this comparison, the features are extracted using the aforementioned db7, db8, sym5 and cubic b-spline wavelet, in addition to the db4 wavelet. In the aforementioned test cases, the features are selected, and the hyper-parameters are optimised using the db4 wavelet for feature extraction. It would, however, not be a true comparison of performance if the features are extracted with the other wavelets while keeping the features and hyper-parameters the same. The feature extraction, feature selection, hyper-parameter tuning, cross-validation and testing processes are hence performed for each wavelet. This ensures that the classifier is optimised for db7, db8, sym4 and cubic b-spline wavelet each. It is this performance that is compared with the performance of the db4 wavelet. It should be noted that the selection of feature selection technique is not done in this case, and the mutual information test is used for feature ranking. The comparison of the precision, recall, and f1-scores can be observed in Figure 7.2.

From Figure 7.2, it can be observed that the performance of the classifier that utilises the db4 wavelet possess the highest scores across the three performance metrics. Hence, it can be concluded from the results that the db4 wavelet is indeed the best choice for the purpose of fault classification in the distribution network in comparison to other wavelets that are used in literature.

7.4 Evaluation of the Data Pre-processing Techniques

Binary segmentation was used to section only the part of the fault waveform that consisted of the fault current/voltage. This process was described in detail in Section 4.1.1. This data processing was applied to all the single-,two-, and three-phase permanent faults prior to the feature engineering step. In the interest of saving computational efforts and for better insights on how the features work, it can be interesting to see if segmentation is a necessary step. Table 7.14 compares the performance scores of the classifier that includes and excludes the segmentation step.

From Table 7.14, it can be observed that the classifier that segments the fault periods before classification performs better than the other, but not by much. An implication of this result is that the feature engineering process itself is able to generate sufficiently discriminating features between the steady state, fault, and extinction states. Hence, if faster classification of faults is desired, it is perhaps not completely necessary to include the segmentation step to direct the

Table 7.14: Comparison of the performance of the classifier with and without the fault segmentation step.

Classifier	Precision (%)	Recall (%)	F1-Score (%)
With Segmentation	95	95	95
Without Segmentation	94	92	93

features in the "right direction" of capturing only the fault state. It should, however, be stated that this thesis does not have an emphasis on the efficiency aspect of computation and this idea is later brought back in Section 8.2 where future directions are discussed.

7.5 Conclusion

In this chapter, the classifier model developed in Chapter 6 is assessed with different test cases. It has been seen that the classifier has an f1-score of 95%, and possesses a strong ability to differentiate between the four classes of stable faults. The results also give way to recommendations for the classifier. First, from the results of test case V in comparison with test case IV, it can be observed that adding synthetic faults to the training data set can improve the performance of the classifier. Second, it was found as shown in Figure 7.2, that the Daubechies-4 wavelet is indeed the right choice for studying electrical power transients. Finally, as explained in Section 7.4, segmenting a fault waveform into just the time period consisting of the fault does not significantly alter classifier performance compared to a case where the segmentation is not performed. The following chapter, Chapter 8 sums up the results by laying out the main conclusions from this thesis.

8 Conclusion and Outlook

During the course of the thesis it was investigated if supervised learning could be used to make important decisions about the nature of faults in the distribution network. This objective fits into the bigger picture of network restoration systems where it is required that a fault is quickly located between two Ring Main Units to isolate the faulted cable and restore power supply to the customers. To this end, determination of fault type is vital to the calculation of the fault location. Besides stable permanent faults, which are easier to identify by the present algorithm used by Alliander, there are some faults that are especially difficult to classify, such as a series of self-extinguishing faults and unstable faults. In the case of the former, this is because a series of short-term extinguishing faults are sometimes recognised as a permanent fault based on the patterns on which they are currently assessed. In the case of unstable faults, the present overly broad definition causes misclassifications of stable faults that have distortions in their initial fault development stages. Hence, it was seen if supervised learning could accurately classify the "easier" stable faults and the concepts were extended to study self-extinguishing and unstable faults.

Stable and unstable faults were handled in two different ways. In the case of stable faults, identifying features were first established using DWT and an algorithm was developed to optimise an SVM classifier such that it used optimal features and hyper-parameters. Additional steps were taken to make the classifier more robust, such as segmenting the fault to contain only the final stages and by adding synthetic faults to provide stronger reference values for the classifier. This classifier was then trained and tested on an existing data set of single-/two-/three-phase permanent and self-extinguishing faults. The key findings in this part of the study were the features that were considered the most differentiating between the classes, and insights as to which mother wavelet performed the best in this fault analysis problem.

The other part of the study was devoted to developing a method for studying single-phase unstable faults using simple decision trees. First, the unstable fault waveforms were studied using Fourier transforms to assess the degree of presence of the 50 Hz fundamental frequency in each cycle over the entire period of the fault waveform. The fault waveforms were then transformed into a function of this Fourier transform component. From here, features were created and were used as an input to simple ensemble decision trees where thresholds for each feature separating stable and unstable behaviour were found. The key findings of thesis are presented in Section 8.1, followed by Section 8.2 that outlines possible avenues for future research.

8.1 Research Questions

The overarching question posed for this thesis is:

In what capacity can modern signal processing and machine learning techniques improve the classification of faults in the distribution network?

The following sub-questions address the key findings in the study:

1. What are the existing machine learning approaches for the detection and classification of faults?

A survey of literature in the fault signal processing and classification fields is covered in Chapter 2. With regards to signal processing, it was found that a majority of the papers published in the past decade use DWT. DWT helps decompose fault waveforms into their constituent frequency bands, thereby enabling the study of faults in the specific frequency band in which the transients occur. It became clear that the selection of the mother wavelet is vital to how the transients are captured in the decomposition process. It was hence found that it is not a matter of which signal processing technique works best for transients, but which wavelet filter was the most suitable for studying the faults. The majority of the papers studied showed that the Daubechies 4 wavelet is the best choice for the study of power system transients. Whether this is indeed the best choice for the study at hand is corroborated in the following research question.

Among different supervised learning classification techniques such as decision tree, k-NN, SVM, and neural network classifiers, SVMs were identified as the choice supervising learning classifier due to their ability to transpose inseparable classes of data to a higher dimensional space where they can be separated by a hyper-plane. It was concluded that the actual study of faults can be facilitated with the use of DWT for fault signal processing and SVMs for classification. Another common theme in the reviewed literature was that the classifiers were trained solely on synthetic faults that were simulated through software such as ETAP or Simulink. This left room for the thesis to study how such classifiers performed on real-world fault data from the distribution network.

2. What criteria can be used to classify a fault as unstable?

In this thesis, unstable faults are studied to see if better criteria can be developed to characterise them. In this pursuit, the lower presence of the fundamental 50 Hz component in transients compared to a steady state signal was considered an interesting aspect to explore. From the results of performing a sliding Fourier analysis across the entire period of fault signals, it was found that there existed significant differences between stable faults and unstable faults in their fundamental frequency Fourier component. Five features were used to develop a scoring scheme by which stable single-phase faults that were misclassified as unstable could be correctly classified as stable faults. These features are explained in Table 4.6. In particular, however, out of the five features identified as potential identifiers of stability, it was found that stability was highly correlated with the most stable period of a signal and the standard deviation of the developed metric F_{ratio} , as validated in Table 5.2 in Section 5.4.

3a. What are the aspects of stable fault waveforms that lend to their differentiation into different fault types?

In Chapter 4, an initial set of 36 features were developed for stable faults based their properties that lend to their classification. These features were either functions of the fault waveforms in the original time domain or were extracted with the use of DWT. The 36 features were narrowed down using feature selection techniques by ranking the features based on their mutual information test scores and choosing the number of features that resulted in the highest cross-validation scores as explained in Chapter 6. The results were 8 features that are functions of the fault signals, such as the RMS, standard deviation, or energy of the zero sequence current/voltage, and detail coefficients in the 1kHz - 2kHz and 500 Hz - 1kHz range. These features can be observed in Table 6.1.

The mother wavelet used for the extraction of the features during signal decomposition was the Daubechies 4 wavelet. In Chapter 7, it was shown that the choice of the Daubechies 4 wavelet is indeed justified compared to other popular wavelet filters used in literature such as the

Daubechies 7/8, Symlet 4, and cubic B-spline wavelets based on the classifier performance scores for each wavelet type. In Chapter 7, the performance of the developed SVM classifier that uses the selected 8 features was assessed on the test set. It was found that the classifier misclassifies three faults in the test set of 60 faults (resulting in an overall accuracy of 95%). Hence, it can be concluded that the 8 features selected, facilitated by the Daubechies 4 wavelet, are sufficient to classify faults with an accuracy of 95%.

3b. How can self-extinguishing faults be differentiated from faults that are directly switched-off by a circuit breaker?

Research question 3a covered the development of a novel method for determining stability. As a particular case of 3a, 3b studies the ability of the Fourier analysis features to distinguish between self-extinguishing and permanent faults (which are faults that are switched off by a circuit breaker). A key difference between self-extinguishing faults and permanent faults is the restrike behaviour of the fault current and voltage waveforms at every zero-crossing (Figure 4.12 for reference). The motivation for this question is to prevent stable permanent faults from being "mistaken" for self-extinguishing faults and to increase the number of correct fault locations sent to the control center for network restoration. It was found that it is indeed possible to differentiate between the two using the developed Fourier analysis features.

A sliding scale for determining stability was developed by considering self-extinguishing faults as instances of instability and permanent faults as examples of stable behaviour. In the stability assessment, the restrike at zero-crossing falls under the envelope of "non-sinusoidal behaviour" that was used as the basis for setting the thresholds for the Fourier analysis features. Hence, each of the five features will have lower values than the thresholds set for the classification of stable faults as explained in Section 5.5. Within the framework of the developed business rules, keeping in the mind the emphasis on avoiding further misclassifications, a series of self-extinguishing faults would either be classified as decidedly unstable, or needing further inspection.

4. What are the methods to evaluate the performance of the machine learning model and how does the developed classifier perform in its ability to distinguish between faults?

This question addresses the performance of the SVM for classifying stable faults in particular. Metrics precision, recall, and f1-score are used to assess classifier performance in this thesis. To maintain objectivity, the performance developed classifier was compared against the "best" and "worst" cases for performance. When tested and trained on synthetic faults alone, the classifier has an accuracy of 100% which is unrealistic in a real-world situation. On the other hand, a baseline for classifier performance is set using a test case where the classifier is trained on synthetic faults and tested on the real-world faults. This test case results in an f1-score of 84%. In the most relevant test case, where the classifier is trained and tested on a combination of the real-world and synthetic faults, the classifier showed an f1-score of 95% - an objectively promising accuracy.

8.2 Avenues for Future Research

It is important to consider the results of this thesis in the context of a number of limitations that have presented themselves in various areas during the study. This section elaborates on these limitations and identifies avenues for future research on this topic.

Data: This thesis improves on the reviewed literature by using a combination of real-world faults from the distribution network and synthetically created faults to develop a classifier through supervised learning. One of the major limiting aspects, however, is the number of reliable real-

world fault data present at hand for the study. The availability of a larger sample space for the faults would provide wider references for the classifier to be trained on, and can capture the different variabilities that are present in fault waveforms. Second, with respect to the determination of unstable behaviour in faults, it would have been possible to determine a non-heuristic technique if more reliable samples were available for the development a supervised learning algorithm. Hence, it would be interesting to see, in both cases of stable and unstable faults, how the classifier would fare as reliable and easy-to-access data becomes the norm with advancements in data engineering.

Classification of unstable faults: Particular to the case of unstable faults, a relatively larger set of reliable data was available only for single-phase unstable faults. In order to diagnose instances of multi-phase instability, a future direction for the study can be to investigate how (and if) the present rules can be extended to multi-phase faults.

Self-extinguishing faults and unstable behaviour: In order to determine a sliding scale between stability and instability, self-extinguishing faults were considered to be an “extreme” case of unstable behaviour, while permanent faults were considered to be the peak of stable behaviour. An implication of this consideration is that true self-extinguishing faults cannot be reliably differentiated from true unstable faults with the current business rules. This thesis focused on the determination of stability for permanent faults as they are the ones that provide accurate fault locations. Hence, this work opens up another avenue for further research: which is to develop a method for distinguishing between unstable faults and self-extinguishing faults.

Regularisation of the classifier: Regularisation is the process of developing a flexible classifier, i.e. a model that has a high generalisation ability. To achieve this, the mutual information feature ranking technique was used along with a hyper-parameter tuning process that involved using an optimisation algorithm to search for useful hyper-parameter combinations over a space. There exist more sophisticated techniques for regularisation such as the L1, L2 or Lasso methods [62] which can be implemented for future use.

Performance metrics: Presently, for the classification of stable faults, relatively simple metrics such as precision, recall, f1-score were used for performance evaluation. It can be investigated if there are other metrics that are more suitable for the evaluation of the classifier.

Computation time: This thesis was an exploratory study to see how well certain signal processing and ML techniques fare with real-world fault data. If this research is to be taken forth for practical implementation it is useful to optimise the run time of the classifier model, an aspect that was not considered in this thesis. One lead is that the data preprocessing method for segmenting the fault may not be as vital as hypothesised, as evidenced by the performance scores in Table 7.14. Similarly, studies can be done in other aspects of the model to check if any processes can be streamlined.

Combining the two (stable and unstable) analyses: This thesis performs two overarching analyses: a check for stability, and another independent check for fault type. In the interest of a self-contained process, it can be useful to combine the two processes into one that performs both of the aforementioned analysis.

Bibliography

- [1] F. Koehler, S. Cobben and F. Provoost, "Self-extinguishing faults in MV cable networks-feasibility study of fault prediction", English, *Renewable Energy Power Quality Journal*, vol. 1, no. 10, pp. 44–49, Apr. 2012, ISSN: 2172-038X. DOI: [10.24084/repqj10.210](https://doi.org/10.24084/repqj10.210).
- [2] B. Das, "Fuzzy logic-based fault-type identification in unbalanced radial power distribution system", *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 278–285, 2006. DOI: [10.1109/TPWRD.2005.852273](https://doi.org/10.1109/TPWRD.2005.852273).
- [3] F. Provoost and W. Van Buijtenen, "Practical experience with fault location in mv cable network", in *CIREDE 2009 - 20th International Conference and Exhibition on Electricity Distribution - Part 1*, 2009, pp. 1–4. DOI: [10.1049/cp.2009.0817](https://doi.org/10.1049/cp.2009.0817).
- [4] F. Provoost, "Intelligent distribution network design", Ph.D. dissertation, Technische Universiteit Eindhoven, 2009. [Online]. Available: <https://research.tue.nl/en/publications/intelligent-distribution-network-design>.
- [5] H. Wolse, G. Geist, B. Hoving, P. Oosterlee and H. Polman, "Experience and tendencies after 40 years outage data registration in the netherlands", *CIREDE - Open Access Proceedings Journal*, vol. 2017, pp. 2279–2282, Oct. 2017. DOI: [10.1049/oap-cired.2017.0251](https://doi.org/10.1049/oap-cired.2017.0251).
- [6] *Alliander annual report 2020 - other non-financial information*, Feb. 2021. [Online]. Available: <https://2020.jaarverslag.alliander.com/verslagen/annual-report-2020/otherinformation9/otherofinanciainformatio6>.
- [7] T. Gu and F. Provoost, "Improved fault location algorithm for MV networks based on practical experience", vol. 2017, Oct. 2017, pp. 1211–1214. DOI: [10.1049/oap-cired.2017.1153](https://doi.org/10.1049/oap-cired.2017.1153).
- [8] T. S. Sidhu and Z. Xu, "Detection of incipient faults in distribution underground cables", *IEEE Transactions on Power Delivery*, vol. 25, no. 3, pp. 1363–1371, 2010. DOI: [10.1109/TPWRD.2010.2041373](https://doi.org/10.1109/TPWRD.2010.2041373).
- [9] W. Stevenson and J. Grainger, *Power System Analysis*. McGraw-Hill Education, 1994, ISBN: 9780070612938.
- [10] T. Gu and F. Provoost, "Improving network performance by recognition and location of self-extinguishing faults", in *22nd International Conference and Exhibition on Electricity Distribution (CIREDE 2013)*, 2013, pp. 1–4. DOI: [10.1049/cp.2013.0728](https://doi.org/10.1049/cp.2013.0728).
- [11] D. C. Robertson, O. I. Camps, J. S. Mayer and W. B. Gish, "Wavelets and electromagnetic power system transients", *IEEE Transactions on Power Delivery*, vol. 11, no. 2, pp. 1050–1058, 1996. DOI: [10.1109/61.489367](https://doi.org/10.1109/61.489367).
- [12] X. Tang, Z. Zhang, Q. Huang and Y. Gong, "Fault location and fault type recognition of power system based on wavelet transform", in *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2019, pp. 689–692. DOI: [10.1109/ISGT-Asia.2019.8881101](https://doi.org/10.1109/ISGT-Asia.2019.8881101).
- [13] M. Dashtdar, R. Dashti and H. R. Shaker, "Distribution network fault section identification and fault location using artificial neural network", in *2018 5th International Conference on Electrical and Electronic Engineering (ICEEE)*, 2018, pp. 273–278. DOI: [10.1109/ICEEE2.2018.8391345](https://doi.org/10.1109/ICEEE2.2018.8391345).

- [14] W. Zhao, Y. Song and Y. Min, "Wavelet analysis based scheme for fault detection and classification in underground power cable systems", *Electric Power Systems Research*, vol. 53, pp. 23–30, Jan. 2000. DOI: [10.1016/S0378-7796\(99\)00033-4](https://doi.org/10.1016/S0378-7796(99)00033-4).
- [15] R. N. Mahanty and P. B. D. Gupta, "Comparison of fault classification methods based on wavelet analysis and ANN", *Electric Power Components and Systems*, vol. 34, no. 1, pp. 47–60, 2006. DOI: [10.1080/15325000691001485](https://doi.org/10.1080/15325000691001485).
- [16] A. Majd, H. Samet and T. Ghanbari, "K-NN based fault detection and classification methods for power transmission systems", *Protection and Control of Modern Power Systems*, vol. 2, Dec. 2017. DOI: [10.1186/s41601-017-0063-z](https://doi.org/10.1186/s41601-017-0063-z).
- [17] H. Livani and C. Y. Evrenosoğlu, "A fault classification method in power systems using DWT and SVM classifier", in *PES T & D 2012*, 2012, pp. 1–5. DOI: [10.1109/TDC.2012.6281686](https://doi.org/10.1109/TDC.2012.6281686).
- [18] W. Li, X. Miao and X. Zeng, "Short circuit fault type identification of low voltage AC system based on black hole particle swarm and multi-level SVM", in *2020 Chinese Automation Congress (CAC)*, 2020, pp. 208–213. DOI: [10.1109/CAC51589.2020.9327638](https://doi.org/10.1109/CAC51589.2020.9327638).
- [19] Z. J. Ren, "Wavelet based analysis of circuit breaker operation", Ph.D. dissertation, Texas A & M University, 2006. [Online]. Available: <https://hdl.handle.net/1969.1/512>.
- [20] S.-c. Kam, S. Nielsen and G. Ledwich, "A circuit-breaker restrike diagnostic algorithm using atp and wavelet transforms", in *AUPEC 2011*, 2011, pp. 1–6. [Online]. Available: <https://eprints.qut.edu.au/53144/>.
- [21] M. Ben Hessine, S. Marrouchi and S. Chebbi, "A fault classification scheme with high robustness for transmission lines using fuzzy-logic system", in *2017 International Conference on Advanced Systems and Electric Technologies (IC ASET)*, 2017, pp. 256–261. DOI: [10.1109/ASET.2017.7983701](https://doi.org/10.1109/ASET.2017.7983701).
- [22] M. Guo, N. Yang and W. Chen, "Deep-learning-based fault classification using hilbert–huang transform and convolutional neural network in power distribution systems", *IEEE Sensors Journal*, vol. 19, no. 16, pp. 6905–6913, 2019. DOI: [10.1109/JSEN.2019.2913006](https://doi.org/10.1109/JSEN.2019.2913006).
- [23] G. James, D. Witten, T. Hastie and R. Tibshirani, *An Introduction to Statistical Learning: with Applications in R*, ser. Springer Texts in Statistics. Springer New York, 2013, ISBN: 9781461471387. [Online]. Available: <https://www.statlearning.com/>.
- [24] C. Liu, Z. Rather, Z. Chen and C. Bak, "An overview of decision tree applied to power systems", *International Journal of Smart Grid and Clean Energy*, pp. 413–419, Jan. 2013. DOI: [10.12720/sgce.2.3.413-419](https://doi.org/10.12720/sgce.2.3.413-419).
- [25] Yong Sheng and S. M. Rovnyak, "Decision tree-based methodology for high impedance fault detection", *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 533–536, 2004. DOI: [10.1109/TPWRD.2003.820418](https://doi.org/10.1109/TPWRD.2003.820418).
- [26] X. Zhu and A. Goldberg, *Introduction to Semi-Supervised Learning*. 2009. DOI: <https://doi.org/10.2200/S00196ED1V01Y200906AIM006>.
- [27] M.-F. Guo, N.-C. Yang and L.-X. You, "Wavelet-transform based early detection method for short-circuit faults in power distribution networks", *International Journal of Electrical Power & Energy Systems*, vol. 99, pp. 706–721, 2018, ISSN: 0142-0615. DOI: <https://doi.org/10.1016/j.ijepes.2018.01.013>.
- [28] 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](https://doi.org/10.1109/TPWRD.2005.844307).
- [29] X. G. Magagula, Y. Hamam, J. A. Jordaan and A. A. Yusuff, "A fault classification and localization method in a power distribution network", in *2017 IEEE AFRICON*, 2017, pp. 1337–1343. DOI: [10.1109/AFRCON.2017.8095676](https://doi.org/10.1109/AFRCON.2017.8095676).

-
- [30] 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*, 2017, pp. 1–6. DOI: [10.1109/PowerAfrica.2017.7991190](https://doi.org/10.1109/PowerAfrica.2017.7991190).
- [31] T. S. Abdelgayed, W. G. Morsi and T. S. Sidhu, "Fault detection and classification based on co-training of semisupervised machine learning", *IEEE Transactions on Industrial Electronics*, vol. 65, no. 2, pp. 1595–1605, 2018. DOI: [10.1109/TIE.2017.2726961](https://doi.org/10.1109/TIE.2017.2726961).
- [32] S. Xiong, Y. Liu, J. Fang, J. Dai, L. Luo and X. Jiang, "Incipient fault identification in power distribution systems via human-level concept learning", *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5239–5248, 2020. DOI: [10.1109/TSG.2020.2994637](https://doi.org/10.1109/TSG.2020.2994637).
- [33] M. Jannati, B. Vahidi and S. H. Hosseinian, "Incipient faults monitoring in underground medium voltage cables of distribution systems based on a two-step strategy", *IEEE Transactions on Power Delivery*, vol. 34, no. 4, pp. 1647–1655, 2019. DOI: [10.1109/TPWRD.2019.2917268](https://doi.org/10.1109/TPWRD.2019.2917268).
- [34] N. S. D. Brito, B. A. Souza and F. A. C. Pires, "Daubechies wavelets in quality of electrical power", in *8th International Conference on Harmonics and Quality of Power. Proceedings (Cat. No.98EX227)*, vol. 1, 1998, 511–515 vol.1. DOI: [10.1109/ICHQP.1998.759961](https://doi.org/10.1109/ICHQP.1998.759961).
- [35] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674–693, 1989. DOI: [10.1109/34.192463](https://doi.org/10.1109/34.192463).
- [36] J. L. Crowley, "A representation for visual information", Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-82-07, Nov. 1981.
- [37] M. Vetterli and C. Herley, "Wavelets and filter banks: Theory and design", *IEEE Transactions on Signal Processing*, vol. 40, no. 9, pp. 2207–2232, 1992. DOI: [10.1109/78.157221](https://doi.org/10.1109/78.157221).
- [38] A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions", *Knowledge and Information Systems*, vol. 61, Dec. 2019. DOI: [10.1007/s10115-019-01335-4](https://doi.org/10.1007/s10115-019-01335-4).
- [39] E. Briscoe and J. Feldman, "Conceptual complexity and the bias/variance tradeoff", *Cognition*, vol. 118, pp. 2–16, 2011. DOI: <https://doi.org/10.1016/j.cognition.2010.10.004>.
- [40] A. Mathur and G. M. Foody, "Multiclass and binary SVM classification: Implications for training and classification users", *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 2, pp. 241–245, 2008. DOI: [10.1109/LGRS.2008.915597](https://doi.org/10.1109/LGRS.2008.915597).
- [41] D. C. Howell, *Statistical methods for psychology*. Cengage Learning Inc, 2020. DOI: [10.2307/2348956](https://doi.org/10.2307/2348956).
- [42] B. Ross, "Calculator for mutual information between a discrete and a continuous data set", *Biophysical Journal*, vol. 106, no. 2, 2014. DOI: [10.1016/j.bpj.2013.11.4438](https://doi.org/10.1016/j.bpj.2013.11.4438).
- [43] G. Simpson, *Elements of information theory*. Willford Press, 2018. DOI: [10.1002/047174882X](https://doi.org/10.1002/047174882X).
- [44] D. Powers, "Evaluation: From precision, recall and f-factor to ROC, informedness, markedness & correlation", *Machine Learning Technology*, vol. 2, Jan. 2008. [Online]. Available: <https://arxiv.org/abs/2010.16061>.
- [45] K. K. S. Suresh, *Electric Circuits and Networks*. Dorling Kindersley (India) Pvt Ltd, 2009, ISBN: 978-8131713907.
- [46] B. Wang, J. Geng and X. Dong, "High-impedance fault detection based on nonlinear voltage-current characteristic profile identification", *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3783–3791, 2018. DOI: [10.1109/TSG.2016.2642988](https://doi.org/10.1109/TSG.2016.2642988).
- [47] C. Rohrbeck, "Detection of changes in variance using binary segmentation and optimal partitioning", 2013. [Online]. Available: <https://www.lancaster.ac.uk/pg/rohrbeck/ResearchTopicI.pdf>.
- [48] C. M. Bishop, *Pattern Recognition and machine learning*. Springer-Verlag New York, 2016, ISBN: 978-0-387-31073-2.
-

- [49] S. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology", *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 660–674, 1991. DOI: [10.1109/21.97458](https://doi.org/10.1109/21.97458).
- [50] M. Kuhn and K. Johnson, *Applied predictive modeling*. Springer, 2013. DOI: [10.1007/978-1-4614-6849-3](https://doi.org/10.1007/978-1-4614-6849-3).
- [51] A. Rojas-Domínguez, L. C. Padierna, J. M. Carpio Valadez, H. J. Puga-Soberanes and H. J. Fraire, "Optimal hyper-parameter tuning of SVM classifiers with application to medical diagnosis", *IEEE Access*, vol. 6, pp. 7164–7176, 2018. DOI: [10.1109/ACCESS.2017.2779794](https://doi.org/10.1109/ACCESS.2017.2779794).
- [52] D. Berrar, "Cross-validation", in *Encyclopedia of Bioinformatics and Computational Biology*, Oxford: Academic Press, 2019, pp. 542–545, ISBN: 978-0-12-811432-2. DOI: <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- [53] M. Dash and H. Liu, "Feature selection for classification", *Intelligent Data Analysis*, vol. 1, no. 1, pp. 131–156, 1997, ISSN: 1088-467X. DOI: [https://doi.org/10.1016/S1088-467X\(97\)00008-5](https://doi.org/10.1016/S1088-467X(97)00008-5).
- [54] M. Pal and G. M. Foody, "Feature selection for classification of hyperspectral data by SVM", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 5, pp. 2297–2307, 2010. DOI: [10.1109/TGRS.2009.2039484](https://doi.org/10.1109/TGRS.2009.2039484).
- [55] T. Rückstieß, C. Osendorfer and P. van der Smagt, "Sequential feature selection for classification", in *AI 2011: Advances in Artificial Intelligence*, D. Wang and M. Reynolds, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, ISBN: 978-3-642-25832-9. DOI: https://doi.org/10.1007/978-3-642-25832-9_14.
- [56] D. W. Aha and R. L. Bankert, "A comparative evaluation of sequential feature selection algorithms", in *Learning from Data: Artificial Intelligence and Statistics V*, D. Fisher and H.-J. Lenz, Eds. New York, NY: Springer New York, 1996, pp. 199–206, ISBN: 978-1-4612-2404-4. DOI: [10.1007/978-1-4612-2404-4_19](https://doi.org/10.1007/978-1-4612-2404-4_19).
- [57] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, "Scikit-learn: Machine learning in Python", *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011. [Online]. Available: <https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>.
- [58] T. Akiba, S. Sano, T. Yanase, T. Ohta and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework", in *Proceedings of the 25rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, ISBN: 978-1-4503-6201-6. [Online]. Available: <https://arxiv.org/abs/1907.10902>.
- [59] M. Grochtmann and K. Grimm, "Classification trees for partition testing", *Software Testing*, vol. 3, 1993. DOI: <https://doi.org/10.1002/stvr.4370030203>.
- [60] A. Althnian, D. AlSaeed, H. Al-Baity, A. Samha, A. B. Dris, N. Alzakari, A. Abou Elwafa and H. Kurdi, "Impact of dataset size on classification performance: An empirical evaluation in the medical domain", *Applied Sciences*, vol. 11, no. 2, 2021, ISSN: 20763417. DOI: [10.3390/app11020796](https://doi.org/10.3390/app11020796).
- [61] A. Graps, "An introduction to wavelets", *IEEE Computer Science Engineering*, vol. 2, pp. 50–61, Feb. 1995. DOI: [10.1109/99.388960](https://doi.org/10.1109/99.388960).
- [62] P. Zhao and B. Yu, "Stagewise lasso", *Journal of Machine Learning Research*, vol. 8, pp. 2701–2726, Dec. 2007. [Online]. Available: <https://www.jmlr.org/papers/volume8/zhao07a/zhao07a.pdf>.

A **Supplementary Figures**

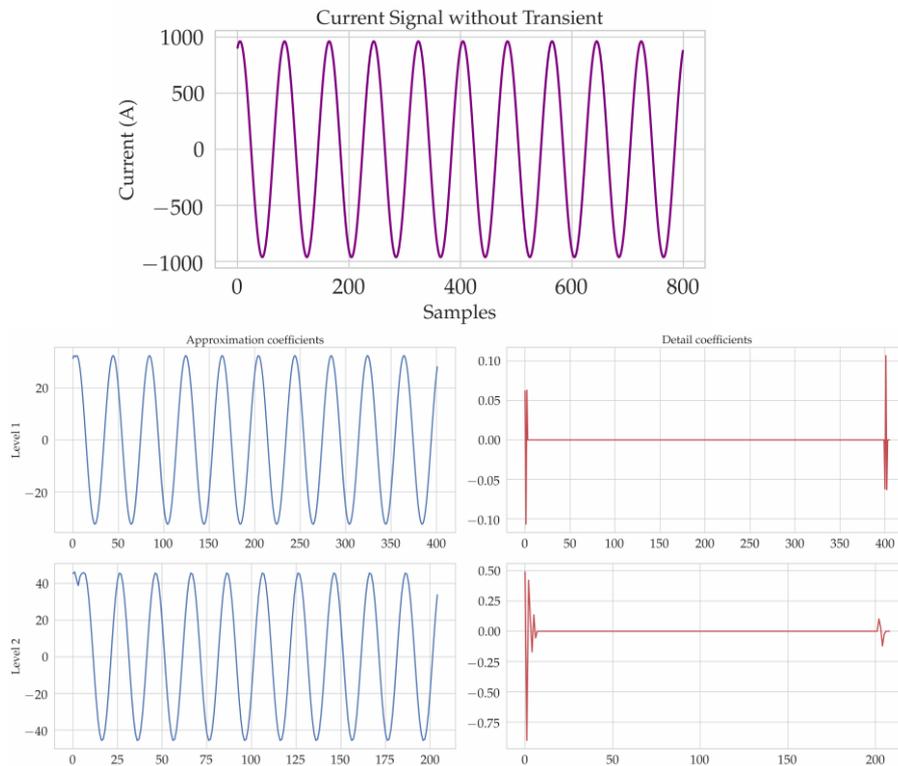


Figure A.1: Signal decomposition results for a waveform that does not possess any transients using the db4 mother wavelet. It can be observed that the detail coefficients are zero in the Level 1 (1-2 kHz) and Level 2 (0.5-1 kHz) frequency bands. The high values of the detail coefficients at the ends of the signal are due to the edge effects at the start and end of the input waveform.

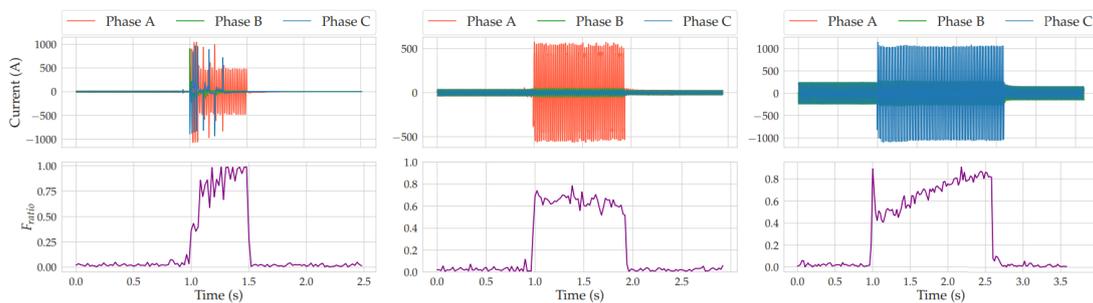


Figure A.2: The variation of F_{ratio} for the faults classified as stable (Figure 5.3 for reference) with the new business rules. It can be observed that F_{ratio} is highly correlated with the perceptible stability of the single-phase faults.

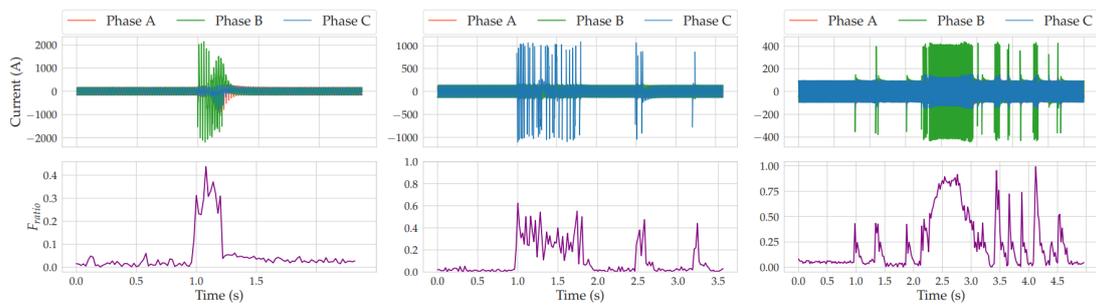


Figure A.3: The variation of F_{ratio} for the faults classified as unstable (Figure 5.4 for reference) with the new business rules. It can be observed that F_{ratio} does not reach as high values as with the examples of stable faults in Figure A.2 and that its variation is able to capture the distortions that are present in the original fault waveforms.

B **Supplementary Results**

Table B.1: Cross-validation scores and hyper-parameter values for the ANOVA-F feature selection technique.

Number of Features	Cross-Validation Score	Hyper-parameters	
		C	γ
0	0	-	-
1	54.11	1000	0.1
2	79.82	1	1
3	80.07	100	1
4	95.36	1000	1
5	96.17	1000	0.1
6	96.17	1000	0.1
7	96.45	1000	0.1
8	97	1000	0.1
9	97.27	1000	0.1
10	97.27	1000	0.1
11	97	1000	0.1
12	96.73	1000	0.1
13	96.45	1000	0.1
14	96.45	1000	0.1
15	96.18	100	0.1
16	96.45	10	1
17	96.44	100	1
18	96.99	100	1
19	96.72	100	1
20	96.44	100	1
21	96.14	100	0.1
22	96.45	10	1
23	96.18	100	0.1
24	96.18	100	0.1
25	96.17	100	1
26	96.45	10	1
27	96.45	10	1
28	96.18	10	1
29	96.17	100	1
30	95.91	100	0.1
31	95.07	1000	0.01
32	94.8	10	1
33	94.8	100	0.1
34	94.8	10	1
35	94.8	100	1
36	94.8	10	1

Table B.2: Cross-validation scores and hyper-parameter values for the mutual information feature selection technique.

Number of Features	Cross-Validation Score	Hyper-parameters	
		C	γ
0	0	-	-
1	81.42	1000	10
2	82.78	1000	10
3	81.95	1000	10
4	94.81	1000	10
5	94.261	1000	1
6	96.17	100	1
7	95.36	100	1
8	95.36	100	1
9	96.18	1000	1
10	96.19	1000	1
11	96.73	100	1
12	97.54	100	1
13	97.27	100	1
14	96.73	1000	0.1
15	96.73	100	1
16	96.73	1000	0.1
17	96.73	1000	0.1
18	96.73	1000	0.1
19	96.73	1000	0.1
20	96.73	1000	0.1
21	96.45	1000	0.1
22	96.45	1000	0.1
23	96.72	1000	0.1
24	96.72	1000	0.1
25	95.09	100	0.1
26	96.18	1000	0.1
27	95.9	1000	0.1
28	95.9	1000	0.1
29	95.9	1000	0.1
30	95.89	100	0.1
31	95.89	1000	0.01
32	95.64	100	1
33	95.57	100	1
34	95.49	1000	0.1
35	95.41	1000	0.1
36	95.34	1000	1