FEDERAL UNIVERSITY OF ITAJUBÁ - UNIFEI
GRADUATE PROGRAM IN
ELECTRICAL ENGINEERING

Robust Classification of Advanced Power Quality Disturbances in Smart Grids

Gabriel Caldas Sardinha de Almeida

Itajubá, November 18, 2021

FEDERAL UNIVERSITY OF ITAJUBÁ - UNIFEI
GRADUATE PROGRAM IN
ELECTRICAL ENGINEERING

Gabriel Caldas Sardinha de Almeida

Robust Classification of Advanced Power Quality Disturbances in Smart Grids

Dissertation submitted to the Graduate Program in Electrical Engineering as part of the requirements to obtain the title of **Master of Science in Electrical Engineering.**

Concentration Area: Electrical Power Systems

Supervisor: Prof. Ph.D. Paulo Fernando Ribeiro

November 18, 2021
Itajubá

Gabriel Caldas Sardinha de Almeida
Robust Classification of Advanced Power Quality Disturbances in Smart Grids/ Gabriel Caldas Sardinha de Almeida. – Itajubá, November 18, 2021-
73 p. : il. (algumas color.) ; 30 cm.

Supervisor: Prof. Ph.D. Paulo Fernando Ribeiro

Dissertation (Master of Science)
Federal University of Itajubá - UNIFEI
Graduate program in electrical engineering, November 18, 2021.

1. Palavra-chave1. 2. Palavra-chave2. I. Orientador. II. Universidade xxx. III. Faculdade de xxx. IV. Título

CDU 07:181:009.3

Gabriel Caldas Sardinha de Almeida

Robust Classification of Advanced Power Quality Disturbances in Smart Grids

Dissertation submitted to the Graduate Program in Electrical Engineering as part of the requirements to obtain the title of **Master of Science in Electrical Engineering.**

Work approved. Itajubá, November 18, 2021:

Prof. Ph.D. Paulo Fernando Ribeiro
Federal University of Itajubá (Supervisor)

Prof. Ph.D. Benedito Donizeti Bonatto
Federal University of Itajubá, Brazil

Prof. Dr. Carlos Augusto Duque
Federal University of Juiz de Fora, Brazil

Itajubá, November 18, 2021

Acknowledgements

I thank my parents for all the unconditional love that I have always received. If I could get here, it was for all the effort they made. Thank you for supporting my decisions, being my two eternal pillars in all phases of my life.

I want to deeply thank Professor Paulo Fernando Ribeiro for all kinds of help provided. With your guidance, I could become a better student, a better professional, but mainly, a better human being. Your advice made me a better person than when I started this journey. I appreciate it. You and professor Antonio C. Zambroni were like fathers to me.

I would also like to thank the professors of UNIFEI, employees, and all fellow students who provided me with the opportunity to hold a master's degree.

Finally, I would like to thank the funding agencies CAPES, CNPq, FAPEMIG, and INERGE for the financial support that made my training possible. This work was carried out with the support of the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financial Code 001.

"Don't count the things you do, do the things that count."

Zig Ziglar

Resumo

A inserção de novos dispositivos na rede, aumento do fluxo de dados, geração intermitente e a informatização massiva aumentaram consideravelmente a complexidade dos sistemas elétricos atuais. Esse aumento resultou em mudanças necessárias, como a necessidade de redes elétricas mais inteligentes para se adaptarem a essa realidade diferente. A nova geração de técnicas de Inteligência Artificial, representada pelo "Big Data", Aprendizado de Máquina ("Machine Learning"), Aprendizagem Profunda e Reconhecimento de Padrões representa uma nova era na sociedade e no desenvolvimento global baseado na informação e no conhecimento. Com as mais recentes Redes Inteligentes, o uso de técnicas que utilizem esse tipo de inteligência será ainda mais necessário. Esta dissertação investiga o uso de processamento de sinais avançado e também algoritmos de Aprendizagem de Máquina para desenvolver um classificador robusto de distúrbios de qualidade de energia no contexto das Redes Inteligentes. Para isso, modelos já conhecidos de alguns problemas de qualidade foram gerados junto com ruídos aleatórios para que o sistema fosse similar a aplicações reais. A partir desses modelos, milhares de sinais foram gerados e a Transformada Wavelet Discreta foi usada para extrair as principais características destas perturbações. Esta dissertação tem como objetivo utilizar algoritmos baseados no conceito de Aprendizado de Máquina para classificar os dados gerados de acordo com suas classes. Todos estes algoritmos foram treinados, validados e por fim, testados. Além disso, a acurácia e a matriz de confusão de cada um dos modelos foi apresentada e analisada. As etapas de geração de dados, extração das principais características e otimização dos dados foram realizadas no software MATLAB. Uma toolbox específica deste programa foi usada para treinar, validar e testar os 27 algoritmos diferentes e avaliar cada desempenho. Todas as etapas do trabalho foram previamente idealizadas, possibilitando seu correto desenvolvimento e execução. Os resultados mostram que o classificador "Cubic Support Vector Machine" obteve a máxima precisão entre todos os algoritmos, indicando a eficácia do método proposto para classificação. As considerações sobre os resultados foram interpretadas, como por exemplo a explicação da performance de cada técnica, suas relações e suas justificativas.

Palavras-chaves: Aprendizado de Máquina. Qualidade de Energia. Redes Inteligentes. Processamento de Sinais. Parâmetros Entrelaçados. Inteligência Artificial.

Abstract

The insertion of new devices, increased data flow, intermittent generation and massive computerization have considerably increased current electrical systems' complexity. This increase resulted in necessary changes, such as the need for more intelligent electrical networks to adapt to this different reality. Artificial Intelligence (AI) plays an important role in society, especially the techniques based on the learning process, and it is extended to the power systems. In the context of Smart Grids (SG), where the information and innovative solutions in monitoring is a primary concern, those techniques based on AI can present several applications. This dissertation investigates the use of advanced signal processing and ML algorithms to create a Robust Classifier of Advanced Power Quality (PQ) Disturbances in SG. For this purpose, known models of PQ disturbances were generated with random elements to approach real applications. From these models, thousands of signals were generated with the performance of these disturbances. Signal processing techniques using Discrete Wavelet Transform (DWT) were used to extract the signal's main characteristics. This research aims to use ML algorithms to classify these data according to their respective features. ML algorithms were trained, validated, and tested. Also, the accuracy and confusion matrix were analyzed, relating the logic behind the results. The stages of data generation, feature extraction and optimization techniques were performed in the MATLAB software. The Classification Learner toolbox was used for training, validation and testing the 27 different ML algorithms and assess each performance. All stages of the work were previously idealized, enabling their correct development and execution. The results show that the Cubic Support Vector Machine (SVM) classifier achieved the maximum accuracy of all algorithms, indicating the effectiveness of the proposed method for classification. Considerations about the results were interpreted as explaining the performance of each technique, its relations and their respective justifications.

Key-words: Advanced Power Quality. Artificial Intelligence. Interlaced Parameters. Machine Learning. Signal Processing. Smart Grids.

Contents

List of Figures 11

List of Tables 13

List of abbreviations and acronyms 14

1 INTRODUCTION 15
1.1 General Context 15
1.2 Objectives 16
1.3 Relevance 17
1.4 Methodology 17
1.5 Dissertation Structure 18

2 THEORETICAL BACKGROUND 19
2.1 State-of-Art 19
2.2 Smart Grid Context 21
2.3 Advanced Power Quality Disturbances 23
2.3.1 Impulsive Transient 25
2.3.2 Oscillatory Transient 26
2.3.3 Voltage Sag 26
2.3.4 Voltage Swell 27
2.3.5 Interruption 27
2.3.6 Harmonic Distortion 28
2.3.7 Voltage fluctuations 30
2.3.8 Interlaced Parameters 30
2.4 Signal Processing Stage 32
2.4.1 Discrete Wavelet Transform 32
2.4.1.1 Coefficients' Entropy 34
2.4.1.2 Coefficients' Energy 34
2.4.2 THD 34
2.5 Machine Learning Basics 35
2.5.1 Decision Trees 37
2.5.2 Naive Bayes Classifiers 38
2.5.3 Support Vector Machines 39
2.5.4 Nearest Neighbour Classifiers 40
2.5.5 Ensemble Classifiers 41

2.5.6 Neural Network Classifiers 42

3 METHODOLOGY 45
3.1 Data Generation 45
3.2 Feature Extraction Stage 46
3.3 Optimization Techniques 47
3.4 Training, Validation and Testing using the Classification Learner toolbox 48

4 RESULTS AND DISCUSSION 50
4.1 Performance and Classification Results 50
4.2 Discussion and Considerations 51

5 CONCLUSIONS 54
5.1 Research Conclusions 54
5.2 Future Works Possibilities 55

BIBLIOGRAPHY 56

APPENDIX 62

APPENDIX A – PUBLISHED PAPERS 63

APPENDIX B – MATLAB CODES 64

APPENDIX C – CONFUSION MATRIX OF ALL ML ALGORITHMS 65

List of Figures

Figure 1.1 – Steps in the development of the PQ classifier 18
Figure 2.1 – Features and technologies of SG 23
Figure 2.2 – Advanced PQ Context 24
Figure 2.3 – Impulsive Transient Disturbance 25
Figure 2.4 – Oscillatory Transient Disturbance 26
Figure 2.5 – Voltage Sag 27
Figure 2.6 – Voltage Swell 28
Figure 2.7 – Momentary Interruption 29
Figure 2.8 – Harmonic Distortion 29
Figure 2.9 – Voltage fluctuations 30
Figure 2.10–Voltage Sag with Harmonics 31
Figure 2.11–Voltage Swell with Harmonics 31
Figure 2.12–Big Picture of Signal Processing methodology used 32
Figure 2.13–Examples of different types of mother wavelets 33
Figure 2.14–Relation between AI and ML 35
Figure 2.15–Framework of ML algorithms 36
Figure 2.16–Learner types and algorithms used in this work 37
Figure 2.17–Example of decision tree model 37
Figure 2.18–A graphical representation of the naive Bayes' model for classification . 38
Figure 2.19–A SVM logic scheme 39
Figure 2.20–A graphical demonstration of the learning algorithm 40
Figure 2.21–An example of ensemble classifier: The random forest model 42
Figure 2.22–Model of a neuron 43
Figure 3.1 – Matrix input data before normalization 47
Figure 3.2 – Matrix input data after normalization 48
Figure 3.3 – Main window of Classification Learner toolbox 48
Figure 3.4 – Train, Validation and Test sets 49
Figure 4.1 – Confusion Matrix of the best classifier (Cubic SVM) 52
Figure 4.2 – ROC Curve of the best classifier (Cubic SVM) 53
Figure C.1 – Confusion Matrix of Medium Neural Network 65
Figure C.2 – Confusion Matrix of Quadratic SVM 65
Figure C.3 – Confusion Matrix of Medium Gaussian SVM 66
Figure C.4 – Confusion Matrix of Fine Gaussian SVM 66
Figure C.5 – Confusion Matrix of Wide Neural Network 66
Figure C.6 – Confusion Matrix of Bilayered Neural Network 67
Figure C.7 – Confusion Matrix of Narrow Neural Network 67

Figure C.8 – Confusion Matrix of Bagged Trees 67
Figure C.9 – Confusion Matrix of Subspace KNN 68
Figure C.10–Confusion Matrix of Linear SVM 68
Figure C.11–Confusion Matrix of Fine KNN 68
Figure C.12–Confusion Matrix of Weighted KNN 69
Figure C.13–Confusion Matrix of Fine Tree 69
Figure C.14–Confusion Matrix of Medium KNN 69
Figure C.15–Confusion Matrix of Boosted Trees 70
Figure C.16–Confusion Matrix of Coarse Gaussian SVM 70
Figure C.17–Confusion Matrix of Cubic KNN 70
Figure C.18–Confusion Matrix of Kernel Naive Bayes 71
Figure C.19–Confusion Matrix of Medium Tree 71
Figure C.20–Confusion Matrix of RUSBoosted Trees 71
Figure C.21–Confusion Matrix of Cosine KNN 72
Figure C.22–Confusion Matrix of Trilayered Neural Network 72
Figure C.23–Confusion Matrix of Gaussian Naive Bayes 72
Figure C.24–Confusion Matrix of Subspace Discriminant 73
Figure C.25–Confusion Matrix of Coarse KNN 73
Figure C.26–Confusion Matrix of Coarse Tree 73

List of Tables

Table 2.1 – Features of investigated PQ disturbances 25
Table 4.1 – Results for each classifier 51

List of abbreviations and acronyms

AC	Alternating Current
AI	Artificial Intelligence
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
DL	Deep Learning
DWT	Discrete Wavelet Transform
GA	Genetic Algorithm
ICT	Information and Communication Technology
IOT	Internet of Things
KNN	K-Nearest Neighbor
LG	Single-line to Ground
LL	Line to Line
LLG	Double-line to Ground
LLL	Three-phase
LLLG	Three-phase to Ground
ML	Machine Learning
MSE	Mean Squared Error
MVMN	Multivariate Multinomial Distribution
PE	Power Electronic
PQ	Power Quality
PSO	Particle Swarm Optimization
PWM	Pulse-width Modulation
RES	Renewable Energy Sources
RMS	Root Mean Square
ROC	Receiver Operating Characteristic
SG	Smart Grid
SOC	State of Charge
SVM	Support Vector Machine
THD	Total Harmonic Distortion
WPT	Wavelet Packet Transform
WT	Wavelet Transform

1 Introduction

1.1 General Context

Massive computerization, insertion of new devices, increased data flow, and intermittent generation have considerably increased current electrical systems' complexity. This increase resulted in necessary changes, such as the need for more intelligent electrical networks to adapt to this different reality. Artificial Intelligence (AI) may allow that these systems can become more robust and reliable.

The use of advanced techniques of AI in power systems is becoming more and more notorious in recent years. The new generation of AI technology represented by Big Data, Machine Learning (ML), Deep Learning (DL), Pattern Recognition, and other methods are closely related to the simulation analysis of large-scale power grids (TANG et al., 2018).

AI represents a new era in society and global development based on information and knowledge. With the recent Smart Grid (SG), the use of techniques that use this type of intelligence will be even more necessary. The AI domain, with advanced ML and cognitive computing capabilities, are a key enabler of unforeseen efficiency capabilities in the context of smart energy grids (ŞERBAN; LYTRAS, 2020).

ML is an area of AI that has emerged as part of the ongoing search for building intelligent machines capable of learning. It is self-learning based on algorithms that mean the system learns from its experience. It uses a statistical learning algorithm that automatically learns and improves without human help (SHARMA; SHARMA; JINDAL, 2021).

One of the essential elements for a good functioning of an electrical system is Power Quality (PQ). Therefore, a high-quality diagnosis of this grid is critical to identify and solve eventual problems. Besides deteriorating grid performance, the loss of PQ also has financial implications, resulting in extra costs related to the additional power consumed (PEREIRA et al., 1998).

In recent years, PQ has become a growing concern for both energy utilities and consumers of the power grid. The integration of distributed generation and renewable energies is one of the major sources of PQ disturbances. The increasing application of switching devices, nonlinear loads, rectifiers, lighting controls, protection, and relaying equipment are also the causes of the PQ disturbances (KHOKHAR et al., 2015). If these disturbances are not identified and classified, they can cause severe damage to distribution systems.

An adequate PQ guarantees the necessary compatibility between consumer equipment and the grid. So, it is an essential aspect of the power system which cannot be neglected (JWG C4.24/CIRED, 2018). In order to improve PQ, the sources and causes of such disturbances must be known before appropriate mitigating actions can be taken.

Distribution companies have devices along with the network that captures waveforms related to various PQ problems. However, due to operational limits such as time, money, qualified labor, and technical knowledge, most of these signals are stored in databases and are not analyzed or classified (WILSON et al., 2020). This loss of information can be vital for a better understanding of electrical distribution systems' condition operations.

Events such as voltage swell, capacitor switching transients, short-term faults, voltage sags, among others, are the leading causes of disturbances measured in the electrical grid. In this way, ML applications become ideal for identifying and classifying these problems in SG, identifying in real-time the thousands of events that occur within a period.

1.2 Objectives

This dissertation investigates the use of advanced signal processing and ML algorithms to create a Robust Classifier of Advanced PQ Disturbances in SG. For this purpose, known models of PQ disturbances were generated in MATLAB with random elements to approach real applications. From these models, thousands of signals were generated with the pattern of these disturbances. Signal processing techniques using Discrete Wavelet Transform (DWT) were used to extract the signal's main characteristics. This research aims to use ML algorithms to classify these data according to their respective features. ML algorithms were trained, validated, and tested. Finally, a comparison between all models was made, based on statistical analysis.

This process was carried out for different models of ML. Also, the accuracy and confusion matrix were analyzed, relating the logic behind the results. The software used was the Classification Learner from MATLAB, in which all stages of the process were developed. Below are some specific goals:

1. Highlight the contribution of combine advanced signal processing techniques with methods based on ML to carry out classification of PQ events in power systems;

2. Show that the use of statistical analysis tools is beneficial when comparing the results of validated models;

3. Bring out the specificity of each tested model and the relationship with its performance;

4. Contextualize modern subjects in the literature and its alignment with applications of the innovative SG context;

5. Develop and validate a robust classifier capable of differentiating ten types of PQ events;

6. Execute a performance analysis of the obtained results;

7. Make data, codes, and methods available in public databases to share with the community of professionals and researchers.

1.3 Relevance

The use of signal processing is recurrent in applications in electrical power systems. The Wavelet Transform (WT) can improve algorithms efficiency dealing with feature extraction for the pattern recognition of PQ disturbances. This tool presents the characteristic of decomposing a signal into different scales, with varying resolution levels from an analyzed signal. This signal processing technique can support AI applications on power systems monitoring and diagnosis since these two subjects present excellent development synergy. In recent times, ML techniques have proven to be effective in numerous applications, including power system studies (ALIMI; OUAHADA; ABU-MAHFOUZ, 2020). Therefore, there are more and more algorithms based on the joining of these elements.

Possibly, pattern recognition will be even more required in future power systems due to the applications' degree of complexity. Studies already demonstrate their potential applications in Internet of Things (IOT) integrated power systems (FARHOUMANDI; ZHOU; SHAHIDEHPOUR, 2021). In this way, ML algorithms play an essential role in overcoming these problems and stand out as a safe and effective technique for possible applications in SG.

1.4 Methodology

At first, an extensive literature review on the international databases and bibliographic references was made to construct the logical basis of the investigation. MATLAB software was used to perform the generation of signals database using mathematical models of the different quality events, extraction stages and optimization techniques. This tool, widely used by researchers, also enabled, through the Classification Learner app (ML specific toolbox), the training, validation, and testing of 27 ML models. Although detection is a vital tool in the PQ area, it was not the focus of our work. In this work, we start from the point where the models have already been detected and focus mainly on the part of disturbances classification.

In the final part of the methodology, proposed analyses and presentation of results with debates aimed at highlighting the work's contribution were conducted. Performance classification results were presented using accuracy and the confusion matrix only for the best cases. However, in the appendix, it is possible to view the analysis of all models. The methodology also provides the code and data used in this work. Fig. 1.1 illustrates the methodology steps.

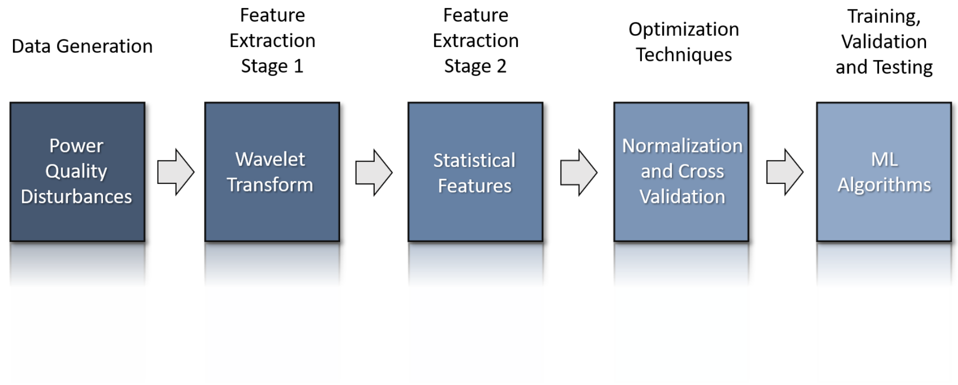

Figure 1.1 – Steps in the development of the PQ classifier.

1.5 Dissertation Structure

This dissertation is divided as follows:

Chapter 2 presents a literature review on pattern recognition for PQ disturbances. The theoretical tools used in this work with advanced signal processing and ML frameworks are presented. Also, the chapter describes the possible future context of SG.

Chapter 3 is directed towards the research's development procedures, going through the data generation, featuring extraction stage, optimization techniques, and ML algorithms design and training steps.

Chapter 4 presents the performance obtained in the ML algorithms test stage and some key considerations.

Finally, chapter 5 brings the research conclusions and the future works possibilities.

2 Theoretical Background

This chapter presents the theoretical background topics necessary for a better understanding of the development of this research. Starting with the state-of-art, exploring the SG context, presenting all advance PQ disturbances covered in this work, passing through the topics involving the Signal Processing Stage, and finally, a ML framework.

2.1 State-of-Art

The use of AI in power systems covers several works. In Oliveira et al. (2017), the authors developed a methodology for microgrid management in islanded conditions using the fuzzy logic methodology. The proposed controller determines what actions will be taken from the input data as renewable-based power, State of Charge (SOC) and renewable power penetration ratio.

Reference Barrios (2015) proposed an energy management system of a microgrid using Genetic Algorithm (GA) composed of two stages of optimization. MATLAB software was used to develop the optimization algorithm, which, according to the author, has shown good capabilities to improve PQ. In Moloi e Yusuff (2020) the authors applied neural networks and wavelet in power system fault diagnostics. The performance of the protection scheme mainly relied on the ability of the algorithm to classify the fault accurately. The Particle Swarm Optimization (PSO) method was also implemented to evaluate the input parameters of the Artificial Neural Network (ANN) classifier.

In Salles et al. (2020) the authors developed a fuzzy logic-based controller for battery energy storage systems and load management to support the operation of a microgrid with photovoltaic (PV) farm generation. The software MATLAB/Simulink was used to simulate the different scenarios. The results demonstrate that fuzzy logic improved the system since it managed the loads and allowed a lower-cost operation properly.

Focused on ML techniques specifically, the literature is also plenty. The work of Cerqueira et al. (2006) presented a method for the detection and classification of PQ disturbances using ANN-based classifiers. The numerical results from computational simulation showed a good performance of the proposed method. Already in Yu e Wang (2009), the authors also proposed a digital system for detection and classification of PQ disturbance, but this time using WT and multiclass Support Vector Machine (SVM).

The study of Chapaloglou et al. (2019) developed a smart energy management system for load smoothing and peak shaving based on a clustering algorithm capable of forecasting the next day's load pattern. The simulation results proved that by applying the proposed algorithm, they could achieve a smoother diesel generator operation and peak shaving.

In Erişti et al. (2013) the author proposes a method to identify and classify several PQ disturbances events using WT based signal processing technique and Least Square SVM. Features like Mean, Skewness, Standard deviation, and Entropy were used to feed the SVM. The authors in Cai et al. (2017) proposed a real-time detection of power system disturbances based on K-Nearest Neighbor (KNN) analysis. The method consists of two stages: offline modeling and online detection.

PQ is also a vital topic of this research. Therefore, it was necessary to highlight some works covering it. Reference (KHOKHAR et al., 2014) reviewed the modeling and simulation of the PQ disturbances due to the exploitation of some types of loads. The PQ disturbances were created by using parametric equations and electrical power distribution system models in MATLAB/SIMULINK environment. The simulation results show that the PQ disturbances created by the two methods are very similar, validating the system.

In S. Salles et al. (2020) the authors presented an alternative for visualizing PQ disturbances through 2-D images from scalograms based on the Continuous Wavelet Transform (CWT). Signals from three different sources were used: mathematical equations, models of transmission and distribution of energy in MATLAB/Simulink, and real signals from a database. The results showed the method's efficiency, showing the interlaced phenomena with simultaneity and the relationship between different signal variation types.

In Aung, Milanovic e Simmons (2004), the authors describes a modular software for assessment and visualization of voltage sag performance. It allows in-depth analysis of the voltage sag performance of the individual buses and the performance of the entire network at different voltage levels. In SALLES (2020) was developed a classifier of PQ performance based on DL algorithms. The author compared each Convolutional Neural Network (CNN) created and achieved high performance.

Focusing on signal processing, the use of WT is historical and constant. Reference (RIBEIRO, 1994) was the first paper to propose using wavelets in PQ problems. The author presented the basic concepts of wavelet theory, investigating its potential for power distortion analysis. In Xu et al. (2006), the author presented some techniques for the analysis and visualization of time-varying waveform distortions. The authors performed a comparative study on four techniques and concluded that the DWT method gives the best visualization for time-varying waveform distortions.

Reference (FRUNT; KLING; RIBEIRO, 2011) proposed using wavelets to analyze fluctuations in load and generation profiles. The authors could determine the fluctuation patterns of the profiles by filtering wavelet components based on their RMS values. They used four different case studies, elaborating on a load profile, wind generation, photovoltaic generation, and the operation of a microgrid, and identified the most relevant frequency detail scaling factors from the analysis. It was shown that the wavelet method could be applied to determine in what time range the essential fluctuations occur in either or both generation and load.

In Gomes et al. (2017) the authors presented the use of time-frequency analysis to identify a time-varying nonlinear load behavior. The technique was applied to a supply current measured at a university building where an elevator had been activated, allowing identifying its behavior at the moment of operation, using Wavelet Packet Transform (WPT). The authors concluded that time-varying harmonics provided by the wavelet decomposition proved effective for determining the behavior of a nonlinear load.

In Souza (2017), the author proposes two PQ monitoring tools focused on microgrids. The first one uses DWT to decompose the waveform at different levels of frequency. The second tool uses WPT to visualize better fundamental frequency and harmonics of greater magnitudes in a given time interval.

The wavelet entropy is also highlighted in some works, like in Zhengyou et al. (2011) where the authors proposed a power system transients classification based on this feature. The method was applicable in power systems together with a neural network classifier and achieved effective classifying result.

All these works make up the state-of-art related to the areas and applications of this work. From them, it is possible to build a solid base for more advanced research to be carried out, contributing to the knowledge of the entire scientific community. In the following sessions, the theoretical basis and research topics will be presented and discussed.

2.2 Smart Grid Context

The drastic transformation in the electric sector and global concern about climate change influenced the necessary modernization of the sector. The strong insertion of new technologies based on IOT and Industry 4.0 allowed the emergence of new concepts, such as the concept of SG. The first official definition of SG was provided by the US Congress (2007), describing with some characteristics as follows:

1. Increased use of digital information and controls technology to improve reliability, security, and efficiency of the electric grid;
2. Deployment and integration of distributed resources and generation, including renewable resources;

3. Deployment of smart technologies (real-time, automated, interactive technologies that optimize the physical operation of appliances and consumer devices) for metering, communications concerning grid operations and status, and distribution automation;

4. Integration of smart appliances and consumer devices;

5. Provision to consumers of timely information and control options.

Other important bodies have their own definition like (International Energy Agency, 2011): "A SG is an electricity network that uses digital and other advanced technologies to monitor and manage the transport of electricity from all generation sources to meet the varying electricity demands of end-users. SGs coordinate the needs and capabilities of all generators, grid operators, end-users, and electricity market stakeholders to operate all parts of the system as efficiently as possible, minimizing costs and environmental impacts while maximizing system reliability, resilience, and stability."

Although there are similarities, to this day, there is still no global definition on the subject. Some key points are the search for a more efficient (energy and economically), reliable and sustainable system. Besides, there is also a strong relationship in the proposal for decentralized systems, small-scale transmission, bidirectional communication, and active prosumers (producer and consumer at the same time). The use of Information and Communication Technology (ICT), Big Data and AI give the basement for the implementation of SG. Fig. 2.1 illustrates the general context and some features present in SG.

The crucial aspect of the SG s is the security concern. Not only physical security (theft of equipment and energy) but also cybersecurity. This is a topic that several authors widely address (XU et al., 2021; HUANG et al., 2019) because, in these new networks, everything is digital and online. In fact, the use of IOT and ML in this area is recurrent (DHARMADHIKARI et al., 2021).

Perhaps we can think that only nations with an advanced economy can benefit from all these technologies and innovations. However, the literature shows different (FADAEENEJAD et al., 2014; Saidani Neffati et al., 2021). In general, the changes required adaptation of the potential prime features of technology, covering niftier equipment integration and internal cross-platform communication adaptable to the context of each developing country. Another aspect of this energy transformation needs to be societal, adapting to the mindset of the policymakers, users, and building general awareness and sense of social responsibility towards clean energy ensuring the sustainability of the planet (MARUF et al., 2020).

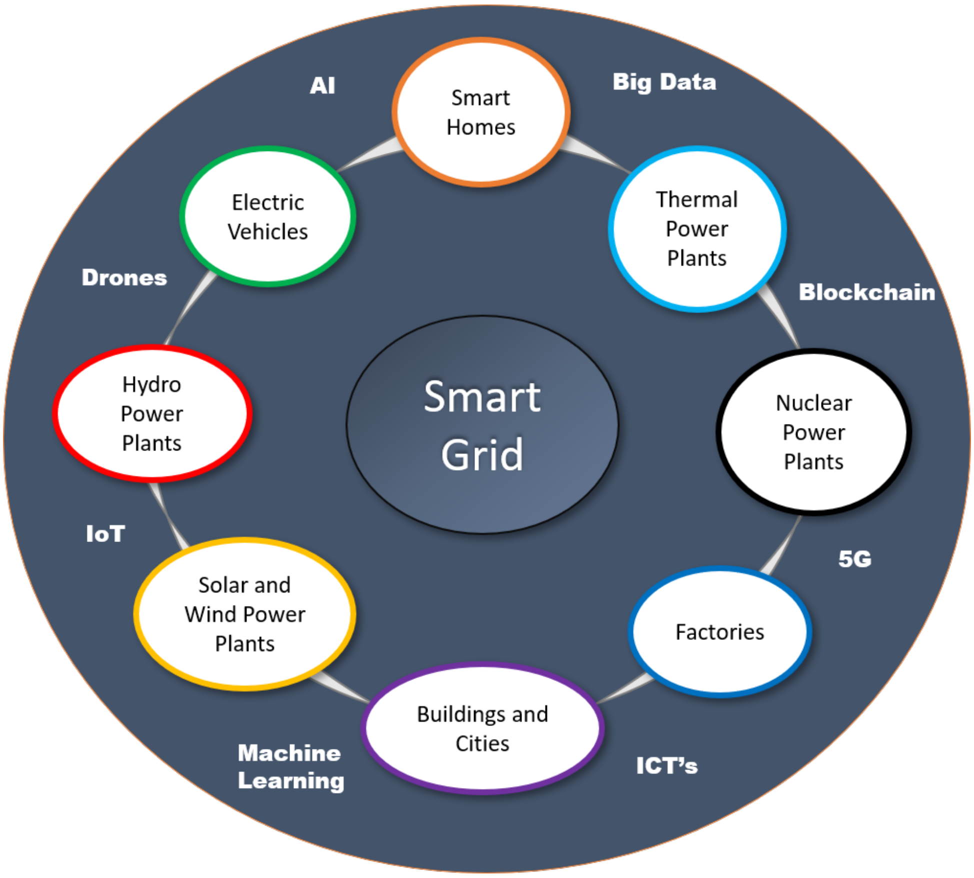

Figure 2.1 – Features and technologies of SG.

In this way, SGs does not exist exclusively in economically advanced countries with a high human development index. Quite the opposite, it is widely possible and viable to implement these technologies in developing economies with proper investments. SGs can also use modern technologies to identify and classification in real-time the thousands of events that occur within a period in order to adopt appropriate mitigating actions.

2.3 Advanced Power Quality Disturbances

The study of PQ is not new. Researches have been dealing with this subject for a long time, like reference IEEE (1994). However, with the modernization of power systems, new problems have arisen. That is why it is necessary to call them advanced PQ disturbances and study them with new signal processing techniques so that they can be understood and classified.

The term PQ refers to a wide variety of electromagnetic phenomena that characterize the voltage and current at a given time and at a given location on the power system (IEEE, 2019). These events can have their source in either the utility or customer wiring system or equipment.

PQ is part of the latest challenges to which the grid is exposed and for which the transition to the SG is necessary. The switching frequency of the converters in wind turbines can causes high-frequency signals flowing into the grid; Solar panels connected to the low-voltage networks will result in overvoltages for example; EV chargers generate harmonics, and the repeated starting of heat pumps can result in visible light flicker (BOLLEN; BAHRAMIRAD; KHODAEI, 2014).

Fig. 2.2 show some of the mainly PQ issues. It is important to stress that these issues generally do not occur separately. They often happen simultaneously, as is the case, of a voltage sag with the presence of harmonics. Because of this, they can be called Interlaced Parameters or interlaced parameters. This is one of the reasons for studying these effects. Upon learning about each parameter, it is possible to model and replicate them.

All of these disturbances have particular waveforms. The electrical waveforms' representation of the electric grid can be compared to the description of the electrocardiogram, representing the automatic function of a human heart, giving insights into its health and function. In the same manner, an evaluation of the electrical signals of a power grid can provide the electrical engineer with the ability to diagnose and predict possible malfunctions of the electric system (RIBEIRO et al., 2013).

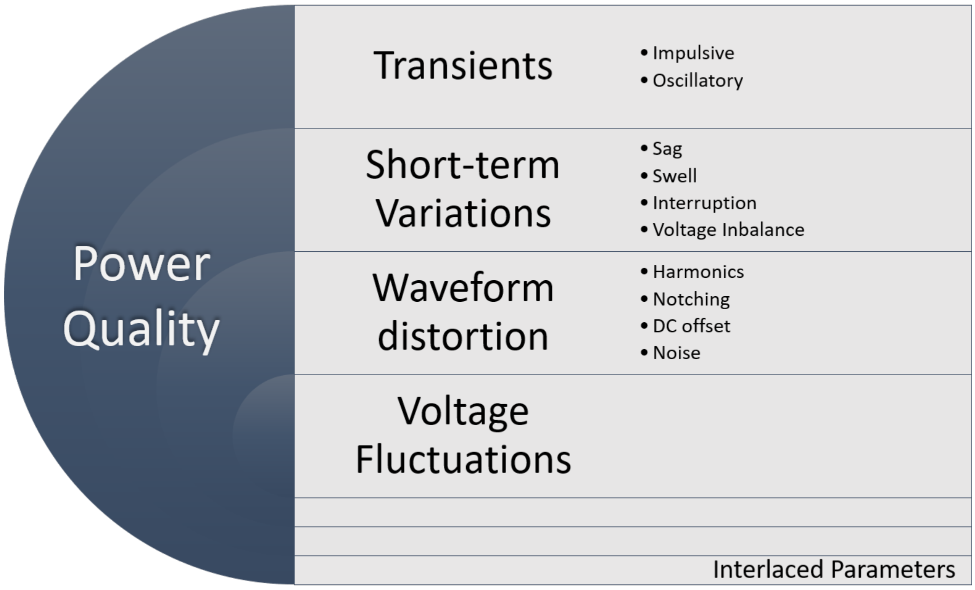

Figure 2.2 – Advanced PQ Context.

This section aims to describe only the events covered by this research. Table 2.1 summarizes the disturbance parameters used in this work. For information on other PQ events, it is advisable to read the standard IEEE (2019).

Table 2.1 – Features of investigated PQ disturbances (Adapted from (IEEE, 2019)).

Categories	**Spectral**	**Duration**	**Magnitude**
Transients			
Impulsive	0.1 ms rise	> 1 ms	
Oscillatory	<5 kHz	0.3-50 ms	0 - 4 pu
Short-term RMS variations			
Sag		< 0.5 - 30 cycles	0.1 - 0.9 pu
Swell		0.5 - 30 cycles	1.1 - 1.8 pu
Interruption		0.5 cycle - 3 sec	< 0.1 pu
Waveform distortion			
Harmonics	0 - 9 kHz	steady state	0–20%
Voltage fluctuations	<25 Hz	intermittent	0.1-7%

2.3.1 Impulsive Transient

Transients usually caused by lightning strikes are known as impulsive. A lightning strike to an overhead line or in the line's neighborhood will lead to a high overvoltage on the line. Such overvoltage often leads to an earth fault or a short-circuit fault (BOLLEN; GU, 2006).

This phenomenon has short-duration, highly damped, non-oscillatory overvoltages, with a few milliseconds or less of duration time. (JWG C4/C6.29, 2016) Due to the high frequencies involved, impulsive transients are damped quickly by impedance circuit elements. Fig. 2.3 illustrate an impulsive transient disturbance.

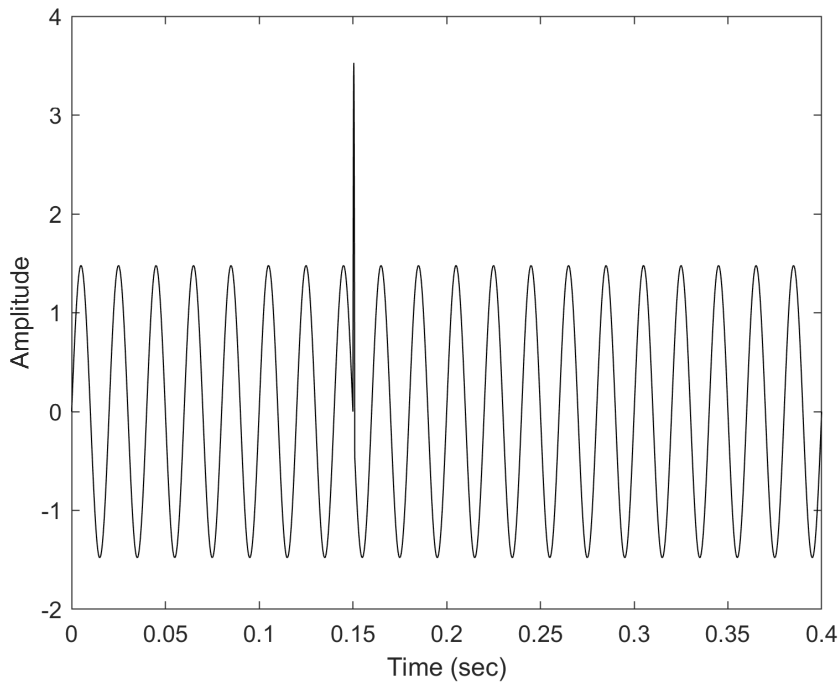

Figure 2.3 – Impulsive Transient Disturbance.

2.3.2 Oscillatory Transient

An oscillatory transient consists of a voltage or current whose instantaneous value changes polarity rapidly multiple times and usually decaying within a fundamental-frequency cycle (IEEE, 2019). The transients related to disconnecting loads, energizing the capacitor bank, and transformers are known as oscillatory transients.

The most severe case from a PQ viewpoint is the energizing of a capacitor. It will lead to an initial change in the voltage waveform toward zero followed by an oscillation with a frequency of a few hundred-hertz (BOLLEN; GU, 2006). The overvoltage is directly proportional to the number of capacitors present in the system. Fig. 2.4 shows an example of oscillatory transient.

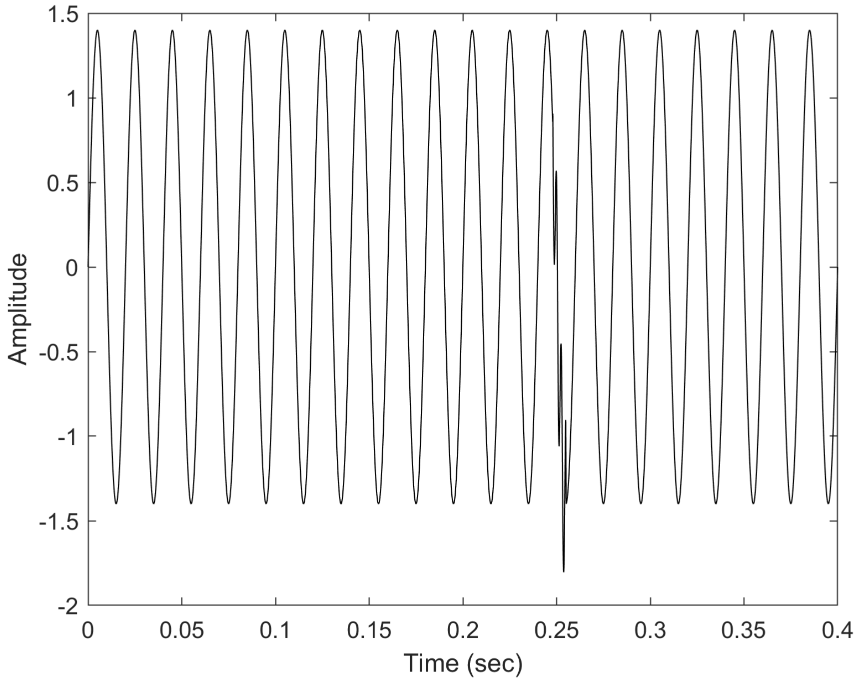

Figure 2.4 – Oscillatory Transient Disturbance.

2.3.3 Voltage Sag

Voltage sags or voltage dips are one of the most common types of PQ problems. They are generally associated with short circuit faults such as Single-line to Ground (LG), Line to Line (LL), Double-line to Ground (LLG), Three-phase (LLL), Three-phase to Ground (LLLG) faults. This problem can also be generated by energizing heavy loads such as energizing large motors. Some of the expected future network functionalities will result in more frequent or longer voltage sags and short interruptions (JWG C4.24/CIRED, 2018).

Voltage sags are short-duration reductions in voltage magnitude. A sag decreases rms voltage to between 0.1 pu and 0.9 pu for durations from 0.5 cycles to 1 min. Normal values are between 0.1 pu and 0.9 pu (IEEE, 2019). The residual voltage (the voltage magnitude during the event) may be anywhere close to zero and close to the nominal voltage (BOLLEN; GU, 2006). Fig. 2.5 shows a typical voltage sag.

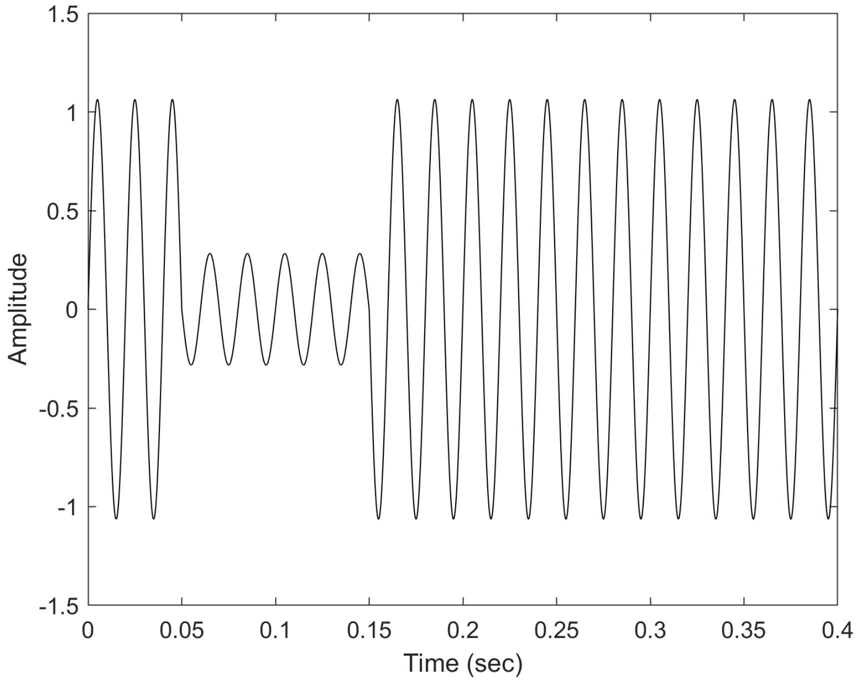

Figure 2.5 – Voltage Sag Disturbance.

2.3.4 Voltage Swell

A short-duration increase in voltage magnitude (JWG C4/C6.29, 2016), generally associated with short-circuit failures in electrical systems. In a fault involving a ground phase, this effect is created in the phase where the fault occurs while the swell is produced in the other phases. This transient also occurs when connecting and disconnecting loads or energizing a capacitor bank, for example.

They are the opposite of sags and much less common. A swell is an increase in rms voltage above 1.1 pu for durations from 0.5 cycles to 1 min, with magnitudes between 1.8 pu (IEEE, 2019). Fig. 2.6 shows an example of voltage swell.

2.3.5 Interruption

Interruptions are defined as a situation in which the voltage magnitude is zero or close to zero. Typical thresholds to detect an interruption are one and 10% of the nominal voltage, for a time of not more than 1-minute (IEEE, 2019).

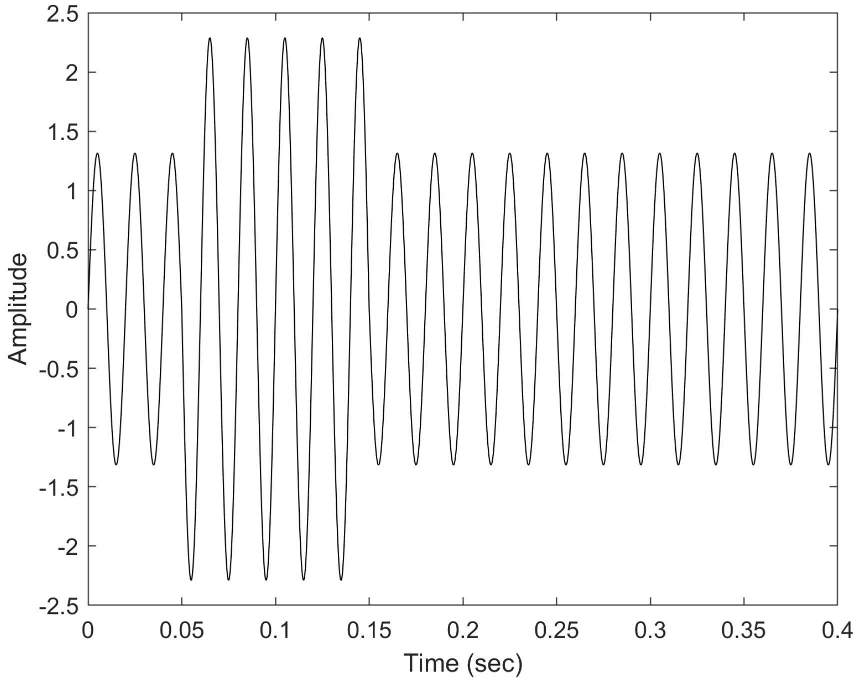

Figure 2.6 – Voltage Swell Disturbance.

Interruptions are due to the disconnection of one or more customers from the power network. The causes for this disconnection can be an opening of a circuit breaker or a fuse due to a short circuit, a negligent operation of a circuit breaker or intentional disconnection of part of the system (BOLLEN; GU, 2006). Fig. 2.7 illustrate an typical interruption event.

2.3.6 Harmonic Distortion

Harmonics are sinusoidal voltages or currents with frequencies that are integers multiple of the system's fundamental frequency. The connection of Renewable Energy Sources (RES) to the electricity network is mainly achieved through the use of Power Electronic (PE) converters, which are sources of harmonic distortion (JWG C4/B4.38, 2019).

The complete harmonic spectrum describes harmonic distortion levels with magnitudes and phase angles of each harmonic component. It is also common to use a single quantity, the Total Harmonic Distortion (THD), as a measure of the practical value of harmonic distortion (DUGAN et al., 2012).

Nonlinear devices in the power system cause harmonic distortion. Rectified input, switching power supplies often used in electronic-based equipment is a major contributor of harmonics in the power system (IEEE, 2019). A voltage source Pulse-width Modulation (PWM) inverter supplied from a six-pulse rectifier is one of the most common configura-

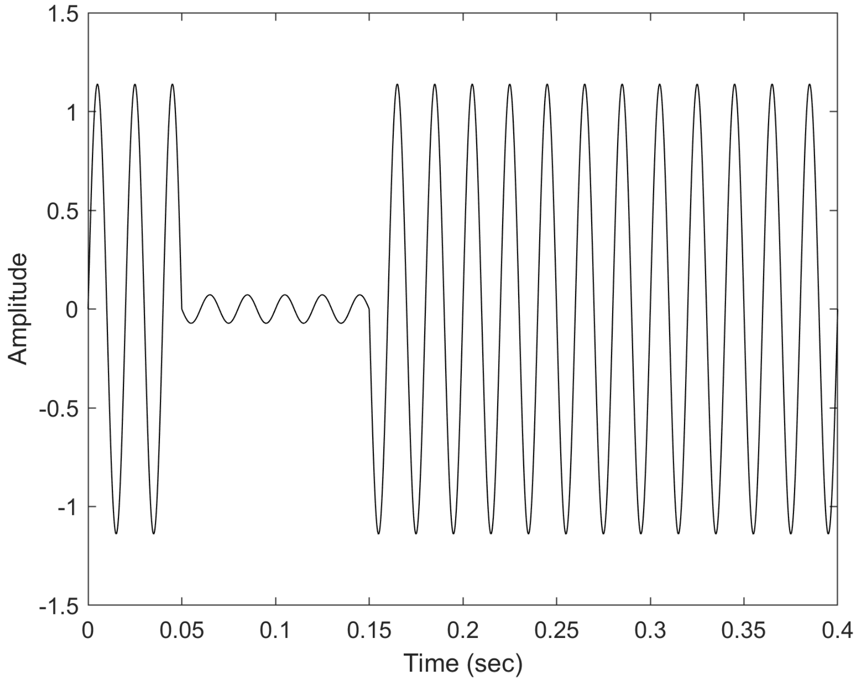

Figure 2.7 – Momentary Interruption.

tions used in variable frequency Alternating Current (AC) drives (JWG C4/B4.38, 2019). Fig. 2.8 shows an example of harmonic distortion.

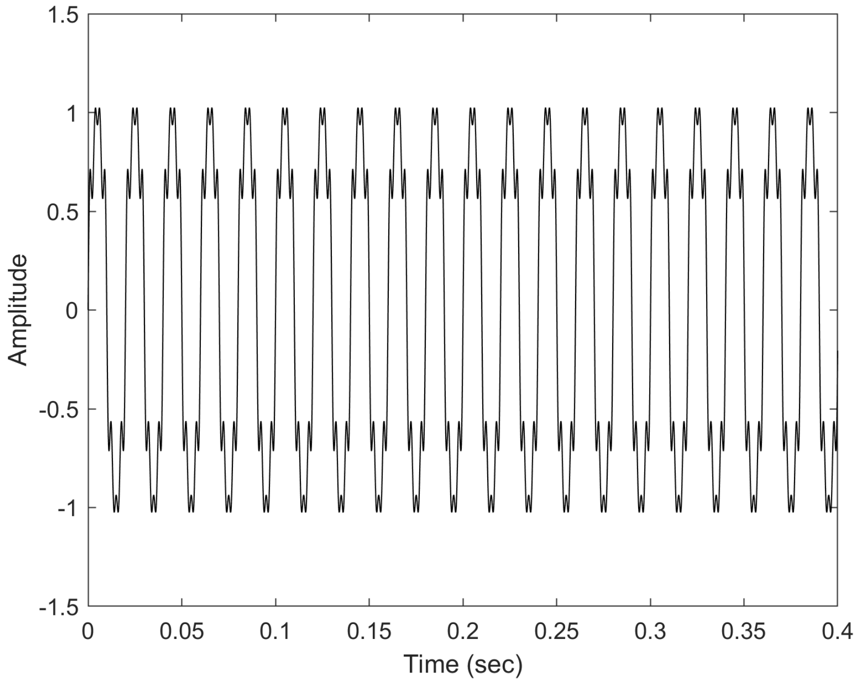

Figure 2.8 – Harmonic Distortion.

2.3.7 Voltage fluctuations

Variations of the voltage envelope or a series of random voltage changes. The magnitude of this Voltage fluctuations does not normally exceed the voltage ranges of 0.95 pu to 1.05 pu. Humans can perceive such voltage fluctuations by changes in lamp illumination intensity (DUGAN et al., 2012).

Loads that exhibit continuous, rapid variations in load current magnitude can cause voltage variations erroneously referred to as "flicker." The term is derived from the impact of the voltage fluctuation on lighting intensity. Voltage fluctuation is an electromagnetic phenomenon, and flicker is an undesirable result of that phenomenon on lighting. They are often confused to the point that the term voltage flicker is used in some documents when the term voltage fluctuation should be used. Such incorrect usage should be avoided (IEEE, 2019).

An essential source of voltage fluctuations is the arc furnace - a large electric oven in which metal is melted. The currents taken by an arc furnace vary at many different time scales (BOLLEN; GU, 2006). An example of a voltage fluctuation is shown in Fig. 2.9.

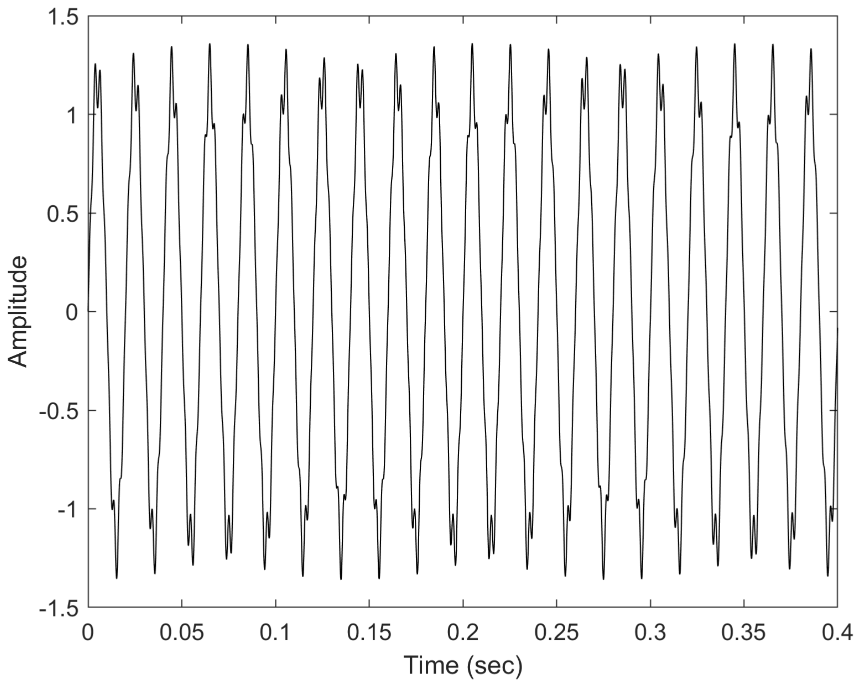

Figure 2.9 – Voltage fluctuations.

2.3.8 Interlaced Parameters

In this work, interlaced parameters were also modeled in order to be classified. The insertion of these parameters occurred so that the results to be obtained are even more similar to what happens in real power systems. Fig. 2.10 illustrates a typical example of voltage sag occurring with harmonics, while Fig. 2.11 shows an example of voltage swell occurring with harmonic distortion.

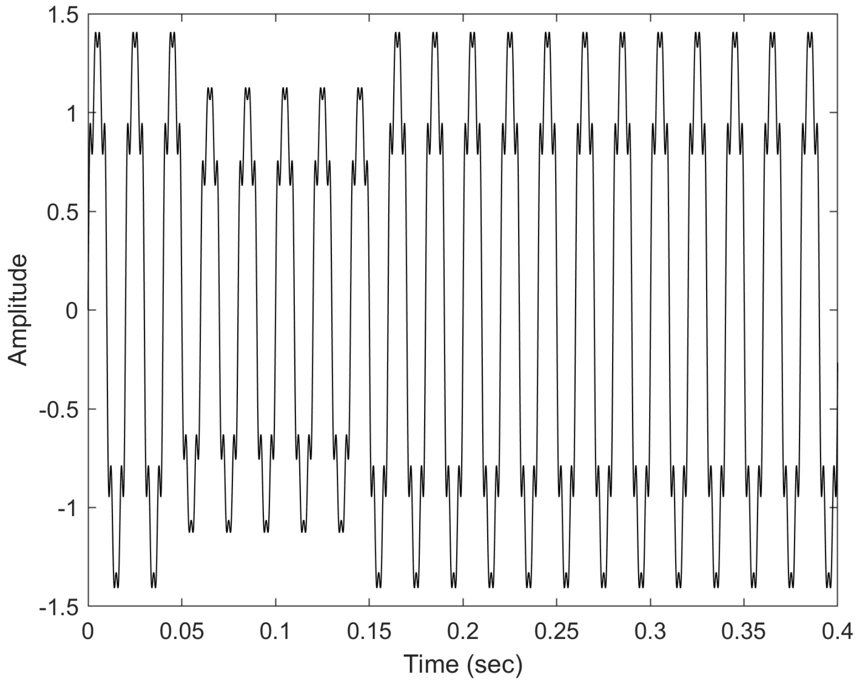

Figure 2.10 – Voltage Sag with Harmonics.

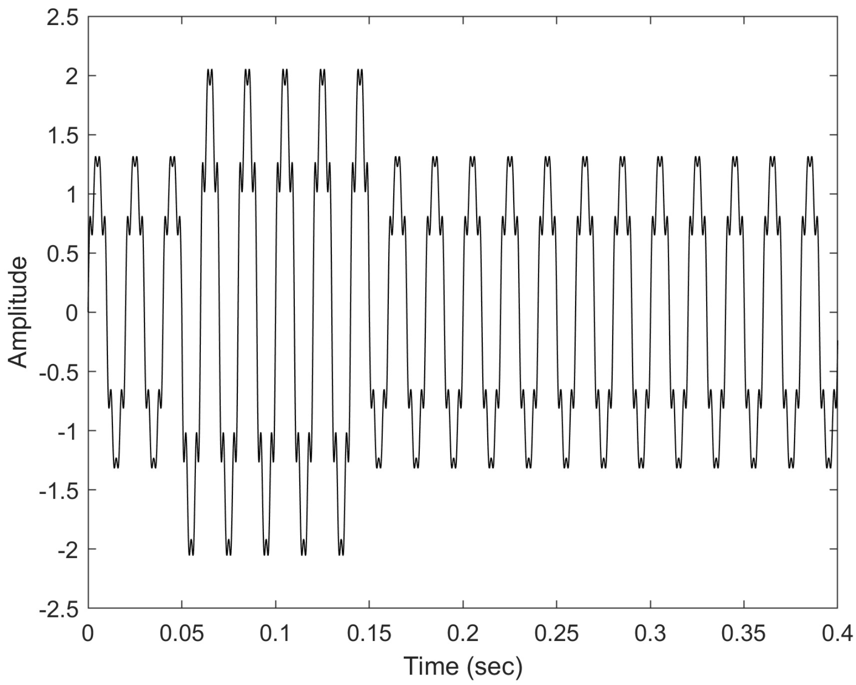

Figure 2.11 – Voltage Swell with Harmonics.

2.4 Signal Processing Stage

In this section, we will present the theoretical basis regarding the signal processing stage that was used in this work. Concepts involving WT, DWT, THD, energy and entropy will be discussed. In this work, the original signal was decomposed with DWT to obtain the coefficients' energy and entropy. Also, THD was extracted to increase the amount of information in the input data matrix of the next steps. Fig. 2.12 shows the big picture of the signal processing methodology used in this work.

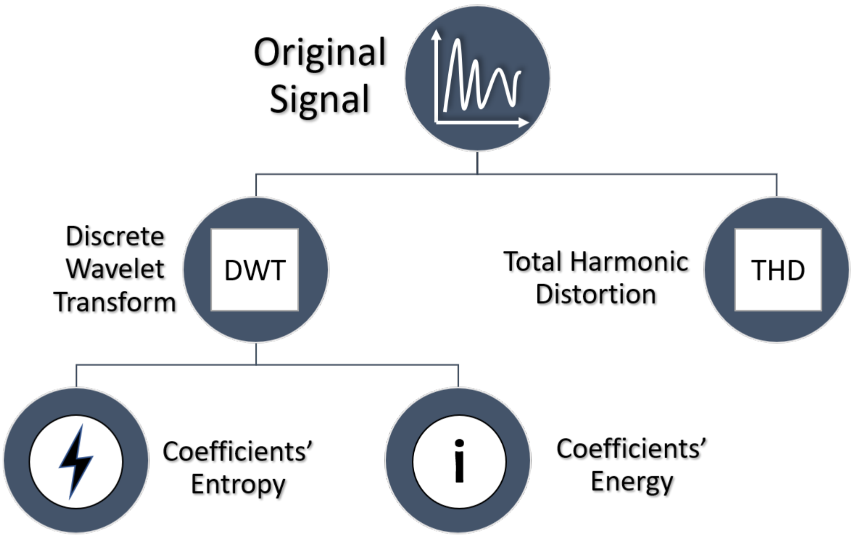

Figure 2.12 – Big Picture of the Signal Processing methodology used.

2.4.1 Discrete Wavelet Transform

The WT is an advanced signal processing tool that performs a significant role in feature extraction for the pattern recognition of PQ disturbances (IBRAHIM; MORCOS, 2002). This tool presents the characteristic of decomposing a signal into different scales, with varying levels of resolution from an analyzed signal (RODRIGUES; TOSTES, 2018). WT also provides a multi-resolution analysis, examining the signal at different time and frequency, separate from the Fourier transform, which only obtains information about the frequency.

Harmonic distortion assumes steady-state conditions and is consequently inadequate to deal with time-varying waveforms (RIBEIRO, 2010). Therefore the use of wavelets can be an alternative to the traditional harmonic analysis. In addition to performing an analysis on the frequency, it also performs an analysis in time, which allows its application in disturbance identification systems, for example.

The signal is decomposed into various scales of a short-term waveform called "mother wavelets" and analyzed. This typical wavelet is a fast decaying oscillating waveform with zero mean value. DWT is a transform derived from a CWT, with the scales and positions discretized while the signal is also discretized. According to (SILVEIRA; M.; RIBEIRO, 2007), DWT is represented as in Equation (2.1):

$$DWT_f^{\Psi}(j,k) = \frac{1}{\sqrt{a_0^j}} \sum_{n=-\infty}^{\infty} f(n)\Psi\left[\frac{n - a_0^j\, k\, b_0}{a_0^j}\right] \tag{2.1}$$

where $j, k, n \in \mathbb{Z}$ and $a_0 > 1$.

In the literature, there are several wavelet functions, such as Morlet, Shannon, Mexican Hat, Meyer, Gabor, and Gaussian, which can be selected according to the nature of the signal to be analyzed (BRONZINI et al., 2007). The wavelets are generated from the mother wavelet by dilations and translations of these parameters, resulting in several wavelet basis functions. Fig. 2.13 shows some examples of mother wavelets.

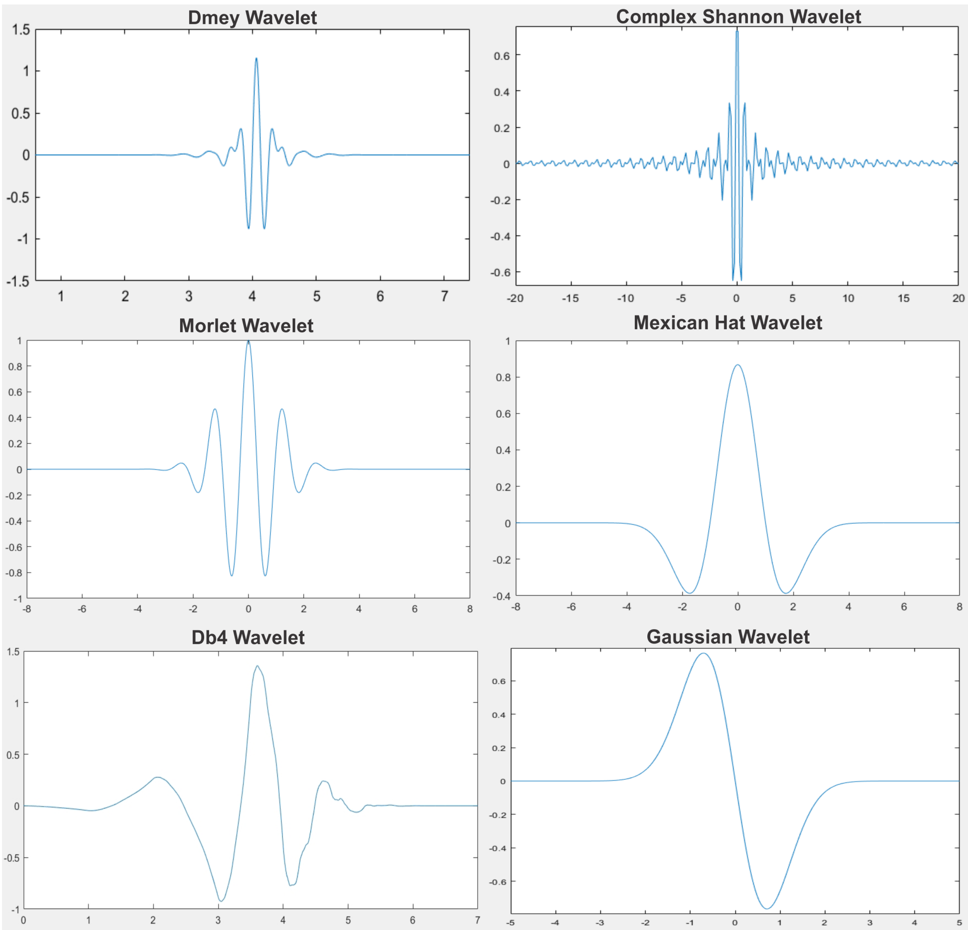

Figure 2.13 – Examples of different types of mother wavelets.

The choice of each wavelet must be consistent with the desired application, making it model the acquired signal accurately. Therefore, a suitable choice of a mother wavelet is not only useful and elegant, but it is also efficient (GALLI; HEYDT; RIBEIRO, 1996). In this work, the "dmey" mother wavelet was used in all PQ disturbs. This choice was due to tests that were carried out, confirming a better performance for this set. However, this mother wavelet is not always the best choice for all jobs involving PQ due to the specificity of each algorithm.

2.4.1.1 Coefficients' Entropy

The Shannon Entropy is an essential quantity in information theory associated with a variable, which can be interpreted as the average level of information inherent in the variable's possible outcomes. According to (COIFMAN; WICKERHAUSER, 1992), this feature can be represented in Equation (2.2):

$$E(s) = -\sum_{n=1}^{N} s_i^2 log(s_i^2) \tag{2.2}$$

where $E(s)$ is the Shannon Entropy, s is the signal and s_i the coefficients of s in an orthonormal basis.

This process resulted in the extraction of three inputs: the entropy of the original signal, the entropy of the approximate coefficients, and the entropy of the detailed coefficients.

2.4.1.2 Coefficients' Energy

The energy of wavelet coefficients gives information about the strength of signals (MathWorks Help Center, 2021d) and is given by the equation:

$$En(s) = \sum_{n=1}^{N} |s_i|^2 \tag{2.3}$$

where $En(s)$ is the Energy, s is the signal and s_i the coefficients of s in an orthonormal basis. This process resulted in the extraction of two inputs: the mean value and the max percentage value of the detailed coefficients energy.

2.4.2 THD

THD is an important parameter used to quantify the level of harmonics in voltage or current waveforms. Along with other parameters, it can be an excellent indicator of the quality of an electrical network, indicating the occurrence of a disturbance according to its level of distortion.

According to IEC 61000-2-2:2002 (2002), THD is defined as the ratio of the Root Mean Square (RMS) value of all the harmonic frequencies of the signal, over the RMS value of its fundamental frequency, where this signal is a measured voltage or current. Equation (2.4) express the THD in percentage of a voltage signal.

$$THD(\%) = \frac{\sqrt{\sum_{n=2}^{\infty} U_n^2}}{U_{rms}}.100 \tag{2.4}$$

where U_n is the RMS value of the harmonic n and U_{rms} is the RMS value of the fundamental frequency.

In MATLAB, the function $thd(x)$ returns the THD of the real-valued sinusoidal signal "x". The THD is determined from the fundamental frequency and the first five harmonics using a periodogram of the same length as the input signal (MathWorks Help Center, 2021f).

2.5 Machine Learning Basics

AI is a term that associates human thinking (decision-making, problem-solving, and learning) and the automation of activities (BELLMAN, 1978). When inserting human thinking into a machine, we can say that this machine has intelligence, even artificial. From this main concept, numerous areas have emerged that implement human characteristics in their algorithms, such as Robotics, Fuzzy Logic, GA, and ML. Fig. 2.14 illustrates the relation between AI and ML.

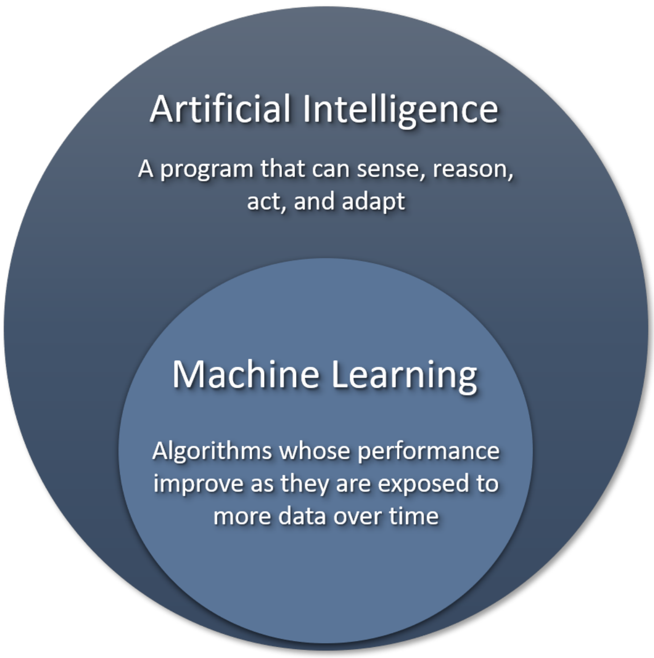

Figure 2.14 – Relation between AI and ML (Adapted from(SINGH, 2018)).

The term "machine learning" was first quoted in (SAMUEL, 1959). After that, several books, articles, and research on the term boosted and defined it as a branch of

AI. ML is characterized by the ability with which machines learn from experience. The main characteristic of this learning is that it is constant, allowing the model to produce correct data based on previously known rules.

ML draws on concepts and results from many fields, including statistics, AI, philosophy, information theory, biology, cognitive science, computational complexity, and control theory (MITCHELL, 1997). Each learner type is based on one of these concepts, and that is why there are so many types of learning. However, there are three main types of learning problems in ML: supervised, unsupervised, and reinforcement learning. Supervised learning was the main learning used in this work.

In supervised learning, the input and target examples are used to create the system's logic. Applications in which the training data contain examples of the input vectors and their corresponding target vectors are known as supervised learning problems (BISHOP, 2006). Among all types of supervised learning problems, two main types stand out: classification, which involves predicting a class label, and regression, which consists of predicting a numerical value.

As this work aims to classify different events from PQ, only ML classification algorithms were used. But the entire area is much more expanded. Fig. 2.15 shows a framework of some ML algorithms, highlighting the ones used in this research.

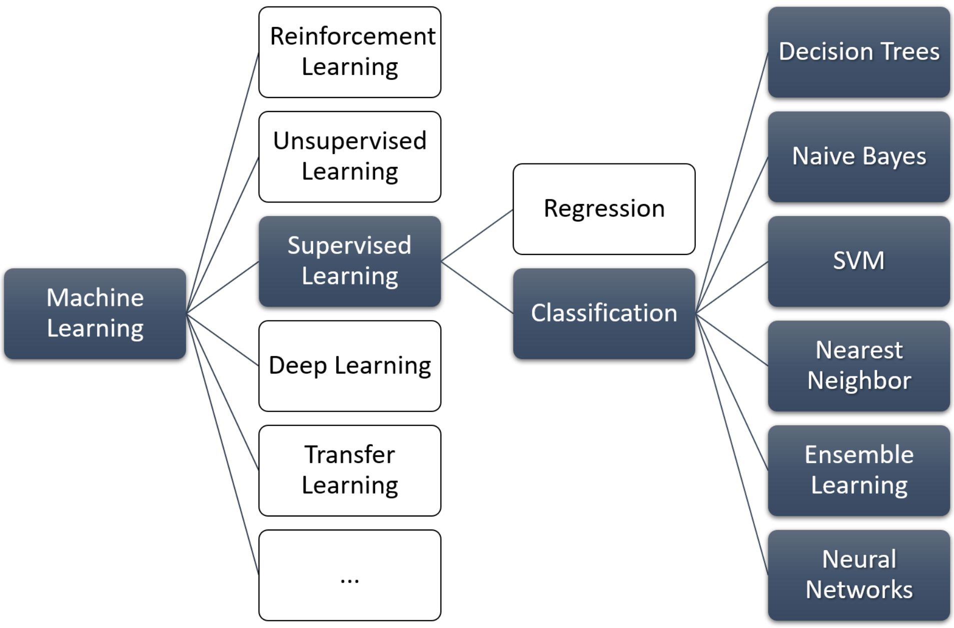

Figure 2.15 – Framework of ML algorithms.

In this work, ML is used to classify PQ disturbances employing six learner types, which will be commented in the next section. The setup of the 27 algorithms and their differentiation will also be presented to relate these characteristics in the future to the results presented in chapter 3. Fig. 2.16 shows all learner types and algorithms used.

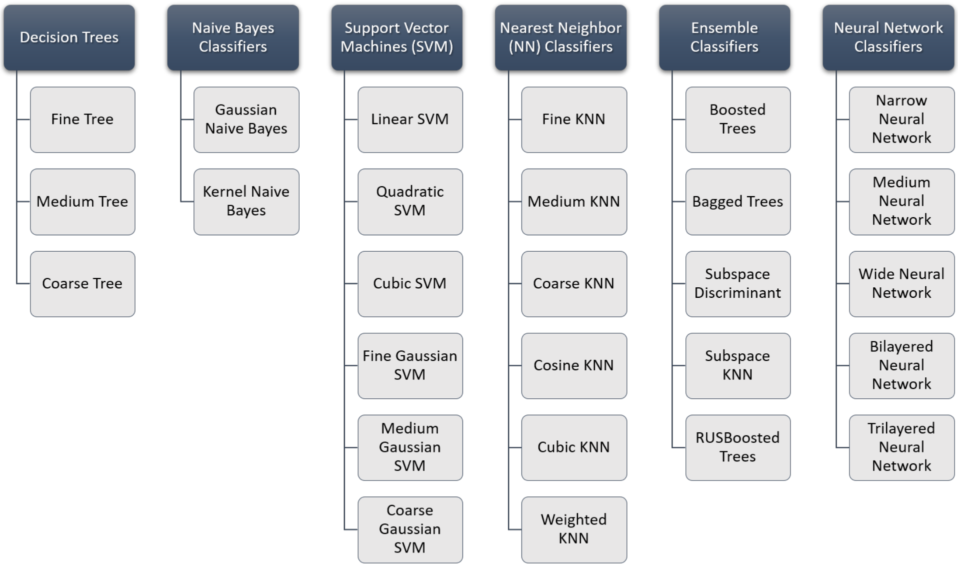

Figure 2.16 – Learner types and algorithms used in this work.

2.5.1 Decision Trees

Decision tree learning is one of the simplest and yet most widely used and practical methods for inductive inference. Decision trees classify instances by sorting them down the tree from the root to some leaf node, which provides the classification of the instance (MITCHELL, 1997). An example of a decision tree model is shown in Fig. 2.17.

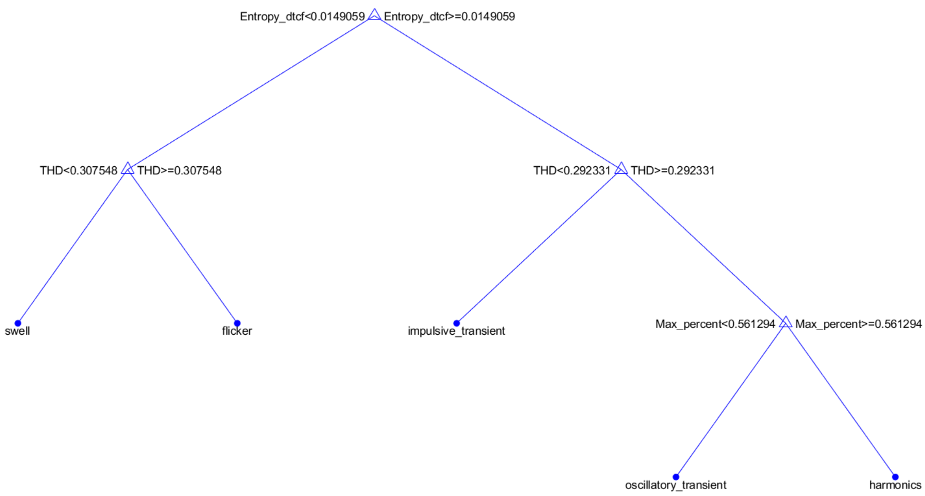

Figure 2.17 – Example of decision tree model.

Each triangle is called a split, and generally, the more splits you have, the more accurate your decision tree will be. The last node of the decision tree, where a decision is made, is called the tree leaves. The following are the settings of the algorithms created that was based on this learning:

- Fine tree: a decision tree with many leaves that makes fine distinctions between classes (maximum number of splits is 100);
- Medium tree: a decision tree of medium flexibility with fewer leaves (maximum number of splits is 20);
- Coarse tree: a simple decision tree with few leaves that makes coarse distinctions between classes (maximum number of splits is 4).

One limitation of decision trees is that the division of input space is based on hard splits in which only one model is responsible for making predictions for any given value of the input variables (BISHOP, 2006).

2.5.2 Naive Bayes Classifiers

It is a classification technique based on Bayes' theorem assuming independence between predictors to simplify the model structure. A Naive Bayes classifier assumes that a particular feature in a class is unrelated to the presence of any other feature.

Fig. 2.18 exemplifies this theorem. Conditioned on the class label z, the components of the observed vector $x = (x1, ..., xD)^T$ are assumed to be independent.

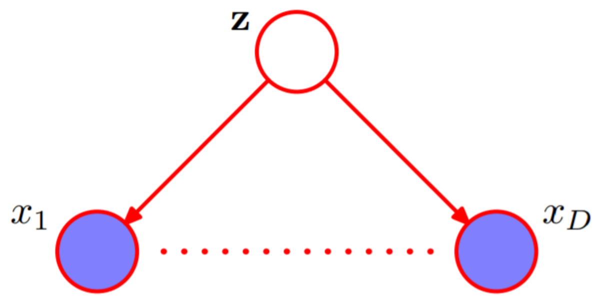

Figure 2.18 – A graphical representation of the naive Bayes' model for classification (Adapted from (BISHOP, 2006)).

The main difficulty applying this method is that they normally require initial knowledge of many probabilities. When these probabilities are not known in advance, they are often estimated based on background knowledge, previously available data, and assumptions about the form of the underlying distributions (MITCHELL, 1997). The algorithms used with this type of learning and its configurations are described below:

- Gaussian Naive Bayes: a naive Bayes classifier that uses Gaussian distribution for numeric predictors and Multivariate Multinomial Distribution (MVMN) for categorical predictors;

- Kernel Naive Bayes: a naive Bayes classifier that uses kernel distribution for numeric predictors and MVMN for categorical predictors.

2.5.3 Support Vector Machines

SVM are a type of ML tool that analyzes data and recognizes patterns or decision boundaries within the dataset used mainly for classification and regression analysis (GHOSH; DASGUPTA; SWETAPADMA, 2019). A SVM classifies data by finding the best hyperplane that separates data points of one class from those of the other class. The best hyperplane for an SVM means the one with the most significant margin between the two classes. This margin means the maximal width of the slab parallel to the hyperplane that has no interior data points. Many planes can separate the two classes, but only one plane can maximize the margin or distance between the classes.

The support vectors are the data points that are closest to the separating hyperplane. In Fig. 2.19, these points are on the boundaries. The location of the support vectors defines the maximum margin hyperplane. Other data points can be moved around freely (so long as they remain outside the margin region) without changing the decision boundary, and so the solution will be independent of such data points (BISHOP, 2006).

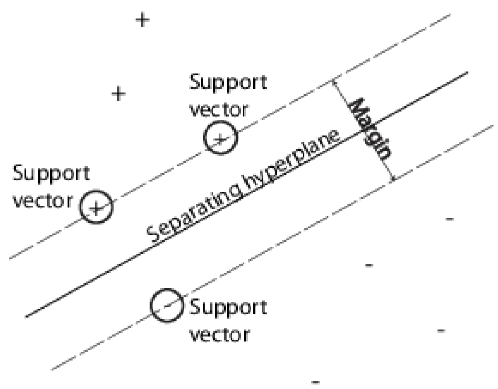

Figure 2.19 – A SVM logic scheme (Adapted from (MathWorks Help Center, 2021b)).

Five algorithms based on SVM learning were selected. The kernel was used to expand the algorithm limits, like the quadratic kernel, the cubic kernel, and the Gaussian kernel. The following are the detailed settings for each model:

- Linear SVM: a support vector machine that makes a simple linear separation between classes, using the linear kernel;

- Quadratic SVM: a support vector machine that uses the quadratic kernel;

- Cubic SVM: a support vector machine that uses the cubic kernel;

- Fine Gaussian SVM: a support vector machine that makes finely-detailed distinctions between classes, using the Gaussian kernel;
- Medium Gaussian SVM: a support vector machine that makes fewer distinctions than a fine gaussian SVM, using the Gaussian kernel;
- Coarse Gaussian SVM: a support vector machine that makes coarse distinctions between classes, using the gaussian kernel.

2.5.4 Nearest Neighbour Classifiers

The learner KNN is a family of techniques that can be used for classification or regression. As a non-parametric learning algorithm, KNN are not restricted to a fixed number of parameters (GOODFELLOW; BENGIO; COURVILLE, 2016). This method categorizes query points based on their distance to points (or neighbors) in a training dataset.

K nearest neighbors is a simple algorithm that can stores all available cases and classifies new cases by a majority vote of its k neighbors. The greater the number of neighbors, the less accurate it will be. Fig. 2.20 shows an example of KNN algorithm.

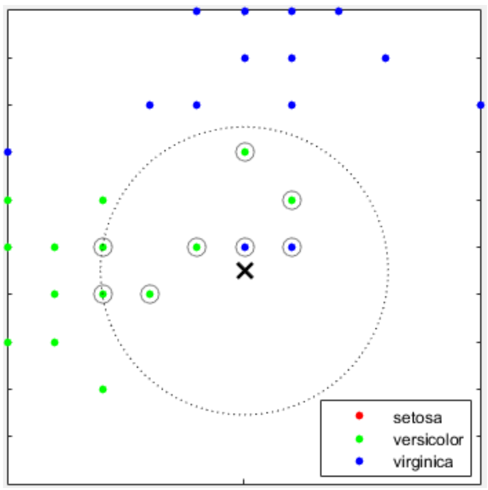

Figure 2.20 – A graphical demonstration of the learning algorithm (Adapted from (MathWorks Help Center, 2021b)).

The Euclidean metric distance was defined as a standard metric. The Cosine, Cubic, and Weighted KNN have their own metrics distances. The algorithms used with this type of learning and its complete setups are described below:

- Fine KNN: a nearest-neighbour classifier that makes finely-detailed distinctions between classes, with the number of neighbors set to 1;
- Medium KNN: a nearest-neighbour classifier that makes fewer distinctions than a fine KNN, with the number of neighbors set to 10;

- Coarse KNN: a nearest-neighbour classifier that makes coarse distinctions between classes, with the number of neighbors set to 100;
- Cosine KNN: a nearest-neighbour classifier that uses the cosine distance metric;
- Cubic KNN: a nearest-neighbour classifier that uses the cubic distance metric;
- Weighted KNN: a nearest-neighbour classifier that uses distance weighting.

The main problem of this method is that it typically considers all instances when retrieving similar training examples from memory. If the target concept depends on only a few of the many available attributes, then the instances that are truly most similar may well be a large distance apart (MITCHELL, 1997).

2.5.5 Ensemble Classifiers

Ensemble classifiers meld results from many weak learners into one high-quality ensemble model (MathWorks Help Center, 2021b). The quality of each one depends on the choice of algorithm. The ensemble method combines multiple models, where each model in the ensemble makes a prediction. A majority vote determines the final prediction among the models.

The Random Forest algorithm is a popular ensemble method used for decision trees. In this method, multiple decision trees are created using bootstrapped datasets of the original data and randomly selecting a subset of variables at each step of the decision tree. The model then chooses the mode of all of the predictions of each decision tree (SHIN, 2020). The bagged trees, for example, uses this method. In Fig. 2.21 is possible to visualize the structure of the Random Forest model.

Ensemble methods train multiple learners to solve the same problem. In contrast to common learning approaches, which try to construct one learner from training data, ensemble methods try to build a set of learners and combine them. Ensemble learning is also called committee-based learning or learning multiple classifier systems (ZHOU, 2012). Much effort is put into what types of weak learners to combine and how to combine them. The following items are the detailed settings for each ensemble classifier:

- Boosted Trees: the model creates an ensemble of medium decision trees using the AdaBoost algorithm. Compared to bagging, boosting algorithms use relatively little time or memory but might need more ensemble members;
- Bagged Trees: a bootstrap-aggregated ensemble of fine decision trees. Often very accurate, but can be slow and memory intensive for large data sets;

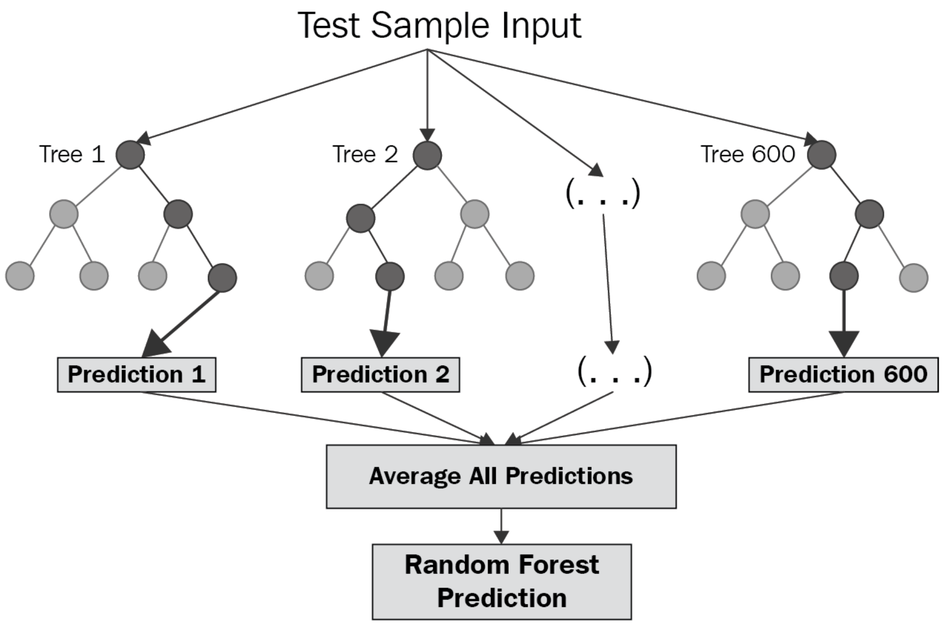

Figure 2.21 – An example of ensemble classifier: The random forest model (Adapted from (PACKT, 2021)).

- Subspace Discriminant: good for many predictors, relatively fast for fitting and prediction, and low on memory usage, but the accuracy varies depending on the data. The model creates an ensemble of Discriminant classifiers using the random subspace algorithm;
- Subspace KNN: the model creates an ensemble of nearest-neighbour classifiers using the Random Subspace algorithm;
- RUSBoosted Trees: the model creates an ensemble of decision trees using the RUS-Boost algorithm, using recommended for skewed data with many observations of one class.

2.5.6 Neural Network Classifiers

The Neural Network is a computational model formed by individual processing units, the artificial neurons. They are interconnected by weights that can be modified according to the quality parameters that evaluate the proximity between the required response and the one obtained (MONTEIRO et al., 2018).

This classifier is an universal approximator that can model any nonlinear function with desired accuracies (WU; ZHANG; CHEN, 2016). The networks are arranged in layers, with the first layer taking in inputs and the last layer producing outputs. The artificial neurons are the central element of an ANN because they are responsible for connecting each layer (HAYKIN, 2001). The model of a neuron is illustrated in Fig. 2.22

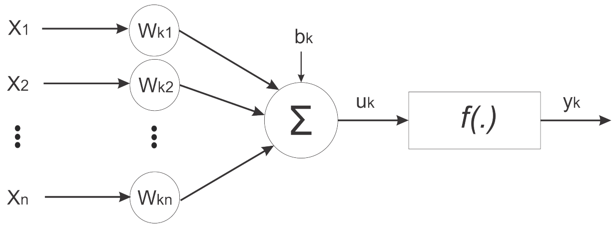

Figure 2.22 – Model of a neuron.

Where the input signals vector $X := [x_1, x_2, ..., x_n]^t, n \in \mathbb{N} = [1, |X|]$, the neurons synaptic weights $W_k := [w_{k1}, w_{k2}, ..., w_{kn}]^t, n \in \mathbb{N} = [1, |X|]$, u_k is the weight's multiplication response with the input signals, b_k is the bias which is an external parameter of the neuron, $f(.)$ is the activation function and y_k is the output response of the neuron.

A back-propagation algorithm based on experimental results is used to train the network. The multiple inputs are applied from previously recorded data to the input layer, each multiplied by a weight, and the product summed (MIRIAM; SEKAR; AMBALAVANAN, 2013). The algorithm updates the network weights so that the Mean Squared Error (MSE) in the network's result is minimized.

Each unit has a set of input links from other units, a set of output links to other units, a current activation level, and a means of computing the activation level at the next step in time, given its inputs and weights (RUSSELL; NORVIG, 1995). The activation function of all trained models was the Rectified Linear Unit or ReLU Layer. The algorithms based on these classifiers are described below:

- Narrow Neural Network: a neural network classifier with one fully connected layer of size 10;
- Medium Neural Network: a neural network classifier with one fully connected layer of size 25;
- Wide Neural Network: a neural network classifier with one fully connected layer of size 100;
- Bilayered Neural Network: a neural network classifier with two fully connected layers, excluding the final fully connected layer for classification;
- Trilayered Neural Network: a neural network classifier with three fully connected layers, excluding the final fully connected layer for classification.

Neural Networks are the most famous and used classification method, with massive growth and popularity in the field. For example, DL algorithms also use this method. For certain types of problems, such as learning to interpret complex real-world sensor data, ANNs are among the most effective learning methods currently known (MITCHELL, 1997).

3 Methodology

This chapter describes the procedures and stages of development. First, all models used to generate PQ signals will be detailed. After, the extraction stage is presented, showing which inputs were extracted from the previous models. Soon after, the procedures adopted for the correct filtering and separation of the data are presented. Finally, the Classification Learner software is detailed, with all steps taken until the test stages.

3.1 Data Generation

In this section, functions were used to generate the disturbances, according to Igual et al. (2018). The chosen sampling frequency was 10kHz, the nominal electrical frequency was set at 50Hz, and all disturbances have 20 cycles. Nine PQ disturbances were simulated: Voltage Sag, Voltage Swell, Momentary Interruption, Harmonic Distortion, Impulsive Transient, Oscillatory Transient, Voltage Sag with Harmonics, Voltage Swell with Harmonics and Voltage Fluctuations.

The healthy condition was also simulated so that the system is also able to identify when the network has no disturbances. The simulation loop for the training group was 500 iterations for each disturbance, each one having little randomness to reduce problems with overfitting. The loop for the test set was 250 iterations for each disturbance.

The Equation 3.1 represents a Voltage Sag disturbance, where α value is between 0.1 and 0.9; A is the amplitude, and $v(t)$ is the desired signal.

$$v(t) = A(1 - \alpha(u(t - t_1) - u(t - t_2)))sin(\omega t - \phi) \tag{3.1}$$

Equation 3.2 describes a Voltage Swell disturbance, where $0.1 \leq \beta \leq 0.8$.

$$v(t) = A(1 + \beta(u(t - t_1) - u(t - t_2)))sin(\omega t - \phi) \tag{3.2}$$

A Momentary Interruption is represent in Equation 3.3, where ρ value is between 0.9 and 1.0.

$$v(t) = A(1 - \rho(u(t - t_1) - u(t - t_2)))sin(\omega t - \phi) \tag{3.3}$$

Harmonic Distortion is represented in Equation 3.4, where α_n value is between

0.05 and 0.15; $n = [3, 5, 7]$.

$$v(t) = A[sin(\omega t - \phi) + \sum_{n=3}^{7} \alpha_n sin(n\omega t - \upsilon_n)] \tag{3.4}$$

The Equation 3.5 represents a Impulsive Transient, where ψ value is between 0.222 and 1.11.

$$v(t) = A[sin(\omega t - \phi) - \psi(e^{-750(t-t_a)} - e^{-344(t-t_a)})((u(t - t_a) - u(t - t_b)))] \tag{3.5}$$

Equation 3.6 describes a Oscillatory Transient, where $-\pi \leq \upsilon \leq \pi$.

$$v(t) = A[sin(\omega t - \phi) + \beta e^{-(t-1)/\tau} sin(w_n(t - t_I) - \upsilon)((u(t - t_{II}) - u(t - t_I)))] \tag{3.6}$$

Voltage Sag with Harmonics is represented in Equation 3.7, where $\alpha_{n'}$ value is between 0.05 and 0.15; $n' = [3, 5]$.

$$v(t) = A(1 - \alpha(u(t - t_1) - u(t - t_2)))[sin(\omega t - \phi) + \sum_{n'=3}^{5} \alpha_{n'} sin(n'\omega t - \upsilon_{n'})] \tag{3.7}$$

The Equation 3.8 represents a Voltage Swell with Harmonics.

$$v(t) = A(1 + \beta(u(t - t_1) - u(t - t_2)))[sin(\omega t - \phi) + \sum_{n'=3}^{5} \alpha_{n'} sin(n'\omega t - \upsilon_{n'})] \tag{3.8}$$

Finally, Voltage Fluctuation disturbance is represented in Equation 3.9, where $0.05 \leq \lambda \leq 0.1$, $w_f = 2\pi f_f$ and $8Hz \leq f_f \leq 25Hz$

$$v(t) = A[1 + \lambda sin(w_f t)]sin(\omega t - \phi) \tag{3.9}$$

3.2 Feature Extraction Stage

After generating and saving the data, the next step is to extract the most important characteristics of each signal. Single-level DWT was applied using the mother wavelet *dmey* to extract this information. The good synergy between wavelet and entropy is stressed in some papers, like in Erişti et al. (2013). This is because wavelet meets the demands of transient signal analysis and entropy is ideal for the measurement of uncertainty (ZHENGYOU et al., 2011). So, the six extracted inputs are:

- Entropy of the original signal;
- Entropy of the approximate coefficients;
- Entropy of the detailed coefficients;
- Mean value of the energy of the detailed coefficients;
- Max percentage value of the energy of the detailed coefficients;
- THD from original signal.

The concatenation process is also necessary to obtain a full matrix with all inputs. Here, every input becomes a column, being the final column of the related event. Fig. 3.1 shows this matrix after the concatenation process and before the normalization process. In this figure, each line contains information about each disturbance.

Matrix

2500x7 table

	1 Entropy_original	2 Entropy_apcf	3 Entropy_dtcf	4 Mean_energy	5 Max_percent	6 THD	7 Event
498	-549.0452	-3.4271e+03	0.1202	8.0926e-05	48.7777	-26.3966	"sag"
499	438.7159	-1.5482e+03	0.1199	8.4844e-05	46.1945	-24.4722	"sag"
500	-243.9977	-2.7049e+03	0.1211	8.5869e-05	45.9647	-24.8155	"sag"
501	695.3827	-1.3829e+03	0.1177	7.8647e-05	49.4663	-29.1690	"swell"
502	-1.2393e+03	-4.4326e+03	0.1198	8.1795e-05	48.1181	-31.2575	"swell"
503	-1.1426e+03	-4.0857e+03	0.1189	8.3697e-05	46.6001	-33.4309	"swell"

Figure 3.1 – Matrix input data before normalization

3.3 Optimization Techniques

The normalization process aims to standardize all samples, with the minimal value corresponding to 0 and the max value corresponding to 1. This process allows the system to be more accurate. The normalization is represented in Equation 3.10, where N is the normalized data, x is the original input data, $max(x)$ and $min(x)$ are maximal and minimal values between all inputs. Fig. 3.2 shows the matrix after the normalization process.

$$N = \frac{x - min(x)}{max(x) - min(x)} \tag{3.10}$$

The last step before the Training Stage is the K-fold cross-validation. In this method, a partition of the dataset is formed by splitting it into k non-overlapping subsets. This procedure is based on the idea of repeating the training and testing computation on different random splits of the original dataset. The method protects against overfitting by partitioning the data set into folds and estimating accuracy on each fold. It also gives a

Matriz_norm

2500x7 table

	1 Entropy_original	2 Entropy_apcf	3 Entropy_dtcf	4 Mean_energy	5 Max_percent	6 THD	7 Event
498	0.8163	0.7841	0.0057	0.0185	0.9344	0.4034	"sag"
499	0.8448	0.8401	0.0067	0.0350	0.8593	0.4787	"sag"
500	0.8581	0.8619	0.0065	0.0367	0.8493	0.4950	"sag"
501	0.6827	0.6502	0.0049	0.0261	0.8929	0.0788	"swell"
502	0.9149	0.9018	0.0012	0.0082	0.9632	0.2077	"swell"
503	0.8169	0.7835	0.0041	0.0120	0.9587	0.2632	"swell"

Figure 3.2 – Matrix input data after normalization

reasonable estimate of the predictive accuracy of the final model trained with all the data (MathWorks Help Center, 2021e). In this work, the dataset was divided into five subsets.

3.4 Training, Validation and Testing using the Classification Learner toolbox

The toolbox presented in MATLAB, named Classification Learner, helped in this step of the work. This software allows you to choose among various ML algorithms to train and validate classification models for binary or multiclass problems (MathWorks Help Center, 2021c). First, the data set from previous stages is inserted into the process. After it is necessary to choose which classifiers will be used and their variations (as presented in Chapter 2.5). Finally, the software performs training and shows the performance of each classifier. Fig. 3.3 illustrates the main software's window.

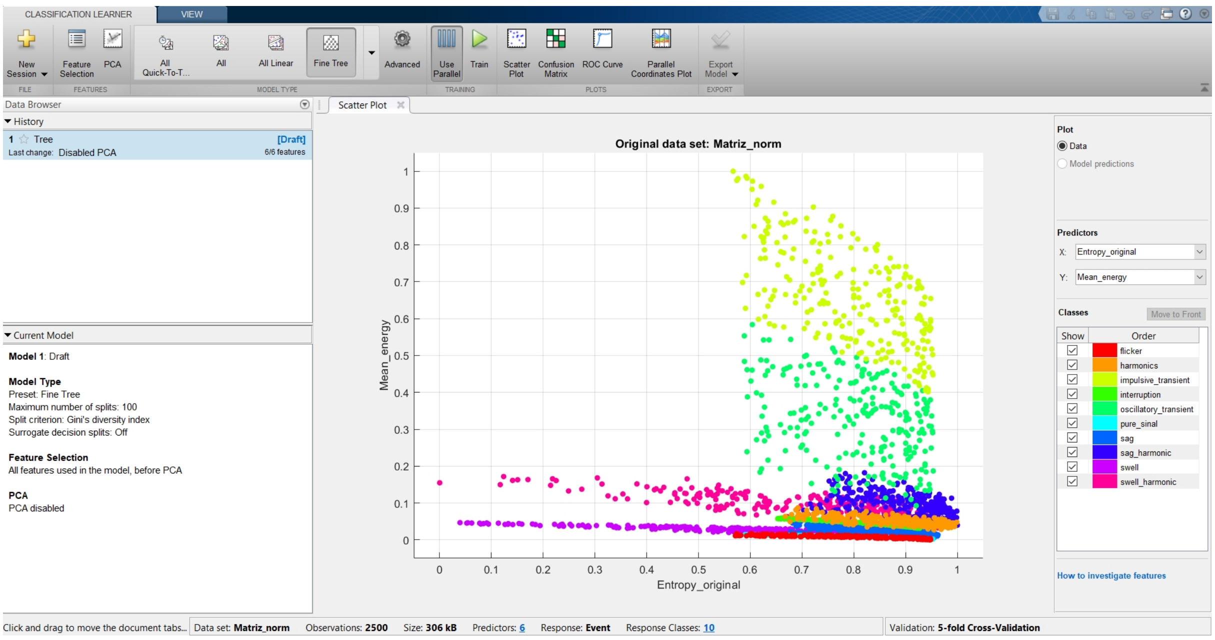

Figure 3.3 – Main window of Classification Learner toolbox

The sets of inputs used are explained in Fig. 3.4. The training and validation process used 5000 inputs, divided into five subsets as a consequence of cross-validation. The test stage contained half of the data from the previous stage (2500 inputs), with these data being new and unknown. The reason for using different data and less than what was trained is precisely to obtain a more accurate assessment of accuracy, simulating what would happen in a real situation.

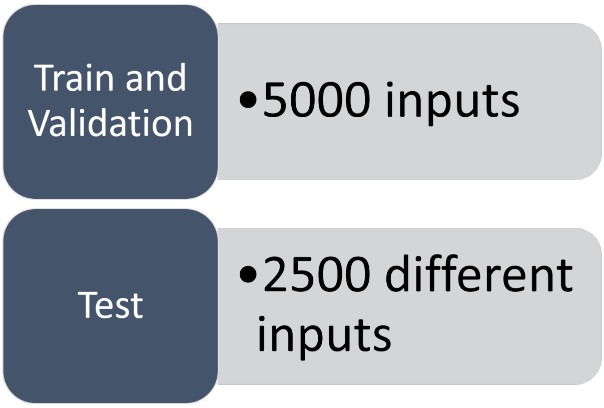

Figure 3.4 – Train, Validation and Test sets

4 Results and Discussion

This chapter presents the performance and classification results with each classification method for the test set. In this way, it is possible to verify the models' validity and compare the classification performance through accuracy. A discussion of the results found is also taken into consideration.

4.1 Performance and Classification Results

A table was prepared with each classifier type, its accuracy, and training time to gather all results. Table 4.1 shows the data for the set of tests after the network has been trained and validated. The accuracy is the percentage of correctly classified observations in relation to the total number of events.

In this work, the objective was to find the classifier with the most significant accuracy, where the accuracy is the correct number of ratings out of the total number of events. Therefore, the training time was not taken into account when choosing the best algorithm. It can be seen from the data that the highest classification accuracy is 98.9%, which was produced by Cubic SVM, followed by Medium Neural Network and Quadratic SVM with 98.6% and the third in rank is the Medium Gaussian SVM with 98.5%.

The confusion matrix of the best classifier (Cubic SVM) is used to illustrate the results in Fig. 4.1. This matrix helps us to understand how the currently selected classifier performed in each class. The numbers in blue color represent the correct predictions, and the remaining numbers are the wrong predictions. Also, the confusion matrix identifies the areas where the classifier has performed poorly, like the $sag - harmonic$ class, with 12 wrong predictions. For a more detailed analysis, the confusion matrix of all classifiers is present in Appendix C, ordered by the highest accuracy.

Another way to understand the classifier performance is by checking the Receiver Operating Characteristic (ROC) curve. The ROC curve shows the true positive rate versus the false positive rate for the currently selected trained classifier. In Fig. 4.2 is possible to visualize the curve from the worst class of the classifier ($sag - harmonic$).

A perfect result with no misclassified points is a right angle to the top left of the plot. A poor result that is no better than random is a line at 45 degrees. The area under curve number is a measure of the overall quality of the classifier. Larger area under curve values indicate better classifier performance (MathWorks Help Center, 2021a). In the case, the classifier performed well, with a large area under the curve and a correct angle in the top left of the plot.

Table 4.1 – Results for each classifier

Classifier	**Classifier type**	**Accuracy (%)**	**Training time (sec)**
Decision Trees	Fine Tree	96.8	12.13
	Medium Tree	95.3	02.21
	Coarse Tree	49.8	01.73
Naive Bayes	Gaussian Naive Bayes	92.6	07.31
	Kernel Naive Bayes	95.6	34.44
SVM	Linear SVM	97.5	13.33
	Quadratic SVM	98.6	09.18
	Cubic SVM	98.9	09.13
	Fine Gaussian SVM	98.4	16.74
	Medium Gaussian SVM	98.5	08.82
	Coarse Gaussian SVM	96.4	08.33
Nearest Neighbor Classifiers	Fine KNN	97.2	03.62
	Medium KNN	96.6	01.91
	Coarse KNN	92.1	01.97
	Cosine KNN	95.2	02.13
	Cubic KNN	96.2	02.87
	Weighted KNN	97.0	01.72
Ensemble Classifiers	Boosted Trees	96.4	32.35
	Bagged Trees	97.8	10.03
	Subspace Discriminant	92.2	09.16
	Subspace KNN	97.7	10.53
	RUSBoosted Trees	95.3	13.61
Neural Network Classifiers	Narrow Neural Network	97.8	38.72
	Medium Neural Network	98.6	39.58
	Wide Neural Network	98.3	83.57
	Bilayered Neural Network	98.2	39.51
	Trilayered Neural Network	93.1	44.91

4.2 Discussion and Considerations

All results achieved an accuracy of at least 92.1%, except for the Coarse Tree that achieved only 49.8% of accuracy. In a general context, the models could classify all events, reaching one of the work's objectives. The use of confusion matrix and ROC Curve brought a positive point to highlight the results since they enable to describe the performance concisely.

The algorithms based on SVM achieved the best results overall. The best classifier was based on a SVM model. In this classifier, the worst class performance was in the $sag-harmonic$ class. This possible happened due to the interlaced parameters mentioned in subsection 2.3.8

The Neural Network Classifiers were the ones that obtained the longest training time among all models. This characteristic coincides with the nature of these classifiers, the slow training speed. The wide neural type was better than the trilayered type, showing

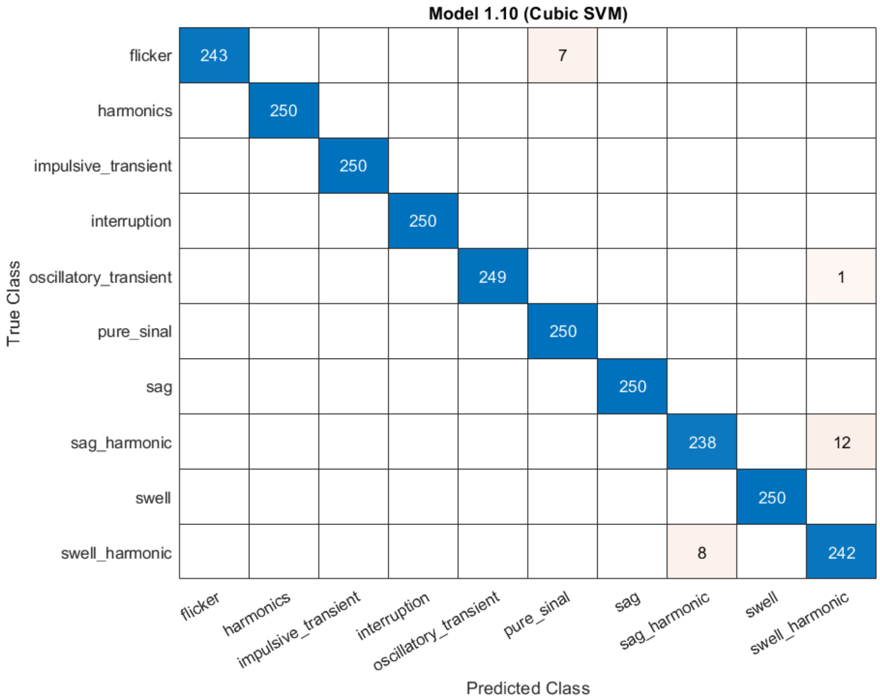

Figure 4.1 – Confusion Matrix of the best classifier (Cubic SVM)

that for this classification problem, increasing the layers of the network only increased the complexity of the system.

Nearest Neighbor models are known for having a low training time but also a lower accuracy. And this is precisely what we can see in Table 4.1. Although they were by far the fastest in training, they failed to perform well in the results, obtaining an average of 95.7% accuracy among all models.

Decision Trees are good generalists but prone to overfitting. And that was precisely what happened in the Coarse Tree model. Naive Bayes classifiers are widely used for text applications, but that is not the context of this work. Maybe, these models did not perform well because the input data of the system were dependent on each other: average energy and maximum energy value, for example. Lastly, the Ensemble classifiers varied the performance in each type, which is a characteristic of this type of model. However, it managed to achieve good accuracy results in general.

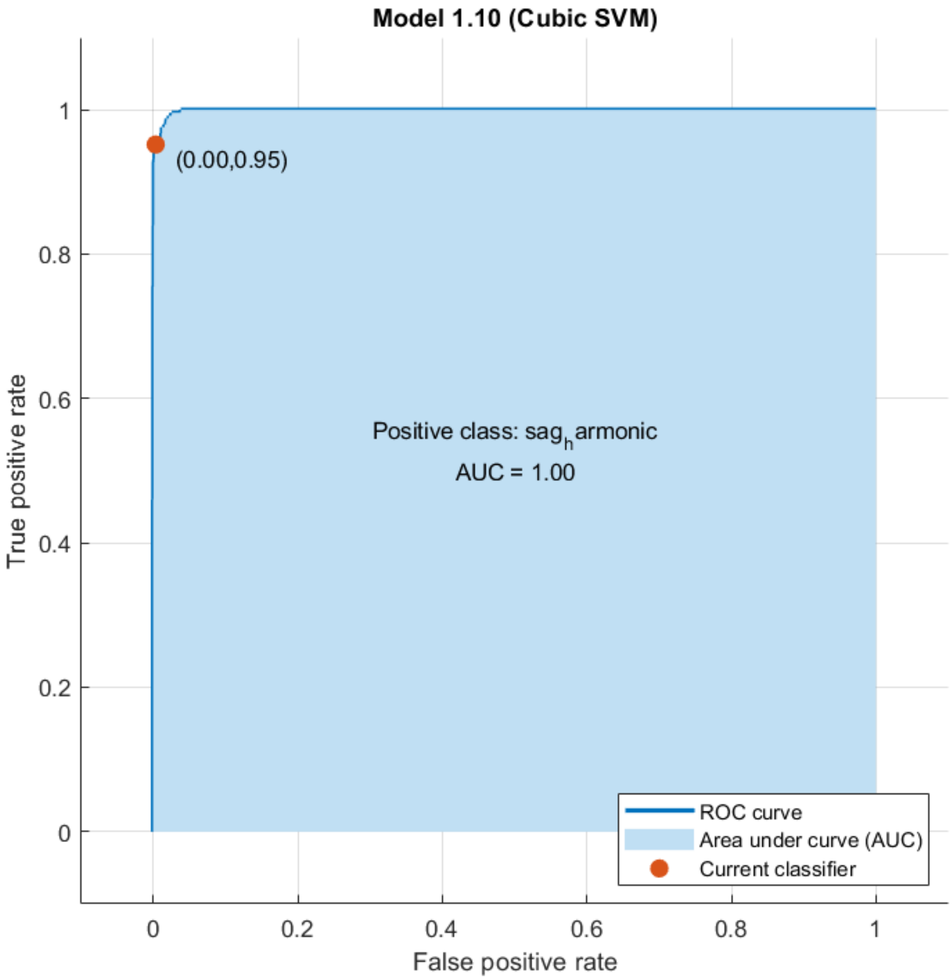

Figure 4.2 – ROC Curve of the best classifier (Cubic SVM)

5 Conclusions

In this chapter, a retrospective of what was developed and the conclusions about the research are presented. The future works possibilities are also highlighted at the end of the chapter.

5.1 Research Conclusions

This work presented the design of a robust classifier of advanced PQ disturbances in the context of SGs. In a first moment, the state-of-art of the application of AI in power systems were discussed, followed by the use of ML techniques, PQ and Signal Processing in this context. All background knowledge necessary for the understanding of this master's dissertation was also highlighted.

The objective of the work was to use DWT to extrude signal patterns from the input data so that later the ML algorithms could correctly classify the events. The stages of data generation, feature extraction and optimization techniques were performed in the MATLAB software. The classification learner toolbox was used for training, validation and testing the 27 different ML algorithms and assess each performance. So, the key topics of this work are:

- Signal generation through mathematical models of the different PQ events;
- Application of DWT method to extract signal parameters (entropy, energy and THD);
- Optimization techniques that improved processing, results and time;
- Training, validation and testing of the different ML models;
- Comparison between accuracy and confusion matrix across all models, proving the effectiveness of the system.

All stages of the work were previously idealized, enabling their correct development and execution. The results show that the Cubic SVM classifier achieved the maximum accuracy of all algorithms, indicating the effectiveness of the proposed method for classification. This method also achieved a reasonable training time. Although, if a higher priority were given to training time, this analysis would need to be different.

Some reasons can explain the justification for the excellent results: Firstly, the six inputs extracted from the signal processing stage were fundamental, proving its ability to handle and map different events. The coefficients' entropy extracted an essential level of information in each input, making possible the model to self-learn the PQ disturbances dynamics.

In addition, choosing the proper mother wavelet was extremely important for this to happen. The normalization of the inputs is also responsible for improving the results, turning them into data that could be better understood by the ML algorithms. Finally, the vast amount of input data allowed all models to be exhaustively trained, validated and tested, avoid overfitting.

The insertion of the SG context was vital to boost current research with what is discussed globally. The interlaced parameters are considered the advanced PQ disturbances and therefore need to be studied more and more. The use of ML algorithms in power systems is increasingly recurrent. Thus, the source code of this work will be available in Appendix B to guarantee a starting point for expanding and exploring this application.

5.2 Future Works Possibilities

This work can be continued, embracing new opportunities for research like the generation of more types of advanced PQ events to be analyzed by the classification system. Also, implementing the system in real-time hardware using data collected within a smart operational grid can be highly beneficial due to the models' minimal training and processing times.

The use of massive datasets extracted from real signals could validate the ML models even better. It should also consider using the DL techniques as features extractors in conjunction with the ML models for classification. CNN are becoming a trend in the area due to their increased accuracy. Future works can implement a signal-noise ratio of 60db so that the model can be even more faithful to what happens in real systems.

Other know ML models could also be used, like linear regression, logistic regression, learning vector quantization and clustering methods. Finally, using other signal processing techniques for feature extraction can be considered, such as CWT, WPT, and Fast-Fourier Transform for example.

Bibliography

ALIMI, O. A.; OUAHADA, K.; ABU-MAHFOUZ, A. M. A Review of Machine Learning Approaches to Power System Security and Stability. *IEEE Access*, v. 8, p. 113512–113531, 2020. ISSN 2169-3536.

AUNG, M. T.; MILANOVIC, J. V.; SIMMONS, P. A. Automated comprehensive assessment and visualization of voltage sag performance. In: *2004 11th International Conference on Harmonics and Quality of Power (IEEE Cat. No.04EX951)*. [S.l.: s.n.], 2004. p. 191–198.

BARRIOS, P. P. V. *Control and energy management system of a microgrid using a genetic algorithm*. 152 p. Tese (Doutorado) — Universidade Estadual de Campinas, 2015.

BELLMAN, R. E. *An Introduction to Artificial Intelligence: Can Computers Think?* San Francisco: Boyd & Fraser Publishing Company, 1978.

BISHOP, C. M. *Pattern Recognition and Machine Learning*. New York: Springer, 2006. ISBN 978-0387-31073-2.

BOLLEN, M. H.; GU, I. Origin of Power Quality Variations. In: _____. *Signal Processing of Power Quality Disturbances*. IEEE, 2006. p. 41–161. ISBN 9780471931300. Disponível em: <https://ieeexplore.ieee.org/document/5224884>.

BOLLEN, M. H. J.; BAHRAMIRAD, S.; KHODAEI, A. Is there a place for power quality in the smart grid? In: *2014 16th International Conference on Harmonics and Quality of Power (ICHQP)*. [S.l.: s.n.], 2014. p. 713–717. ISSN 2164-0610.

BRONZINI, M. et al. Power system modal identification via wavelet analysis. In: *2007 IEEE Lausanne Power Tech*. [S.l.: s.n.], 2007. p. 2041–2046.

CAI, L. et al. Real-Time Detection of Power System Disturbances Based on k -Nearest Neighbor Analysis. *IEEE Access*, v. 5, p. 5631–5639, 2017. ISSN 2169-3536.

CERQUEIRA, A. S. et al. Digital System for Detection and Classification of Power Quality Disturbances. *IEEE Latin America Transactions*, v. 4, n. 5, p. 345–352, 2006. ISSN 1548-0992 VO - 4.

CHAPALOGLOU, S. et al. Smart energy management algorithm for load smoothing and peak shaving based on load forecasting of an island's power system. *Applied Energy*, v. 238, p. 627–642, 2019. ISSN 0306-2619. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0306261919301035>.

COIFMAN, R. R.; WICKERHAUSER, M. V. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, v. 38, n. 2, p. 713–718, 1992. ISSN 1557-9654.

DHARMADHIKARI, S. C. et al. A smart grid incorporated with ML and IoT for a secure management system. *Microprocessors and Microsystems*, v. 83, p. 103954, 2021. ISSN 0141-9331. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0141933121001332>.

DUGAN, R. C. et al. *Electrical Power Systems Quality*. 2nd. ed. [S.l.]: McGraw-HillCompanies, 2012. 525 p. ISBN 978-0071761550.

ERIşTI, H. et al. Optimal feature selection for classification of the power quality events using wavelet transform and least squares support vector machines. *International Journal of Electrical Power & Energy Systems*, v. 49, p. 95–103, 2013. ISSN 0142-0615. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0142061513000276>.

FADAEENEJAD, M. et al. The present and future of smart power grid in developing countries. *Renewable and Sustainable Energy Reviews*, v. 29, p. 828–834, 2014. ISSN 1364-0321. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1364032113006126>.

FARHOUMANDI, M.; ZHOU, Q.; SHAHIDEHPOUR, M. A review of machine learning applications in IoT-integrated modern power systems. *The Electricity Journal*, v. 34, n. 1, p. 106879, 2021. ISSN 1040-6190. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1040619020301718>.

FRUNT, J.; KLING, W. L.; RIBEIRO, P. F. Wavelet Decomposition for Power Balancing Analysis. *IEEE Transactions on Power Delivery*, v. 26, n. 3, p. 1608–1614, 2011. ISSN 1937-4208 VO - 26.

GALLI, A. W.; HEYDT, G. T.; RIBEIRO, P. F. Exploring the power of wavelet analysis. *IEEE Computer Applications in Power*, v. 9, n. 4, p. 37–41, 1996. ISSN 1558-4151 VO - 9.

GHOSH, S.; DASGUPTA, A.; SWETAPADMA, A. A Study on Support Vector Machine based Linear and Non-Linear Pattern Classification. In: *2019 International Conference on Intelligent Sustainable Systems (ICISS)*. [S.l.: s.n.], 2019. p. 24–28.

GOMES, R. O. et al. Experimental identification of a non-Linear load behaviour using time-varying harmonics. In: *XII CBQEE*. [S.l.: s.n.], 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. 800 p.

HAYKIN, S. *Neural Networks: Principles and Practice*. 2. ed. Porto Alegre: Bookman, 2001. ISBN 978-0-13-147139-9.

HUANG, G. et al. Cyber-Constrained Optimal Power Flow Model for Smart Grid Resilience Enhancement. *IEEE Transactions on Smart Grid*, v. 10, n. 5, p. 5547–5555, 2019. ISSN 1949-3061.

IBRAHIM, W. R. A.; MORCOS, M. M. Artificial intelligence and advanced mathematical tools for power quality applications: a survey. *IEEE Transactions on Power Delivery*, v. 17, n. 2, p. 668–673, 2002. ISSN 1937-4208 VO - 17.

IEC 61000-2-2:2002. *Electromagnetic compatibility (EMC) - Part 2-2: Environment - Compatibility levels for low-frequency conducted disturbances and signalling in public low-voltage power supply systems*. 2002.

IEEE. IEEE Recommended Practice for Monitoring Electric Power Quality. *IEEE Std 1159-1995*, p. 1–80, nov 1994.

IEEE. IEEE Recommended Practice for Monitoring Electric Power Quality. *IEEE Std 1159-2019 (Revision of IEEE Std 1159-2009)*, p. 1–98, 2019.

IGUAL, R. et al. Integral mathematical model of power quality disturbances. In: *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*. [S.l.: s.n.], 2018. p. 1–6.

International Energy Agency. *Smart Grid Technology Road Map*. [S.l.], 2011. Disponível em: <https://www.iea.org/reports/technology-roadmap-smart-grids>.

JWG C4.24/CIRED. *Power Quality and EMC Issues with Future Electricity Networks*. [S.l.]: CIGRE, 2018. ISBN 978-2-85873-421-4.

JWG C4/B4.38. *Network Modelling for Harmonic Studies*. [S.l.]: CIGRE, 2019.

JWG C4/C6.29. *Power Quality Aspects of Solar Power*. [S.l.]: CIGRE, 2016.

KHOKHAR, S. et al. MATLAB/Simulink based modeling and simulation of power quality disturbances. In: *2014 IEEE Conference on Energy Conversion (CENCON)*. IEEE, 2014. p. 445–450. ISBN 978-1-4799-4848-2. Disponível em: <https://ieeexplore.ieee.org/document/6967545/>.

KHOKHAR, S. et al. A comprehensive overview on signal processing and artificial intelligence techniques applications in classification of power quality disturbances. *Renewable and Sustainable Energy Reviews*, v. 51, p. 1650–1663, 2015. ISSN 1364-0321. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1364032115007157>.

MARUF, M. H. et al. Adaptation for sustainable implementation of Smart Grid in developing countries like Bangladesh. *Energy Reports*, v. 6, p. 2520–2530, 2020. ISSN 2352-4847. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2352484720313214>.

MathWorks Help Center. *Assess Classifier Performance in Classification Learner*. 2021. Disponível em: <https://www.mathworks.com/help/stats/assess-classifier-performance.html>.

MathWorks Help Center. *Choose Classifier Options*. 2021. Disponível em: <https://www.mathworks.com/help/stats/choose-a-classifier.html>.

MathWorks Help Center. *Classification Learner App*. 2021. Disponível em: <https://www.mathworks.com/help/stats/classification-learner-app.html?s{_}tid=CRUX{_}>.

MathWorks Help Center. *Energy for 1-D Wavelet*. 2021. Disponível em: <https://www.mathworks.com/help/wavelet/ref/wenergy.html>.

MathWorks Help Center. *Select Data and Validation for Classification Problem*. 2021. Disponível em: <https://www.mathworks.com/help/stats/select-data-and-validation-for-classification-problem.html>.

MathWorks Help Center. *Total harmonic distortion*. 2021. Disponível em: <https://www.mathworks.com/help/signal/ref/thd.html>.

MIRIAM, E. J.; SEKAR, S.; AMBALAVANAN, S. Artificial Neural Network technique for predicting the lifetime and performance of lead-acid battery. *IRACST – Engineering Science and Technology: An International Journal*, v. 3, n. 2, p. 393–401, 2013.

MITCHELL, T. M. *Machine Learning*. [S.l.]: McGraw-HillCompanies, 1997. 432 p. ISBN 0070428077.

MOLOI, K.; YUSUFF, A. A. A Wavelet-Neural Network-Based Technique for Fault Diagnostics in Power System. In: *2020 7th International Conference on Soft Computing Machine Intelligence (ISCMI)*. [S.l.: s.n.], 2020. p. 131–135. ISSN 2640-0146.

MONTEIRO, L. F. R. et al. Determination of Renewable Generation Operation with the Aid of the ANN. In: *2018 13th IEEE International Conference on Industry Applications (INDUSCON)*. [S.l.: s.n.], 2018. p. 375–380. ISBN VO -.

OLIVEIRA, D. Q. et al. A fuzzy-based approach for microgrids islanded operation. *Electric Power Systems Research*, v. 149, p. 178–189, 2017. ISSN 0378-7796. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378779617301669>.

PACKT. *Decision tree-based ensemble methods*. 2021. Disponível em: <https://subscription.packtpub.com/book/big{_}data{_}and{_}business{_}intelligence/9781789132212/2/ch02lvl1sec21/decision-tree-based-ense>.

PEREIRA, F. et al. An analysis of costs related to the loss of power quality. In: *8th International Conference on Harmonics and Quality of Power. Proceedings (Cat. No.98EX227)*. Athens, Greece: IEEE, 1998. v. 2, p. 777–782. ISBN 0-7803-5105-3. Disponível em: <http://ieeexplore.ieee.org/document/760141/>.

RIBEIRO, P. F. Wavelet transform: An advanced tool for analyzing non-stationary harmonic distortion in power systems. In: *IEEE ICHPSVI*. Bologna, Italy: [s.n.], 1994. p. 365–369.

RIBEIRO, P. F. Visualization of Time-Varying Waveform Distortions with Wavelets. In: *Time-Varying Waveform Distortions in Power Systems*. IEEE, 2010. p. 175–186. ISBN 9780470746745. Disponível em: <http://ieeexplore.ieee.org/document/6168985>.

RIBEIRO, P. F. et al. (Ed.). *Power Systems Signal Processing For Smart Grids*. Chichester, United Kingdom: John Wiley and Sons Ltd, 2013. ISBN 9781118639283. Disponível em: <http://doi.wiley.com/10.1002/9781118639283>.

RODRIGUES, C. E. M.; TOSTES, M. E. d. L. Characterization of supraharmonics using the wavelet packet transform. In: *2018 18th International Conference on Harmonics and Quality of Power (ICHQP)*. [S.l.: s.n.], 2018. p. 1–6. ISBN 2164-0610 VO -.

RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence: a modern approach*. New Jersey: Prentice Hall, 1995. 946 p. ISBN N 0-13-103805-2.

S. Salles, R. et al. Visualization of Quality Performance Parameters Using Wavelet Scalograms Images for Power Systems. In: *Anais do Congresso Brasileiro de Automática 2020*. sbabra, 2020. Disponível em: <https://www.sba.org.br/open{_}journal{_}systems/index.php/sba/article/vie>.

Saidani Neffati, O. et al. Migrating from traditional grid to smart grid in smart cities promoted in developing country. *Sustainable Energy Technologies and Assessments*, v. 45, p. 101125, 2021. ISSN 2213-1388. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2213138821001351>.

SALLES, R. d. S. *The use of advanced signal processing and deep learning for pattern recognition in integrated metrics of quality performance: a smart grid application.* 85 p. Tese (Doutorado) — Federal University of Itajubá, 2020. Disponível em: <https://repositorio.unifei.edu.br/jspui/handle/123456789/2327>.

SALLES, R. S. et al. Fuzzy Logic-Based Controller for BESS and Load Management in a Microgrid Economic Operation. In: *2020 IEEE PES Transmission Distribution Conference and Exhibition - Latin America (T D LA)*. [S.l.: s.n.], 2020. p. 1–6. ISSN 2472-9639.

SAMUEL, A. L. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, v. 3, n. 3, p. 210–229, 1959.

SHARMA, N.; SHARMA, R.; JINDAL, N. Machine Learning and Deep Learning Applications-A Vision. *Global Transitions Proceedings*, v. 2, n. 1, p. 24–28, 2021. ISSN 2666-285X. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666285X21000042>.

SHIN, T. *All Machine Learning Models Explained in 6 Minutes.* 2020. Disponível em: <https://towardsdatascience.com/all-machine-learning-models-explained-in-6-minutes-9fe30ff6776a>.

SILVEIRA, P. M.; M., S.; RIBEIRO, P. F. Using Wavelet decomposition for Visualization and Understanding of Time-Varying Waveform Distortion in Power System. In: *VII CBQEE*. [S.l.: s.n.], 2007.

SINGH, S. *Cousins of Artificial Intelligence*. 2018. Disponível em: <https://towardsdatascience.com/cousins-of-artificial-intelligence-dda4edc27b55>.

SOUZA, M. F. Z. de. *Modelagem e simulação integrada ao processamento e diagnóstico do desempenho de parâmetros elétricos no contexto de redes inteligentes*. 147 p. Tese (Doutorado) — Federal University of Itajubá, 2017. Disponível em: <https://repositorio.unifei.edu.br/jspui/handle/123456789/943>.

TANG, Y. et al. Framework for artificial intelligence analysis in large-scale power grids based on digital simulation. *CSEE Journal of Power and Energy Systems*, v. 4, n. 4, p. 459–468, 2018. ISSN 2096-0042.

US Congress. *Energy Independence and Security Act of 2007*. 2007. 12 p. Disponível em: <https://www.smartgrid.gov/document/title{_}xiii{_}energy{_}independence{_}and{_}securit>.

WILSON, A. J. et al. Automated Identification of Electrical Disturbance Waveforms within an Operational Smart Power Grid. *IEEE Transactions on Smart Grid*, p. 1, 2020. ISSN 1949-3061 VO -.

WU, J.; ZHANG, C.; CHEN, Z. An online method for lithium-ion battery remaining useful life estimation using importance sampling and neural networks. *Applied Energy*, v. 173, p. 134–140, 2016. ISSN 0306-2619. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0306261916304846>.

XU, J. et al. Bayesian adversarial multi-node bandit for optimal smart grid protection against cyber attacks. *Automatica*, v. 128, p. 109551, 2021. ISSN 0005-1098. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0005109821000716>.

XU, J. et al. Some Techniques for the Analysis and Visualization of Time-varying Waveform Distortions. In: *2006 38th North American Power Symposium*. [S.l.: s.n.], 2006. p. 257–261. ISBN VO -.

YU, X.; WANG, K. Digital System for Detection and Classification of Power Quality Disturbance. In: *2009 Asia-Pacific Power and Energy Engineering Conference*. [S.l.: s.n.], 2009. p. 1–4. ISBN 2157-4847 VO -.

ZHENGYOU, H. et al. Study of a new method for power system transients classification based on wavelet entropy and neural network. *International Journal of Electrical Power & Energy Systems*, v. 33, n. 3, p. 402–410, 2011. ISSN 0142-0615. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0142061510001882>.

ZHOU, Z.-H. *Ensemble Methods: foundations and algorithms*. [S.l.]: CRC Press, 2012.

ŞERBAN, A. C.; LYTRAS, M. D. Artificial Intelligence for Smart Renewable Energy Sector in Europe—Smart Energy Infrastructures for Next Generation Smart Cities. *IEEE Access*, v. 8, p. 77364–77377, 2020. ISSN 2169-3536.

Appendix

APPENDIX A – Published Papers

This appendix lists the work associated with the author's master's degree program financed by the Coordination for the Improvement of Higher Education Personnel - Brazil (CAPES) - Financial Code 001.

- Almeida, G.C.S.; Souza, A.C.Z.; Ribeiro, P.F. A Neural Network Application for a Lithium-Ion Battery Pack State-of-Charge Estimator with Enhanced Accuracy. Proceedings 2020, 58, 33. https://doi.org/10.3390/WEF-06915
- Almeida, G.C.S.; Souza, A.C.Z. ; Ribeiro, P.F. A Neural Network Application for a Lithium-ion Battery Pack State-of-Charge Estimator with Enhanced Accuracy. In: The First World Energies Forum, 2020, Roma. The First World Energies Forum.
- Salles, R.S.; Almeida, G.C.S.; Silva, L.R.M.; Duque, C. A; Ribeiro, P.F. Visualization of Quality Performance Parameters Using Wavelet Scalograms Images, In: XXIII Congresso Brasileiro de Automática, Nov. 2020.
- Salles, R.S.; Almeida, G.C.S.; Fuly, B. I. L.; Souza, A.C.Z.; Ribeiro, P. F. Fuzzy Logic-Based Controller for BESS and Load Management in a Microgrid Economic Operation, 2020 IEEE PES Transmission and Distribution Conference and Exhibition - Latin America (TD LA), 2020, pp. 1-6, doi: 10.1109/TDLA47668.2020.9326102.
- Almeida, Gabriel C. S.; Salles, R. S. ; Silva, M. N. S. ; Zambroni, A. C. S. ; F. Ribeiro, Paulo . The Need of Normative Technologies for Smart Living Cities. In: Antonio Carlos Zambroni de Souza; Maarten J. Verkerk; Paulo Fernando Ribeiro. (Org.). Interdisciplinary and Social Nature of Engineering Practices. 1ed.Switzerland: Springer International Publishing, 2021, v. 61, p. 1-438.

APPENDIX B – MATLAB Codes

All MATLAB codes used in this work can be found at:

- <https://bit.ly/3h4fUZg>

APPENDIX C – Confusion Matrix of all ML algorithms

In this appendix are the confusion matrix of the remaining ML algorithms for the tests set, ordered by the highest accuracy.

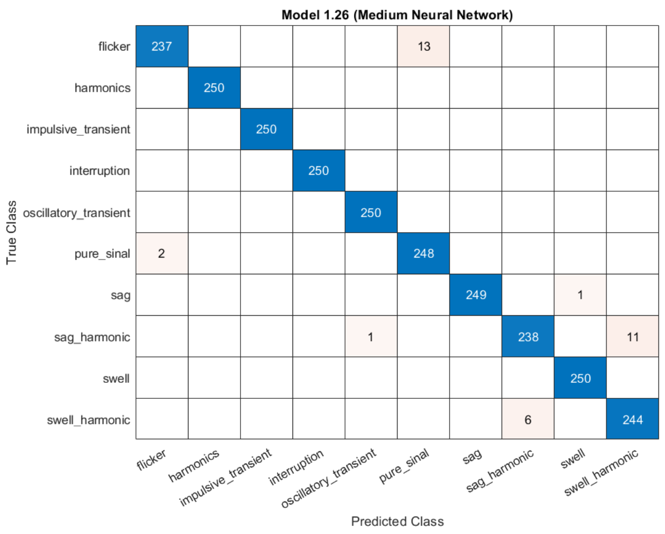

Figure C.1 – Confusion Matrix of Medium Neural Network

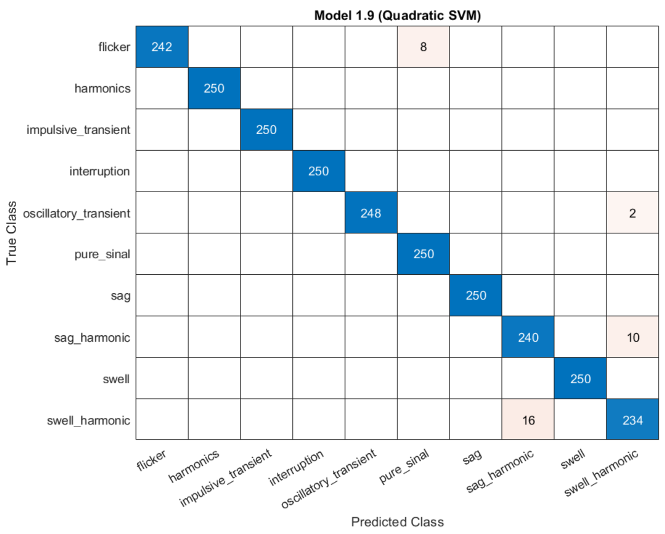

Figure C.2 – Confusion Matrix of Quadratic SVM

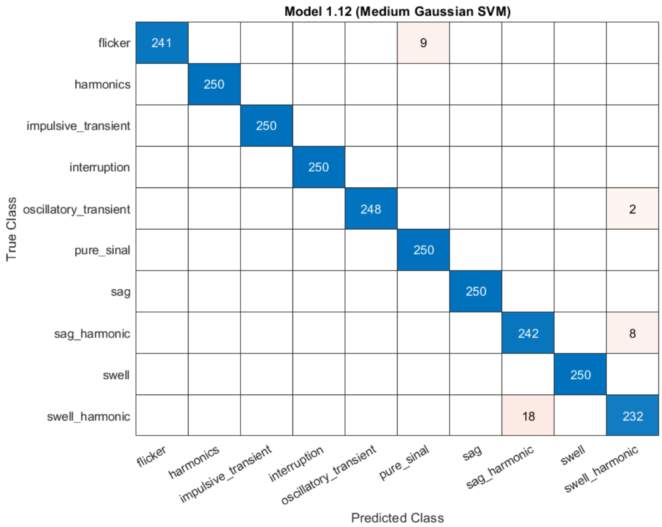

Figure C.3 – Confusion Matrix of Medium Gaussian SVM

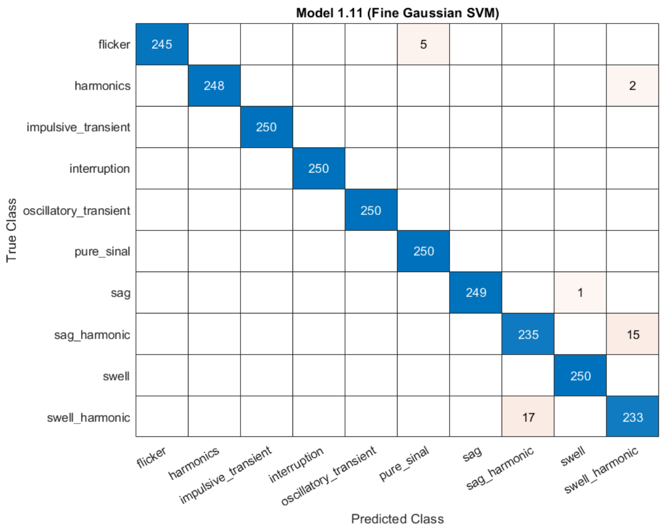

Figure C.4 – Confusion Matrix of Fine Gaussian SVM

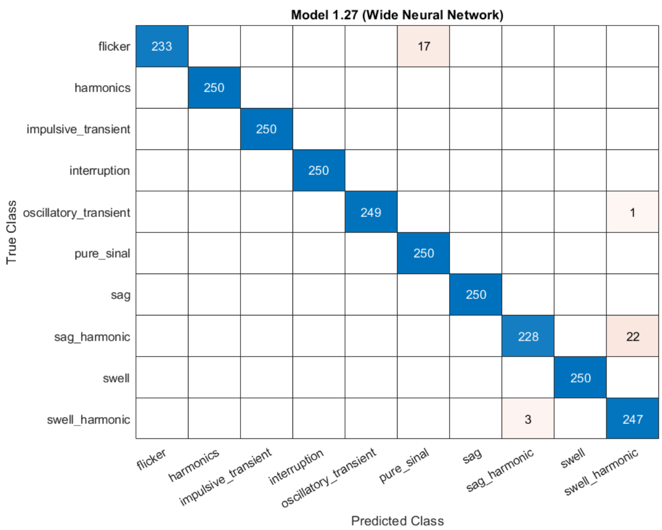

Figure C.5 – Confusion Matrix of Wide Neural Network

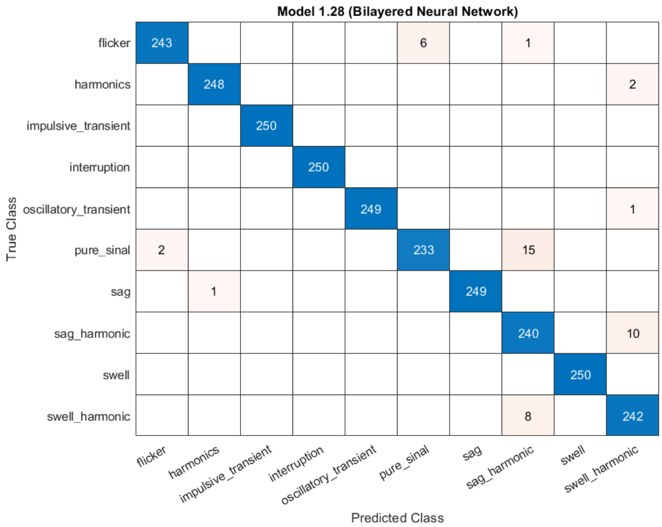

Figure C.6 – Confusion Matrix of Bilayered Neural Network

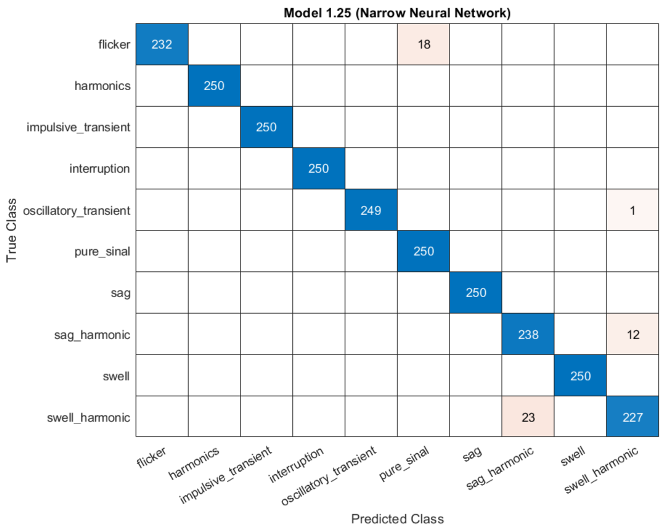

Figure C.7 – Confusion Matrix of Narrow Neural Network

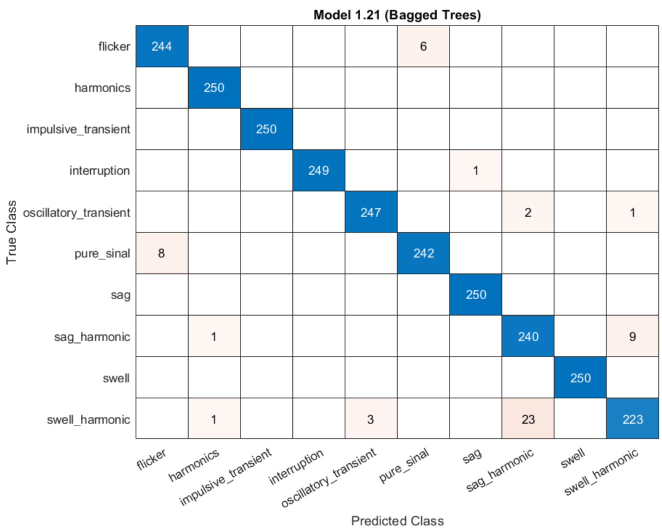

Figure C.8 – Confusion Matrix of Bagged Trees

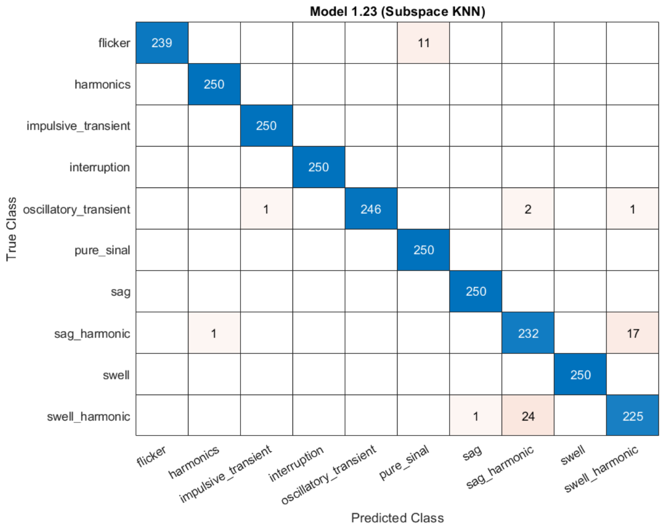

Figure C.9 – Confusion Matrix of Subspace KNN

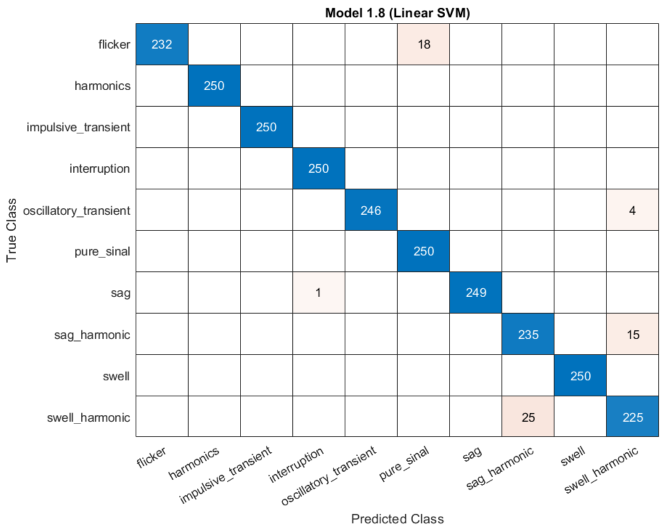

Figure C.10 – Confusion Matrix of Linear SVM

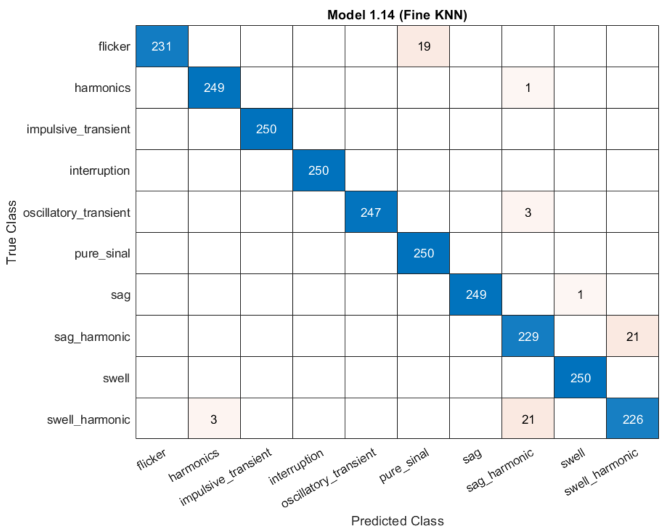

Figure C.11 – Confusion Matrix of Fine KNN

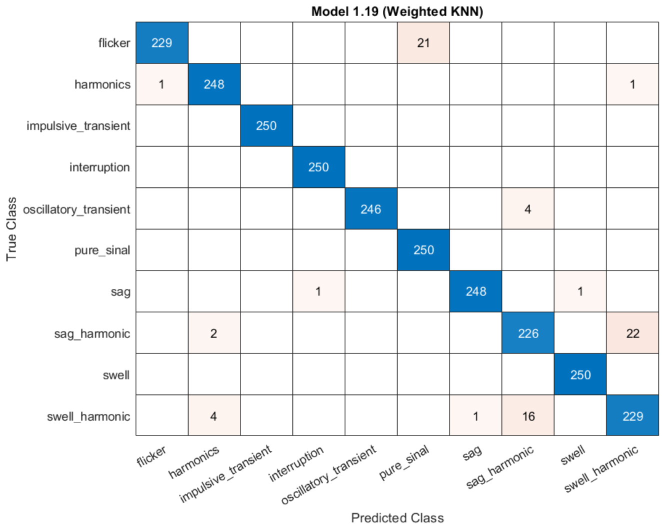

Figure C.12 – Confusion Matrix of Weighted KNN

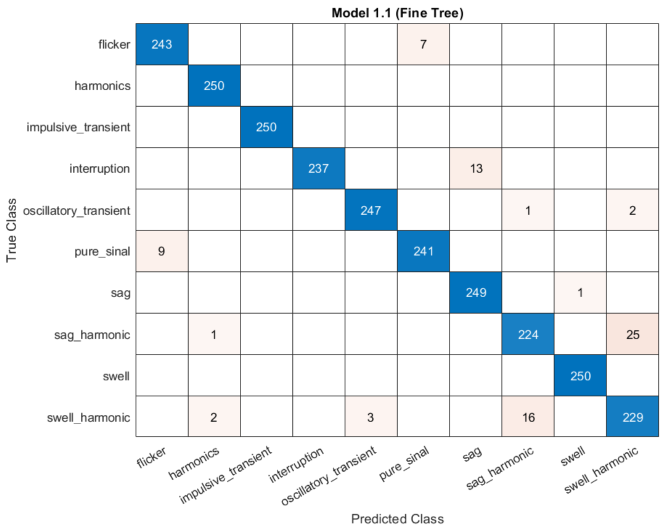

Figure C.13 – Confusion Matrix of Fine Tree

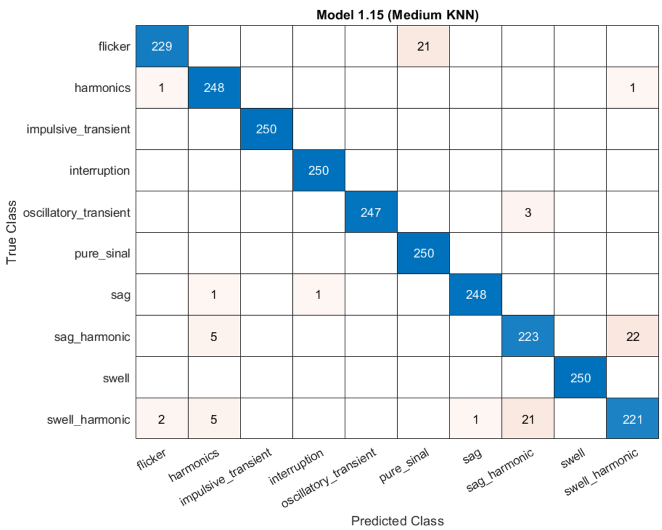

Figure C.14 – Confusion Matrix of Medium KNN

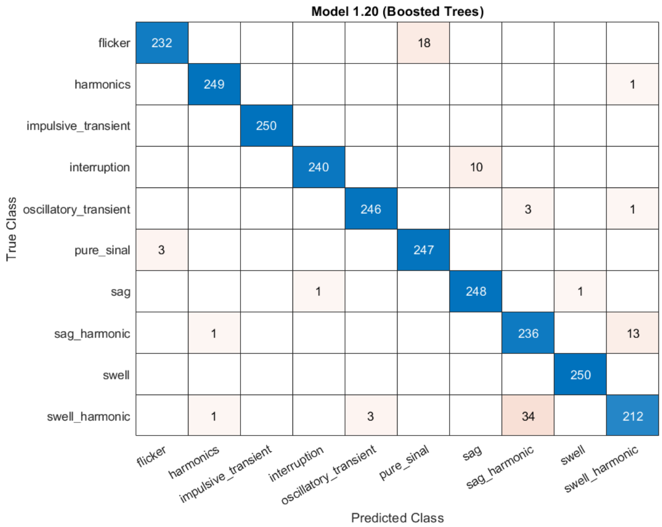

Figure C.15 – Confusion Matrix of Boosted Trees

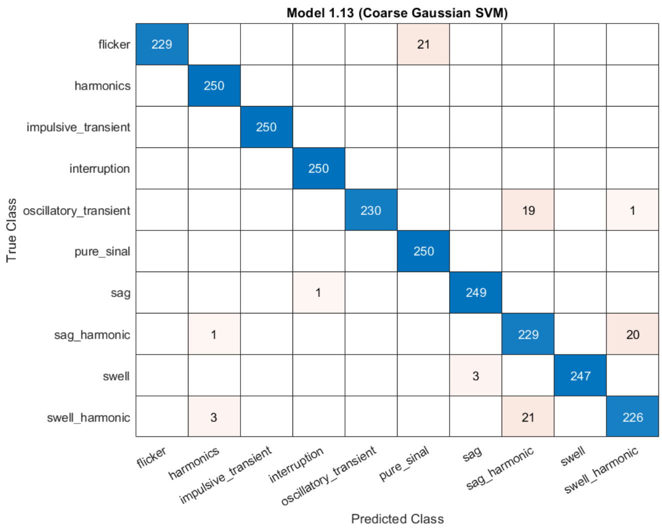

Figure C.16 – Confusion Matrix of Coarse Gaussian SVM

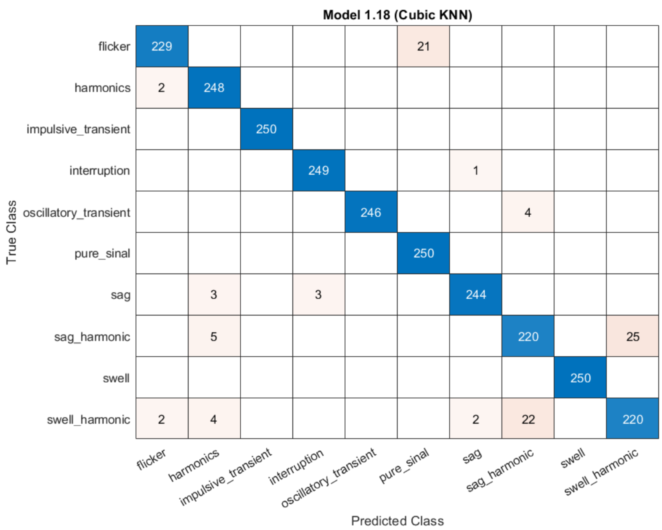

Figure C.17 – Confusion Matrix of Cubic KNN

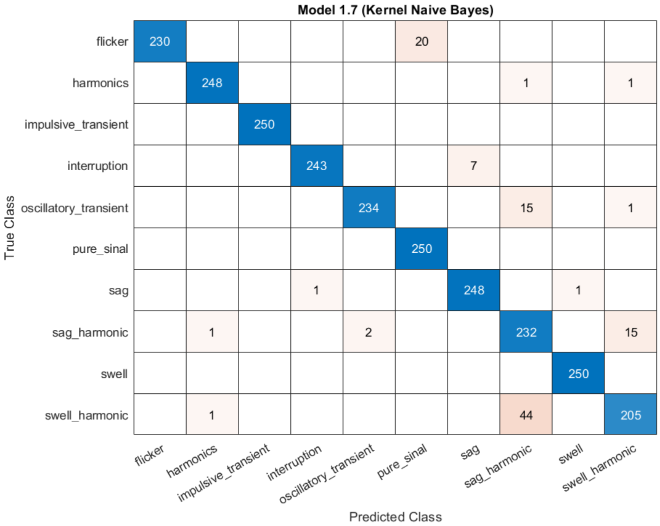

Figure C.18 – Confusion Matrix of Kernel Naive Bayes

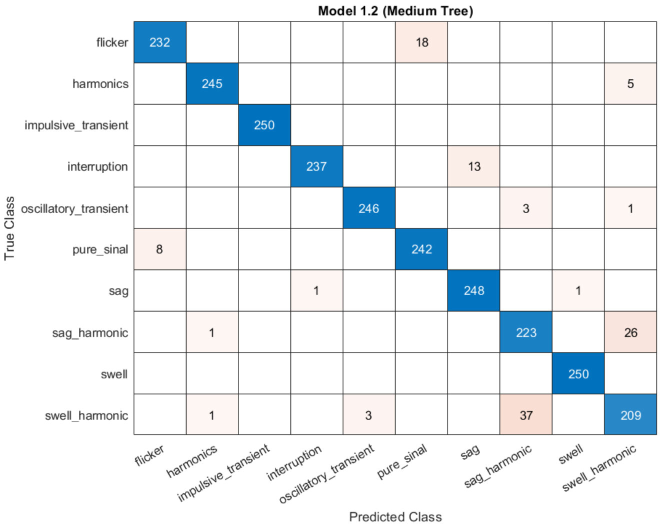

Figure C.19 – Confusion Matrix of Medium Tree

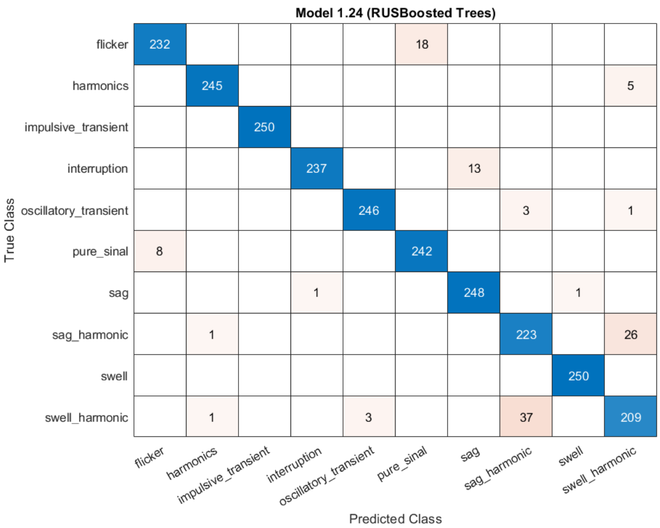

Figure C.20 – Confusion Matrix of RUSBoosted Trees

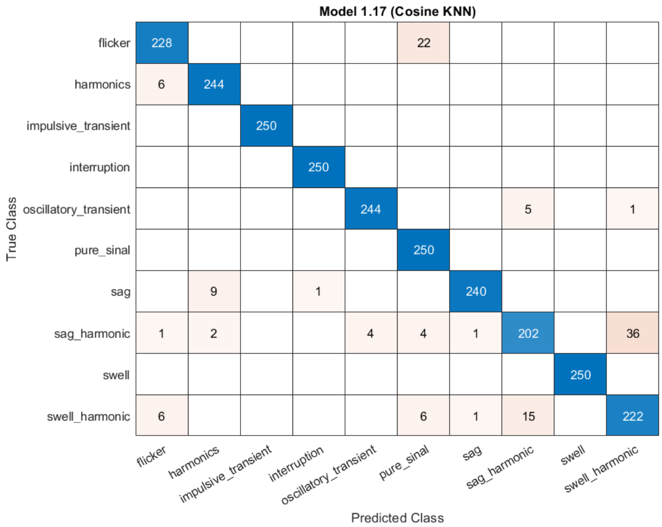

Figure C.21 – Confusion Matrix of Cosine KNN

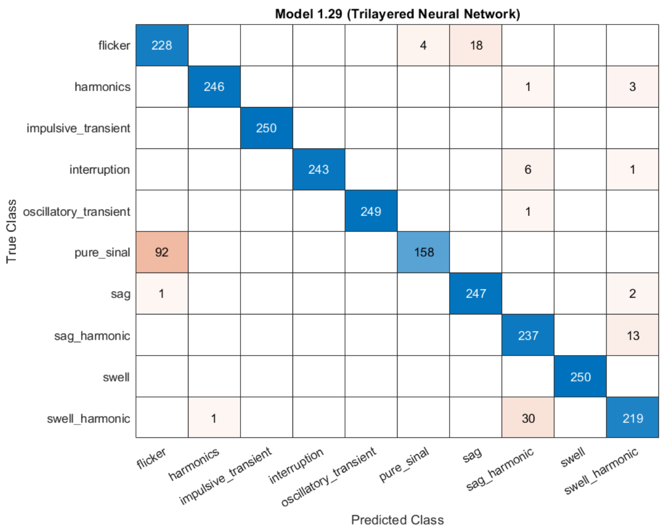

Figure C.22 – Confusion Matrix of Trilayered Neural Network

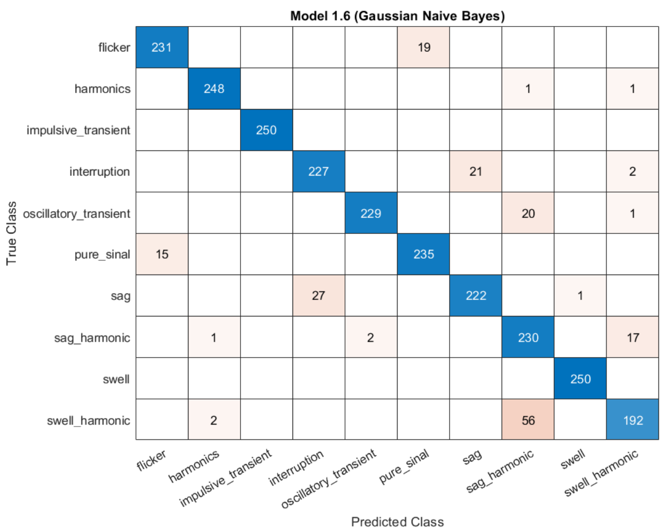

Figure C.23 – Confusion Matrix of Gaussian Naive Bayes

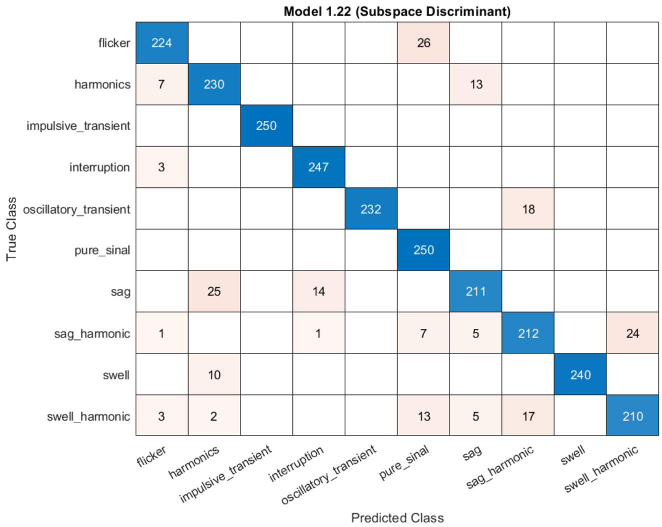

Figure C.24 – Confusion Matrix of Subspace Discriminant

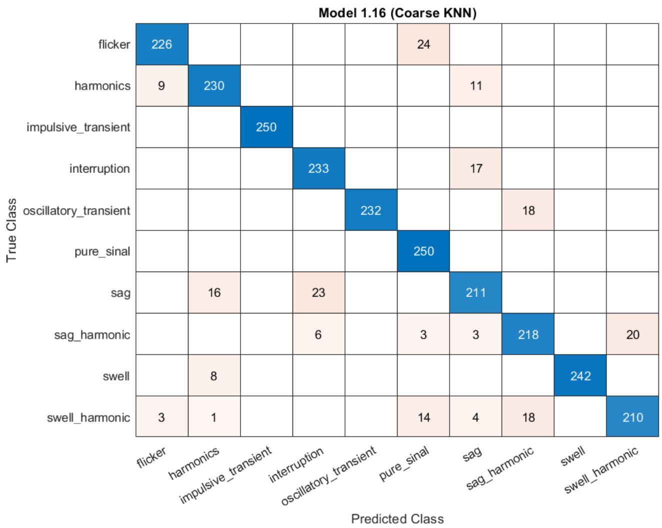

Figure C.25 – Confusion Matrix of Coarse KNN

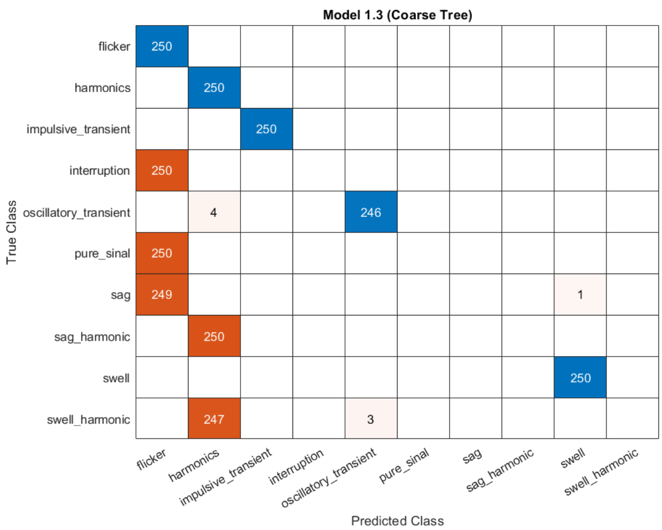

Figure C.26 – Confusion Matrix of Coarse Tree