Zaid Zaid

VOLTAGE DISTURBANCE DETECTION AND CLASSIFICATION IN DISTRIBUTION NETWORKS USING SIMULATED EVENTS

A data-driven framework using realistic load profiles and machine learning

Master's Thesis
Faculty of Information Technology and Communication Sciences (ITC)
Professor Sami Repo
Dr. Mehdi Attar
November 2025

ABSTRACT

Zaid Zaid: Voltage Disturbance Detection and Classification in Distribution Networks using Simulated Events
Master's Thesis
Tampere University
Master's Degree Program in Electrical Engineering, Smart Grids
November 2025

The increasing penetration of distributed energy resources (DERs) such as photovoltaic (PV) systems, electric vehicles (EVs), and heat pumps has made modern low-voltage (LV) distribution networks more dynamic and prone to voltage disturbances. These events, arising from rapid fluctuations in load or generation, can cause voltage deviations that complicate voltage control coordination. This thesis presents a data-driven, simulation-based framework for detecting and classifying voltage disturbances in distribution networks, with the aim of supporting coordinated voltage control (CVC) strategies.

A detailed interconnected HV–MV–LV grid model was developed in MATLAB to simulate disturbance scenarios under realistic time-series load profiles. Synthetic voltage and current data were generated by injecting diverse disturbance types—such as load surges and upstream voltage sags—under varying network conditions. Key monitoring indicators, including voltage dips, current magnitudes, and statistical variations of node voltages, were extracted as features for classification. Multiple machine learning algorithms, including decision trees, support vector machines (SVM), k-nearest neighbors (k-NN), and ensemble methods, were trained and evaluated on their ability to distinguish both the type and origin of disturbances (HV, MV, or LV).

Results demonstrate that machine learning models can reliably identify the disturbance type and origin with high accuracy when trained on sufficiently resolved monitoring data. The analysis further highlights how data resolution, feature selection, and disturbance characteristics influence detection reliability. The proposed framework establishes a foundation for integrating intelligent disturbance detection and classification into real-time coordinated voltage control (CVC) systems, thereby improving grid resilience and operational efficiency.

Keywords: Voltage disturbances, Distribution networks, Machine learning, Disturbance detection, Disturbance classification, Simulation-based framework

USE OF AI IN THESIS

I have utilised AI tools in my thesis:

☐ No
☑ Yes

The AI tools utilized in my thesis and their purposes are described below:

Names and versions of AI tools:
- ChatGPT (version: GPT-4.0)

The use of AI tools in this thesis served multiple purposes, primarily aimed at enhancing the clarity, professionalism, and coherence of the academic writing. These tools were strategically employed throughout the research process to ensure the quality and readability of the thesis.

Sections where AI tools were used: Almost all the section

I acknowledge that I am fully responsible for the entire content of my thesis, including the parts generated by AI, and accept accountability for any violations of ethical standards in publications.

Preface

This master's thesis was carried out as part of the Master of Science program in Electrical Engineering specialization in Smart Grids at Tampere University. It was commissioned by the Unit of Electrical Engineering at Tampere University, under the supervision of Professor Sami Repo.

I would like to express my sincere gratitude to my supervisor, Professor Sami Repo, for his exceptional guidance, insightful suggestions, and continuous encouragement throughout the course of this thesis. I am especially thankful for his patience, availability for discussions, and valuable feedback that consistently improved the quality of this work. It has been a privilege to learn under his supervision, and the experience has greatly shaped my research approach and professional perspective. I also thank Dr. Mehdi Attar for his valuable suggestions and collaboration during the thesis work.

Lastly, I extend heartfelt thanks to my family and friends for their constant encouragement and understanding during this endeavor. My deepest gratitude goes to my family for the continuous encouragement that made it possible for me to pursue my studies.

I hope this thesis will contribute to ongoing research and development in the area of intelligent power system operation and serve as a useful reference for future studies in voltage disturbance detection and classification.

Tampere, 10 November 2025

Zaid Zaid

Table of Contents

List of Figures

List of Tables

LIST OF SYMBOLS AND ABBREVIATIONS

ADMS	Advanced Distribution Management System
AMI	Advanced Metering Infrastructure
ANN	Artificial Neural Network
B	Line Charging Susceptance
Bb	Shunt Susceptance at the Bus
CNN	Convolutional Neural Network
CVC	Coordinated Voltage Control
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
DER	Distributed Energy Resource
DSO	Distribution System Operator
DT	Decision Tree
DWT	Discrete Wavelet Transform
EMD	Empirical Mode Decomposition
EN 50160	European Standard for Voltage Characteristics of Electricity Supplied by Public Distribution Systems
EV	Electric Vehicle
Gb	Shunt Conductance at the Bus
GW	Gigawatt
HHT	Hilbert–Huang Transform
HV	High Voltage
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IMF	Intrinsic Mode Function
k-NN	k-Nearest Neighbors
kW	Kilowatt
kWh	Kilowatt-hour
LDA	Linear Discriminant Analysis
LSTM	Long Short-Term Memory
LV	Low Voltage
MATLAB	Matrix Laboratory
ML	Machine Learning
MV	Medium Voltage
MVAr	Megavolt-ampere Reactive
MW	Megawatt
OLTC	On-Load Tap Changer
P	Active Power
PCA	Principal Component Analysis

PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
Pd	Active Power Demand
Pg	Active Power Generation
PF	Power Factor
PMU	Phasor Measurement Unit
PQ	Power Quality / Load Bus Type (PQ Bus)
Ps	Phase Shift Angle
PV	Photovoltaic / PV Bus Type
Q	Reactive Power
Qd	Reactive Power Demand
Qg	Reactive Power Generation
Qg_min	Minimum Reactive Power Limit
Qg_max	Maximum Reactive Power Limit
RMS	Root Mean Square
R	Line Resistance
RTU	Remote Terminal Unit
SCADA	Supervisory Control and Data Acquisition
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine
THD	Total Harmonic Distortion
Θ	Voltage Angle
V	Voltage Magnitude
VQMS	Voltage Quality Monitoring System
WT	Wavelet Transform
X	Line Reactance
μPMU	Micro Phasor Measurement Unit

1. INTRODUCTION

1.1. Motivation

The rapid growth of distributed energy resources (DERs)—such as rooftop photovoltaic (PV) systems, electric vehicles (EVs), and heat pumps—has made low-voltage (LV) distribution grids more dynamic and less predictable, increasing the complexity of accurate forecasting and control. While these technologies are essential for a sustainable energy future, they also increase the likelihood of voltage disturbances due to sudden changes in local load or generation. These voltage anomalies can compromise power quality, affect equipment longevity, and lead to suboptimal control actions.

In modern distribution networks, timely and accurate detection of these disturbances—and especially their origin—is critical. Not all disturbances should trigger the same control response. Events caused by local load or PV fluctuations within the LV grid require immediate attention from local voltage control systems. In contrast, disturbances that originate from the medium-voltage (MV) or high-voltage (HV) levels are typically outside the control scope of the LV grid and are better handled by upstream mechanisms such as on-load tap changers (OLTCs). To prevent miscoordination, the LV control systems must be able to distinguish between internal and external disturbances.

This thesis develops a simulation-based framework using realistic time-series load profiles to detect and classify voltage disturbances, with the goal of supporting future real-time grid control, particularly by identifying their type (e.g., load surge, upstream sag etc) and origin (LV, MV, or HV). This classification forms a critical decision-making input for coordinated voltage control (CVC) systems and supports a hierarchical, time-graded control strategy across the grid.

1.2. Objectives

The primary objective of this thesis is to develop a robust, simulation-based algorithm for detecting and classifying voltage disturbances in low-voltage (LV) distribution networks. The algorithm is designed to identify both the type (e.g., load surge, external disturbance etc) and the origin (internal LV, MV-originated, or HV-originated) of each disturbance. To support this goal, the work involves building a detailed HV–MV–LV grid model, simulating realistic disturbance scenarios using time-series load data, and analyzing the resulting voltage and current behavior.

A significant component of the work is the creation of labeled synthetic data, where realistic load profiles are used as a basis, and disturbances are artificially injected into the network under varying conditions. From these simulations, multiple monitoring indicators—such as voltage dips, line current, standard deviation of bus voltages, and transformer current peaks—are extracted and used as features for classification. A set of machine learning models is then trained and evaluated, including decision trees, support vector machines (SVM), k-nearest neighbors (k-NN), and ensemble methods. The classifiers are assessed on their ability to correctly identify both the origin and the specific nature of voltage disturbances, under different monitoring configurations.

The objective is not only to design an accurate classifier, but also to understand how data resolution, feature selection, and event characteristics affect detection reliability. This enables a practical foundation for future implementation in real-time control systems, particularly for enhancing the decision-making logic in Coordinated Voltage Control (CVC) frameworks.

1.3. Research Questions

To achieve these objectives, the thesis will address the following research questions:

- What monitoring indicators are most effective for detecting voltage disturbances in LV grids?
- How can monitoring data be used to determine whether a disturbance originated in the LV grid or propagated from upstream networks (MV or HV)?
- What classification techniques are most effective for distinguishing internal versus external disturbances?
- What is the minimum data resolution and monitoring infrastructure needed to enable reliable detection under realistic constraints?

1.4. Scope

The scope of this thesis includes the development of a MATLAB-based grid simulation environment that models an interconnected HV-MV-LV system. The focus is on simulating voltage-related disturbances, extracting relevant features from load flow data, and developing a classification algorithm capable of determining the disturbance origin. This is a purely simulation-based study and does not involve Real-time control implementation or Hardware testing.

The scope of the thesis is narrowed to focus exclusively on the detection and classification of voltage disturbances, while other disturbance types (e.g., harmonics, frequency deviations, and transient instabilities) are excluded. The following activities are included within the scope:

- Development of a grid model comprising HV, MV, and one detailed LV grid using simulation tools.
- Simulation of a range of voltage disturbance scenarios through time-series load flow analysis, including variations in local demand and generation, as well as upstream voltage deviations.
- Analysis of load flow results to identify monitoring indicators suitable for voltage disturbance detection.
- Development and testing of classification methods to determine the type and origin of each disturbance and evaluation of their performance.

2. Types and Standards of Voltage Disturbances

Voltage disturbances are deviations from nominal voltage conditions and represent a key threat to power quality in electrical networks [1]. In low-voltage (LV) distribution grids—especially those with high penetration of photovoltaic (PV) generation and dynamic load behavior—such disturbances can lead to customer complaints, equipment malfunctions, and unnecessary triggering of protection or control systems [2]. The growing complexity of LV networks, along with tighter regulatory requirements, necessitates precise detection, classification, and coordinated control response to various voltage anomalies.

This chapter reviews the key types of voltage disturbances typically encountered in distribution systems and the associated international standards that define their operational thresholds. These definitions form the foundation for the detection logic, simulation modeling, and classification framework developed in this thesis.

2.1. Classification of Voltage Disturbances

Voltage disturbances are broadly categorized based on duration, magnitude, and waveform characteristics. The standards most commonly used for classification include IEEE Std 1159, which defines disturbance categories and their duration/magnitude ranges, and EN 50160, which specifies the permissible limits of supply voltage quality under normal operating conditions. Together, these standards provide both the classification framework and the operational thresholds [3][4].

a) Short-Duration Voltage Variations

These events range from a fraction of a cycle up to one minute.

- **Voltage Sags (Dips):** RMS voltage drops to between 0.1 and 0.9 pu, lasting from 0.5 cycles to 1 minute. Often caused by motor starting or internal faults.
- **Voltage Swells:** RMS voltage rises to between 1.1 and 1.8 pu, similarly lasting up to 1 minute. Typically associated with load rejection or capacitor switching.
- **Interruptions:** Complete loss of voltage. Classified as:
    - *Instantaneous* (0.5–30 cycles),
    - *Momentary* (30 cycles–3 seconds),
    - *Temporary* (3 seconds–1 minute).

b) Long-Duration Voltage Variations

These persist for over a minute and often indicate sustained issues.

- **Undervoltages:** RMS levels between 0.8 and 0.9 pu, often caused by overloading or poor tap changer settings.
- **Overvoltages:** RMS values above 1.1 pu, typically due to load shedding or high local generation (e.g., PV peaks).

c) Transients

High-frequency disturbances occurring for microseconds to milliseconds.

- **Impulsive Transients:** Sudden unidirectional deviations, often due to lightning or fault clearing.
- **Oscillatory Transients:** Damped waveform oscillations that may arise from capacitor bank switching, control-induced resonances, or interactions between network components.

d) Waveform Distortions

These include non-sinusoidal steady-state anomalies:

- **Harmonics:** Integer multiples of the fundamental frequency, caused by inverters and nonlinear loads.
- **Interharmonics:** Non-integer multiples, often from variable-speed drives.
- **Notching and Noise:** Local waveform damage from power electronics or electromagnetic interference.

e) Voltage Imbalance

Voltage imbalance occurs when the magnitudes or phase angles of a three-phase system are not equal. This condition is often caused by uneven distribution of loads across phases, faults, or asymmetrical wiring. Imbalance can lead to equipment overheating, reduced efficiency, and increased losses in the system.

f) Voltage Fluctuations and Flicker

These are rapid, repetitive changes in voltage that can lead to perceptible flicker in lighting systems. Flicker severity is often measured using the Pst index defined in EN 50160.

2.2. Standards for Power Quality

2.2.1. IEEE Standard 1159

This standard defines various power quality disturbances and provides recommended practices for monitoring voltage events. It serves as a reference for duration and magnitude thresholds used in the thesis to generate labeled disturbance categories.

2.2.2. EN 50160

The European standard outlines permissible variations in supply voltage under normal operating conditions. Key specifications include:

- **Voltage Magnitude:** ±10% of nominal for 95% of each week.
- **Frequency:** 50 Hz ±1% for 99.5% of time.
- **THD:** ≤8% in LV networks.
- **Flicker (Pst):** ≤1.0 for 95% of the time.

2.3. Implications for CVC Coordination

The accurate classification of these disturbances has operational importance. Coordinated Voltage Control (CVC) refers to hierarchical strategies that integrate voltage regulation actions across LV, MV, and HV levels [5]. It integrates local controllers such as on-load tap changers (OLTCs), capacitor banks, and distributed generation reactive power controllers into a unified strategy. The goal is to ensure system-wide voltage stability and compliance with standards (e.g., EN 50160), while minimizing control conflicts and optimizing the use of available resources. In this context:

- **LV-origin disturbances** require **fast CVC response**,
- **MV-origin disturbances** may justify **delayed CVC**, letting MV-level control act,
- HV-origin disturbances are typically ignored by CVC, since recovery is handled by upstream controllers such as OLTCs. However, in the unlikely case that the expected OLTC action is delayed or does not occur, the CVC framework should include a fallback strategy, such as temporary local regulation at the LV/MV level or escalation to operator intervention.

Hence, disturbance types and their origin are not merely technical concepts—they directly impact the coordination of grid control hierarchy. Voltage disturbances play a critical role in defining power quality. The classification presented by IEEE 1159 and EN 50160 forms the foundation for both simulation and analysis in this thesis. Understanding these categories is essential for training classifiers that detect the nature and origin of disturbances. By translating these standard definitions into simulation actions and machine learning features, this thesis bridges the gap between theory and practical grid intelligence.

3. Detection Techniques for Voltage Disturbances

Voltage disturbances, such as sags, swells, transients, and interruptions, have a significant impact on power quality and the reliable operation of modern distribution systems. With the growing penetration of distributed energy resources (DERs), particularly photovoltaic (PV) generation in low-voltage (LV) networks, identifying and responding to these disturbances promptly is increasingly critical. This chapter provides a comprehensive literature review of detection methods for voltage disturbances and discusses their relevance and application in the context of this thesis work, which focuses on real-time disturbance detection and classification in LV grids.

Before discussing detection methods, it is important to outline the monitoring infrastructure that enables disturbance detection in LV grids. Real-time monitoring typically relies on devices such as smart meters, power quality analyzers, and micro-PMUs (μPMUs) [6]. These differ widely in data resolution (from sub-cycle sampling to minute-level averages), update frequency, and communication capabilities [7]. Depending on the system design, measurements may remain local to the device, be aggregated at the edge (e.g., in substation controllers), or be transmitted to central/cloud platforms. Communication delays and data losses are common challenges, which directly affect the feasibility of real-time detection and control coordination. In the literature, “real-time LV grid monitoring” usually refers to infrastructures capable of delivering sub-second to few-second updates with sufficient accuracy to capture disturbances such as sags, swells, and transients.

In this thesis, no real monitoring system is directly utilized. Instead, simulated time-series data is used to represent realistic measurement conditions. The design assumption is conceptually aligned with the Integrated Automation for Distribution Grids and DERs (INGA) project monitoring system [8], which provides high-resolution, low-latency data. However, compared to most currently deployed LV monitoring systems (which typically offer slower sampling and limited availability), INGA system represents a significantly more advanced capability. Thus, the data used in this thesis should be viewed as optimistic relative to present-day monitoring practice, but illustrative of what advanced monitoring would enable. Table 1 summarizes the main types of monitoring devices used in LV distribution networks and highlights their data resolution, update frequency, and typical application. These characteristics directly influence the suitability of each system for disturbance detection and real-time control.

Table 1. *Comparison of monitoring devices in power systems by resolution, frequency, availability, and application.*

Monitoring Device	Data Resolution	Update Frequency	Availability	Typical Use
Smart Meters (AMI)	15 min – 1 hour averages	15 min – 1 hour	Local (customer) → Utility	Billing, long-term load profiling, limited PQ data
Power Quality Analyzers	Sub-cycle (0.2–1 ms)	Continuous or event-based	Local or aggregated	Detailed PQ monitoring (sags, swells, harmonics)
Micro-PMUs (μPMUs)	120–240 samples per cycle	30–120 Hz	Local + Edge/Cloud	High-resolution synchrophasors, dynamic event capture
Substation Controllers	Cycle-level (10–20 ms)	50–100 Hz	Edge	Aggregation, fast voltage control actions
SCADA/RTUs	Seconds – minutes	1–10 s typical	Centralized (Utility SCADA)	Supervisory monitoring, slow grid management

While traditional signal processing methods remain widely used, machine learning (ML) techniques are becoming more prominent due to their flexibility, adaptability, and ability to work with noisy or non-stationary data [9]. In this thesis, detection is not treated as an isolated event-detection problem but rather as a component of a broader machine-learning-based classification pipeline, where detection contributes to real-time decision-making for control coordination.

3.1. Traditional Detection Methods

3.1.1. Threshold-Based Detection

The most commonly used approach in practical systems is threshold-based detection. In this method, voltage or current is continuously monitored and compared against predefined thresholds. For example, a voltage sag is detected if the RMS voltage drops below 90% of the nominal value for a specified duration (e.g., 0.5 cycles to 1 minute) [10].

This method is computationally efficient and easy to implement, especially in systems with limited processing capabilities. However, its limitations include poor noise tolerance, rigidity in changing grid conditions, and sensitivity to measurement placement. Despite these drawbacks, this technique remains widely deployed in industry-grade monitoring devices and standards such as IEEE 1159 and EN 50160 [11].

3.1.2. RMS and Duration Analysis

The RMS (Root Mean Square) method extends the simple threshold-based approach by calculating RMS values over sliding time windows rather than relying on instantaneous measurements. This enables the detection process to account for both the magnitude and duration of a disturbance. While threshold-based detection simply flags when a signal crosses a predefined limit, RMS and duration analysis can distinguish between very short, transient voltage dips and longer, sustained undervoltages. This makes it more robust in differentiating disturbance categories, although it still suffers from limited noise tolerance and sensitivity to window length selection [12].

3.1.3. Symmetrical Component Analysis

Symmetrical component analysis decomposes three-phase voltages or currents into positive, negative, and zero-sequence components [13]. This method is particularly effective in detecting unbalanced conditions, asymmetrical faults, or phase voltage deviations [14]. Since many LV disturbances involve phase imbalance due to uneven load distribution or faults, symmetrical components provide a structured way to capture such anomalies. However, accurate application requires synchronized three-phase measurements, which may not always be available at LV monitoring points [15].

3.1.4. Kalman Filtering and Adaptive Filters

Kalman filters and other adaptive filtering approaches provide dynamic tracking of voltage signals in noisy environments [16]. These methods estimate the system state in real time and can separate disturbances from measurement noise. Their main advantage lies in robustness against noise and adaptability to changing operating conditions. However, they require careful tuning and involve higher computational overhead compared to simple threshold methods. In practice, Kalman filters are more common in research settings than in widespread LV monitoring deployment [17].

3.2. Advanced Signal Processing Techniques

3.2.1. Short-Time Fourier Transform (STFT)

STFT provides time-frequency information by analyzing the frequency content within short time windows. In general, this method is effective in detecting frequency-based disturbances such as harmonics and transients. However, since harmonic analysis is beyond the scope of this thesis, the focus here is on the potential of STFT for transient detection. STFT also suffers from a fixed resolution trade-off: using a narrow window improves time resolution but blurs frequency details, while a wider window improves frequency resolution at the cost of time precision [18].

3.2.2. Wavelet Transform (WT)

Wavelet analysis allows multi-resolution signal decomposition, ideal for identifying sharp transients, flickers, or localized disturbances [19]. Many studies have used Discrete Wavelet Transform (DWT) to extract features and detect abnormalities with high temporal precision [20]. Due to the computational overhead and the goal of minimizing implementation complexity, these signal processing-heavy approaches were not adopted in this thesis. The monitoring frequency (1-minute resolution) and lack of high-frequency waveform data further reduce their applicability.

3.2.3. Hilbert-Huang Transform (HHT) / Empirical Mode Decomposition (EMD)

The Hilbert-Huang Transform (HHT), often implemented through Empirical Mode Decomposition (EMD), provides an adaptive, data-driven approach to analyzing non-stationary signals [21]. It decomposes a signal into intrinsic mode functions (IMFs), which are then analyzed using Hilbert spectral methods. This technique is effective for capturing oscillatory transients, flicker, and other disturbances with non-stationary characteristics. Its flexibility makes it superior to STFT in certain cases, but it is computationally expensive and highly sensitive to noise, limiting its deployment in real-time LV monitoring [22].

3.2.4. Prony Analysis

Prony analysis models a signal as a sum of damped exponential functions, making it particularly suitable for identifying oscillatory transients and resonance phenomena [23]. It allows precise estimation of frequency, damping factors, and amplitude of transient components. Although effective for waveform decomposition, Prony analysis requires high-resolution time-domain data and is sensitive to noise and model order selection

[24]. As a result, it is more common in transmission system studies and offline analysis than in real-time LV grid applications.

3.2.5. Statistical and Derivative-Based Detection

In this thesis, disturbance detection is facilitated by the extraction of statistical indicators from node voltages and line currents. The selection of these indicators is motivated by their ability to capture disturbance signatures in a computationally efficient and interpretable way. Features such as:

- **Maximum voltage dip across monitored LV buses**: captures the most severe local effect of a disturbance, directly reflecting sag or undervoltage severity.
- **Standard deviation of voltage among buses**: quantifies spatial voltage imbalance and helps distinguish local disturbances (high variability) from upstream events (more uniform impact).
- **Derivative or rate of voltage change at critical nodes**: highlights sudden variations, making it suitable for detecting fast disturbances such as transients or switching events.
- **Transformer peak loading (as disturbance proxy):** reflects stress on supply points, which often increases during disturbances and can serve as an indirect indicator.

These features were chosen because they balance interpretability, computational efficiency, and relevance to disturbance categories defined in IEEE 1159 and EN 50160, making them particularly suited for machine learning models that require structured numerical input." These features are calculated for each simulation timestep and used in supervised learning classifiers to determine the class and origin of the disturbance. Further details and justifications for their selection is presented in chapter 6 and onwards.

3.3. Detection in the Context of ML-Based Classification

Unlike traditional disturbance detection systems that isolate the detection stage, this thesis integrates detection within a machine learning pipeline. Each timestep is treated as a sample, and its feature vector captures the state of the grid. The classifier then determines:

- Whether a disturbance is present,
- What type it is (internal load, PV trip, MV-origin, HV-origin),
- Whether the origin is internal (CVC should act) or external (CVC should block).

This integrated approach improves accuracy and relevance in modern smart grids, where disturbances are often subtle and multi-source.

3.4. Summary

Detection of voltage disturbances is a critical function in modern grid operations, enabling real-time control and protection. In contrast to traditional threshold-based methods, modern monitoring and control systems in LV grids increasingly use feature-based detection approaches. These systems extract statistical indicators (such as voltage deviations, standard deviation across buses, or rate of change) and use them as structured inputs for machine learning models. This approach is needed because disturbances in smart grids are often subtle, multi-source, and masked by noise, making simple thresholds insufficient. In this thesis, this concept is realized through the feature-engineering framework described in Section 3.2.5, which provides the inputs for the ML-based classification pipeline detailed in Chapter 6.

4. Classification Techniques for Voltage Disturbance Detection

4.1. Introduction

Voltage disturbances in low-voltage (LV) distribution grids, such as voltage sags, swells, transients, and overvoltages, pose a significant challenge for both operational reliability and the coordination of control systems like Coordinated Voltage Control (CVC). As power systems evolve with the increasing integration of photovoltaic (PV) generation and fluctuating load demands, disturbances are becoming more frequent, less predictable, and harder to distinguish by conventional threshold-based monitoring. PV fluctuations can create rapid voltage sags or swells, while sudden load changes introduce irregular transients and imbalance conditions. These dynamics increase the risk of misclassification or delayed response if only manual or offline analysis is used. Therefore, the need for accurate, automated, and real-time classification of disturbances has become imperative to ensure timely and appropriate control actions.

This chapter presents a critical review of the techniques used for voltage disturbance classification, with an emphasis on their relevance to modern LV grid operation. Traditional methods, including rule-based and statistical approaches, are discussed alongside contemporary machine learning (ML) and hybrid methods. Special attention is given to classifiers applied in this thesis work, particularly decision trees, support vector machines (SVM), k-nearest neighbors (k-NN), and ensemble models trained on grid-simulated features. These models enable real-time inference on the origin of a disturbance, distinguishing between internal and external events. In this context, an internal disturbance refers to anomalies that originate within the LV network itself (e.g., load fluctuations, local PV generation trips), for which CVC should take corrective action. An external disturbance, on the other hand, originates upstream in the MV or HV network (e.g., voltage deviations from the main grid), where corrective action is managed by higher-level controllers such as OLTCs, and the CVC response should be blocked or the unlikely case that the expected OLTC action is delayed or does not occur, the CVC framework should include a fallback strategy. This distinction is critical for ensuring that voltage control actions are applied only when effective and necessary.

While supervised learning models such as Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), Decision Trees (DT), and Artificial Neural Networks (ANN) have been well established for decades, their application to voltage disturbance classification has gained traction more recently. This shift is driven by two main factors:

1. **Increased computational resources**, which make it feasible to train and deploy these models in real-time applications.
2. **Greater data availability** in modern power systems, thanks to advanced monitoring devices and widespread simulation frameworks, which provide the labeled datasets required for training.

In this context, these classifiers have been successfully trained on both synthetic and historical disturbance data to learn patterns from voltage and current profiles. Once trained, they can perform near real-time classification, distinguishing not only disturbance types but also their likely origin [25][26].

4.2. Traditional Classification Methods

4.2.1. Rule-Based Systems

Historically, disturbance detection in power systems relied heavily on deterministic, rule-based methods using thresholds on voltage magnitudes and durations. For example, a voltage sag is identified when the RMS voltage drops below 90% of the nominal for a specified duration [27]. While rule-based systems are simple, interpretable, and easy to implement on embedded devices, their effectiveness deteriorates under variable operating conditions, such as those caused by cloud-induced PV fluctuations or simultaneous load increases in smart homes. In addition to basic single-variable thresholding on RMS voltage, the literature also reports multi-variable rule-based approaches, which combine indicators such as voltage, current, frequency deviation, and rate-of-change to improve detection accuracy [28]. These methods extend the rule-based concept by adding logical conditions (e.g., if *voltage sag* **and** *high current* → classify as disturbance), but they still suffer from limited adaptability to unseen operating conditions compared to data-driven techniques. For example, *Dandea & Grigoras (2023)* combine PV penetration, load evolution, and generation output in their expert system for LV networks. *Yalman et al. (2022)* use both voltage and current magnitude changes across phases to locate voltage sags and identify disturbance types. These methods extend the simple threshold approach but still face limitations under highly variable conditions, motivating more flexible data-driven classification pipelines.

In the context of this thesis, rule-based detection alone would be insufficient to differentiate between internal and external events. This is because simple voltage thresholds cannot capture the spatial distribution of disturbances across the network or account for dynamic operating conditions. For example, both a local PV trip and an upstream MV voltage dip may produce similar magnitude changes at the LV bus, even though their required control responses are different. Furthermore, hierarchical CVC systems rely on time-grading and coordination logic to avoid conflicting actions across LV, MV, and HV levels. Rule-based methods lack the contextual awareness to incorporate such logic, which requires information beyond local thresholds, such as disturbance origin and propagation characteristics.

4.2.2. Statistical Techniques

Statistical classification techniques extend rule-based logic by using descriptive measures (e.g., mean, standard deviation, skewness) derived from voltage time-series data. Linear models such as Principal Component Analysis (PCA) or Linear Discriminant Analysis (LDA) have been used to reduce dimensionality and improve interpretability [29]. However, many statistical approaches rely on simplifying assumptions that do not always hold in practice. For example, they often assume that the data is stationary (statistical properties such as mean and variance do not change over time) and that classes are linearly separable (disturbances can be separated by simple decision boundaries). In real-world grids, these assumptions are rarely valid: voltage disturbances can be highly dynamic (non-stationary), last only for a very short duration, or occur simultaneously with other events, making them difficult to separate with purely linear statistical methods.

In the present thesis work, descriptive statistics such as voltage deviation, standard deviation across LV buses, and changes in bus voltage rates are explicitly computed during simulation to build a richer, ML-compatible feature set for classification.

4.3. Machine Learning Techniques

Machine learning (ML) has become increasingly important for voltage disturbance classification due to its ability to learn from complex datasets, generalize across diverse operating conditions, and adapt to noisy or overlapping signals. Unlike traditional threshold-based methods, ML approaches can incorporate multiple features simultaneously and optimize classification boundaries automatically, making them well-suited for LV distribution grids with high penetration of distributed generation [30].

4.3.1. Supervised Learning Models

Supervised learning methods have gained prominence for disturbance classification due to their adaptability, generalization ability, and capacity to learn from complex datasets. These models are trained on labeled disturbance events (e.g., sags due to load spikes or overvoltages from PV injection) and can recognize similar patterns in new data.

Styvaktakis, Bollen, and Gu (2002) presented one of the earliest approaches to event classification using RMS voltage measurements instead of waveform data. Their method relied on rule-based logic and thresholding, and it was able to classify common events such as interruptions, voltage sags, and swells. However, because only RMS snapshots were used, the approach had limitations in detecting overlapping or subtle events. This study is important because it shows that even low-resolution RMS features can support classification, although with reduced sensitivity compared to waveform-based methods.

Building on simulation-based approaches, Mozaffari et al. (2022) developed a real-time detection and classification framework for power quality disturbances. Their work used simulated disturbance features extracted from voltage time-series data, applying sequential detection and multi-hypothesis classification. The results demonstrated faster detection and higher accuracy compared with baseline detectors, confirming that simulated data can be useful for developing and benchmarking PQ classifiers. However, their study did not address the specific challenge of distinguishing disturbance origin across LV, MV, and HV grids.

Zhang et al. (2020) explored ensemble methods, specifically Random Forest classifiers, applied to a set of power quality indices such as sags, Total Harmonic Distortion (THD), and flicker severity. They reported classification accuracies exceeding 97%, demonstrating that ensembles are more robust than individual classifiers under noisy or overlapping conditions. Their findings align with a broader trend in the literature: ensembles provide higher generalization ability, making them attractive for real-world PQ monitoring applications.

Finally, Caicedo et al. (2023) published a systematic review of real-time PQ disturbance detection and classification methods. Their survey mapped out the different features and approaches used in the field, noting that while waveform-based methods dominate, fewer works rely on RMS or statistical features combined with simulated data. The review also highlighted that most existing studies focus on detecting disturbance type but do not address the classification of disturbance origin in the context of hierarchical control.

These studies show that supervised methods can achieve high accuracy (>90–97%) when well-chosen features are provided. Decision Trees and k-NN offer simplicity and interpretability, SVM provides robustness in high-dimensional spaces, and ensembles consistently deliver the best generalization.

4.3.2. Unsupervised Learning

Unsupervised learning methods, such as k-means and DBSCAN, are typically clustering-based and have been applied in power quality research for exploratory analysis and anomaly detection [31]. Unlike supervised learning, which requires pre-labeled data and assigns each feature vector to one of several predefined classes, clustering does not rely on labels. The beauty of clustering is that it does not necessarily force every data point into an existing class. Instead, it can generate new clusters for previously unseen patterns, making it particularly effective in detecting anomalies or outliers that fall outside normal operating conditions.

This property is especially valuable in power systems, where unexpected or rare disturbances may not be represented in training data. For example, if a disturbance produces features that deviate significantly from known sags, swells, or flickers, clustering methods can isolate them as a new group, signaling the presence of an unfamiliar event.
However, for the deterministic classification tasks required in coordinated grid control (CVC), clustering methods are less suitable. They provide no direct mapping to operational categories such as LV-origin vs. MV-origin disturbances, and manual interpretation or post-labeling is needed to assign practical meaning to clusters. As such, while unsupervised methods are highly valuable for anomaly detection, exploratory analysis, and pre-labeling for supervised training.

4.4. Deep Learning Approaches

Deep learning models like Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have demonstrated high performance in disturbance classification tasks, particularly when applied to large datasets of waveform recordings or spectrograms [32]. The effectiveness of these models strongly depends on the availability of sufficient training data, since the number of learnable parameters is typically very high. A common rule of thumb is that the dataset should contain at least an order of magnitude (≈10×) more samples than the number of parameters in the network; otherwise, the risk of overfitting becomes prohibitive, and the model fails to generalize. This requirement explains why deep learning approaches are often limited to studies with access to extensive PQ monitoring datasets.

In this thesis, the input data is generated through simulated disturbance scenarios rather than collected from real-world monitoring systems. While this approach provides controlled and well-labeled samples for training and evaluation, the overall dataset size and diversity are significantly smaller compared to large-scale monitoring databases. This limitation further restricts the feasibility of applying deep learning models, which require extensive and heterogeneous training data to achieve generalization.

Secondly, Deep learning approaches such as CNNs and LSTMs typically operate on raw waveforms or spectrograms, where they can automatically learn discriminative features from harmonic content, transient signatures, or temporal dependencies. In contrast, the present thesis uses simulation-based RMS values sampled at discrete time steps, which lack the fine-grained temporal and spectral information embedded in waveform data. This makes waveform-driven deep learning methods less relevant for the scope of this work.

Thirdly, it is worth noting that once trained, neural networks can be computationally efficient at inference, which makes them attractive for embedded deployment. However, their training phase demands significant computational resources and large datasets, and their lack of interpretability can be problematic for grid operation contexts that require transparent decision-making. For this reason, the thesis emphasizes lightweight, explainable classifiers such as decision trees, SVMs, K-NN and ensembles, which better align with the objectives of CVC coordination.

4.5. **Hybrid and Modular Approaches**

Hybrid methods combine physical modeling and machine learning to exploit both domain knowledge and data-driven flexibility. For instance, some studies have combined rule-based disturbance detection with ML classifiers to improve robustness against noise and changing operating conditions [33]. In the context of this thesis, the hybrid idea is reflected in the modular pipeline: disturbance features are extracted from power-flow simulations (a physics-based model), and classification is carried out using ML algorithms. This approach leverages the interpretability of physical modeling while benefiting from the adaptability of machine learning.

The thesis implements a modular structure where the dataset is generated through physical modeling (via power flow simulations), and relevant features are selected from these results for ML-based classification. This architecture allows easy replacement of components (e.g., testing different classifiers) and reflects hybrid best practices.

4.6. Summary

This chapter has reviewed traditional, statistical, and machine learning approaches for disturbance classification in distribution networks. Traditional rule-based and threshold-based methods remain simple and interpretable, but they show clear limitations under noisy or dynamic grid conditions. Statistical techniques extend this by incorporating descriptors such as mean, variance, or skewness of bus voltages, offering richer representations but still depending on manual feature engineering.

Supervised learning methods, including decision trees, k-NN, SVM, and ensemble models, have become the most widely applied in the literature. When combined with RMS, statistical, or PQ index features, these models regularly achieve classification accuracies above 90–97%, with ensembles providing the strongest robustness under noise. Deep learning approaches (CNNs, LSTMs) push accuracy even further (>95–98%) when waveform datasets are available, but their reliance on large labeled data and reduced interpretability limits their immediate applicability in LV grid CVC contexts. Hybrid and modular approaches strike a balance by combining physical insights with ML adaptability, improving robustness while preserving interpretability and design flexibility.

Taken together, the literature demonstrates that supervised learning methods, when coupled with well-chosen features, are now sufficiently powerful for real-time voltage disturbance classification. Deep learning and hybrid methods represent promising directions as richer datasets and monitoring infrastructures become available. Within this context, the thesis adopts a modular, simulation-based approach that leverages statistical features from power-flow results and applies lightweight supervised classifiers. This design choice reflects both the current state of the art in the literature and the practical requirements of disturbance origin classification in LV grids for CVC coordination. The comparative evaluation of different analytical and machine learning techniques used in this research is presented in Table 2.

Table 2. *Summary of techniques with corresponding pros, cons, and their application in this thesis.*

Technique	Pros	Cons	Use in Thesis
Rule-Based	Simple, explainable	Rigid, poor generalization	Not used
Statistical	Fast, interpretable	Linear assumptions, limited depth	Partial
Decision Trees	Fast, interpretable, edge-suitable	May overfit on noisy data	Main model
SVM	Robust, generalizable	Complex for large datasets	Used
k-NN	Simple, effective	Slow with large data	Used
Ensemble Models	Robust, accurate	Higher compute	Used
Deep Learning (CNN, LSTM)	High accuracy, automatic features	Data-intensive, black-box	Not used

5. Monitoring and Data Acquisition in Low-Voltage Grids

The increasing penetration of distributed energy resources (DERs), electric vehicles, and responsive loads in low-voltage (LV) distribution grids has made voltage quality monitoring more critical than ever. Voltage disturbances—such as sags, swells, transients, and harmonics—can impair the operation of local control schemes, trip PV inverters, or lead to excessive wear of equipment. Effective detection and classification of such disturbances require accurate and timely monitoring of electrical quantities across the LV network. This chapter explores the academic foundations and technological trends in LV grid monitoring, focusing on the measurable indicators, data acquisition constraints, and how these influence the feasibility of voltage disturbance detection.

5.1. The Role of Monitoring in Voltage Quality Management

Monitoring in LV grids serves two primary purposes: [34] ensuring voltage quality for end-users in compliance with standards such as EN 50160 and IEEE 1159, and [21] providing feedback for local voltage control systems such as coordinated voltage control (CVC). With increasing variability from photovoltaic (PV) generation and electrified loads, the likelihood of voltage deviation events—both localized and propagated from upstream grids—has increased significantly [34][35].

While traditional distribution networks relied on periodic manual inspections or billing-grade data from smart meters, modern LV grids require continuous and fine-grained observation of electrical states to detect and respond to rapidly occurring disturbances. The ability to distinguish between internal (LV-originated) and external (MV/HV-originated) events depends not only on the disturbance type but also on the availability, resolution, and accuracy of data provided by the real-time monitoring system.

5.2. Measurable Quantities in LV Grid Monitoring

Voltage and current are the primary physical quantities measured in LV grid monitoring. These are further processed to derive secondary indicators:

- **Voltage magnitude (RMS)** and **angle**
- **Current flow** in feeders and laterals
- **Active (P) and reactive power (Q),** including their **direction of flow**
- **Power factor (PF)**

- **System frequency**
- **Total harmonic distortion (THD)**
- **Voltage unbalance**, especially in three-phase networks

Among these, voltage magnitude and current flow are directly used in this thesis to simulate and classify disturbances. Specifically, features such as maximum voltage dip, standard deviation, and line current peaks are extracted from time-series voltage/current data in the simulation framework to train a machine learning classifier.

These indicators are typically derived from RMS values calculated over short cycles (e.g., 10–20 ms). However, in practical monitoring devices, such high-resolution RMS values are often not directly available to users. Instead, devices usually provide aggregated data (e.g., averages over 100 ms or 1 s). In this thesis, the data is simulated at a 1-second resolution. The underlying load profiles are originally available as hourly time series over one year (8760 values). To obtain second-level resolution for disturbance simulations, Piecewise Cubic Hermite Interpolation (PCHIP) was applied to generate smooth intermediate values without introducing artificial oscillations [36]. The details of Piecewise Cubic Hermite Interpolation (PCHIP) is in section 5.5. For PV units, no external time-series data were used; instead, PV generation was modeled as static injections at designated buses, consistent with the defined case studies.

5.3. Sensor Types and Monitoring Architectures

While this thesis does not implement real-world sensors, it is important to understand the general capabilities of available monitoring devices in practice:

- **Smart Meters**: Widely deployed at consumer endpoints, smart meters typically provide aggregated data at 15–60-minute intervals to centralized data storage, mainly for billing and consumption monitoring [37]. However, many devices are also capable of offering near real-time measurements through local communication interfaces. Early generations of smart meters—commonly those installed before 2013—had a relatively coarse voltage measurement resolution (typically around 1 V steps), which limited their usefulness for detailed power quality analysis.

  Recent developments, however, have significantly improved both measurement resolution and reporting frequency, enabling finer-granularity monitoring suitable for power quality and state estimation studies [38] [39]. In Finland, most utilities are expected to complete smart-meter upgrades by 2028, further enhancing data availability and accuracy in LV network monitoring [40].

- **Phasor Measurement Units (PMUs)** and **Micro-PMUs (μPMUs)**: Provide high-resolution synchronized measurements (30–240 samples/second), but are rarely deployed in LV networks due to cost and complexity [41].
- **Voltage Quality Analyzers**: Portable or semi-permanent devices, typically deployed at substations or feeders for short-term campaigns. They provide Class A measurements per IEC 61000-4-30, which makes them suitable for verifying compliance with voltage quality standards or investigating specific PQ complaints. However, their scope is usually localized and temporary.
- **Voltage Quality Monitoring Systems (VQMS):** Permanently installed infrastructures consisting of distributed monitoring devices connected via communication networks. Unlike analyzers, VQMS are designed for continuous, system-wide supervision of voltage quality. They enable utilities to track long-term trends, detect disturbances in near real time, and feed data into control or protection schemes.
- **Supervisory Control and Data Acquisition (SCADA)** systems: Typically limited to MV or HV networks, with sampling rates of seconds.

In this thesis, the idealized full-coverage virtual monitoring scenario has been assumed, where voltage and current at all buses and lines are available for each simulation time step. This assumption reflects a future-proof or research-oriented testbed, enabling isolation of classification accuracy from monitoring limitations.

5.4. Resolution and Sampling Rate: Impact on Detection

Resolution is a critical factor in the successful detection and classification of disturbances. Real-world systems may collect data at:

- **15-min to 1-hour intervals** (Smart meters – too coarse for event detection)
- **1-minute resolution** (Advanced metering or event-triggered recorders – minimal baseline for disturbance tracking)
- **Sub-second resolution** (PMUs – ideal but rarely available in LV)

In addition to time-resolution, the measurement resolution of monitoring devices also plays a critical role. For example, many smart meters provide voltage values with only 1 V granularity, which may be insufficient to capture subtle variations relevant for disturbance detection. In contrast, advanced PQ monitoring devices can report with millivolt-level precision, enabling more accurate characterization of small deviations [42].

It is also important to recognize that such fine resolution introduces sensitivity to the measurement chain, including instrument transformers and current sensors, whose intrinsic errors may become significant at millivolt scales [43]. Moreover, update frequency—that is, how often new RMS or averaged values are produced—can differ from the internal sampling rate of the device. For example, while the internal sampling ratio of modern smart meters may be relatively high, the reported values may still be updated every 1–10 seconds depending on configuration.

Thus, both the temporal resolution (how often measurements are taken) and the value resolution (how precisely quantities are recorded) jointly determine the effectiveness of disturbance detection.

5.5. Load Data Interpolation Using Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)

In this study, realistic load profiles were required at a one-second resolution to simulate time-domain voltage disturbance events in the distribution network model. However, the available load data was recorded on an hourly basis, representing 8,760 data points for one year. Direct use of hourly data would not capture the rapid fluctuations and short-duration transients that characterize voltage disturbances. Therefore, an interpolation technique was applied to obtain continuous load data with a sampling frequency suitable for dynamic simulation.

Among various interpolation approaches, the **Piecewise Cubic Hermite Interpolating Polynomial (PCHIP)** method was selected. PCHIP constructs a series of cubic polynomials between each pair of adjacent data points, ensuring that both the function values and the first derivatives are continuous across intervals. Unlike standard cubic spline interpolation, which can introduce oscillations and overshoots between widely spaced data points, PCHIP preserves the shape and monotonicity of the original data [36], [44]. This makes it particularly well-suited for electrical load data, where abrupt or non-physical fluctuations (e.g., negative or excessively large load values) are unrealistic and undesirable.

Mathematically, for two consecutive data points $(x_i, y_i) and \left(x_{\{i+1\}}, y_{\{i+1\}}\right)$, PCHIP defines a local cubic Hermite polynomial as:

$$H_{i(x)} = h_{\{00\}(t)} y_i + h_{\{10\}(t)}(x_{\{i+1\}} - x_i) m_i + h_{\{01\}(t)} y_{\{i+1\}} + h_{\{11\}(t)}(x_{\{i+1\}} - x_i) m_{\{i+1\}} \quad (5.1)$$

Where $t = \frac{x-xi}{xi+1-xi}$ is the normalized distance between the two points, and m_i and m_{i+1} represent the slopes (first derivatives) at the corresponding data points. The four basis functions are defined as:

$$h_{\{00\}(t)} = 2t^3 - 3t^2 + 1 \tag{5.2}$$

$$h_{\{10\}(t)} = t^3 - 2t^2 + t \tag{5.3}$$

$$h_{\{01\}(t)} = -2t^3 + 3t^2 \tag{5.4}$$

$$h_{\{11\}(t)} = t^3 - t^2 \tag{5.5}$$

The variables and parameters used in the above equations are defined as follows:

$x_i, x_{\{i+1\}}$: Consecutive data points in the time (or input) domain.

$y_i, y_{\{i+1\}}$: Corresponding measured or simulated load (or output) values

$m_i, m_{\{i+1\}}$: Slopes (first derivatives) at the data points

t: Normalized interpolation variable between 0 and 1

$h_{\{00\}(t)}, h_{\{10\}(t)}, h_{\{01\}(t)}, h_{\{11\}(t)}$: Cubic Hermite basis functions defining the polynomial shape over one interpolation interval

The slopes m_i are computed such that the interpolant remains monotonic wherever the data is monotonic, using weighted harmonic means of adjacent finite differences. This ensures that if the load data between y_i and y_{i+1} is non-decreasing (or non-increasing), the interpolated curve will not introduce artificial reversals or oscillations [45], [46].

In the context of this research, PCHIP interpolation was employed to generate second-by-second load values for the entire year, resulting in 8,760×3,600=31,536,000 samples. To maintain computational feasibility while preserving temporal resolution, a focused window of 2.43 hours (146 minutes) was selected for high-resolution analysis.

This time span provides a representative segment for evaluating the disturbance detection and classification framework under realistic dynamic conditions, without overwhelming computational resources. However, it is acknowledged that the selected window does not encompass the full variability of load behavior across seasons or differing photovoltaic (PV) generation levels. For instance, load data captured during summer periods may include strong PV-induced fluctuations, while winter profiles are typically dominated by heating demand with limited PV activity. As a result, models trained on data from a single time window may not generalize optimally to all seasonal conditions.

Despite this limitation, the restricted dataset serves as a proof of concept for validating the proposed data-driven methodology. The emphasis here is on demonstrating the framework's capability to process high-resolution synthetic measurements and to identify disturbance features effectively, rather than on achieving universal model generalization. Future work could expand the analysis to include multi-seasonal datasets to improve robustness and transferability.

Figure 1 shows a representative segment of the dataset after applying Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) interpolation to the original hourly system-load data. The interpolated curve provides a smooth and continuous per-second profile that preserves the overall shape of the measured load without introducing oscillations or overshoot. This example demonstrates how PCHIP effectively refines coarse temporal data while maintaining the physical realism of the underlying load behavior.

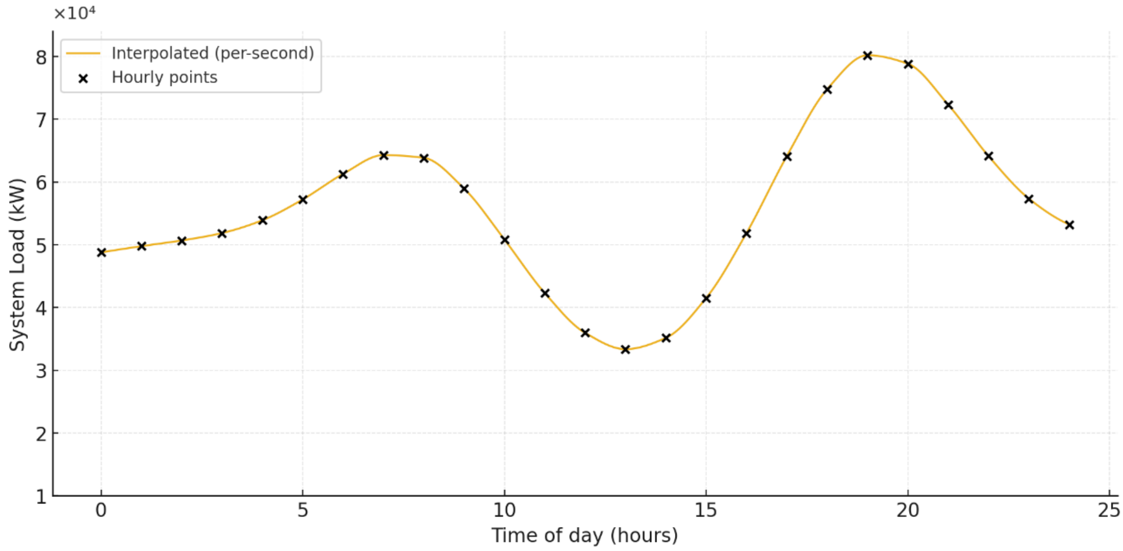

Figure 1. *Illustrative example of dataset segment generated by PCHIP interpolation*

5.6. Monitoring Coverage and Its Role in Classification

The classifier developed in this thesis relies on feature vectors composed of:

- Per-bus voltage dips
- Per-line current magnitudes
- Bus-level voltage variability (e.g., standard deviation)
- Transformer loading metrics (e.g., peak current)

These features are directly derived from full-grid virtual measurements, emulating ideal observability. However, in real-world deployments, such coverage is rare. A limited number of sensors or their suboptimal placement can significantly reduce classification accuracy and increase the chance of **false positives** or **false negatives**.

Although this thesis assumes full visibility to isolate the effectiveness of the classifier logic, future research could incorporate sensor placement optimization or state estimation methods to account for partial observability [47].

5.7. Summary

Monitoring in LV grids forms the empirical basis for any voltage quality analysis or control strategy. This chapter reviewed the measurable quantities, sampling considerations, sensor types, and limitations of real-world monitoring systems. It justified the assumptions made in this thesis—particularly the use of full-coverage, 1-minute interval virtual measurements—as a controlled environment for training and validating a disturbance origin classifier. While not reflective of current widespread practice, such assumptions enable focused analysis of classification accuracy under ideal data conditions, setting the stage for future real-world adaptation.

6. Research Methodology and Implementation

6.1. Introduction

This chapter presents the full implementation of the simulation and classification framework developed for the detection and classification of voltage disturbances in low-voltage (LV) distribution networks, with external disturbances originating from the medium- and high-voltage (MV/HV) grid levels. The system was implemented entirely in MATLAB and encompasses the modeling of a radial distribution grid, simulation of diverse disturbance scenarios, dynamic power flow analysis, feature engineering, and the deployment of supervised machine learning algorithms for automated classification under simulated conditions.

The focus is on creating a reproducible and modular simulation pipeline, enabling both experimentation with disturbance scenarios and systematic training of disturbance classifiers. Distinct from approaches that rely solely on synthetic data, this framework integrates real or statistically representative annual load profiles with controlled disturbance injection, yielding a hybrid dataset that supports both realism and experimentation. The implementation supports both binary (origin-based) and multi-class (type-based) classification tasks. The following sections describe, in sequence, the core technical elements of the system: data acquisition, grid configuration, simulation dynamics, disturbance modeling, load flow computation, feature engineering, machine learning workflow, and result visualization. The overall methodology adopted in this study is illustrated in Figure 2. As shown, the process begins with data input and grid structure setup, followed by simulation of power flow and disturbances. Relevant electrical features are then extracted and used for classifier training in both binary and multi-class configurations. The final step involves model evaluation using performance metrics and confusion matrices, with the results stored for analysis.

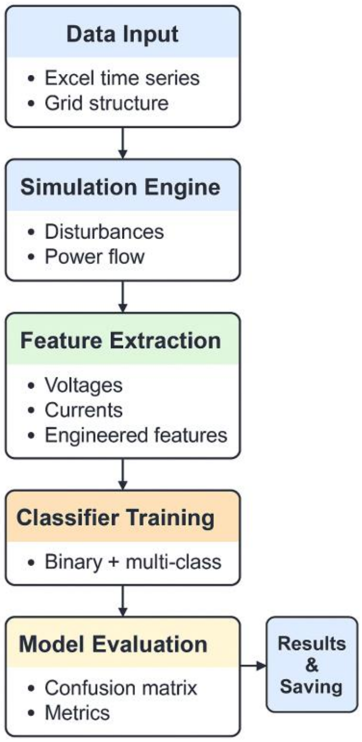

Figure 2. *Overview of the proposed methodology and data processing stages.*

6.2. Justification for Simulation-Based Research

Simulating power systems is a widely accepted approach in electrical engineering due to the complexity, safety concerns, and regulatory constraints involved in testing real disturbances on live grids. This methodology leverages simulation to safely replicate a wide range of scenarios, with realistic load behavior and disturbance dynamics.

The simulation framework offers several advantages. It enables controlled testing without endangering equipment or violating regulatory constraints, ensures repeatability by allowing identical scenarios to be replayed under different classifier settings, and provides the flexibility to test different grid configurations, control responses, and disturbance types. Furthermore, the high time-resolution used (1-second intervals) allows the system to capture transient phenomena that are critical for accurate classification but are rarely available in field data.

While the disturbances themselves are synthetic, the load profiles are based on real consumption behavior, providing a meaningful basis for simulation. However, it is important to acknowledge a key limitation: the original profiles are available only as hourly averages (8,760 points/year). To obtain a higher temporal resolution, these data were interpolated to 1-second intervals using piecewise cubic interpolation (PCHIP). This interpolation enables the simulation of short-term disturbance dynamics but does not guarantee that the synthetic 1-second values faithfully represent real-world second-to-second variations in consumption.

Furthermore, the load profiles used here represent aggregated consumption from groups of similar customers, rather than measurements from individual consumers. As such, the profiles capture average trends but not the variance or random fluctuations typically present in individual load behavior. Consequently, the resulting synthetic data should be interpreted as representative of *mean load dynamics* suitable for controlled scenario testing, rather than a full stochastic reproduction of real distribution-level variability. Therefore, the results should be viewed as an indicator of disturbance classification feasibility under controlled synthetic conditions, rather than an exact reflection of field measurements.

6.3. Research Design Overview

The study employs a quantitative, simulation-driven experimental design. The process begins with the acquisition of real-world hourly load profile data for 2018, representing various consumer types and obtained from an official source [48]. This was followed by the modeling of a 17-bus radial distribution network in MATLAB, composed of a single low-voltage grid under observation, connected through a medium-voltage feeder to a high-voltage source. The network includes residential and PV-connected buses, and was designed to support both internal disturbance injection in the LV grid and external disturbance propagation from the MV/HV side. Disturbances were artificially injected into the model to reflect both internal (LV) and external (MV/HV) disturbance conditions, as detailed in this chapter. This enabled dynamic behavior simulation and classification based on disturbance origin.

These simulations generate system-wide voltage and current values that are synchronized within each time step, since the load-flow solution provides a consistent snapshot of the entire grid at a given second. Across time, these snapshots are aligned to the same simulation clock, resulting in a coherent time series of synchronized measurements. From this raw data, relevant statistical and electrical features are extracted to form the input vectors for supervised machine learning algorithms. The design enables control over data quality, disturbance variability, and system repeatability—making it well-suited for training robust classifiers in environments where real-world disturbance data is sparse or unavailable. This simulation-based approach specifically addresses two recurring limitations in power system research: the difficulty of obtaining labeled field data for disturbances, and the challenge of testing under controlled yet varied operational conditions.

6.4. Data Acquisition and Preprocessing

After import into MATLAB, the data were preprocessed for consistency and usability, including date format conversion and scaling of power levels to match realistic LV, MV, and HV network parameters. The load distribution across different voltage levels is illustrated in Figure 3. Although missing or anomalous values were not explicitly handled, the dataset was assumed clean based on its official source. Active power values were normalized to a 1 MW system base, used as the per-unit reference for all load and generation quantities.

No time-varying PV generation profile was included; PV output was modeled as a constant injection at specified buses. This simplification was adopted for two main reasons. First, the objective of this study is to evaluate disturbance detection and classification performance, rather than to analyze PV generation variability or grid-balancing effects. Using a constant PV injection isolates the influence of voltage disturbances from the stochastic variations caused by solar irradiance changes, allowing for clearer attribution of classifier performance to disturbance events themselves. Second, incorporating realistic PV dynamics would require additional meteorological data (e.g., solar irradiance and temperature) and high-frequency PV generation profiles, which were not available at the same temporal resolution as the load dataset. Therefore, a steady PV injection provides a computationally efficient and methodologically consistent representation of distributed generation for this study.

As discussed in chapter 5, The load dataset was interpolated to a one-second resolution using piecewise cubic Hermite interpolation (PCHIP). The second-wise resolution is essential for capturing short-lived voltage disturbances—such as sags, swells, or switching transients—that occur over milliseconds to seconds. These events' onset, magnitude, and recovery would be invisible in hourly or minute-level data. Thus, interpolation enables second-level simulation where needed without requiring year-long high-resolution modeling. High-resolution data also benefit machine learning classifiers for disturbance detection. Fine-grained time series reveal critical transient patterns in voltage, current, and power—such as dips, spikes, and rapid changes—that are vital for accurate classification.

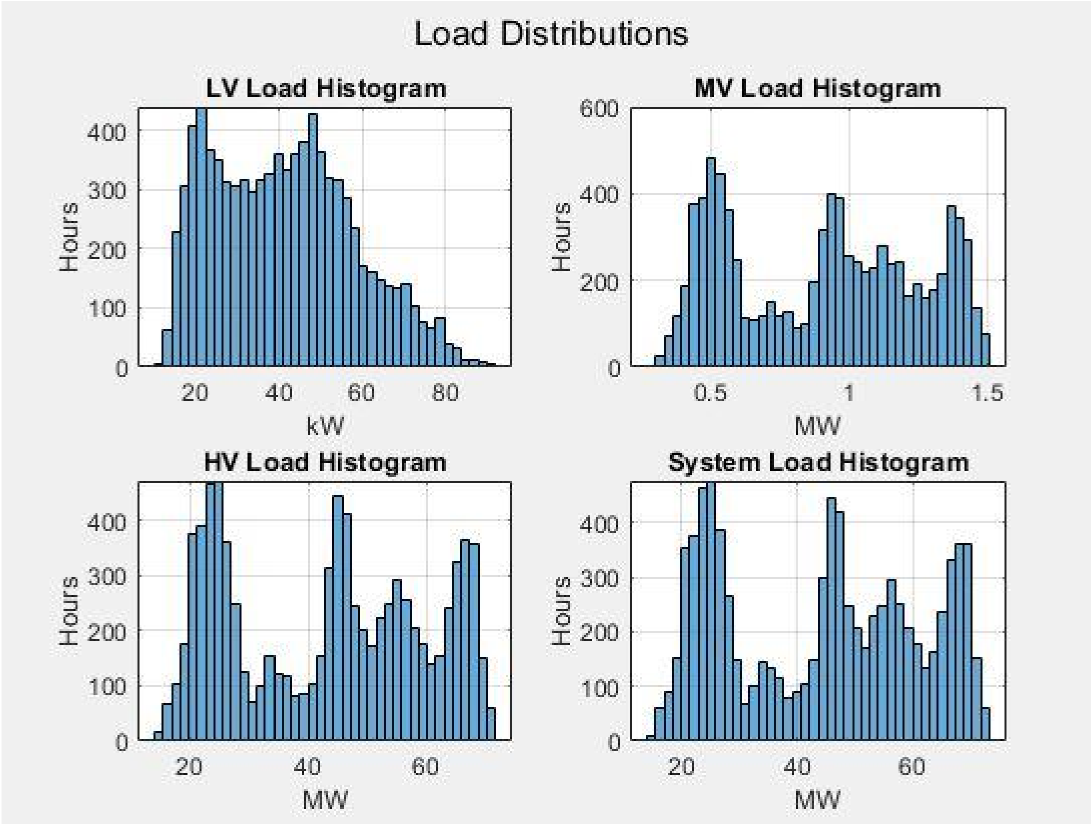

Figure 3. *Load distribution histograms for different voltage levels (LV, MV, HV) and overall system load.*

The computed annual mean, median, and standard deviation values reveal the magnitude and variability of the demand at each network level. The low-voltage (LV) network exhibits the smallest magnitude of load (mean = 41.2 kW), while the high-voltage (HV) network dominates the overall energy demand (mean = 43.4 MW). The medium-voltage (MV) level bridges the two, with an average of approximately 0.91 MW.

6.5. Grid Topology and Modeling Approach

The modeled network comprises a 17-bus radial system representing a semi-urban low-voltage (LV) grid connected to a medium-voltage (MV) feeder, with a simplified high-voltage (HV) source included to emulate external disturbances. Although the complete network allows realistic upstream–downstream interactions, analysis focuses solely on the LV grid.

The topology includes residential and commercial loads, PV-connected buses, an MV–LV transformer interface, and the HV source for simulating upstream voltage disturbances. Bus and line parameters are expressed in per-unit values, with impedance and voltage levels selected to reflect typical distribution network characteristics. While not directly taken from IEEE or DSO models, these parameters adhere to standard utility-scale simulation practices. The illustration is shown in Figure 4.

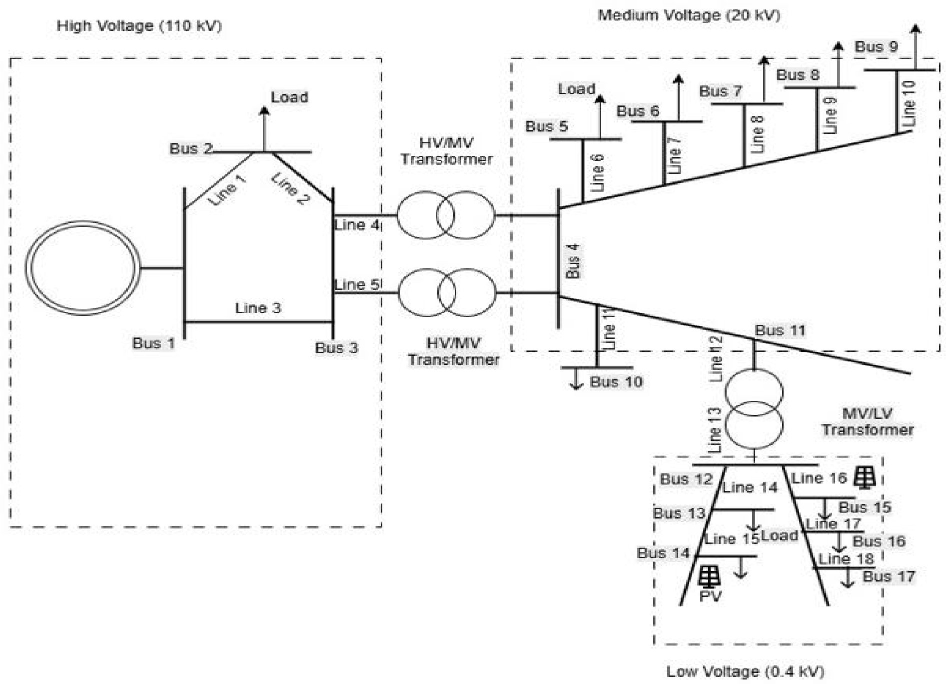

Figure 4. *Modelled Network*

The structural definitions of the modeled network components are summarized in **Appendix A**. These include the detailed organization of the bus and line matrices that define the electrical parameters of the system, such as node voltages, power injections, and branch impedances. These matrix structures form the basis of the power flow and disturbance simulations implemented in this study.

To reduce complexity, the system is modeled using a balanced single-phase approximation. Despite this simplification, it effectively captures voltage variations and line loading behavior under both normal and disturbed conditions, enabling accurate disturbance detection analysis. The modeled system is systematically categorized across three voltage levels as following:

- **High Voltage (HV)**: Buses 1–3 (nominal: 110 kV)
- **Medium Voltage (MV)**: Buses 4–11 (nominal: 20 kV)
- **Low Voltage (LV)**: Buses 12–17 (nominal: 400 V)

The topology is hierarchical, reflecting real-world distribution systems:

- Buses 1 to 3 represent transmission-level interconnection points.
- HV-to-MV transformation occurs between buses 3 and 4 via two transformers.
- MV feeders (buses 5–9 and 10–11) distribute power further into the network.

- A transformer at bus 11 connects the MV system to the LV network starting at bus 12, which then branches into two LV feeders: buses 13–14 and buses 15–17.

This structure allows clear separation and control of load allocation, line characteristics, and analysis granularity across voltage levels. The detailed modeling of the grid components is presented below:

6.5.1. Line Parameters

The grid's line parameters are modeled using standard cable data [49]–[51]. For each line, the resistance and reactance per unit length (in ohms/km) were calculated based on conductor type and cable specifications. Then, the total R and X values were determined for the actual physical lengths of the lines. These are converted into per-unit (p.u.) values using the base impedance derived from the system base power and nominal voltages [52].

6.5.2. Load Profiles and their Mapping

Load assignment plays a central role in system realism and classifier performance. The system uses 14 annual load profiles [48], each consisting of hourly energy consumption values for 8,760 hours (1 year). These profiles represent different consumer types (residential, commercial, industrial) with annual energy totals ranging from 1,000 kWh to 6,000,000 kWh. The original profiles were normalized to a per-unit base during preprocessing to maintain proportional energy relationships while allowing consistent comparison across voltage levels.

These profiles were mapped to buses across the LV, MV, and HV networks based on relative energy demand:

- **LV mapping**: Low-consumption profiles, representing residential, agricultural, and small commercial users, were assigned to LV buses. This simulates real-world load distribution in residential neighborhoods.
- **MV mapping**: Mid-range profiles such as commercial facilities, light industries and aggregated residential areas were scaled and mapped to MV buses, representing commercial feeders or light industrial zones.
- **HV mapping**: The highest-demand profiles such as businesses and industrial consumers were scaled and assigned to HV buses to simulate bulk loads or substation exits.

6.5.3. Load Scaling and Profile Conversion

The load profiles used in this study represent consumer power consumption in kilowatts (kWh/h) at hourly intervals across a full year (8,760 data points per profile). The resulting second-wise kW values are then converted into per-unit (p.u.) values for compatibility with the power flow solver. Using a base power of 1 MVA (or 1000 kW), the per-unit load at each time step is calculated. This per-unit scaling ensures numerical consistency in the simulation, supports correct impedance matching across voltage levels, and facilitates accurate voltage and current magnitude computation throughout the grid.

6.5.4. Load Assignment Summary and Balance

After all mappings and scalings, the total load per voltage level is as follows:

a) **LV Load Summary:**
   - **Number of Customers**: 11
   - **Total Energy**: ~230,000 kWh
   - **Average Load**: 41.21 kW (whole LV grid)
   - Distributed across buses 12–17 in proportion to profile size and location.

b) **MV Load Summary:**
   - **Total Energy**: ~7.96 GWh
   - **Average Load**: 908.79 kW (whole MV grid)
   - Assigned to buses 4–10, representing urban feeders or industrial clusters.

c) **HV Load Summary:**
   - **Total Energy**: ~380 GWh
   - **Average Load**: 43.38 MW (whole HV grid)
   - Assigned to buses 1–3, modeling bulk supply points.

Figure 5 illustrates the annual load variation across the HV, MV, and LV networks. Each curve represents the aggregated load at the respective voltage level after all mappings were applied. The seasonal trend is evident, with higher loads during winter months and reduced demand in summer, corresponding to typical heating-related consumption patterns. The smooth weekly and daily variations demonstrate that the mapped profiles collectively reproduce realistic temporal dynamics suitable for simulation-based disturbance analysis.

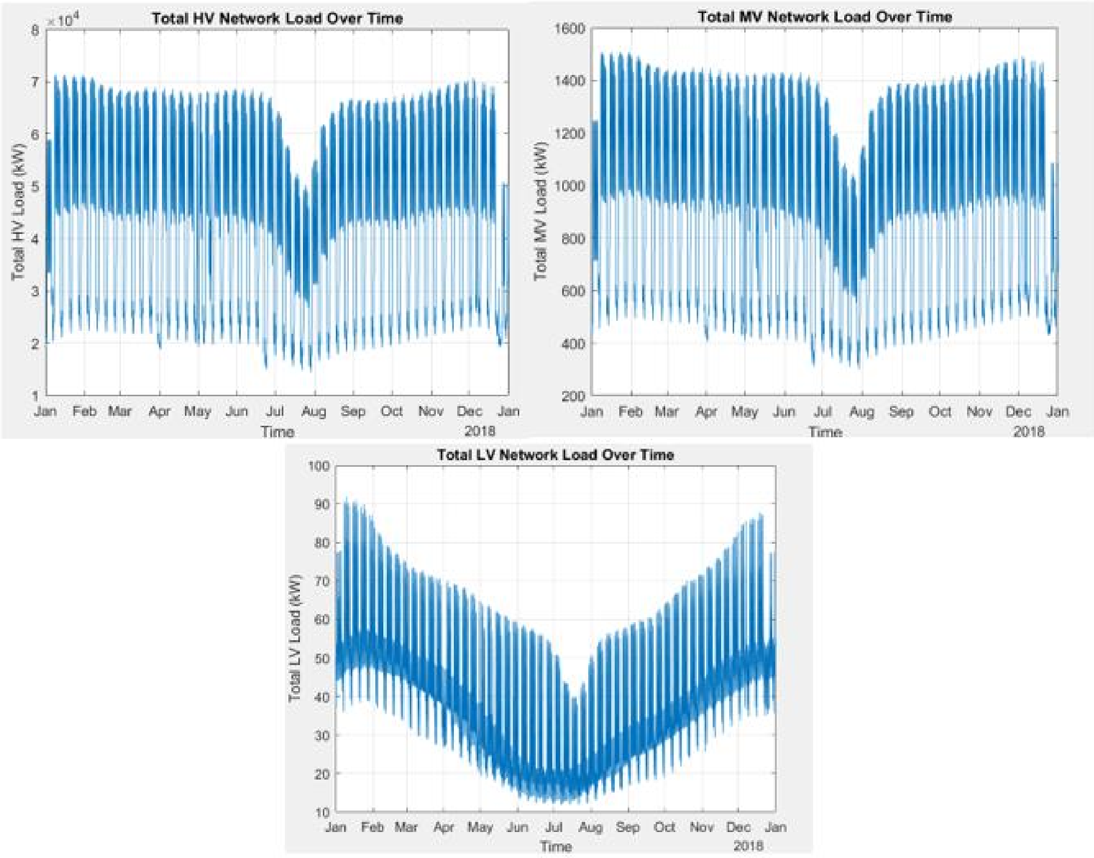

Figure 5. *Load assignment on HV, MV and LV*

6.5.5. Transformer Parameter Finalization and System Validation for LV Disturbance Classification

The reliable simulation of a power system, particularly for the purpose of disturbance detection and classification in the low-voltage (LV) domain, depends critically on the accurate modeling of key components — transformers being one of the most important. In this context, the transformer parameters — including per-unit resistance (R), reactance (X), and tap-changing ratios — were finalized through careful analysis, iterative tuning, and validation using power flow simulation results and system behavior [53]–[55].

The selected per-unit values (R = 0.005–0.01 p.u., X = 0.05–0.10 p.u.) align with typical ranges reported in literature and standard transformer data [54], [55]. The tap ratios were fixed at 1.00 to isolate disturbance effects without voltage regulation influence. The modeled grid includes three transformers across two voltage interfaces:

- **Two HV–MV transformers (in parallel)** connecting the 110 kV high-voltage transmission network to the 20 kV medium-voltage network.
- **One MV–LV transformer** connecting the 20 kV medium-voltage network to the 400 V low-voltage distribution system.

The main transformer parameters employed in the network model are summarized in Table 3. As shown, the system includes two high-voltage to medium-voltage transformers and one medium-voltage to low-voltage transformer, each characterized by their respective impedance values and tap ratios. These parameters are essential for accurate load-flow in the simulated grid. The per-unit base for all calculations is set to 1 MVA, which simplifies transformer modeling and enables consistent scaling of load and line parameters across the system.

Table 3. *Transformer data including voltage levels, impedance parameters, and tap ratios used in the study.*

Transformer	Voltage Levels	Resistance (R)	Reactance (X)	Tap Ratio
HV–MV Transformer #1	110 kV → 20 kV	0.010	0.100	1.00
HV–MV Transformer #2	110 kV → 20 kV	0.010	0.100	1.00
MV–LV Transformer	20 kV → 0.4 kV	0.005	0.050	1.00

a) Rationale Behind Parameter Selection

- **Resistance (R):** The selected resistance values for both HV–MV and MV–LV transformers are small but non-zero, avoiding unrealistic lossless behavior while maintaining physical plausibility. Typical R values in practical transformers are low due to the copper windings, and values such as 0.005–0.01 p.u. on a 1 MVA base reflect such characteristics.
- **Reactance (X):** The dominant impedance in transformers is reactance, often in the range of 0.05–0.15 p.u. The selected values (0.05 p.u. for MV–LV and 0.1 p.u. for HV–MV) are aligned with practical engineering data and ensure appropriate voltage drop.
- **Tap Ratios = 1.00:** Although real-world transformers may include tap changers to maintain voltage within bounds, in this study, tap changers are disabled (set to a ratio of 1.00) to isolate the effects of external and internal disturbances without introducing artificial voltage compensation.

The finalized transformer parameters — R, X, and tap ratios — have been carefully chosen and validated to align with both engineering best practices and the specific goals of voltage disturbance classification in the LV network. The values ensure numerical stability, simulation accuracy, and physical realism, forming a reliable foundation for subsequent disturbance injection, feature extraction, and classifier training phases.

6.5.6. Reactive Power Control and PV Inverter Modeling

To simulate realistic smart inverter behavior, a voltage-dependent reactive power control strategy (Q(V) droop) was implemented at PV-connected buses. Reactive power output was dynamically adjusted based on local voltage measurements: voltages below 0.95 p.u. triggered maximum reactive injection, while those above 1.05 p.u. resulted in absorption. Intermediate values invoked a linear droop response. At extreme over-voltage levels exceeding 1.10 p.u., PV active power was curtailed, and reactive support was withdrawn. These per-step evaluations enabled detailed simulation of inverter behaviour and its effect on voltage stability, following principles inspired by standards such as IEEE 1547—thereby improving the realism of simulated responses used for disturbance origin classification.

Importantly, this control logic was not embedded directly into the load-flow solver itself. The original load-flow algorithm retained its constant-voltage PV bus formulation, where each PV node attempts to maintain a reference voltage (Vref) as long as the required reactive power lies within its limits (Qmin–Qmax). When limits are reached, the bus would traditionally switch from PV to PQ mode. In this implementation, however, the Q(V) droop controller was applied externally in a supervisory loop: the solver was executed to obtain bus voltages, the droop law was evaluated to compute updated reactive power commands, and then the bus data were modified (sometimes forcing PV nodes to PQ) before re-running the solver. This approach ensured that the inverters' reactive support followed the intended droop characteristic, while still relying on the standard load-flow routine for the underlying network solution. The implemented Q(V) droop characteristic is defined by *Equation (6.1)*, while the corresponding droop slope m is computed from *Equation (6.2)* based on the inverter's reactive capability and voltage bandwidth. The maximum and minimum reactive power limits, determined by the inverter's rated power and power factor constraint, are expressed in *Equation (6.3)*. Together, these relationships ensure a continuous and physically consistent Q–V control behavior across the entire operating range.

$$Q(V) = \begin{cases} Q_{\max} & \text{if } V < V_{\min} \\ m(V - V_{\text{ref}}) & \text{if } V_{\min} \le V \le V_{\max} \\ Q_{\min} & \text{if } V > V_{\max} \end{cases} \tag{6.1}$$

- V: voltage at PV bus
- $Q(V)$: reactive power output
- m: droop slope
- $V_{\text{ref}}, V_{\min}, V_{\max}$: thresholds

The slope m is chosen to guarantee a smooth transition between the two limits, and is derived from the inverter's reactive capability and the voltage band width:

$$m = \frac{Q_{\max} - Q_{\min}}{V_{\min} - V_{\max}} \tag{6.2}$$

For example, assuming a power factor constraint of 0.95, the maximum reactive capability is given by:

$$Q_{\max} = P \cdot \tan(\cos^{-1}(0.95)), \quad Q_{\min} = -Q_{\max} \tag{6.3}$$

where P is the inverter's rated active power. With thresholds Vmin=0.95 p.u., Vmax = 1.05 p.u., and Vref= 1.0 p.u., the slope ensures that the full reactive range is utilized between the under- and over-voltage limits.

6.6. Disturbance Modeling and Implementation

In real-world distribution networks, disturbances arise from various internal and external causes, such as PV disconnections, voltage sags, or abrupt load variations. To emulate these phenomena realistically, seven disturbance classes were defined and systematically introduced into the simulation environment, representing typical operating anomalies in low-, medium-, and high-voltage networks.

The implementation supports both randomized and reproducible disturbance scenarios. Using a fixed random seed ensures consistent timing and sequencing of events across runs, which is essential for supervised learning. Each disturbance is characterized by four parameters: start time, duration (30–300 seconds), class type, and spatial location (bus index or upstream node).

During simulation, the system dynamically modifies load and generation setpoints for internal events or upstream bus voltages for external ones. All events are isolated in time and space to maintain clear labeling, and system parameters are reset to baseline after each event.

Each simulation run includes ten disturbances, covering all seven classes, sampled at one-second resolution. This process produces thousands of labeled time steps per run, resulting in approximately 6,000 balanced feature vectors across eight classes (seven disturbance types plus normal operation). The final dataset was divided into training (70%), validation (20%), and test (10%) subsets for model development and evaluation.

6.6.1. Disturbance Activation Mechanism

During each simulation time step, the code checks whether a disturbance is active by comparing the current simulation time t with predefined start and end times. If a disturbance is active, its corresponding case is executed. The affected bus is selected based on the disturbance type, and system parameters are modified accordingly. Disturbance states are tracked using current_label and current_origin, which record the type and source (internal vs. external) of each event.

The simulation defines seven types of disturbances (cases 0 through 6), each representing a unique operating condition. These are discussed in detail below.

a) Case 0: External disturbances

This disturbance simulates an external disturbance impacting the distribution network. It reduces the voltage at a representative bus to emulate the impact of a sag originating outside the modeled LV area. Specifically, the voltage at selected bus is reduced by 20%. Although simplistic, this approach effectively introduces a system-wide voltage stress without targeting any specific generation or load assets. It serves as a generic representation of upstream issues, such as transmission-level disturbances or feeder outages. it does not explicitly model the fault current contribution within the grid components where the disturbance occurs.

b) Case 1: Load Increase

This disturbance models a sudden increase in customer demand. It randomly selects a bus and scales up its active power by a factor between 1.5 and 2.0. If the bus initially has zero load, a small base load is added first to prevent multiplication by zero: This simulates events such as multiple large appliances turning on or electric vehicle (EV) charging clusters activating unexpectedly.

c) Case 2: PV Power Reduction

This disturbance simulates a partial derating of a PV inverter, potentially due to shading, over-temperature, or curtailment. The inverter's active power (Pg) is reduced by a random ratio between 20% and 80%. The reactive power (Qg) is recalculated based on a fixed power factor (PF = 0.95), and if the remaining power falls below a threshold (e.g., 200 kW), the bus type is changed from PV (type 2) to PQ (type 3). This disturbance closely mimics real-world behavior where inverters reduce output due to operational limits or system commands.

d) Case 3: PV Trip

This case models the complete disconnection of a PV inverter, which may occur due to protection trips, hardware failure, or grid-code violations (e.g., overvoltage). The inverter's active and reactive power outputs are set to zero, and the bus is switched from PV type (constant-voltage generator bus) to PQ type (load bus). This ensures that the bus no longer enforces a voltage setpoint after the inverter is disconnected. In practice, this bus then behaves like a normal demand node. A very small constant load is added if necessary to avoid the numerical instability that can occur when a bus is left with no injections at all. This approach guarantees that the power-flow solver converges consistently while correctly representing the inverter trip as a loss of generation.

e) Case 4: Multi-Customer Load Spike

In this event, multiple LV buses (excluding PV buses) are randomly selected and subjected to a simultaneous load increase. Each selected bus experiences a load growth of 20% to 70%. This simulates scenarios such as heatwave-induced AC usage or coincidental appliance activation, leading to network-wide stress.

f) Case 5: MV-Origin Disturbance

This disturbance selects a random MV bus and applies a sudden load change by scaling its active power by 0.5 to 0.9. It mimics upstream voltage fluctuations or reconfiguration events.

g) Case 6: HV-Origin Disturbance

Here, all HV buses are subjected to a simulated voltage sag. The voltage magnitude is directly set to 0.9 p.u., representing a disturbance propagating from the transmission level. This event is treated as external and can significantly affect the downstream voltage profile. This value is severe enough to produce observable downstream impacts in the LV network while still maintaining overall system solvability. In future extensions, the sag magnitude could be randomized (e.g., within 0.85–0.95 p.u.) to introduce variability across simulation runs, but in this study a fixed 0.9 p.u. level was used to ensure consistent labeling and comparability.

h) Case 7: Normal Condition

If no disturbance is active at a given time step, the system operates under baseline load and generation conditions, and the corresponding sample is labeled as "normal" (Class 7). In this state, no disturbances are introduced; instead, bus voltages, line currents, and other variables are determined solely by the natural balance of demand, generation, and network impedances. This provides a clean separation between undisturbed operating points and disturbed conditions, ensuring consistent labeling for dataset construction.

6.6.2. Disturbance Classification: Internal vs. External

To support supervised learning, each disturbance is labeled as internal (originating within LV area) or external (from MV or HV side). This helps the classifier learn both the type and origin of disturbances based on voltage and current responses. By introducing diverse and physically meaningful disturbances at varying locations and magnitudes, the simulation creates a rich dataset suitable for training classification models. This robust methodology allows for development of intelligent disturbance detection strategies.

6.7. Load Flow Solver and Dynamic Analysis

To calculate bus voltages and line currents under different operating and disturbance scenarios, load flow (power flow) analysis is essential. In this thesis, the Newton-Raphson method was chosen for its robustness and compatibility with the simulation objectives. It was implemented in MATLAB as part of the custom simulation routine, balancing computational efficiency with model flexibility. The network operating point was computed using a Newton–Raphson based power flow routine, executed sequentially at one-second resolution to generate a quasi-static time-series of system states. The solver determined bus voltages, angles, and branch currents, with appropriate base values assigned for each voltage level.

At each step, the resulting bus voltages and branch currents were stored for subsequent feature extraction. From these raw electrical states, both direct variables (bus voltages, line currents) and derived indicators (e.g., voltage spread across LV buses, rate of voltage change at a representative bus, transformer current) were calculated. Together, these time-indexed outputs provided the foundation for constructing the feature set used in the classification framework.

6.7.1. Line and Transformer Current Calculation

Accurate calculation of current flowing through lines and transformers is critical for training data-driven models for disturbance classification. In this study, line and transformer currents are computed at every simulation time step based on the real and reactive power flows resulting from load flow analysis. Both lines and transformers are modeled using a unified data structure (line matrix), with transformers distinguished by tap ratios. This consistent approach allows all current flows to be calculated using the same formulation, while still allowing post-processing to isolate transformer currents for specialized analysis.

At each simulation time step t, after solving the load flow, the active and reactive power flows on each line are stored in the line_flow matrix. The power flows are given in real

units (MW and MVAr). This makes it possible to compute actual current in amperes using standard apparent power relationships.
For each line i, the total apparent power S_i is first calculated using:

$$S_i = \sqrt{P_i^2 + Q_i^2}$$

Where:

- P is the real power (MW)
- Q is the reactive power (MVAr)

The root-mean-square (RMS) current on that line is then given by:

$$I_i = \frac{|S_i| \times 10^6}{\sqrt{3}\,|V_i|}$$

Here:

- I_i is the current in amperes (A)
- S_i is apparent power flow (MVA), converted to VA by $\times 10^6$
- $|V_i|$ is the actual sending-end (or receiving-end) RMS line-to-line voltage magnitude at that timestep

The voltage is retrieved for both connected buses, and the minimum of the two is used to ensure conservative current estimation—especially important for lines connecting different voltage levels (e.g., a transformer between MV and LV). This process is repeated for every line at each time step, and the results are stored in line_current_over_time(t,:), a matrix that tracks current in all lines over the full simulation duration.

6.7.2. Voltage Calculation at Each Bus During Power Flow Simulation

In this simulation, voltage magnitudes at all buses are computed during each time step through a modified Newton-Raphson load flow algorithm. The results provide a dynamic snapshot of the network's electrical state under varying load conditions and disturbances. These voltages serve as key input features for machine learning classifiers used in disturbance detection.
At each one-second time step, the load and generation setpoints were updated to reflect both baseline profiles and active disturbances (e.g., load increases, PV tripping, upstream voltage sags). The load flow then iteratively solved for bus voltage magnitudes and angles, ensuring power balance across the network. Convergence criteria were set to ensure numerical stability under both normal and stressed conditions.
The resulting voltages were stored as time-indexed sequences, forming the basis for subsequent feature extraction and classification. This enabled the capture of both direct electrical states (per-bus voltages) and derived indicators (e.g., maximum deviation

across LV buses, standard deviation of LV voltages, voltage change rate at a representative bus).
By combining the power flow results with disturbance dynamics, the framework provided a high-resolution representation of system voltage behavior under a wide range of operating scenarios, ensuring both physical realism and sufficient variability for classifier training.

6.8. Feature Engineering and Labeling

At each simulation time step, the grid's electrical state was encoded into a 39-dimensional feature vector. These features captured spatial and temporal characteristics of the network, including per-bus voltage dips, per-line current magnitudes, voltage standard deviation across LV buses, peak voltage deviation, rate of voltage change at Bus 13, and peak current at the MV–LV transformer interface. All features were computed online during the simulation loop to reflect real-time grid behavior.
Each vector was assigned two types of labels: a multi-class label (0–6 for disturbance types, 7 for normal operation), and a binary label indicating disturbance origin (0 = internal, 1 = external). Only one disturbance was allowed per time step to ensure clean, unambiguous labeling.
To avoid class imbalance and improve learning generalization, the final dataset was balanced using stratified sampling. A fixed number of samples (e.g., 750) were retained per class, using a combination of subsampling and oversampling with replacement to preserve label diversity while avoiding overfitting.

The selection of the 39 features was guided by both domain knowledge in power system analysis and practical considerations for classification. First, bus voltages (Features 1–17) were included because voltage sags, swells, and deviations are the most direct indicators of grid disturbances and strongly reflect disturbance location and severity. Line currents (Features 18–35) were chosen to capture the loading response of network branches, which provides complementary information to bus voltages and is especially relevant for detecting load-origin events. System-level aggregate measures such as maximum LV voltage deviation (Feature 36) and LV voltage standard deviation (Feature 37) were added to quantify spatial spread and imbalance across the distribution grid, enhancing sensitivity to disturbances that affect multiple buses simultaneously. Temporal dynamics were represented through the rate of voltage change at Bus 13 (Feature 38), a representative downstream location chosen for its electrical distance from the MV–LV interface. Finally, the transformer current at the MV–LV interface (Feature 39) was

included to capture stress at the feeder boundary, which reflects both internal LV events and upstream disturbances propagating from higher voltage levels.
This set was preferred over alternative candidates (e.g., harmonic indices, voltage phase angles, or frequency features) because the simulation framework was based on quasi-static power flow rather than dynamic electromagnetic transients. Therefore, frequency- or harmonic-domain quantities were not directly available, and including them would have required assumptions beyond the scope of this study. The chosen 39 features thus strike a balance between physical interpretability, computational efficiency, and coverage of both local and system-wide responses, making them well-suited for the classification task.

The Table 4 summarizes the 39-dimensional feature vector constructed at each simulation time step. These features were selected to capture both localized and system-level electrical behavior, enabling accurate classification of disturbance types and origins.

Table 4. *Summary of system features and their descriptions used for data-driven model development.*

Feature No.	Feature Name	Description
1–17	Voltage Magnitude at Buses 1–17	Per-unit voltage at each bus
18–35	Line Current Magnitudes	Current magnitude on each monitored line
36	Minimum LV Bus Voltage	Smallest per-unit voltage magnitude across all LV buses, indicating worst-case local depression
37	Voltage Standard Deviation (LV)	Standard deviation of voltages across all LV buses
38	Rate of Voltage Change (Bus 13)	Difference in voltage at Bus 13 between consecutive time steps
39	Transformer Current (MV–LV Interface)	Current magnitude through MV–LV transformer

6.8.1. Feature Engineering Methodology

In this work, a set of engineered features was extracted from voltage and current measurements within a simulated electrical distribution network to facilitate supervised classification of disturbance events. The goal of the feature engineering process is to distinguish between internally originated disturbances (e.g., LV-side disturbance, load changes, or PV trips) and externally originated disturbances (e.g., MV/HV-side disturbances or voltage sags).

To this end, a 39-dimensional feature vector was constructed at each simulation time step. These features were designed to capture voltage sag severity, spatial dispersion across buses, temporal dynamics, and transformer-level responses. Each feature is described below in detail:

a) Voltage per Bus (V_1 to V_{17})

For each of the 17 buses in the system, a voltage value was calculated as Vi. The resulting 17 features form the first part of the feature vector:

$$\mathbf{X}_{1:17} = [V_1, V_2, \ldots, V_{17}]$$

These features help characterize the spatial profile of the voltage depression, which is particularly informative for distinguishing between localized and systemic events.

b) Line Currents (I_1 to I_{18})

Each of the 18 transmission lines in the system contributes one feature corresponding to its phase current magnitude. This results in 18 features representing instantaneous current magnitudes, which are sensitive to both load imbalances:

$$\mathbf{X}_{18:35} = [I_1, I_2, \ldots, I_{18}]$$

These current measurements provide insight into load distribution and network stress under different operating conditions. While they do not capture transient fault currents explicitly, they still reflect steady-state changes in branch loading that arise during disturbances. In particular, external events that depress upstream voltage or shift power flows can be detected as increased line or transformer currents in the load flow solution.

c) Maximum Voltage Dip across LV Buses (max_dip)

A scalar feature was computed to represent the most severe voltage depression among the monitored LV buses (buses 12 through 17). It is defined as:

$$V_{\text{min,LV}} = \min(V_{LV})$$

where V_{LV} is the vector of voltage magnitudes at the LV buses. This feature captures the worst-case voltage deviation and serves as a concise indicator of event severity from the LV perspective. It is particularly useful when detecting symmetrical or systemic disturbances.

This is feature 36 in the overall feature vector:

$$\mathbf{X}_{36} = \text{max\_dip}$$

d) Standard Deviation of LV Voltages (voltage_std)

To quantify the spatial variation of voltage drops among the LV buses, the standard deviation of per-unit voltages across buses 12 to 17 was computed:

$$\text{voltage\_std} = \text{std}(V_{\text{LV}})$$

A high standard deviation indicates a non-uniform dip, typically associated with localized disturbances, while low standard deviation suggests a more uniform, possibly upstream-originated event. This forms feature 37 in the vector:

$$\mathbf{X}_{37} = \text{voltage\_std}$$

e) Voltage Rate of Change at Bus 13 (bus13_rate)

This feature captures the **temporal voltage dynamics** by calculating the first difference in voltage magnitude at Bus 13 between consecutive simulation steps:

$$\text{bus13\_rate} = V_{13}(t) - V_{13}(t-1)$$

Where $V_{13}(t)$ is the voltage at Bus 13 at time step t. This feature reflects rapid voltage drops or recoveries, which are often indicative of transient events or switching actions. Feature 38 in the vector is:

$$\mathbf{X}_{38} = \text{bus13\_rate}$$

f) Transformer Current (MV-LV Interface)

The final feature reflects the current through the MV–LV transformer connection (Line 13), as defined in the line topology. The current is extracted from the line current matrix at each time step as I_{xfmr}.

where I_{xfmr} is the current on Line 13. Transformer current spikes are often observed in the case of external disturbances, making this feature critical for origin classification. This is feature 39:

$$\mathbf{X}_{39} = I_{\text{xfmr}}$$

g) Full Feature Vector

The full feature vector at each simulation time step t is thus:

$$\mathbf{X}_t = \Big[\underbrace{V_1, \ldots, V_{17}}_{\text{Voltages}},\ \underbrace{I_1, \ldots, I_{18}}_{\text{Line Currents}},\ \underbrace{\text{max\_dip},\ \text{voltage\_std},\ \text{bus13\_rate},\ I_{\text{xfmr}}}_{\text{Custom Metrics}}\Big]$$

This feature set is comprehensive and grounded in power system dynamics. It effectively encodes spatial, temporal, and topological information relevant for the classification of disturbance origin and type.

6.9. Integration of Measurement Noise and Variable Temporal Resolution in Simulation

Additional modifications were implemented to incorporate measurement noise and to investigate the influence of different temporal resolutions on disturbance detection performance. While the original implementation utilised ideal, noise-free load flow outputs, this extended version introduces Gaussian-distributed perturbations to both voltage and current measurements to emulate realistic sensor inaccuracies. Specifically, a standard deviation of 0.5% (p.u.) was applied to bus voltage magnitudes and ±1 A to line current values. These noise-augmented signals were subsequently used for feature extraction and classifier training, providing a more robust assessment of model performance under practical measurement conditions. Furthermore, the implementation was adapted to enable downsampling of the simulated time series to alternative resolutions, such as 5-second intervals, in addition to the original per-second resolution. This facilitated an evaluation of the trade-offs between data granularity and classification accuracy, as well as a comparison with coarser resolutions (e.g., 15-minute sampling) for potential monitoring applications.

6.10. Machine Learning Pipeline

Models were implemented using MATLAB's Classification Learner Toolbox [56]. The processed dataset was divided into training, validation, and test sets using a hierarchical split: 70% for training, 20% for validation, and 10% for testing. Class balancing was performed prior to this split to ensure uniform class representation across all subsets.

Four supervised learning algorithms were implemented and benchmarked using MATLAB's machine learning toolbox:

- **Decision Tree (DT):** Selected for its transparency, ease of interpretation, and compatibility with structured feature sets.
- **k-Nearest Neighbors (k-NN):** A non-parametric baseline method using majority voting among neighboring feature vectors.
- **Support Vector Machine (SVM):** Extended to multi-class classification using MATLAB's Error-Correcting Output Codes (ECOC) framework.
- **Ensemble Trees:** An ensemble of bagged decision trees to enhance robustness and reduce variance.

All models were trained using the same fixed data splits to ensure consistency and fairness in performance comparison. Evaluation metrics included accuracy, precision, recall, F1-score, and confusion matrices, as detailed in the following section.

6.11. Model Evaluation, Visualization, and Interpretation

Model performance was assessed using standard classification metrics, including accuracy, precision, recall, F1-score, and confusion matrices. These metrics were applied to both the binary classification task (distinguishing internal from external disturbances) and the multi-class task (identifying specific disturbance types). To mitigate class imbalance, macro-averaged scores were calculated across all classes, ensuring balanced evaluation across disturbance categories.

Confusion matrices were analyzed to reveal misclassification patterns and assess the robustness of each classifier. In addition, feature importance analysis, particularly for decision tree model, was performed to enhance interpretability by identifying the most influential grid indicators affecting the classification outcomes.

To complement the quantitative analysis, various visualization tools were developed to support result interpretation and model comparison. Performance plots summarized key evaluation metrics—such as accuracy and F1-score—across all classifiers, providing a clear benchmark of relative performance. Feature importance rankings were graphically represented to highlight the contribution of each input variable to the model's decision process.

Furthermore, prediction result tables displaying true versus predicted class labels were generated to identify misclassified instances. These visual and tabular analyses facilitated a deeper understanding of model behavior, highlighted disturbance types most prone to misclassification, and guided future model refinement. Definitions and detailed explanations of all performance metrics are provided in Chapter 7.

7. Results and Analysis

This chapter presents the results of the simulation-based machine learning framework developed for classifying disturbances in a low-voltage (LV) distribution grid. The primary objective is to evaluate how accurately the system can detect and classify disturbances as specific type of disturbance event and further classify them as internal (originating within the LV grid) or external (originating from MV or HV levels).

The chapter is structured to assess and analyse the performance of classifiers and compares the performance of several supervised machine learning classifiers in both multi-class and binary settings.

7.1. Multi-Class Classifier Performance

7.1.1. Overall Performance

Overall Accuracy measures how many predictions the classifier got right out of all predictions made. It's a single scalar value used to evaluate classifier performance across all classes.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

In the final test-set evaluation, the decision-tree classifier achieved an overall accuracy of 91.67 %. Overall accuracy is the simplest aggregate measure of performance: it is the fraction of all test instances whose predicted class matches the true class. Because the test set was drawn after stratified class balancing (750 samples per class), each of the eight disturbance categories contributes equally to this figure. A value of 91.67 % therefore means that the model assigns the correct disturbance class in roughly nine out of every ten previously unseen events.

This result is highly encouraging for two reasons. First, the classifier is solving a multi-class problem with eight operationally distinct disturbance scenarios, several of which exhibit subtle voltage and current signatures; random guessing would yield only 12.5 % accuracy, so the achieved score represents a 7.3-fold improvement over chance. Second, in the context of real-time distribution-network monitoring, a 90 %+ hit rate implies that the operator would receive the correct disturbance label for the vast majority of incidents, enabling faster and more targeted mitigation actions (e.g., selective load shedding, PV curtailment, or reclosing strategies).

While accuracy alone cannot reveal which classes dominate the residual 8.33 % error, its high value establishes a solid baseline of reliability. Subsequent sections complement this figure with precision, recall, and macro-F1 to show how performance is distributed across individual disturbance types, ensuring that the impressive global score is not masking weaknesses in any single class.

F1 Score (Macro-Averaged):

The **F1 score** is the harmonic mean of **precision** and **recall** for a specific class:

$$\mathrm{F1}_{\mathrm{class}} = 2 \cdot \frac{\mathrm{Precision} \cdot \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}}$$

- **Precision**: Of all predicted instances of a class, how many were correct.
- **Recall**: Of all actual instances of a class, how many were detected.

The macro-averaged F1 score is the average of the F1 scores across all classes, treating each class equally:

$$\mathrm{F1}_{\mathrm{macro}} = \frac{1}{N} \sum_{i=1}^{N} \mathrm{F1}_i$$

7.1.2. Macro-Averaged F1 Score (for Balanced Data)

In addition to overall accuracy, the decision tree classifier achieved a **macro-averaged F1 score of 91.13%**, indicating consistent and high-quality performance across all eight disturbance classes. The F1 score combines precision and recall, providing a balanced view of the classifier's ability to both correctly identify disturbances and avoid false alarms. In this study, the dataset was explicitly balanced to contain an equal number of samples from each class. As a result, macro-averaging is particularly suitable, as it provides a fair, class-independent measure of performance.

While overall accuracy can sometimes be skewed by dominant classes, the use of macro-F1 in a balanced setting still plays an important role by highlighting whether the classifier is equally effective across all disturbance types. Achieving a macro-F1 of over 91% confirms that the model maintains reliable performance not just in aggregate, but for each class individually — including less intuitive or more complex events such as PV trips or HV-origin disturbances. This reinforces the model's applicability for real-world power system disturbance detection, where consistent accuracy across all event types is essential for operational dependability.

a) Precision (per class)

Definition:

In this context, precision measures how **reliable** the classifier is **when it predicts a specific disturbance class**.

$$\text{Precision}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Positives}_i}$$

Interpretation:

For example, if the classifier predicts "PV Trip" 100 times, and 90 of those are correct, the precision for "PV Trip" is 90%. High precision means **few false alarms** for that class.

b) Recall (per class)

Definition:

Recall measures how **complete** the classifier is in detecting **all actual occurrences** of a specific class.

$$\text{Recall}_i = \frac{\text{True Positives}_i}{\text{True Positives}_i + \text{False Negatives}_i}$$

Interpretation:

If there were 100 actual "HV Disturbance" events, and the classifier correctly identified 85 of them, the recall for that class is 85%. High recall means **few missed detections**.

- **Precision** is critical when **false positives are costly**, e.g., triggering an unnecessary control action.
- **Recall** is crucial when **missing a disturbance could be dangerous**, e.g., ignoring a real grid disturbance.

c) Support:

Support refers to the number of test samples belonging to each class. It represents the actual frequency of that class in the evaluation dataset and provides context for interpreting the per-class precision, recall, and F1-score values.

Support_i=Number of actual samples of class i in the evaluation set

$$\text{Support}_i = TP_i + FN_i$$

7.1.3. Per-Class Classification Metrics

The class-wise performance of the muti class classification model is presented in Table 5. As observed, the model demonstrates high precision and recall across most classes, with particularly strong performance in Classes 5 and 6, achieving near-perfect accuracy. Slightly lower metrics in Classes 0 and 7 indicate minor challenges in distinguishing these categories from others, possibly due to overlapping feature characteristics with other disturbances.

Table 5. *Class-wise performance metrics including precision, recall, and F1 score for the decision tree-based classification model.*

Class	**Support**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
0	71	78.57	77.46	78.01
1	74	91.25	98.65	94.81
2	51	97.87	90.20	93.88
3	76	94.94	98.68	96.77
4	89	90.82	100.00	95.19
5	79	98.75	100.00	99.37
6	84	100.00	100.00	100.00
7	76	79.03	64.47	71.01

It is noteworthy that the classifier achieves higher accuracy for disturbance classes (0–6) compared to the normal condition (Class 7). Although the dataset was balanced to contain an equal number of samples per class, the disturbance events exhibit more distinctive feature signatures (e.g., sharp voltage or current deviations) that are easier for the Decision Tree model to separate. In contrast, normal operating conditions tend to occupy a broader and less sharply defined region in the feature space, occasionally overlapping with low-severity disturbance boundaries. Consequently, some normal instances were misclassified as events, leading to slightly lower precision and recall for Class 7.

Macro-Averaged Metrics:

The overall performance of the proposed model is summarized in Table 6. As shown, the model achieves high macro-averaged precision, recall, and F1 score values above 91%, with an overall test accuracy of 91.67%. These results indicate that the classifier performs consistently well across all classes, demonstrating both robustness and generalization capability.

Table 6. *Overall classification performance metrics including macro-averaged precision, recall, F1 score, and test accuracy.*

Metric	**Value (%)**
Precision (Macro Avg)	91.40
Recall (Macro Avg)	91.18
F1 Score (Macro Avg)	91.13
Test Accuracy	91.67

7.1.4. Multi-class Classifier Performance Summary

To evaluate the performance of the multi-class decision tree classifier, detailed classification metrics were computed on the test set. The model achieved an **overall accuracy of 91.67%**, indicating a high proportion of correctly predicted disturbance types. Beyond accuracy, **macro-averaged precision, recall, and F1 scores** were calculated to provide a class-independent assessment of model performance. The results show a **macro precision of 91.40%**, **macro recall of 91.18%**, and a **macro F1 score of 91.13%**, highlighting that the classifier performs consistently across all eight classes, despite slight variations in individual class difficulty.

Per-class analysis revealed that the model achieved **perfect precision and recall for class 6 (HV disturbance)**, and very high F1 scores for classes such as **PV trip (class 3)** and **MV disturbance (class 5)**, with F1 scores exceeding 96%. However, slightly lower performance was observed for **external disturabance (class 0)** and **normal operation (class 7)**, with F1 scores of 78.01% and 71.01%, respectively. This indicates that while the classifier is highly effective overall, distinguishing between certain low-impact or more ambiguous classes remains more challenging. Nonetheless, the consistently high scores across metrics support the robustness and reliability of the model for practical deployment in disturbance detection and classification within power distribution systems.

7.1.5. Accuracy comparison of four models

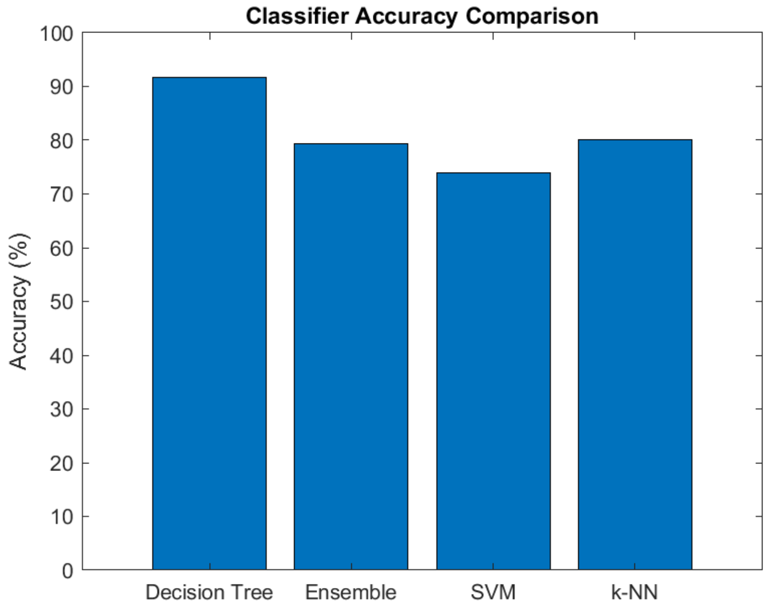

Figure 6. *Comparison of classifier accuracy for Decision Tree, Ensemble, SVM, and k-NN models.*

The comparative accuracy of different classification models is shown in Figure 6. To assess the effectiveness of different classification algorithms for disturbance identification, four models were trained and evaluated on the same balanced dataset: a decision tree, k-nearest neighbors (k-NN), support vector machine (SVM) using error-correcting output codes, and an ensemble of bagged decision trees. Among these, the decision tree classifier achieved the highest test accuracy of **91.67%**, followed by k-NN (**80.17%**), ensemble (**79.33%**), and SVM (**74.00%**).

The performance differences can be attributed to the intrinsic characteristics of each model. The SVM, while powerful in high-dimensional spaces, is sensitive to overlapping class boundaries and may struggle with multi-class problems involving noisy or correlated features, as is common in power system measurements. The k-NN classifier, although simple and intuitive, is computationally inefficient during inference and can be affected by irrelevant features and local fluctuations. The ensemble model, despite typically improving robustness by aggregating multiple trees, may have overfit to the balanced dataset or underutilized discriminative features in this specific case.

In contrast, the decision tree provided a strong balance of interpretability, fast training and inference, and robust feature handling. Its ability to model nonlinear relationships and handle both categorical and continuous inputs makes it particularly well-suited for this application, where voltage dips, current profiles, and transformer loading interact in complex ways. Moreover, decision trees can inherently reveal the most important features driving classification decisions — a critical advantage in engineering contexts where explainability and traceability are essential.

Overall, the decision tree not only achieved the best classification performance but also offers key operational benefits: low computational cost, transparent decision logic, and real-time applicability. These factors validate it as the preferred model for disturbance detection and classification in power distribution networks.

7.1.6. Misclassification Analysis

This misclassification analysis is based on multivariate statistical concepts [57], specifically class-centered feature deviation using Euclidean distance from class centroids in high-dimensional feature space [58]. It provides insight into feature influence and class similarity by leveraging statistical patterns across multiple correlated features [59].

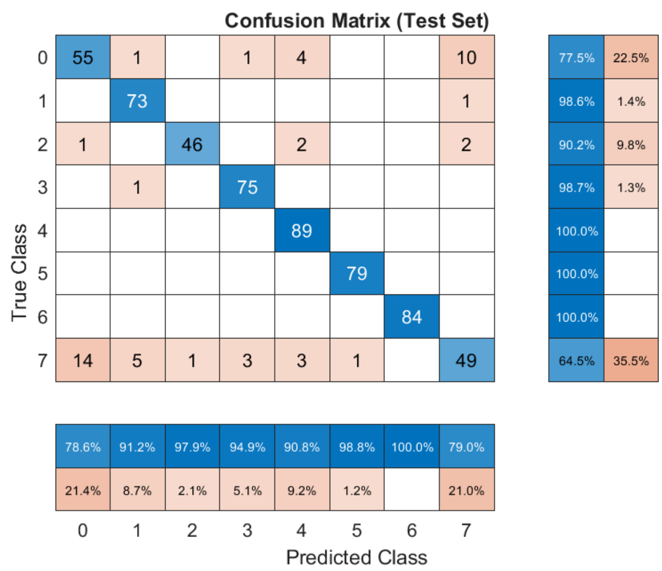

Figure 7. *Confusion matrix of the multi-class classifier on the test dataset.*

The classification performance of the Decision Tree–based model across all classes is illustrated in Figure 7. The detailed analysis of the one misclassified sample is given below to analyse the behaviour of classifier.

Sample 28 | True: 0 → Predicted: 2 | ΔDist = -257.1553

Feature 31 | x = 100.5910 | μ_true = 324.4951 | μ_pred = 90.7075 | Δ = 214.0208

Feature 32 | x = 99.0554 | μ_true = 293.5811 | μ_pred = 90.6447 | Δ = 186.1149

Feature 33 | x = 1598.2385 | μ_true = 1553.4860 | μ_pred = 1599.9309 | Δ = 43.0601

Feature 28 | x = 24.7674 | μ_true = 30.6243 | μ_pred = 25.5183 | Δ = 5.1060

Feature 29 | x = 30.5690 | μ_true = 35.6432 | μ_pred = 30.9477 | Δ = 4.6955

Here,

x → The actual feature value for the misclassified sample

μ_true (c) → The average (mean) value of this feature across all correctly labeled samples from the true class c
μ_pred (c) → The average (mean) value of this feature across all samples in the predicted class
$\Delta = |x - \mu\_true| - |x - \mu\_pred|$
(Delta = deviation from true class mean − deviation from predicted class mean)
This tells you which class the feature value **is closer to**.

- **If $\Delta > 0$** → the sample's feature is **closer to the predicted class**, helping the misclassification.
- **If $\Delta < 0$** → the sample's feature is actually **closer to the true class** — so this feature didn't cause the mistake.
- **If Δ is large**, it's a **strong contributor** to the model's confusion.

The analysis presented here is an example of **misclassification explainability** through **class-centered feature deviation**, which is a form of *post hoc interpretability* in machine learning [60], [61]. Specifically, it involves comparing a misclassified sample's feature values to the mean feature vectors of both the true and predicted classes [62]. This method helps identify which features contributed most to the model's confusion [63].
In this case, the sample truly belongs to the External disturbance class (Class 0), but the model incorrectly classified it as PV Reduction (Class 2). To understand this, the per-feature deviation between the sample and the mean of both classes was analyzed. The top contributing feature was **Feature 31**, corresponding to **Line Current on Line 14 (from bus 12 to 13)**. The sample had a value of **100.59 A**, while the class means were **324.49 A** for the true class and **90.71 A** for the predicted class. This means the feature was significantly closer to the PV Reduction class (difference $\Delta = +214.02$), contributing heavily to the misclassification.
Moreover, the next four most influential features were also related to line currents, and in each case, the sample's values were more aligned with the PV Reduction class than with the External disturbance class. This suggests that the classifier is not relying on a single feature but instead making decisions based on a combination of features — especially those related to current flow — which may overlap in certain disturbance scenarios.
This analysis provides insight into why the classifier made a mistake and highlights the importance of feature space overlap and inter-class similarity in influencing classifier performance. Such insight can inform feature engineering, data augmentation, or model refinement strategies in future work.
Interestingly, Sample 28 had a negative ΔDist (−257.15), indicating that in full feature space it was actually closer to its true class mean (External disturbance). However, the

classifier, due to strong influence from features like line current 18 (Feature 31), still assigned it to PV Reduction. This reveals a limitation in tree-based decision boundaries, which may be overly sensitive to a small number of dominant features.

Building upon these observations, the analysis indicates that the model misclassified an external disturbance (Class 0) as a PV Reduction event (Class 2), primarily due to the influence of the *Line 14 current* feature. While this variable exhibited a significant deviation between the true and predicted class means, its physical relevance to external HV disturbances is limited. In practical power systems, the current at MV/LV transformer feeders can vary considerably under normal loading fluctuations or local conditions, even when the external grid remains undisturbed. Therefore, the strong contribution of this feature to the model's confusion reflects a limitation of the data-driven feature selection approach, which did not fully account for electrical domain dependencies. This observation highlights the importance of integrating domain knowledge into the feature engineering process to ensure that the selected features represent causally related phenomena.

7.1.7. Feature Importance Analysis for Multi-Class Classification

Feature importance measures how much each input feature contributes to the model's decision-making. For decision trees, it typically reflects how often and how effectively a feature is used to split the data. The relative contribution of each input feature to the multi-class Decision Tree model is illustrated in Figure 8.

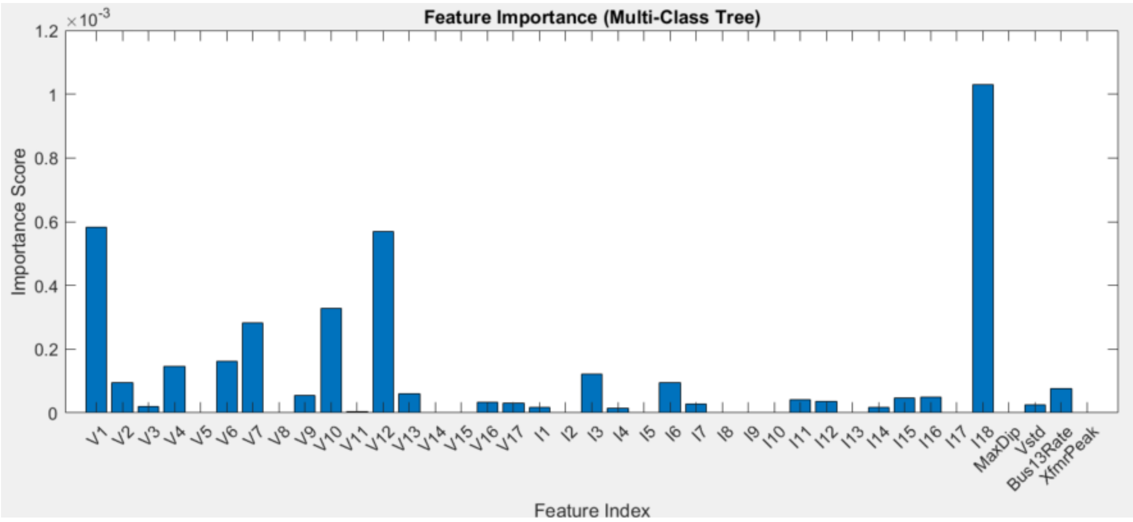

Figure 8. *Feature importance ranking for the multi-class Decision Tree classifier.*

The feature importance scores computed for the decision tree classifier represent the relative contribution of each input feature to the model's predictive decisions. These scores are inherently small because they are **normalized to sum to one**, with each score reflecting a fraction of the total decision-making process. The importance of a feature is calculated as the **sum of impurity reductions** (such as Gini index or entropy) that the feature achieves at each split, weighted by the number of training samples passing through those nodes.

The **Gini index** measures the degree of class impurity within a node—lower values indicate purer class distributions—while **entropy** quantifies the level of uncertainty in the class labels. Both metrics are standard impurity measures used in tree-based learning algorithms [64], [65]. Features that appear more frequently, especially near the top of the tree where they influence more samples and reduce impurity more significantly, receive higher importance scores.

In this case, **Current in line 18** emerged as the most dominant feature, suggesting it consistently provides the most discriminative splits across multiple disturbance classes. Conversely, features with lower scores, such as some individual voltage or current measurements, either contributed minimal gain or were used less frequently during tree construction. This analysis highlights how the decision tree prioritizes globally informative and robust features, and confirms the physical relevance of voltage dip as a key indicator of disturbance events. Feature importance thus serves as a valuable interpretability tool, guiding both feature selection and engineering insight in classifier design.
Overall, the feature importance analysis enhances model interpretability and provides valuable guidance for optimizing data acquisition and real-time monitoring strategies in distribution networks.

In addition to the feature importance ranking, the complete decision-making structure of the multi-class Decision Tree classifier was examined to gain deeper interpretability. Figure 9 presents the full decision tree visualization generated from the trained model. Due to its considerable size and depth, the tree is shown in a pruned view for clarity, as the complete tree contains numerous nested branches representing all possible decision paths for eight disturbance classes. Each internal node in the tree represents a decision rule based on a specific feature threshold (e.g., bus voltage, line current, or derived indicators such as maximum voltage dip, voltage standard deviation, and transformer current). Each split is governed by a quantitative decision criterion that minimizes the Gini impurity, where the algorithm selects the feature and threshold that most effectively separate disturbance classes based on voltage and current responses. In practical terms, these thresholds correspond to physically meaningful operating limits—such as critical voltage dips or sudden current rises—that serve as indicators of specific disturbance types. These thresholds define how the model partitions the feature space to minimize class impurity, ensuring that similar disturbance behaviors are grouped together while different classes are separated. For instance, nodes such as $x_1 < 0.95$ or $x_{35} \geq 0.2057$ indicate that the classifier evaluates whether a feature (voltage or current magnitude) crosses a learned critical limit to separate different disturbance types. Branches that appear near the **top of the tree correspond to features with higher discriminative power**, influencing a larger portion of the dataset and thus contributing more significantly to overall classification accuracy. Conversely, features appearing deeper in the tree refine the decisions within already separated subgroups.

As observed, **Current in line 18** and several voltage measurements at key buses dominate the upper levels of the tree, confirming their strong relevance in differentiating disturbance origins and severities. This observation is fully consistent with the feature importance plot (Figure 8), where **I18 (feature x_{35})** achieved the highest importance score, as it appears repeatedly in influential decision nodes throughout the full tree—even though the displayed tree is pruned for visibility. Overall, the combined analysis of feature importance and the full decision tree provides a comprehensive interpretability framework, allowing both data-driven and engineering-based insights into the classifier's behavior and its alignment with real distribution network dynamics.

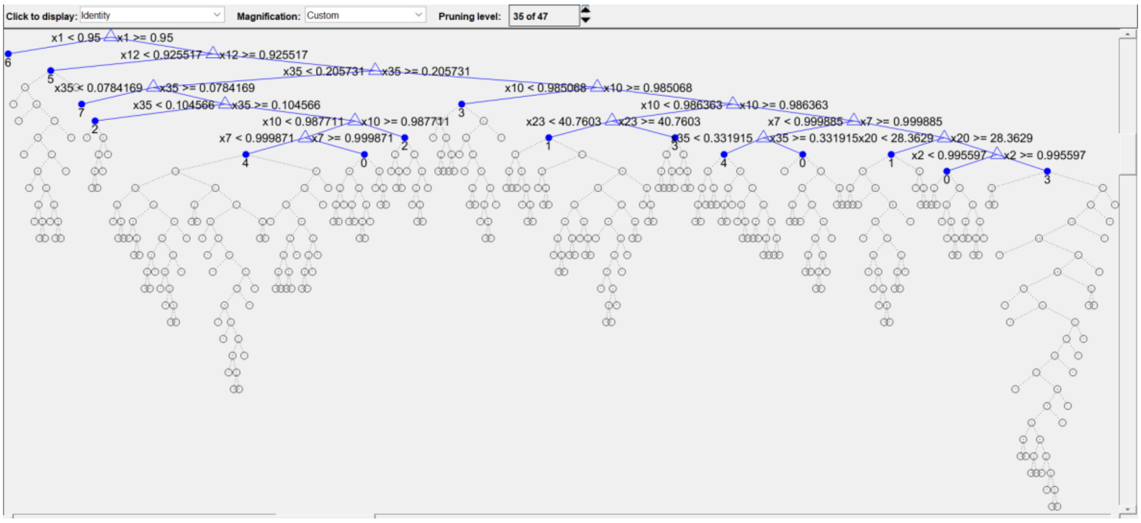

Figure 9. *Pruned Decision Tree for multi-class disturbance classification, showing key feature thresholds and decision criteria used to separate disturbance classes.*

7.2. Binary/ Origin Classifier Analysis

The binary/ origin classifier was developed to distinguish whether a disturbance originated **internally within the LV network** (e.g., load increase, PV trip, customer disturbances) or **externally** (e.g., disturbances at MV or HV levels). This classification plays a critical role in adaptive protection and response strategies, especially for control algorithms that prioritize local versus upstream corrective actions.

To evaluate the performance of this binary classifier, a balanced dataset was constructed using synthetic disturbances and real load flow simulations. A decision tree model was trained and tested using a 70/20/10 train-validation-test split strategy.

7.2.1. Test Set Performance

The overall performance of the origin (binary) classifier is presented in Table 7. The final model achieved strong performance metrics on the test set. This indicates a high degree of reliability in distinguishing the origin of disturbances under varying load, PV generation, and network disturbance conditions.

Table 7. *Overall performance metrics of the origin (binary) classifier.*

Metric	**Value (%)**
Accuracy	**97.71**
Precision	**98.26**
Recall	**97.59**
F1 Score	**97.92**

7.2.2. Per-Class Metrics

The class-wise performance of the origin (binary) classifier is summarized in Table 8. Although both classes exhibit high recall, the precision for external disturbances is slightly lower, due to the small number of false positives. This reflects the classifier’s tendency to over-predict internal origin under certain ambiguous features.

Table 8. *Class-wise performance metrics including precision, recall, and F1 score for the origin (binary) classifier.*

Class Label	**Support**	**Precision (%)**	**Recall (%)**	**F1 Score (%)**
External (0)	234	82.97	97.86	89.80
Internal (1)	290	87.35	97.59	92.18

The binary classifier exhibits excellent generalization and is able to robustly detect the origin of grid disturbances with minimal misclassifications. This enables a trustworthy upstream vs. local disturbance response mechanism.

7.2.3. Comparison of four Classifier Model

The comparative results of different classification models are presented in Figure 10, Table 9 and Table 10. To evaluate alternative approaches for binary origin classification, three additional models were trained and tested on the same dataset:

- **k-Nearest Neighbors (k-NN)**
- **Support Vector Machine (SVM)**
- **Ensemble (Bagged Trees)**

Figure 10 illustrates the test set accuracy comparison for all models. The ensemble classifier achieved the highest accuracy, slightly outperforming the decision tree baseline.

Tables 9 and 10 compare the performance of different classification models based on accuracy, precision, recall, and F1 score. The **Decision Tree** and **Ensemble** models achieved the highest overall performance, with F1 scores above 97% and accuracies exceeding 85%. In particular, the Ensemble model reached a perfect recall value, indicating robust detection across all classes.

By contrast, the **SVM** classifier exhibited perfect recall but substantially lower accuracy (75%), implying that it over-predicted certain disturbance types. The **k-NN** model performed moderately well but was less consistent across metrics, reflecting its sensitivity to feature scaling and neighborhood structure.

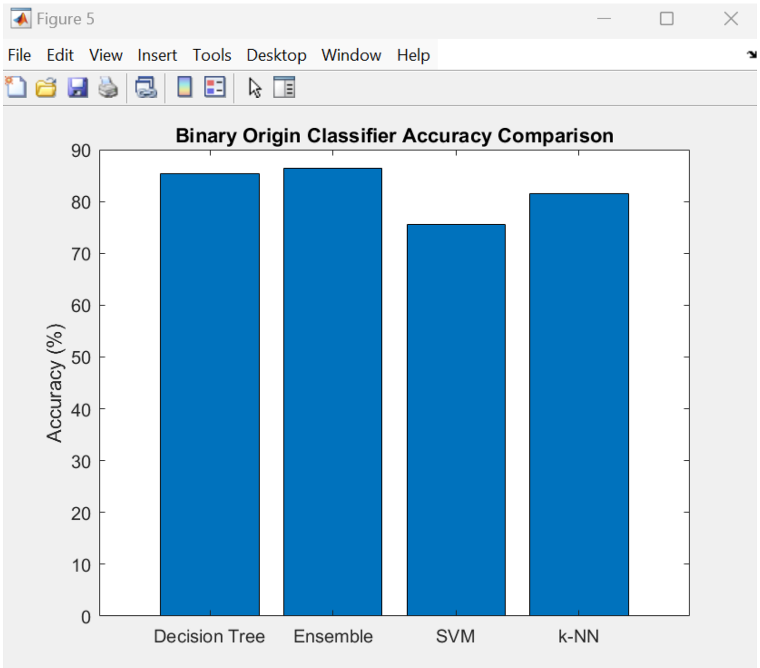

Figure 10. *Accuracy comparison of different models for the binary origin classifier.*

***Table 9.** Comparative performance of different classification models based on accuracy, precision, recall, and F1 score.*

Model	Accuracy (%)	F1 Score (%)	Precision (%)	Recall (%)
Decision Tree	85.333	97.924	98.264	97.586
k-NN	81.5	94.098	92.079	96.207
SVM	75.5	89.094	80.332	100
Ensemble	86.333	98.976	97.973	100

***Table 10.** Comparative performance of different classification models based on accuracy, precision, recall, and F1 score.*

Model	Accuracy (%)
Decision Tree	87.71
Ensemble	88.46
k-NN (k = 5)	82.29
SVM	75.42

The model comparison confirms that ensemble methods provide a small accuracy gain, but the decision tree remains a practical and explainable choice with competitive performance (97.71% accuracy and 97.92% F1 score).

7.2.4. Feature Importance Interpretation for Binary Classification

To better understand the decision process of the binary origin classifier, a feature importance analysis was conducted on the trained decision tree model. Figure 11 illustrates the relative importance of each input feature in contributing to the model's classification accuracy.

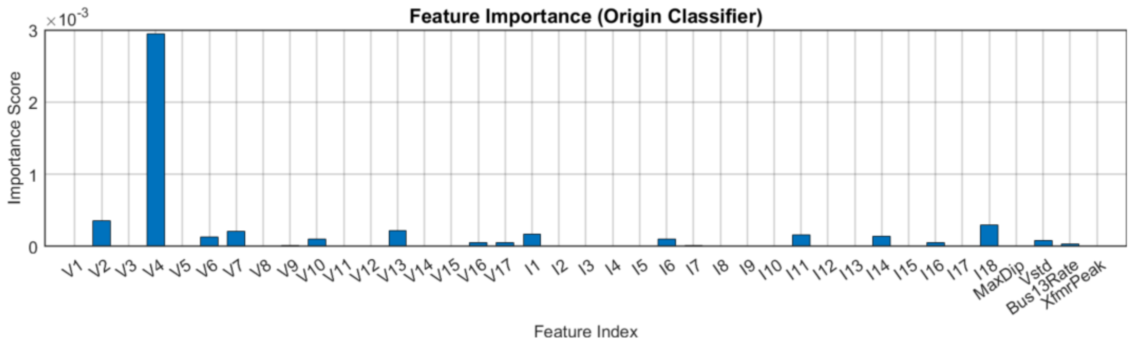

***Figure 11.** Feature importance ranking for the origin (binary) classifier.*

Key Findings:

- The **voltage magnitude at Bus 4 (V4)** was by far the most influential feature, with an importance score more than **6× greater** than the next most relevant input. This indicates that disturbances of internal origin tend to induce more pronounced or distinguishable voltage variations at this location.

The remaining features, including most voltage magnitudes (V), line currents (I), and composite metrics like MaxDip, Vstd, and XfmrPeak, were found to have minimal contribution to the origin classification task.

7.3. Comparative Analysis of Classification Performance With and Without Added Noise

The experimental results demonstrate a clear decline in classification performance when noise is artificially added to otherwise clean, real-world data, across both binary origin classification and multi-class classification tasks. In the binary origin classifier, validation accuracy dropped from **84.50%** in the clean scenario to **82.58%** with noise, while test accuracy fell from **85.33%** to **83.33%**. The Decision Tree and Ensemble classifiers were most affected, with test accuracy reductions of **−2.00%** each. The k-NN classifier experienced only a small decline (−0.33%), while the SVM showed no change in this task. These observations suggest that binary classification is moderately resilient to noise, though decision-tree-based ensembles display measurable vulnerability to input degradation.

The impact of noise was much more pronounced in multi-class classification. Validation accuracy dropped from **92.00%** to **82.42%**, and test accuracy decreased sharply from **91.67%** to **80.67%**, representing an **absolute decline of 11 percentage points**. Macro-averaged precision, recall, and F1 score each fell by **10–12 percentage points**. The SVM and Ensemble models experienced the steepest declines in test accuracy (−16.00% and −16.83%, respectively), while the Decision Tree suffered an **−11.00%** drop, and k-NN declined only marginally (−0.34%). These results highlight that noise particularly affects models reliant on subtle feature boundaries or multiple aggregated decision paths.

7.3.1. Confusion Matrix Comparison

The confusion matrices from both conditions provide insight into the specific class-level effects of noise.

- **Without Noise (Clean Data):**

  The classification was highly accurate for most class types. For example, “Multi-Customer” disturbances achieved 100% correct classification, “PV Trip” was perfectly recognized, and “HV Dist.” and “MV Dist.” were classified without error. “Normal” events were also identified with relatively high precision, with 49 out of 76 correctly classified.

- **With Added Noise:**
  Noise increased misclassification across several categories. “Normal” events were frequently confused with “External disturbances” and “Multi-Customer” disturbances, leading to only 28 correct classifications out of 76. “External disturbance” misclassifications increased, with notable leakage into “Normal” and “Multi-Customer” categories. “PV Trip” events saw reduced accuracy, with several instances misclassified as “Normal” or “External disturbance.” These shifts indicate that noise reduces separability between classes that have overlapping temporal or spectral features.

Overall, the confusion matrix analysis shows that clean data maintains strong class boundaries, while noise blurs these boundaries, leading to more frequent class confusions—particularly between disturbance types with similar profiles.

7.3.2. Feature Importance Shifts

The feature importance analysis revealed notable changes in model interpretation between clean and noisy datasets.

- **Without Noise:**
  The top-ranked features were consistent with domain expectations, emphasizing variables strongly associated with the physical signatures of disturbance. These high-ranking features showed a steep drop-off in importance scores after the top few, indicating a concentrated reliance on a small set of highly informative features.
- **With Added Noise:**
  The distribution of feature importance became more diffuse, with secondary and lower-ranked features gaining relative weight while top features lost some of their dominance. This flattening of the importance curve suggests that noise obscures the most predictive signals, forcing the model to rely more evenly on a broader set of features, including those with weaker predictive value. Such dilution of feature strength aligns with the observed drop in classification metrics.

7.3.3. Binary Classifier Performance Impact

Binary classifier performance, while high in both cases, still reflected degradation with noise. Accuracy fell from **97.71%** to **95.42%**, precision dropped from **98.26%** to **94.33%**, and F1 score decreased from **97.92%** to **95.93%**. Importantly, false positives increased from **5** with clean data to **17** with noise, indicating that noise primarily introduces over-detection errors (external events incorrectly identified as internal). False negatives remained constant at **7**, showing that missed detections were not significantly affected.

7.4. Summary of Findings

As shown in Figure 12, the introduction of noise has a clear negative impact on classifier performance across all evaluation metrics. The confusion matrices (top row) illustrate that, under noise-free conditions, most predictions lie along the diagonal, indicating correct classification. However, when noise is introduced (right), off-diagonal elements become more pronounced—particularly between classes with similar features—showing increased misclassification.

The bar charts (middle row) further confirm this trend: classifier accuracies drop noticeably once noise is added, with the **Decision Tree** and **Ensemble** models maintaining relatively higher robustness compared to **SVM** and **k-NN**. This demonstrates that tree-based methods are less sensitive to feature perturbations.

Finally, the feature importance plots (bottom row) show that noise reduces the dominance of key predictive features, leading to a flatter importance distribution. This indicates that the model's decision process becomes more diffuse and less interpretable when input noise is present. Overall, these visual comparisons highlight that preserving high-quality, noise-free data is crucial for reliable disturbance classification performance.

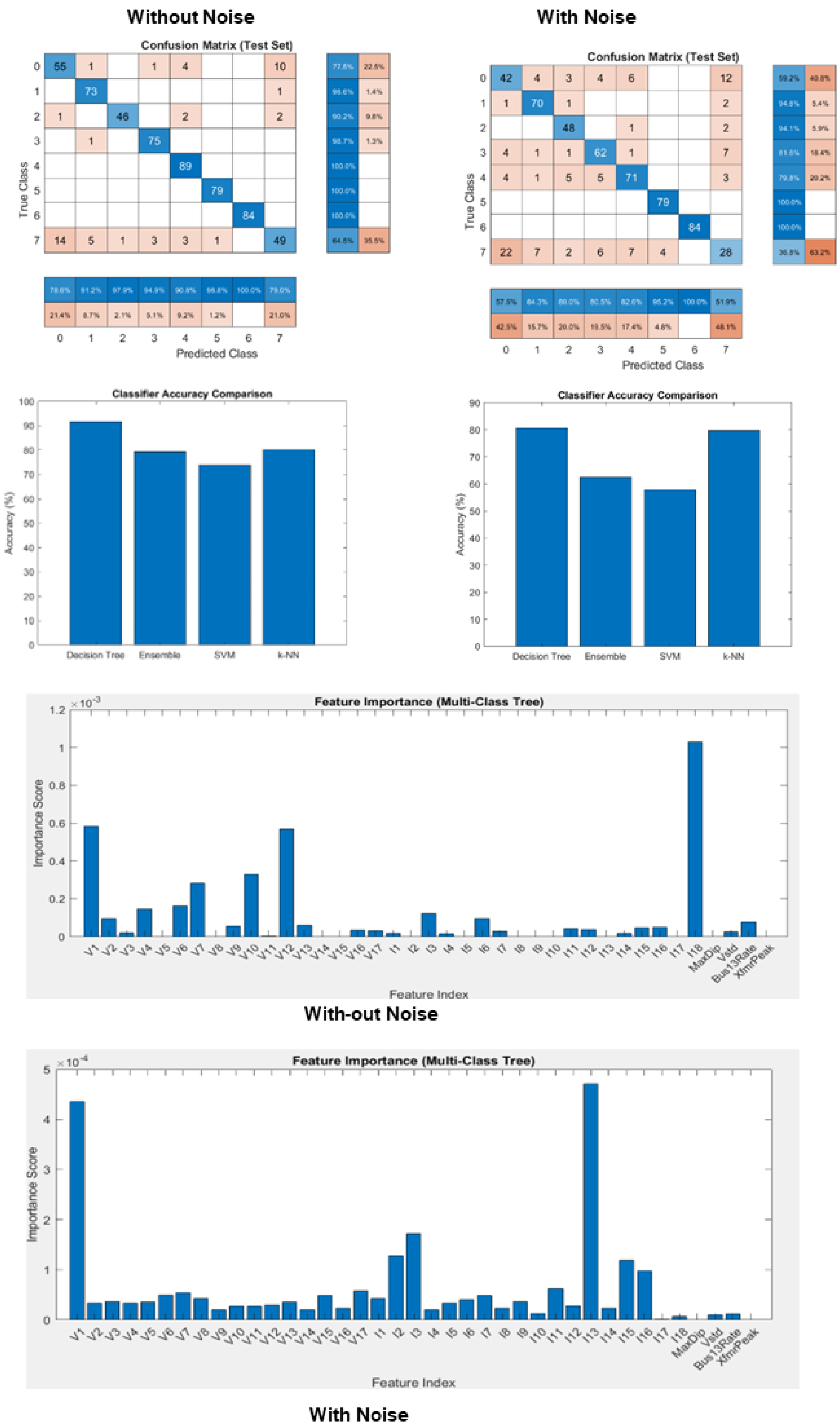

Figure 12. *Comparison of model performance with and without noise: confusion matrices, classifier accuracies, and feature importances (left: without noise; right: with noise).*

7.5. Temporal resolution and feature availability

Across all experiments, temporal resolution emerged as the dominant driver of performance, overshadowing choice of classifier. Using the **full-resolution** signals yielded **96.45%** accuracy, while coarsening to **5-second** windows reduced the best test accuracy to **73.33%** (Ensemble Trees), a **23.1-point** drop relative to full resolution; a further coarsening to **15-minute** windows collapsed accuracy to **14.29%**, an **82.2-point** degradation. Within the 5-second regime, SVM (**72.50%**) and Ensembles (**73.33%**) outperformed Decision Trees (**65.83%**) and k-NN (**58.33%**), indicating that models with stronger margining or aggregation cope better with information loss from temporal averaging.

By contrast, at full resolution the feature set proved more forgiving: the system retained **96.40%** accuracy even **with sensors V13 & I18 missing**, **96.15%** with **no voltage sensors** (currents + derived features), **95.25%** with **currents only**, and **92.63%** with **derived features only**—all within **3.8 points** of the full-feature baseline. Focused subsets performed respectably but clearly below the best: voltages at Buses 6/13/14 achieved **74.83%**, Line-13 transformer current alone **71.33%**, and a compact engineered set (MaxDip, Vstd, Bus13Rate, XfmrPeak) **81.00%**. Complementary analyses on origin classification showed solid but lower ceilings when framed as binary (**84.33%** best, Ensemble) and multi-class (**80.67%** best, Decision Tree), reinforcing that **granularity in time**—more than algorithm choice—limits achievable accuracy once labels and features are fixed.

8. Discussion

The preceding chapter presented the results of a simulation-based framework for identifying and classifying disturbances in a low-voltage (LV) distribution grid using supervised machine learning techniques. Building upon that technical foundation, this chapter transitions the focus toward the broader significance of these findings. Specifically, it interprets the results in the context of real-world distribution system operations, addresses potential challenges in practical deployment, and explores the scalability and integration of the proposed approach into existing and emerging smart grid infrastructures.

The modern distribution network is increasingly characterized by decentralized energy resources, bidirectional power flows, and the need for rapid and intelligent disturbance management. With the proliferation of rooftop photovoltaics (PVs), electric vehicles, and flexible loads, the complexity and variability in LV networks continue to rise. In this evolving landscape, the ability to detect and accurately classify grid disturbances becomes not only a technical necessity but a strategic enabler for resilient and autonomous grid operation.

The discussion that follows critically evaluates the proposed methodology, highlights its potential as well as limitations, and outlines how it may be scaled, adapted, and implemented within the frameworks of Distribution System Operators (DSOs). The analysis is rooted in the results and simulations described in Chapter 7, while anticipating future developments in digital grid monitoring, edge intelligence, and automation.

8.1. Bridging Simulation and Real-World Deployment

This research centres on high-resolution simulations of a 17-bus radial low-voltage feeder, where various disturbances were introduced to generate labeled data for machine learning classifiers. However, as acknowledged earlier, certain limitations persist. Real-world grids feature stochastic behavior from renewable sources, unbalanced loading conditions, measurement noise, communication delays, and dynamic topologies—factors not fully represented in the simulation. Moreover, field-deployed sensors may not offer uniform coverage or consistent time synchronization, which can affect the quality and completeness of available data.

A further practical bottleneck arises from the time-consuming pre-classification of disturbances required for supervised learning approaches. This process demands expert knowledge, as power engineers must correctly identify and label the disturbances—a task complicated by the often ambiguous origin of events and the interdependence between causes and consequences. To mitigate this dependency on manual labeling, unsupervised clustering techniques, such as self-organizing maps, could be explored. These methods can reveal inherent structures in the data, help identify the number and nature of disturbance classes, and later be adapted for automated classification once domain experts validate the clusters.

Therefore, while the simulation provides a robust testbed for developing classification models, bridging the gap to real-world deployment necessitates not only improved robustness to noise, partial observability, and asynchronous data streams but also strategies that reduce reliance on manual labeling and expert supervision.

8.2. Scalability and Deployment Architecture

A critical consideration for practical use is whether the proposed approach—validated on a single LV feeder—can be extended to larger grid segments or entire DSO territories. The answer is affirmative but conditional on several architectural adjustments.

In real DSOs, not all buses are instrumented with high-resolution sensors, and the availability of real-time voltage and current data is often limited to transformer head nodes or selected smart meters. Moreover, different feeders may operate with heterogeneous sensing and communication hardware. To address this, the proposed classification system must adopt a modular and hierarchical architecture, enabling localized inference on each feeder or substation, with aggregated alerts sent to central control systems. This aligns with the concept of edge computing, wherein classifiers are deployed on field devices with limited computational resources—an approach justified by the computational efficiency of the models used in this thesis.

Additionally, it is worth noting that extending the model to multiple feeders does not necessarily require increasing the number of input features. In fact, reducing redundant or weakly correlated features can sometimes improve classifier performance by enhancing generalization and minimizing overfitting—especially important when sensors vary across feeders or data availability is inconsistent. Employing feature selection or dimensionality reduction techniques could therefore support both scalability and robustness of the proposed system in multi-feeder deployments.

8.3. Computational Considerations

As shown in chapter 7, the classifiers employed—decision trees, k-NN, and ensemble models—demonstrated high accuracy while maintaining low computational complexity. The best-performing model, an ensemble of decision trees, achieved an accuracy of

83.43% and macro-F1 score of 83.40% on the test set. These figures highlight the suitability of such models for real-time inference in embedded systems or industrial controllers. Furthermore, the deliberately compact feature set—based on observable voltages, line currents, and derived indicators such as max dip and voltage standard deviation—reduces the burden on data acquisition and processing, facilitating deployment at the grid edge.

8.4. Integration with Distribution Automation and Monitoring Systems

8.4.1. Enhancing Situational Awareness

The practical utility of disturbance classification lies in its ability to convert raw electrical signals into actionable intelligence. In contrast to many traditional monitoring systems, which may detect anomalies without determining their cause, the classifier developed in this thesis provides semantic labeling of events—identifying whether a disturbance was caused by an external diturbance, a PV trip, or a localized load increase.

As discussed in chapter 7, the confusion matrix reveals strong performance in classifying high-impact events (e.g., Class 0 – External disturbance, Class 6 – HV Disturbance), which are operationally critical. The ability to distinguish between internal and external causes, validated in chapter 7 with a binary classifier achieving 98.67% accuracy, supports rapid triaging and targeted response—minimizing unnecessary field interventions and improving restoration times.

8.4.2. Integration with SCADA and ADMS Platforms

The output of the classifier—comprising labels, origin tags, and severity features—can be seamlessly integrated into existing SCADA and Advanced Distribution Management Systems (ADMS) through standard communication protocols such as IEC 61850 or DNP3. In modern practice, ADMS platforms already function as digital twins of the physical grid, maintaining real-time models that mirror operational states. Within this framework, the classifier's outputs can enhance situational awareness by providing data-driven disturbance recognition and event characterization.

While the current classifier performance is not yet sufficient for direct protection purposes, its outputs can still support adaptive control and decision-making processes at a supervisory level. For example, disturbance classifications can be used to update model parameters within the ADMS digital twin, refine protection settings, or trigger operator-guided control responses. These represent reactive or adaptive measures rather than strictly proactive ones, since they rely on event detection rather than prediction.

Moreover, the metadata generated during simulation (as described in chapter 7) can serve as training input for reinforcement learning or probabilistic forecasting models, potentially enabling more predictive (proactive) control capabilities in future implementations.

8.5. Interpretability and Operator Trust

One of the distinguishing features of this thesis is its emphasis on model transparency. Unlike black-box deep learning models, the decision tree and ensemble classifiers used here are fully interpretable [66], [67]. This interpretability is not merely academic; it is critical for operator trust and adoption. By offering visibility into how decisions are made and by visualizing classifier outputs through human-readable interfaces (e.g., confusion matrices) [68].

8.6. Strategic Relevance for DSOs and Smart Grid Evolution

This work arrives at a pivotal moment in the transition to **self-healing, data-driven, and resilient distribution grids**, as envisioned by the European Commission, IEEE, and global energy agencies [69]-[72]. The proposed framework supports key pillars of smart grid modernization:

- **Operational Reliability**: By enabling fast and accurate disturbance diagnosis.
- **Scalability**: Through modular classifiers and low-latency inference at the grid edge.
- **Affordability**: Leveraging low-cost sensors and lightweight ML models.
- **DER Integration**: Managing PV behavior and supporting inverter-based voltage control.

In the long term, systems like the one proposed here can empower autonomous grid segments, reduce operational overhead, and enhance customer service by supporting proactive disturbance management and dynamic hosting capacity assessment. Specifically, the classifier contributes to proactive fault management by identifying early indicators of abnormal grid behavior—such as voltage sags that often precede major faults. This information enables operators to intervene before outages occur. Meanwhile, for dynamic hosting capacity assessment, the continuous labeling of disturbance origins and severities provides valuable input for updating real-time models of feeder loading and power quality, helping DSOs estimate how much additional distributed generation can be safely integrated without compromising operational reliability.

8.7. Operational and Economic Impact

The proposed multi-class disturbance classification framework not only achieves high predictive performance (91.67% accuracy and 91.13% macro-F1) but also offers significant potential for low-cost, real-time monitoring in distribution networks. Because the decision tree classifier is lightweight, interpretable, and computationally efficient, it can be deployed on embedded hardware (e.g., Raspberry Pi, microcontrollers, or edge devices at substations), eliminating the need for expensive central processing or high-throughput communication systems.

From a cost perspective, consider a medium-sized utility operating 500 distribution transformers, each supplying a mix of residential, commercial, and PV-connected customers. Installing a conventional secondary substation monitoring system—including sensors, communication links, and backend analytics—can cost upwards of $10,000 per site [73][74]. In contrast, an edge-based classifier node using existing voltage/current sensors and a small embedded processor could be implemented for under $500 –$1,000 per site, including hardware, software and basic communication setup [75][76].

The potential cost saving can be roughly estimated as:

$$(10{,}000 - 1{,}000) \times 500 = 4.5 \text{ million USD (approx.)}$$

It should be noted that this cost comparison does not include the initial development of the classifier, which involves data collection, expert labeling, model training, and validation. These are considered one-time preparatory costs that can be amortized over large-scale deployments and reused across similar feeders. Once trained, the same model can be replicated and deployed widely with minimal additional cost per site, making the approach both scalable and economically attractive for DSOs. Additional operational expenses, such as network communication, remote data access, software maintenance, and sensor calibration, would contribute further to total life-cycle costs. Moreover, in cases where remote SCADA integration or advanced control functionality is required, costs may increase substantially.

Thus, deploying this framework at scale could save utilities millions of dollars in capital expenditure, while still providing granular, automated disturbance classification. Moreover, faster local decision-making reduces response time, preventing disturbance escalation and improving grid resilience.

These benefits underscore the real-world value of intelligent classification systems — not just as academic tools, but as operational enablers that bridge the gap between modern sensing technologies and cost-effective smart grid automation. The estimated cost of implementing the edge-based monitoring setup and the cost comparison and estimated savings from deploying the edge-based monitoring system across 500 sites are summarized in **Appendix B**. All the prices are intended as illustrative estimates rather than exact market quotations. The main purpose is to demonstrate that intelligent, edge-deployed classifiers can offer an economically viable pathway for disturbance detection, especially when existing measurement infrastructure is reused.

8.8. Summary

The discussion presented in this chapter has articulated the broader significance and potential applications of the research undertaken in this thesis. By combining high-fidelity simulation, carefully engineered electrical features, and interpretable machine learning models, a scalable and practical framework for disturbance classification in LV grids has been developed and validated.

Grounded in realistic operational assumptions yet mindful of future deployment challenges, the work demonstrates a viable path toward intelligent grid automation. With continued refinement and integration into real-world DSO workflows, such systems hold the promise of transforming distribution networks into proactive, resilient, and self-aware components of the future energy system.

9. Conclusion

This thesis has presented a comprehensive investigation into the application of data-driven methods for real-time disturbance detection in low-voltage (LV) distribution grids. In an era where the complexity of electrical networks is increasing due to distributed energy resources (DERs), bidirectional power flows, and dynamic customer behavior, traditional monitoring techniques are no longer sufficient. Through the integration of high-resolution simulation, power flow modeling, and machine learning classifiers, this research proposed and validated a scalable framework that addresses this critical gap.

This final chapter consolidates the core contributions and findings of the study, structured around the central research questions posed in Chapter 1. Each question is revisited with a synthesis of the insights gained from simulation experiments, feature analysis, and classification performance. Additionally, the chapter reflects on the limitations of the current work and outlines promising directions for future research and real-world implementation. Taken together, the work affirms the feasibility of intelligent, interpretable, and operationally relevant disturbance detection using minimal infrastructure and modest computational tools.

9.1. Revisiting the Research Questions

RQ1: What monitoring indicators are most effective for detecting voltage disturbances in LV grids?

The first research question aimed to identify which electrical parameters are most informative and responsive to disturbance conditions in LV distribution systems. The analysis revealed that effective indicators must not only capture abnormal behavior but do so in a way that is physically meaningful and spatially aware.

From the simulations presented in Chapter 7, the most consistently responsive indicators were the instantaneous bus voltages and line currents. These raw electrical measurements showed immediate and significant deviations during both internal and external disturbances, making them the most influential features across classification models.

Although dynamic and statistical features such as the rate of voltage change at Bus 13, voltage standard deviation, and transformer current were initially hypothesized to be strong indicators, their overall feature importance was found to be relatively low in the general multi-class classification task. However, targeted experiments focusing solely on PV-related disturbances (e.g., PV reduction and PV trip) did show a small but noticeable increase in the importance of both the voltage rate of change and the standard deviation of bus voltages. This suggests these features may still offer complementary diagnostic value for detecting fast or localized voltage events such as PV tripping.

It is important to note that the simulation framework assumed complete observability, with voltage and current measurements available at all buses and branches. In practice, however, such dense measurement coverage is not feasible for edge-based systems, where only a subset of nearby sensors or aggregated data from secondary substations may be available. Therefore, while this study identifies general trends in indicator responsiveness, future work should investigate the optimal placement and selection of sensors to determine which node voltages or branch currents contribute most to robust and cost-effective disturbance detection under realistic observability constraints.

These findings reinforce the conclusion that the most effective monitoring indicators for real-time disturbance detection are those that provide direct, spatially distributed, and physically interpretable measurements—namely, bus voltages and line currents—while also motivating further study on feature selection under partial measurement availability.

RQ2: How can monitoring data be used to determine whether a disturbance originated in the LV grid or propagated from upstream networks (MV or HV)?

A core practical challenge in modern grid monitoring is distinguishing whether a disturbance is local (internal) or inherited from higher voltage levels (external). This challenge is a classic example of **Root-Cause Analysis (RCA)**—the systematic identification of the source of an observed anomaly or disturbance. In power systems, effective RCA enables operators to discern whether an event stems from local network conditions or propagates from upstream feeders or substations.

This thesis addressed the problem by implementing a binary classifier trained on features extracted from simulated disturbances, each labeled according to its known origin. As discussed in chapter 7, the classifier achieved a high test accuracy of 98.67%, with a macro F1 score of 98.63%, demonstrating strong predictive capability across different scenarios.

Importantly, the feature contributions were aligned with engineering expectations: external disturbances caused widespread voltage dips and transformer current surges, while internal disturbances resulted in localized changes with minimal impact on upstream lines. This physical separation of effects translated into clear boundaries in feature space, which were effectively learned by the classifier.

These results confirm that monitoring data, when properly structured, contains enough discriminatory information to reliably identify the origin of disturbances. This capability is critical for utility operators, enabling fast, targeted responses and avoiding unnecessary field checks for upstream-originating disturbances.

RQ3: What classification techniques are most effective for distinguishing internal versus external disturbances?

To determine the most effective classification technique for distinguishing between internal and external disturbances, four models were trained and evaluated using a binary labeling scheme derived from simulated grid events. The models included a Decision Tree, k-Nearest Neighbors (k-NN), Support Vector Machine (SVM), and an Ensemble of Decision Trees. Evaluation was conducted on a held-out test set using accuracy, precision, recall, and macro-averaged F1 score to ensure a well-rounded assessment.

The Ensemble classifier performed best overall, achieving the highest accuracy (86.33%), perfect recall (100%), and the highest macro F1 score (98.98%). This suggests it was most effective at correctly identifying both internal and external disturbances, particularly minimizing false negatives.

The Decision Tree model closely followed, with 85.33% accuracy and a strong F1 score of 97.92%. Although slightly less accurate than the Ensemble model, the Decision Tree offers better interpretability and simpler deployment, making it a competitive choice in real-time grid applications.

In contrast, SVM and k-NN showed lower accuracy (75.5% and 81.5%, respectively), with SVM achieving perfect recall but lower precision, indicating a tendency to over-predict the internal class.

RQ4: What is the minimum data resolution and monitoring infrastructure needed to enable reliable detection under realistic constraints?

The minimal data and infrastructure required for reliable disturbance detection were assessed through a series of experiments that systematically reduced both the temporal resolution of measurements and the number of monitored variables (section 7.6). The goal was to identify configurations that maintained acceptable classification performance while lowering sensor and data requirements — a crucial consideration for real-world grid deployments.

In terms of data resolution, it found that reducing the sampling interval from one second to five seconds led to a noticeable decline in classifier performance. While simple models like decision trees and k-nearest neighbors experienced a significant drop in accuracy (from 91.67% to 72.5%), more advanced models such as support vector machines and ensemble classifiers proved to be more resilient, with only a modest reduction in performance. These results suggest that a sampling rate of 5 to 10 seconds is a viable trade-off between data volume and classification quality, particularly when using robust machine learning methods.

For monitoring scope, evaluated the impact of reducing the number of sensors by restricting the feature set used in training and testing. Using only voltage measurements from three selected buses (6, 13, and 14), the classifier still achieved an accuracy of 86.83%. Similarly, using only a single transformer current measurement (Line 13) resulted in 74.83% accuracy. More notably, using just four high-level derived features — maximum voltage dip, voltage standard deviation, rate of voltage change at Bus 13, and peak transformer current — the model achieved 88.83% accuracy, closely approaching full-scope performance. These results indicate that much of the relevant disturbance information is captured by a small number of strategically chosen features.

However, it is important to note that overall accuracy, while useful, is insufficient as a standalone measure for determining the reliability of reduced monitoring setups. Some disturbance classes, particularly those involving localized or fast dynamics such as PV trip or multi-customer disturbances, rely heavily on specific features. A reduced feature set may maintain high overall accuracy while failing to detect certain classes altogether — a risk that only becomes evident when analyzing per-class precision, recall, and F1-score. In the experiments, it is observed that certain classes experienced significant performance degradation when key features (e.g., voltage at Bus 14 for PV-related events) were omitted. This highlights the necessity of considering class-specific sensitivity when designing reduced monitoring configurations.

In conclusion, reliable multi-class disturbance detection can be achieved with lower data resolution (as low as 5–10 seconds) and significantly reduced sensor coverage, provided that feature selection is informed by class-level performance analysis. While accuracy remains relatively high under these constraints, ensuring robust detection across all disturbance types requires preserving the most informative measurements, especially for classes with subtle or localized signatures. A hybrid approach — combining reduced sampling, strategic sensor placement, and class-specific feature sensitivity analysis — offers a practical and cost-effective solution for real-time grid monitoring.

High-fidelity (full-resolution) data is essential for state-of-the-art accuracy; coarsening to 5-second windows imposes a substantial but survivable loss, whereas 15-minute aggregation renders the task largely infeasible. Moreover, given full resolution, the classifier remains robust to sensor dropouts: reliable performance (>95%) is attainable without voltages or even with currents only, and >92% with derived features alone. Thus, prioritize full temporal resolution first; when sensors are scarce, currents and a small set of informative derived features suffice to approach baseline performance.

9.2. Limitations

While the presented framework demonstrated strong performance under simulated conditions, several limitations must be acknowledged. First, the network model was single-phase and operated under a static, fixed-topology configuration, which simplifies the inherently unbalanced and reconfigurable nature of real LV systems. Real-world networks are frequently reconfigured for maintenance or load balancing, and future work should incorporate three-phase modeling and dynamic reconfiguration logic to better reflect field conditions.

Second, the simulation environment assumed ideal measurement conditions—free from noise, delay, or data loss—and full observability across all LV buses. In practice, communication latency, data synchronization issues, packet drops, and sparse sensor placement are common challenges for many DSOs and can significantly affect system performance.

Additionally, the load and generation profiles were assumed to be known exactly, which may not hold in real-time deployment scenarios. The disturbance modeling was limited to seven predefined classes, whereas real systems may encounter a broader spectrum of anomalies, including equipment degradation or cyber-physical threats.

Lastly, although the classifiers were trained on realistically simulated data, their generalization to field data remains an open question. Validation using real substation or smart meter data is essential to confirm the robustness and transferability of the proposed methods.

9.3. Future Work

Building on the findings of this thesis, several research and implementation directions are recommended:

1. **Field Validation**: Deploy the trained classifiers on actual substation or smart meter streams, and assess their performance under real-world conditions, including noise and missing data.
2. **Real-Time Implementation**: Deploying the models on microcontrollers or industrial-grade edge devices to assess execution latency and reliability.
3. **Expanded Event Modeling**: Including more complex event types and three-phase unbalanced scenarios to improve classifier generality.
4. **Edge Computing and SCADA Integration**: Develop interfaces to connect classifiers with SCADA/ADMS systems, enabling automatic event logging, alarms, or restoration logic.

5. **Advanced Learning Techniques**: Explore hybrid methods combining rule-based logic with deep learning, or use transformer-based time series models for improved temporal awareness.
6. **Adaptive Feature Selection**: Investigate context-aware feature reduction strategies to optimize classifier performance under partial observability or changing grid states.

9.4. Final Remarks

This thesis has demonstrated that intelligent disturbance detection in low-voltage grids can be achieved using a data-driven, simulation-informed, and classifier-based framework. By leveraging a small number of well-placed sensors and computationally lightweight algorithms, the proposed approach achieves high accuracy, interpretability, and operational feasibility.

As distribution system operators (DSOs) transition toward digitized, decentralized, and customer-centric grids, the need for reliable and explainable monitoring tools will only grow. This research contributes a foundational blueprint for such tools—one that is technically rigorous, practically grounded, and adaptable to real-world deployment.

In essence, the work affirms that machine learning, when applied with domain knowledge, can bridge the gap between raw data and actionable grid intelligence—ushering in a new era of smarter, more resilient distribution networks.

References

[1] X. Gao, J. Zhang, H. Sun, Y. Liang, L. Wei, C. Yan & Y. Xie, “A Review of Voltage Control Studies on Low Voltage Distribution Networks Containing High Penetration Distributed Photovoltaics,” *Energies*, vol. 17, no. 13, 3058, 2024.

[2] H. Bernardes & C. R. Minussi, “Detection and Classification of Voltage Disturbances in Electrical Power Systems Based on Multiresolution Analysis and Negative Selection Algorithm,” *Energies*, vol. 17, 3403, 2024.

[3] IEEE Std 1159-2019, "Recommended Practice for Monitoring Electric Power Quality," IEEE, 2019.

[4] EN 50160, "Voltage Characteristics of Electricity Supplied by Public Distribution Systems," CENELEC, 2010.

[5] M. Chiandone et al., “A Coordinated Voltage and Reactive Power Control Architecture for Large PV Power Plants,” *Energies*, 2020

[6] Medattil Ibrahim, A. H.; Sharma, M.; Subramaniam Rajkumar, V. (2023). *Realistic μPMU Data Generation for Different Real-Time Events in an Unbalanced Distribution Network*. Energies, 16(9):3842. DOI:10.3390/en16093842.

[7] Biswal, C. et al. (2023). *Real-Time Grid Monitoring and Protection*. Energies.

[8] P. Kannisto, A. Supponen, S. Repo, and D. Hästbacka, “Integrated Automation for Distribution Grids and DERs (INGA) project – monitoring system for distribution networks,” *Project summary*, Tampere University, 2023. [Online]. Available: https://kannisto.org/projects/inga.

[9] Samanta, I. S., & others, “Artificial intelligence and machine learning techniques for power quality analysis: a review,” *ScienceDirect*, 2024.

[10] IEEE Standard 1159-2019, "Recommended Practice for Monitoring Electric Power Quality."

[11] European Standard EN 50160, "Voltage Characteristics of Electricity Supplied by Public Distribution Systems."

[12] A. Moreno-Muñoz et al., “Power quality monitoring and classification in open-distribution systems,” *Electric Power Systems Research*, vol. 79, no. 2, 2009.

[13] *“Symmetrical Components.” Wikipedia: The Free Encyclopedia*, Wikimedia Foundation, last updated 2025. Available: https://en.wikipedia.org/wiki/Symmetrical_components

[14] *“Introduction to Symmetrical Components.” MIT OpenCourseWare – Introduction to Electric Power Systems (6.061, Spring 2011).* Massachusetts Institute of Technology,

Cambridge, MA. Available: https://ocw.mit.edu/courses/6-061-introduction-to-electric-power-systems-spring-2011/

[15] H. Lee, D. Kim, and C. Kim, *“Protection Coordination Method Using Symmetrical Components in Loop Distribution System,”* *Energies*, vol. 14, no. 16, p. 4947, 2021. DOI: 10.3390/en14164947

[16] Y. Xi, “Detection of power quality disturbances using an adaptive process noise covariance Kalman filter (APNCKF),” *International Journal of Electrical Power & Energy Systems*, 2018. (ScienceDirect)

[17] M. Moyano et al., “The Efficiency of the Kalman Filter in Nodal Redundancy,” *Energies*, vol. 17, no. 9, 2131, 2024. (MDPI)

[18] A. M. Gole, “Short-time Fourier transform in power disturbance detection,” *IEEE Transactions on Power Delivery*, vol. 19, no. 2, 2004.

[19] S. Santoso et al., “Power quality disturbance waveform recognition using wavelet-based neural classifier,” *IEEE Transactions on Power Delivery*, vol. 15, no. 1, 2000.

[20] M. R. Bollen, “Understanding Power Quality Problems,” *IEEE Press*, 2000.

[21] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998. DOI: 10.1098/rspa.1998.0193

[22] D. C. Bowman and J. M. Lees, “The Hilbert–Huang Transform: High-Resolution Spectral Analysis for Non-Stationary Data,” *Seismological Research Letters*, vol. 84, no. 6, pp. 1074–1080, 2013. DOI: 10.1785/0220130055

[23] J. F. Hauer, C. J. Demeure, and L. L. Scharf, “Initial results in Prony analysis of power system response signals,” *IEEE Transactions on Power Systems*, vol. 5, no. 1, pp. 80–89, Feb. 1990. DOI: 10.1109/59.49090

[24] R. Bouaynaya, D. Shin, and D. Metaxas, “Coding Prony’s method in MATLAB and Python,” *BMC Bioinformatics*, vol. 19, no. 1, article 473, 2018. DOI: 10.1186/s12859-018-2473-y

[25] Jindal, A., Kumar, M., & Singh, G. K. (2012). "A review on classification of power quality disturbances in electric power systems." *International Journal of Engineering Science and Technology*, 4(3), 739–750.

[26] Ghosh, A., & Ledwich, G. (2002). "Power quality enhancement using custom power devices." *Springer*.

[27] IEEE Std 1159-2019 – Power Quality Disturbances Terminology.

[28] A. Rodríguez, J. A. Aguado, F. Martín, J. J. López, F. Muñoz & J. E. Ruiz, “Rule-based classification of power quality disturbances using S-transform,” *Electric Power*

Systems Research, vol. 86, pp. 113–121, 2012.
[29] A. Abur and A. Gómez-Expósito, "Power System State Estimation: Theory and Implementation", CRC Press, 2004.
[30] I. S. Samanta, S. Dey, A. Sinha, and A. K. Chattopadhyay, "Artificial intelligence and machine learning techniques for power quality analysis: a review," *Electric Power Systems Research*, vol. 235, p. 110617, 2024. DOI: 10.1016/j.epsr.2024.110617 — *Available on ScienceDirect (Elsevier).*
[31] Jain, A. K., et al. "Data clustering: a review", ACM Computing Surveys, 1999.
[32] Zhang, Y., et al. "Deep learning-based classification of power quality disturbances", IEEE TSG, 2019.
[33] Dash, P. K., et al. "Power quality disturbance data compression, detection, and classification using wavelet transform", IEEE TPE, 2000.
[34] Bollen, M.H.J., "Understanding Power Quality Problems: Voltage Sags and Interruptions," IEEE Press, 2000.
[35] CENELEC, "EN 50160: Voltage Characteristics of Electricity Supplied by Public Distribution Systems", 2010.
[36] F. N. Fritsch and R. E. Carlson, "Monotone Piecewise Cubic Interpolation," *SIAM Journal on Numerical Analysis*, vol. 17, no. 2, pp. 238–246, 1980.
[37] Depuru, S.S.S.R., Wang, L., Devabhaktuni, V., "Smart Meters for Power Grid—Challenges, Issues, Advantages and Status", *Renewable and Sustainable Energy Reviews*, 2011.
[38] Aidon, *RJ12 HAN Interface – Aidon Smart Energy Devices (Technical Specification Sheet)*, Aidon Oy, 2023.
[39] Landis+Gyr, *Meet Revelo™ – The IoT Grid Sensing Meter (v4.4)*, Landis+Gyr, Jan. 2020.
[40] Helen Electricity Network Ltd., "Information on switching to 15-minute metering," Helen.fi, 2024, [Online]. Available: https://www.helensahkoverkko.fi/en/current-events/news/information-on-switching-to-15-minute-metering.
[41] Stewart, E.M., et al., "Using Micro-Synchrophasors for Distribution System Applications", *IEEE Transactions on Smart Grid*, 2014.
[42] International Electrotechnical Commission, *IEC 61000-4-30: Electromagnetic Compatibility (EMC) – Testing and Measurement Techniques – Power Quality Measurement Methods*, 3rd ed., Geneva, Switzerland: IEC, 2015.
[43] International Electrotechnical Commission, *IEC 62053-22: Electricity Metering Equipment – Particular Requirements – Static Meters for Active Energy (Classes 0.2 S and 0.5 S)*, Geneva, Switzerland: IEC, 2003.

[44] MathWorks, “pchip — Piecewise Cubic Hermite Interpolating Polynomial,” *MATLAB Documentation*, 2024. [Online]. Available: https://www.mathworks.com/help/matlab/ref/pchip.html

[45] SciPy Developers, “PchipInterpolator,” *SciPy v1.12.0 Reference Guide*, 2024. [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.PchipInterpolator.html

[46] *SLATEC Common Mathematical Library*, “Subroutine PCHIM — Monotone Piecewise Cubic Hermite Interpolation,” *Netlib Repository*, 1993. [Online]. Available: https://www.netlib.org/slatec/pchip/pchim.f

[47] Baran, M.E., Kelley, A., "State Estimation for Real-Time Monitoring of Distribution Systems", *IEEE Transactions on Power Systems*, 2009.

[48] *Updating National Type User Definitions and Principles for Utilization in Network Service Charges Reviews*, The Energy Authority, Finland, Feb. 2019.

[49] Nexans, *“Overhead line conductors – Aluminium Conductor Steel Reinforced (ACSR) ‘Duck’,”* Product Data Sheet, Nexans, 2024.

[50] ABB, *“AHX-W185 Aluminium Cable – 24 kV Medium-Voltage Cable Specifications,”* Technical Catalogue, ABB Oy, 2023.

[51] Prysmian Group, *“AXMK Low-Voltage Power Cable 0.6/1 kV,”* Product Data Sheet, Prysmian Group, 2023.

[52] IEC 60228:2019, *Conductors of Insulated Cables*, IEC, Geneva, Switzerland, 2019.

[53] J. D. Glover, T. J. Overbye, and M. S. Sarma, *Power System Analysis and Design*, 7th ed., Boston, MA, USA: Cengage Learning, 2022.

[54] P. Kundur, *Power System Stability and Control*, New York, NY, USA: McGraw-Hill, 1994.

[55] IEC 60076-1:2011, *Power Transformers – Part 1: General*, International Electrotechnical Commission, Geneva, Switzerland, 2011.

[56] The MathWorks, Inc., *MATLAB Classification Learner Toolbox*, Natick, MA, USA: The MathWorks, Inc., 2025. [Online]. Available: https://www.mathworks.com/products/classification-learner.html

[57] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed., New York, NY, USA: Wiley-Interscience, 2001.

[58] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York, NY, USA: Springer, 2006.

[59] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., New York, NY, USA: Springer, 2009.

[60] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed., 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

[61] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, Aug. 2016, pp. 1135–1144.

[62] C. M. Bishop, *Pattern Recognition and Machine Learning*, New York, NY, USA: Springer, 2006.

[63] S. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

[64] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and Regression Trees*, CRC Press, 1984.

[65] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed., Springer, 2009.

[66] C. Molnar, *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, 2nd ed., 2022. [Online]. Available: https://christophm.github.io/interpretable-ml-book/

[67] F. Doshi-Velez and B. Kim, “Towards a rigorous science of interpretable machine learning,” *arXiv preprint arXiv:1702.08608*, 2017.

[68] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why should I trust you?: Explaining the predictions of any classifier,” *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, 2016, pp. 1135–1144.

[69] European Commission, *“Smart Grid Task Force: Vision and Roadmap for Europe’s Smart Grids,”* European Commission Directorate-General for Energy, Brussels, 2023. [Online]. Available: https://energy.ec.europa.eu

[70] IEEE PES, *“IEEE Smart Grid Vision for 2030 and Beyond,”* IEEE Power & Energy Society, New York, NY, 2022.

[71] International Energy Agency (IEA), *“Digitalization and the Future of Energy,”* IEA, Paris, France, 2022. [Online]. Available: https://www.iea.org

[72] CIGRÉ Technical Brochure 831, *“Resilient Distribution Systems: Concepts and Practices,”* International Council on Large Electric Systems, 2021.

[73] European Commission, *Cost–Benefit Analysis of Smart Metering and Smart Grid Deployment*, EUR 28770 EN, 2017.

[74] P. Bertoldi et al., “Smart Grid Projects Outlook 2017: Facts, Figures and Trends in Europe,” *JRC Science for Policy Report*, Publications Office of the European Union, 2017.

[75] S. Bhela, K. Mahmud, and G. Town, “Edge-computing framework for distribution grid

monitoring using low-cost IoT devices," *IEEE Access*, vol. 9, pp. 45312–45324, 2021.
[76] K. Mahmud and G. Town, "A review of microcontroller-based low-cost monitoring systems for smart distribution networks," *Energies*, vol. 14, no. 12, pp. 1–20, 2021.

Appendices

Appendix A – Network Model Data Structures

This appendix presents the detailed structural definitions of the network model components used in the simulation framework. These include the **bus matrix** and **line matrix**, which together define the electrical characteristics and interconnections of the modeled network.

Table 11. *Structure and Parameter Definitions of the Bus Matrix*

Column	Label	Description
1	bus_no	Bus number
2	V	Voltage magnitude (p.u.)
3	Θ	Voltage angle (degrees)
4	Pg	Active power generation (p.u.)
5	Qg	Reactive power generation (p.u.)
6	Pd	Active power demand (p.u.)
7	Qd	Reactive power demand (p.u.)
8	Gb	Shunt conductance at the bus
9	Bb	Shunt susceptance at the bus (p.u.)
10	bus_type	1 = Slack, 2 = PV, 3 = PQ
11	Qg_max	Minimum voltage limit (p.u.)
12	Qg_min	Maximum voltage limit (p.u.)

Table 12. *Structure and Parameter Definitions of the Line Matrix*

Column	Label	Description
1	from_bus	Sending-end bus number
2	to_bus	Receiving-end bus number
3	R	Line resistance (p.u.)
4	X	Line reactance (p.u.)
5	B	Line charging susceptance (p.u.)
6	Tap_ratio	Tap ratio of transformer
7	Phase_shift	Phase shift angle (degrees)

Appendix B: Cost Comparison and Deployment Feasibility

B1. Bill of Materials (Edge-Based Classifier Node)

Table 13. *Approximate component-wise cost estimation for the edge computing system.*

Component	Approx.Cost (USD)	Notes
Raspberry Pi 4 (4GB RAM)	$75	Edge computing unit (runs Python-based ML models or compiled C code)
Voltage/Current Sensors	$100	Assume sensors already installed or low-cost CT/PT
Data Acquisition Interface	$50	e.g., ADC module or Modbus-compatible unit
Communication Module	$75	LTE/5G modem or Ethernet interface for data transfer
SD Card, Power Supply, Case	$50	Enclosure + boot storage
Software Deployment	$0	Open-source Python ML stack (Scikit-learn / TensorFlow Lite)
Total (per site)	**≈ $350–$700**	Scalable and replicable

B2. Cost Comparison: Conventional vs. Edge-Based Monitoring

Table 14. *Cost comparison between conventional substation monitoring and the edge-based system.*

Monitoring System	Cost per Site (USD)	For 500 Sites	Total Cost
Secondary Substation Monitoring System	$10,000	500	$5,000,000
Edge-Based Classifier Node	$700	500	$250,000
Estimated Savings	—	—	**$4,650,000**

Note: *The above comparison includes hardware, but excludes classifier development costs, operational expenses, such as network communication, remote data access, software maintenance, sensor calibration, remote SCADA integration or advanced control functionality. The above-mentioned figures are intended as illustrative estimates rather than exact market quotations.*