



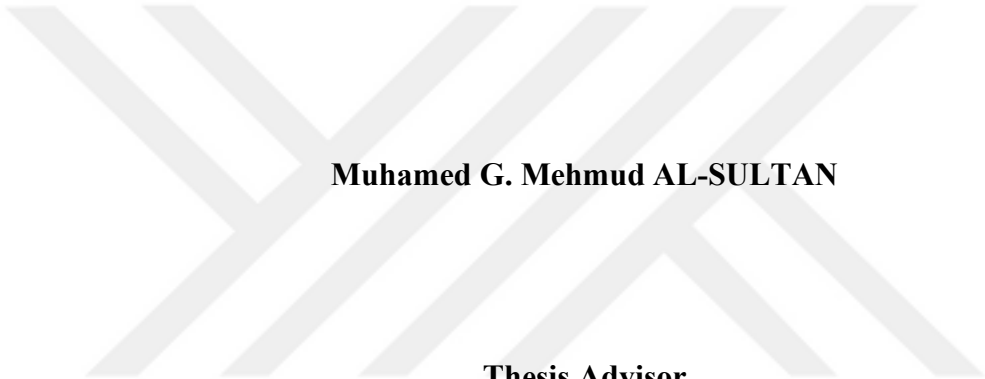
A NEW MODEL FOR FAULT IDENTIFICATION APPROACH IN POWER TRANSMISSION LINES

**2025
PhD THESIS
COMPUTER ENGINEERING**

Muhamed G. Mehmud AL-SULTAN

**Thesis Advisor
Assist. Prof. Dr. İsa AVCI**

**A NEW MODEL FOR FAULT IDENTIFICATION APPROACH IN POWER
TRANSMISSION LINES**



Muhamed G. Mehmud AL-SULTAN

**Thesis Advisor
Assist. Prof. Dr. İsa AVCI**

**T.C
Karabuk University
Institute of Graduate Programs
Department of Computer Engineering
Prepared as
Ph. D. Thesis**

**KARABUK
January 2025**

I certify that in my opinion the presented thesis that has been submitted by Muhamed G. Mehmud AL-SULTAN titled “A NEW MODEL FOR FAULT IDENTIFICATION APPROACH IN POWER TRANSMISSION LINES” is fully adequate in scope and quality as a thesis for the degree of Ph. D.

Assist. Prof. Dr. İsa AVCI

.....

Thesis Advisor, Department of Mechanical Engineering

This thesis is accepted by the examining committee with a unanimous vote in the Dept. of Mechanical Engineering as a Ph.D. of Science thesis. January 19, 2025.

Examining Committee Members (Institutions)

Signature

Chairman : Assist.Prof.Dr. Nehad T.A. RAMAHA (KBU)

.....

Member : Assist. Prof. Dr. İsa AVCI (KBU)

.....

Member : Assist. Prof. Dr. Bilal YOUSFI (KBU)

.....

Member : Assist. Prof. Dr. Ali HAMİTOĞLU (ISU)

.....

Member : Assist. Prof. Dr. Murat KOCA (VYYU)

.....

The degree of Ph.D. of Science by the thesis that has been submitted was approved by the Administrative Board of the Institute of Graduate Programs, Karabuk University.

Assoc. Prof. Dr. Zeynep ÖZCAN

.....

Director of the Institute of Graduate Programs



"I declare that all the information that has been presented in this thesis was gathered and presented by ethical principles and academic regulations and I have according to requirements of those regulations and principles that were cited all those which don't originate in this work too."

Muhamed G. Mehmud AL-SULTAN

ABSTRACT

Ph.D. Thesis

A NEW MODEL FOR FAULT IDENTIFICATION APPROACH IN POWER TRANSMISSION LINES

Muhamed G. Mehmud AL-SULTAN

**Karabuk University
Institute of Graduate Programs
Department of Computer Engineering**

Thesis Advisor:

Assist. Prof. Dr. Isa AVCI

January 2025, 127 pages

The rapid adoption of electric vehicles (EVs) and the increasing integration of fast-charging stations (FCS) into modern power networks have introduced significant challenges to system stability and reliability. Critical issues, such as voltage sag (VS), high impedance faults (HIF), and low impedance faults (LIF), disrupt system operations and can cause severe equipment damage. This study proposes a novel hierarchical protection framework tailored for an IEEE-14 bus power system with eight concurrently operating FCS supported by photovoltaic (PV) systems and the main grid. The framework employs a hybrid machine learning approach: a two-layer autoencoder (TLA) for feature extraction, Long Short-Term Memory (LSTM) networks for initial fault categorization into VS, HIF, and LIF, and Extreme Gradient Boosting (XGBoost) for refining fault classification into subcategories, including single-line-to-ground (LG), line-to-line-to-ground (LLG), three-phase (LLL), and three-phase-to-ground (LLL) faults. The proposed methodology achieved

99.8% accuracy in voltage sag detection, 96% classification accuracy for high impedance faults (HIF), and 91% for low impedance faults (LIF). The LSTM model exhibited near-perfect precision (1.00 for VS, 0.96 for LIF) and recall rates (1.00 for VS, 0.94 for LIF), ensuring robust detection capabilities. Furthermore, the framework offers enhanced computational efficiency, scalability, and robustness to noise, making it a reliable solution for real-time fault detection and classification. This work provides a foundation for advancing intelligent fault management systems in modern power grids with high FCS penetration, addressing both current and future challenges.

Key Words : Transmission faults detection and classification, Voltage sag, Fast charging station, Autoencoder, LSTM, XGBoost.

Science Code : 91408

ÖZET

Doktora Tezi

GÜÇ İLETİM HATLARINDA ARIZA TANIMLAMA İÇİN YENİ BİR MODEL

Muhamed G. Mehmud AL-SULTAN

Karabük Üniversitesi

Lisansüstü Eğitim Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:

Dr. Öğr. Üyesi İsa AVCI

Ocak 2025, 127 sayfa

Elektrikli araçların (EV) hızlı benimsenmesi ve hızlı şarj istasyonlarının (FCS) modern güç şebekelerine artan entegrasyonu, sistem kararlılığı ve güvenilirliği açısından önemli zorluklar ortaya çıkarmıştır. Voltaj düşmesi (VS), yüksek empedanslı arızalar (HIF) ve düşük empedanslı arızalar (LIF) gibi kritik sorunlar, sistemin işleyişini aksatmakta ve ciddi ekipman hasarlarına yol açabilmektedir. Bu çalışma, sekiz hızlı şarj istasyonunun (FCS) fotovoltaik (PV) sistemler ve ana şebeke desteğiyle eşzamanlı olarak çalıştığı IEEE-14 otobüs güç sistemi için özel olarak tasarlanmış yeni bir hiyerarşik koruma çerçevesi önermektedir. Çerçeve, karma bir makine öğrenimi yaklaşımı kullanmaktadır: Özellik çıkarımı için iki katmanlı bir otoenkoder (TLA), voltaj düşmesi (VS), yüksek empedanslı arızalar (HIF) ve düşük empedanslı arızalar (LIF) gibi başlangıç arıza kategorileri için Long Short-Term Memory (LSTM) ağları ve tek hat-toprak (LG), hat-hat-toprak (LLG), üç fazlı (LLL) ve üç fazlı-toprak (LLLG) arızaları gibi alt kategorilere sınıflandırma için Extreme Gradient Boosting

(XGBoost). Önerilen yöntem, voltaj düşmesi tespitinde %99,8, yüksek empedanslı arızaların sınıflandırılmasında

%96 ve düşük empedanslı arızaların sınıflandırılmasında %91 doğruluk oranı elde etmiştir. LSTM modeli, voltaj düşmesi için 1.00, düşük empedanslı arızalar için 0.96 hassasiyet ve voltaj düşmesi için 1.00, düşük empedanslı arızalar için 0.94 hatırlama oranlarıyla neredeyse kusursuz bir performans sergilemiştir. Ayrıca, çerçeve geliştirilmiş hesaplama verimliliği, ölçeklenebilirlik ve gürültüye karşı dayanıklılık sunarak gerçek zamanlı arıza tespiti ve sınıflandırması için güvenilir bir çözüm sağlamaktadır. Bu çalışma, yüksek FCS penetrasyonuna sahip modern güç şebekelerinde zeki arıza yönetim sistemlerinin geliştirilmesi için bir temel oluşturarak hem mevcut hem de gelecekteki zorluklara çözüm sunmaktadır.

Anahtar Kelimeler : İletim arızalarının tespiti ve sınıflandırılması, Voltaj düşmesi, Hızlı şarj istasyonu, Otoenkoder, LSTM, XGBoost.

Bilim Kodu : 91408

ACKNOWLEDGMENT

In the name of Allah, the Most Gracious, the Most Merciful, all praise is due to Allah, whose guidance and blessings have been my source of strength and perseverance throughout this challenging yet rewarding Ph.D. journey.

I dedicate this dissertation, first and foremost, to my beloved parents, whose unwavering love, sacrifices, and vision for my future laid the foundation of my academic pursuits. They instilled in me the value of education and ensured that I was always given the best opportunities to learn and grow. Without their support and prayers, this milestone would not have been possible.

To my dear wife, my constant source of strength and encouragement, this achievement is as much yours as it is mine. Your patience, understanding, and unwavering belief in my abilities have been the bedrock of my perseverance. In moments of despair, you gave me hope, and in moments of uncertainty, you provided me with clarity. I am forever grateful for your presence in my life.

This work is also dedicated to the cherished memory of friends and relatives who lost their lives during the tragedies of war. Their absence is deeply felt, but their legacy lives on in our hearts, reminding us to strive for a brighter and more peaceful future.

I extend my sincere gratitude to my friends and colleagues whose unwavering support, intellectual exchange, and camaraderie were invaluable throughout this journey. My deepest appreciation goes to my supervisor, Assist. Prof. Dr. Isa AVCI, for their exceptional guidance and steadfast support, and to my thesis follow-up committee for their insightful advice and commitment, which have profoundly shaped the course of this dissertation. To all who contributed to this achievement in any capacity, your support has been a cornerstone of my success, and I am truly thankful.

CONTENTS

	<u>Page</u>
APPROVAL.....	ii
ABSTRACT.....	iv
ÖZET.....	vi
ACKNOWLEDGMENT.....	viii
CONTENTS.....	ix
LIST OF FIGURES	xiv
LIST OF TABLES	xv
SYMBOLS AND ABBREVIATIONS INDEX.....	xvi
 PART 1	 1
INTRODUCTION	1
1.1. BACKGROUND.....	1
1.1.1. Significance of Fault Identification in Power Transmission Lines	2
1.1.2. Symmetrical Components and Fault Types.....	3
1.2. PROBLEM STATEMENT	5
1.3. OBJECTIVES	5
1.4. KEY CONTRIBUTIONS.....	6
1.5. OUTLINES OF THE PROJECT.....	7
 PART 2	 8
THEORETICAL BACKGROUND.....	8
2.1. INTRODUCTION.....	8
2.2. FUNDAMENTALS OF ELECTRIC TRANSMISSION SYSTEMS.....	8
2.2.1. Basic Components and Operation	10
2.2.2. Challenges Posed by Integration on Grid Stability.....	12
2.3. FAULT TYPES AND THEIR CAUSES	13
2.3.1. Line to Ground Faults.....	16
2.3.2. Line-to-Line Faults	17

	<u>Page</u>
2.3.3. Double Line to Ground (Earth) Faults.....	18
2.3.4. Symmetrical Three-Phase Faults	19
2.3.5. Open Conductor Faults	20
2.4. FAULT DETECTION TECHNIQUES.....	20
2.4.1. Conventional Methods.....	21
2.4.2. Advanced Techniques.....	22
2.5. FAULT CLASSIFICATION METHODS	23
2.5.1. Rule-Based Systems	24
2.5.2. Machine Learning Approaches.....	25
2.6. INTEGRATION OF RENEWABLE ENERGY SOURCES	25
2.6.1. Solar Power.....	26
2.6.2. Wind Power	27
2.7. FAST CHARGING STATIONS.....	28
2.7.1. Technologies and Standards	29
2.7.2. Impact on Power System Operations.....	30
2.8. FAULT MODELING IN RENEWABLES AND FCS.....	30
2.8.1. Dynamic Modeling of PV, Wind, And FCS.....	31
2.8.2. Load Fluctuations and Renewable Intermittency	33
PART 3	35
MACHINE LEARNING AND DEEP LEARNING OVERVIEW	35
3.1. OVERVIEW OF MACHINE LEARNING AND AI TECHNIQUES	36
3.1.1. Advances In Artificial Neural Networks	37
3.1.2. Evolution Of Deep Learning Techniques	38
3.1.3. Motivation For Deep Reinforcement Learning	39
3.2. FUNDAMENTALS OF ARTIFICIAL NEURAL NETWORKS.....	40
3.2.1. Neuron Models and Activation Functions.....	41
3.2.3. Feedforward and Backpropagation Algorithms.....	42
3.3. ADVANCED DEEP LEARNING ARCHITECTURES	43
3.3.1. Convolutional Neural Networks (CNNs) and Applications	44
3.3.2. Recurrent Neural Networks (RNNs) and Their Variants	44
3.3.3. Generative Adversarial Networks (GANs).....	45

	<u>Page</u>
3.4. TRAINING AND OPTIMIZATION IN DEEP LEARNING.....	46
3.4.1. Loss Functions and Regularization.....	46
3.4.2. Optimization Techniques: SGD, Adam, RMSprop	47
3.4.2.1. Stochastic Gradient Descent (SGD).....	47
3.4.2.2. Adam Optimizer.....	48
3.4.2.3. RMSProp.....	49
3.4.3. Hyperparameter Tuning and Cross-Validation.....	49
3.4.3.1. Hyperparameter Tuning.....	49
3.4.3.2. Cross-Validation	50
3.5. DEEP REINFORCEMENT LEARNING	51
3.5.1. Introduction to Reinforcement Learning	51
3.5.2. Q-Learning and Policy Gradient Methods.....	52
3.5.3. Deep Q-Networks (DQN) and Their Variants.....	53
3.6. DUAL-LAYER AUTOENCODER (DLA)	53
3.6.1. DLA Architecture	54
3.6.2. Efficiency of DLA in Handling Large Datasets	55
3.6.3. Feature Extraction and Dimensionality Reduction in DLA	55
3.7. LONG SHORT-TERM MEMORY (LSTM).....	56
3.7.1. LSTM's Architecture.....	57
3.7.2. LSTM's Ability to Capture Temporal Dependencies.....	57
3.7.3. LSTM's Application to Sequential Fault Classification.....	58
3.8. EXTREME GRADIENT BOOSTING (XGBoost).....	59
3.8.1. XGBoost Architecture	60
3.8.2. Xgboost's Advantages for Multi-Class Classification	61
3.8.3. Xgboost's Suitability for Refining Fault Detection Outputs	61
PART 4	63
LITERATURE REVIEW.....	63
4.1. INTRODUCTION.....	63
4.2. IMPORTANCE OF FAULT DETECTION AND CLASSIFICATION IN TRANSMISSION POWER LINES	65
4.3. HISTORICAL OVERVIEW OF FAULT DETECTION TECHNOLOGIES.....	66

	<u>Page</u>
4.4. SIGNAL PROCESSING-BASED FAULT DETECTION TECHNIQUES ..	68
4.4.1. Wavelet Transform.....	70
4.4.2. Discrete Wavelet Transform	70
4.5. AI-BASED FAULT DETECTION TECHNIQUES	72
4.6. HYBRID APPROACHES.....	74
4.7. CHALLENGES AND LIMITATIONS IN FAULT DETECTION AND CLASSIFICATION.....	74
4.8. COMPARATIVE ANALYSIS OF DETECTION TECHNIQUES.....	76
4.9. ML/DL ADVANCES IN FAULT DETECTION	77
 PART 5	 81
SYSTEM DESCRIPTION	81
5.1. POWER SYSTEM MODEL	81
5.2. DESCRIPTION OF FAULT CASES.....	83
5.3. AUTOENCODER FOR FEATURE EXTRACTION.....	85
5.4. BASELINE LSTM FOR CLASSIFICATION AND DETECTION.....	87
 PART 6	 90
PROPOSED METHODOLOGY	90
6.1. DESCRIPTION OF PROPOSED METHODOLOGY	92
6.2. STEPS IN METHODOLOGY	94
6.3. FEATURES EXTRACTION USING TWO-LAYER AUTOENCODER	95
6.4. VOLTAGE SAG, LIF AND HIF DETECTION AND CLASSIFICATION USING LSTM	96
6.5. SYMMETRICAL AND UNSYMMETRICAL FAULTS DETECTION AND CLASSIFICATION USING XGboost.....	96
 PART 7	 98
RESULTS AND DISCUSSION	98
7.1. ANALYSIS OF VARIOUS FAULT CASES	98
7.1.1. Impact Assessment of High Impedance Faults.....	98
7.1.2. Impact Assessment of Low Impedance Faults	100

	<u>Page</u>
7.1.3. Impact Assessment of Voltage Sag Due To Fcs.....	101
7.2. PERFORMANCE ANALYSIS OF DEEP LEARNING-BASED MODEL FOR DETECTION AND CLASSIFICATION OF FAULTS	102
7.3. PERFORMANCE ANALYSIS OF SYMMETRICAL AND UNSYMMETRICAL FAULT DETECTION AND CLASSIFICATION MODEL	107
7.4. STATE-OF-THE-ART COMPARISION	111
 PART 8	 114
CONCLUSIONS.....	114
 REFERENCES.....	 116
 RESUME	 127

LIST OF FIGURES

	<u>Page</u>
Figure 2. 1. Typical electric power system [20].....	9
Figure 2. 2. Classification of transmission line faults [32].	16
Figure 2. 3. The line to ground fault with Z_f fault impedance [33].	17
Figure 2. 4. Line to line fault with Z_f fault impedance [33].	18
Figure 2. 5. Line to line to ground fault with z_f fault impedance [33].....	19
Figure 2. 6. Three-phase fault with z_f fault impedance [33].....	19
Figure 2. 7. Open conductor fault [33].....	20
Figure 3. 1. A brief guide to AI and machine learning methods [51].	37
Figure 4. 1. Discrete Wavelet Transform.....	71
Figure 5. 1. Modified test system.....	83
Figure 5. 2. Description of fault cases.	84
Figure 6. 1. Proposed methodology.	91
Figure 7. 1. HIFs at 3 to 3.5 seconds with 0° inception angle and $100\ \Omega$ resistance.	100
Figure 7. 2. LIFs at 3 to 3.5 seconds with 0° inception angle and $10\ \Omega$ resistance.	101
Figure 7. 3. Voltage sag at 2 to 2.5 seconds due to FCSs.	102
Figure 7. 4. Performance metrics of LSTM.	105
Figure 7. 5. Confusion matrix of LSTM. 7.5 (a) Train data 7.5 (b) Test data.	105
Figure 7. 6. Performance metrics of XGBoost for HIF.....	108
Figure 7. 7. Confusion matrix of XGBoost for HIF. 7.7 (a) Train data 7.7 (b) Test data.	109
Figure 7. 8. Performance metrics of XGBoost for LIF.	110
Figure 7. 9. Confusion matrix of XGBoost for LIF. 12 (a) Train data 12 (b) Test data	110

LIST OF TABLES

	<u>Page</u>
Table 2. 1. Transmission line fault cause codes and outage frequency [29].....	14
Table 4. 1. Literature Review.....	78
Table 5. 1. Description of PV system	82
Table 5. 2. Description of fault parameter.	85
Table 7. 1. Hyperparameters of autoencoder, LSTM, and XGBoost.....	103
Table 7. 2. Comparison of the proposed scheme with other methods.	112

SYMBOLS AND ABBREVIATIONS INDEX

SYMBOLS

c_t	: Memory Cell State
f_t	: Forget Gate
i_t	: Input Gate
o_t	: Output Gate
W_e	: Encoder Weight
b_e	: Encoder Bias
W_d	: Decoder Weight
b_d	: Decoder Bias
t	: Latent Vector
σ	: Activation Function
Γ_r	: Reference Irradiation (W/m ²)
γ_r	: Reference Cell Temperature (°C)
I_{sc}	: Short Circuit Current at Reference Conditions (A)
V_{oc}	: Open Circuit Voltage at Reference Conditions (V)
I_{mpp}	: Current at Maximum Power Point (A)
V_{mpp}	: Voltage at Maximum Power Point (V)
R_f	: Fault Resistance (Ω)
S_i	: State of Circuit Breaker (on or off)
ω	: Weight
b	: Bias
P_{max}	: Maximum Power Output (kW)
V_s	: Voltage Sag
Δt	: Time Interval

ABBREVIATIONS

AI	: Artificial Intelligence
ML	: Machine Learning
DL	: Deep Learning
FD	: Fault Detection
EVs	: Electric Vehicles
FCS	: Fast Charging Station
HIF	: High Impedance Faults
LIF	: Low Impedance Faults
LSTM	: Long Short-Term Memory
RMS	: Root Mean Square
SVM	: Support Vector Machine
XGBoost	: Extreme Gradient Boosting
RNN	: Recurrent Neural Network
PMUs	: Phasor Measurement Units
VS	: Voltage Sag
FIA	: Fault Inception Angle ($^{\circ}$)
SNR	: Signal-to-Noise Ratio
FPR	: False Positive Rate
TPR	: True Positive Rate
DNN	: Deep Neural Network
HOS	: Higher-Order Statistics
GNN	: Generalized Neural Network
TLA	: Two-layer autoencoder
LG	: single-line-to-ground faults
LLG	: line-to-line-to-ground
LLL	: Three-phase faults
LLLG	: three-phase-ground faults
VMD	: Variational mode decomposition

PART 1

INTRODUCTION

1.1. BACKGROUND

Minimizing the risk of power outages often represents the core goal of electrical grid regions. As power systems grow in complexity and renewable energy sources and fast charging stations (FCS) for electric vehicles become more common, the new requirements for stability and reliability become overwhelming. Blackouts cause enormous financial damages to utility agencies, industrial plants, and commercial customers and threaten the operations of critical infrastructures [1], [2]. Faults such as short circuits, insulator breakdowns, and environmental disturbances (e.g., thunderstorms) form an active part of causing a fault, especially for transmission systems (mainly low-voltage networks) [3]. In addition to traditional shunt faults, modern power networks face progressive problems, like voltage sag (VS) and high-impedance faults (HIF), which deteriorate power quality and hinder the system [4]. For example, HIF frequently occurs through the contact of tree branches with electrical conductors or damaged cables and is difficult to detect due to small fault currents that make conventional protection schemes ineffective [5].

These issues become even more challenging with the widespread development of sophisticated intelligent grid systems largely driven by the growing and diversifying use of available renewable energy sources. When the output from centralized generation units is closely linked to remote and widely spread load centers or when demand spikes in densely populated urban areas, the rapid implementation of high-voltage transmission solutions becomes crucial. Additionally, the swift expansion of charging infrastructure and the widespread adoption of electric vehicles necessitates the development of extensive rapid-charging networks capable of accommodating substantial and fluctuating power demands. This requirement places stringent and

complex demands on secondary phase substations and requires primary transmission line voltages to achieve levels of 400 kV. It is also important to promote the ongoing advancement and application of innovative amplitude regulation techniques to ensure the entire grid system's operational resilience, stability, and efficiency [6], [7].

In literature and technical standards, three distinct levels of fault protection are generally identified: primary, secondary, and tertiary. At the secondary level, the power system's design emphasizes minimizing the occurrence of faults. As a result, protective devices are intentionally installed to reduce their activation frequency. Currently, sophisticated control techniques are employed to enhance the tensile strength of electrical power transmission lines and to decrease the likelihood of short circuits during substantial electrical energy transfers between the grid and fast-charging stations. This control system manages the fast-charging process to stay within predefined limits. Nevertheless, in rare and unforeseen incidents, such as impedance and voltage sags in line areas, lower-level protection devices may be bypassed while backup protection devices are activated. The processes involved in these activations include relay mechanisms, methodologies for fault location, zoning strategies, and the incorporation of nano-grid designs [8].

1.1.1. Significance of Fault Identification in Power Transmission Lines

Continuous, reliable, and efficient operation of high-voltage transmission lines is a critical need, especially with the increased grid usage for supporting intermittent renewable energy and the planned increase in the number of fast-charging stations to support the emergent market for electric vehicles. In power network protection, fault identification is essential in isolating damaged assets for maintenance or repair. Fast, precise, and reliable fault localization is the central theme in minimizing the time required to restore power to the system and further preventing the risk of the cascading failure of the bulk power system. By developing an intelligent multi-level structure, we can diagnose these three types of faults (i.e., VS, HIF, and LIF) occurring in parallel in the power transmission network. Results derived from a detailed case study provide proof of concept predictions of the approach. This original hierarchical fault protection

system has vast potential implications for improving resilience and supporting active power systems using renewable energy [9], [10].

1.1.2. Symmetrical Components and Fault Types

Symmetrical components are a mathematical framework for examining unbalanced three-phase power systems, particularly those experiencing faults. The fundamental concept involves breaking down the unbalanced three-phase voltages or currents into three distinct sets of balanced phasors:

- **Positive-sequence components:** These components align with the original system's phase sequence (e.g., A-B-C) and denote the balanced portion of the system.
- **Negative-sequence components:** These components exhibit the reverse phase sequence (e.g., A-C-B), arising due to the system's unbalanced nature during faults.
- **Zero-sequence components:** These components are in phase with one another and represent the common part shared by all three phases, playing a critical role in analysing ground faults.

Faults occurring in power transmission lines can be categorized into two main groups:

- **Symmetrical faults:** These faults impact all three phases uniformly, preserving the system's balance. The most prevalent symmetrical fault type is the three-phase fault (LLL), where all three conductors are short-circuited together.
- **Unsymmetrical faults:** These faults affect phases unevenly, resulting in an unbalanced system. Typical unsymmetrical faults include:
 - **Single line-to-ground fault (SLG):** A fault occurring between a single conductor and the ground.
 - **Line-to-line fault (LL):** A fault between two conductors.
 - **Double line-to-ground fault (LLG):** A fault affecting two conductors and the ground.

The analysis of power systems extensively utilises symmetrical components, which significantly enhance system evaluation efficiency and protection strategies' design. This process begins with selecting a reference phase from the available options: phase 'a', characterized by counterclockwise rotation; phase 'b', rotating counterclockwise by 120° relative to phase 'a'; or phase 'c', rotating counterclockwise by 240° . A vector corresponding to the chosen reference phase is then established at the origin, serving as the foundation for further analysis. Impedances are subsequently incorporated based on the rotational attributes of each phase. The resultant phasor defines the symmetrical component associated with the designated phase [11].

In power networks with considerable fast-charging infrastructure or substantial load generation adjacent to substations, intricate issues such as current infeed and localized grid compensation can arise. Additionally, renewable energy sources introduce distinct challenges in faults situations. Some production technologies might not possess built-in fault tolerance, which could result in complete or partial output unavailability during these events. This constraint requires thoroughly examining protection strategies to ensure system stability and reliability [12].

Two-phase-to-ground faults, two-phase faults, single-phase-to-ground faults, and asymmetric fault scenarios are intended to replicate instances of voltage dips and spikes that disrupt voltage levels and load current or cause imbalances in power systems. Integrating high-penetration, distributed fast-charging infrastructure and renewable energy resource configurations leads to variations in power flow, voltage reduction, and impedance faults. Additionally, connecting fast-charging infrastructure and renewable energy sources to the power grid results in high-voltage vectors that exceed the voltage levels typically present in the power grid. This injection can potentially cause issues for a Voltage Source Converter employed in a low voltage direct current fast charger, affecting both the coupling and decoupling of grid faults during pre-charge and operational phases, which may necessitate taking the fast-charging infrastructure or part of its load offline. Similarly, faults occurring within renewable energy resources and the low-voltage bus can have analogous consequences compared to those within medium- or high-voltage configurations [13], [14].

1.2. PROBLEM STATEMENT

The proliferation of fast-charging stations (FCS) of electric vehicles (EVs) and renewable energy sources (RES), such as photovoltaic (PV) systems, poses new and unique challenges to contemporary power networks. Although they encourage environmental friendliness, these novelties introduce shifting dynamic load shocks and sporadic power flows, which reduce grid viability. Major concerns about operational continuity and equipment reliability are voltage sag (VS), high-impedance faults (HIF), and low-impedance faults (LIF) [6]. Up to 90% of reported disturbances come from voltage sag alone, which can devastate industrial and commercial systems that utilize sensitive devices, including programmable logic controllers (PLC) and variable frequency drives (VFD).

Existing approaches for fault detection and classification methods based on signal processing and machine learning algorithms [15] do not perform well for real-time applications. They are not scalable, robust to noise, or adaptive enough to tackle the complex and changing landscape represented by the increasing penetration of FCSs on a grid combined with grid-injected renewables. These shortcomings hinder the validity and robustness of power systems, with extended fault recovery times and amplified operational threats.

To meet these challenges, a scalable, efficient solution for detecting and classifying faults in real time needs to be developed. This will allow the system to respond adequately and ensure adequate grid performance under different operating conditions [16]. The solution must navigate the needs of modern power grid networks while minimizing disruptions and ensuring grid reliability.

1.3. OBJECTIVES

This study bridges a critical gap in managing faults and voltage sags under dynamic grid conditions by proposing a new hierarchical model for a fault identification approach in power transmission lines. This framework enhances the resilience and

reliability of power systems and provides a foundation for integrating intelligent fault management solutions in future smart grids [15].

The key objectives of this study are outlined below:

- To propose a hierarchical fault protection framework for power networks, including low-voltage transmission systems with fast-charging stations (FCS) and photovoltaic (PV) systems, ensuring real-time applicability and robustness.
- To identify and classify the sources of voltage sag, distinguishing between routine occurrences (e.g., charging at FCS) and abnormal events (e.g., faults) in the power network.
- To develop a two-layer autoencoder for feature extraction to efficiently analyze voltage sag, low impedance faults (LIF), and high impedance faults (HIF), which involve sequential and non-linear signals. Additionally, long-short-term memory (LSTM) networks can be employed to classify and detect disturbances, ensuring effective handling of these challenges.
- To utilize XGBoost to classify and detect symmetrical and asymmetrical faults, leveraging autoencoder-based feature extraction to accurately target LIF and HIF faults.

1.4. KEY CONTRIBUTIONS

This study addresses critical fault detection and classification challenges for power systems integrating fast-charging stations (FCS) and photovoltaic (PV) systems. Bridging gaps in existing methods provides a scalable, efficient, and robust solution suitable for real-time applications. The key contributions include:

- **Robust Fault Protection Framework:** Developing a framework for detecting and classifying voltage sag (VS), high-impedance fault (HIF), and low-impedance fault (LIF) with a hierarchical structure that proves its potential for applications in dynamic grid settings and real-time manner.
- **Advanced Feature Extraction and Sequential Classification:** Utilizing a two-layer auto-encoder with Long Short-Term Memory (LSTM) networks

facilitates feature extraction efficiently and precise fault diagnosis from complex time-series data.

- **Enhanced Fault Categorization.** Using Extreme Gradient Boosting (XGBoost), the framework classifies faults into symmetrical and asymmetrical categories and further abstracts them to improve accuracy, ensuring reliable and precise performance across diverse fault scenarios.
- **Scalable and Adaptive Solution.** The proposed approach fills the gap of traditional methods and offers computational efficiency, noise resilience, and adaptability for modern power grids.

1.5. OUTLINES OF THE PROJECT

The Dissertation begins with a detailed description of the distribution system modifications in PART 1: Introduction. The theoretical background of the methodologies used is presented in PART 2. PART 3 gives an overview of artificial neural networks and deep learning fundamentals. The literature review was presented in PART 4. Following this, PART 5 presents the core principles, and a comprehensive explanation of the techniques employed in the proposed methodology, including autoencoders, Long Short-Term Memory (LSTM) networks, and XGBoost. PART 6 elaborates on the proposed technique, encompassing the complete system architecture, fault detection, and classification processes using XGBoost. Additionally, the results of the LSTM classifier are presented through various figures and graphs in PART 7, providing a clear visualization of the model's performance. And finally, the conclusion of the work is presented in PART 8.

PART 2

THEORETICAL BACKGROUND

2.1. INTRODUCTION

Modern power systems are marked by their ability to integrate remote renewable energy sources, which require high-capacity transmission power systems to distribute their output smoothly and minimize losses. In particular, this is true for transport facilities, including electric vehicle fast charging stations. Increasing the capacity of transmission power systems also means not threatening power system reliability, which is determined by factors such as increasing faults due to equipment aging and many other causes. Detecting and accurately classifying power system faults has proved critical for service continuity. Nowadays, there is research using vibration and current sensors and signal processing techniques to improve the classification of faults, but it is focused on distribution power systems. The present document comprehensively describes the most common fault types in transmission power systems and the reasons for their occurrence. It highlights the growing importance of Distributed Energy Storage Systems to support traditional energy storage devices and improve reliability in the presence of monitoring systems, with vibration sensors also used for fast, automated detection and classification of faults [17].

2.2. FUNDAMENTALS OF ELECTRIC TRANSMISSION SYSTEMS

In exploring the complex relationship between power systems and renewable energy sources, particularly emphasizing photovoltaic systems alongside both onshore and offshore wind turbines, it is crucial to examine several vital elements thoroughly. These include power oscillations, gradual voltage fluctuations often accompanied by imbalances, disturbances in voltage stability, frequency stability, and the overall dependability of operations and system integrity. Such investigations must also include

short circuit analysis with a clear classification specifically directed at home energy generation setups and an in-depth evaluation of the positioning and functionality of fast charging stations within the continuously advancing framework of smart grid technology [18].

This research aims to identify and categorize faults within electric transmission power systems, essential for dynamic security evaluations, especially in response to disturbances, thereby improving real-time situational awareness. A critical challenge in this domain lies in addressing issues of online unbalancing, complexity, high dimensionality, and uncertainties that complicate the identification of damaged components. Adopting innovative big data methodologies coupled with sophisticated data mining and pattern recognition techniques is necessary to tackle these challenges. These methods are designed to accurately interpret voltage and current behavior signals during the intricate process of rapid discrete wavelet transform (DWT). The overarching goal is to secure and optimally manage fast charging stations capable of effectively charging zero-emission battery electric vehicles in an environmentally sustainable manner [19].

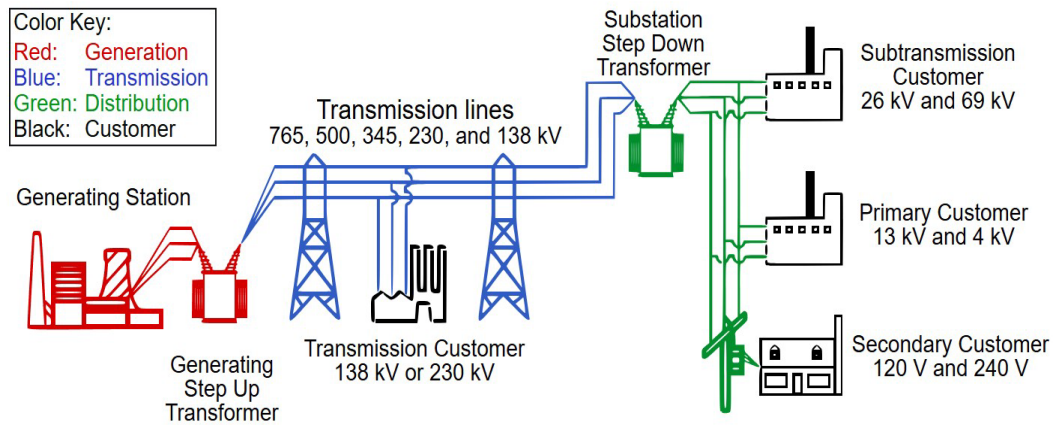


Figure 2. 1. Typical electric power system [20]

Within electric power transmission systems, as shown in Figure 2.1, a diverse range of scenarios that could potentially jeopardize the safety and reliability of the system are diligently monitored. Today's grid systems comprise an array of integrated units, including Flexible AC Transmission Systems (FACTS), Static Var Compensators (SVCs), wind farms, Off-Peak Generation (OPGs), and various capacities of Ultra-

High Voltage (UHV), High Voltage (HV), and Extra High Voltage (EHV) power lines. Therefore, it is essential to thoroughly investigate how wind farms linked to UHV and HV lines impact the stability of the overall power system during their functional phases. This inquiry focuses on detecting and classifying faults within electric transmission power systems to strengthen security assessments and refine control center operations. That means closely observing voltage and current signatures of fast DWT in real-time, enabling the quick verification of outcomes for the temporal identification of power system failure, which means a big data capacity. A summary of data to ensure awareness and good coordination of hierarchical protection systems [21].

In addition, this work will also shine a light on the importance of distribution systems, not only for electric transport, but also on the need for setting up fast charging stations for electric vehicles, which play an important and enabling role in making a transition towards sustainable mobility. Lastly, a comprehensive analysis will present a rapid multi-step High Voltage Direct Current (HVDC) Power/Energy Storage (P/ES) system, with an emphasis on crucial aspects including frequency response, effective electricity transmission, low-cost valve arrangement techniques, as well as fire and transient protection to promote stability and improve overall security [22].

2.2.1. Basic Components and Operation

The grid consists of a very complex network which integrates distributed generation and modern energy storage technologies to meet the evolving needs and challenges of our energy system. Various transmission lines, classified in different categories according to voltage levels, further explain this complex system. Another distinguishing factor of high-voltage transmission lines is that they can stretch over a distance of greater than 100 kilometres, while the range for medium-voltage lines is around 10 to 100 kilometres. In contrast, low-voltage lines usually have a limited range and only cover a distance of less than 10km. All the transmission line categories are critical links in delivering the generated electrical power to its end-users, such as distribution substations and ultra-fast charging stations attached to EHV lines for enhanced energy transfer [23].

Furthermore, advanced technologies have been actively applied to further facilitate power flow management with reinforcement at the system level, such as flexible AC transmission systems. These cutting-edge technologies comprise fundamental systems such as phase-angle voltage controllers for precise voltage control, series capacitors, which greatly enhance transmission performance, phase shifters that deftly adjust active power flow as necessary, and static var compensators that provide critical reactive power supply. Key needs of power solutions that are custom build meet the tailoring to fulfil unique operational needs, this will optimize performance along energy network [24]. Moreover, the role of HVDC (high-voltage direct current) transmission lines cannot be overlooked when discussing the efficient transfer of electrical energy over long distances with low losses. The establishment of shore supply stations at generation sites, especially offshore, could become more relevant as the pursuit of a prosperous, powerful and adaptive framework of energy for a more sustainable future has gained momentum. This evolution is dedicated further to the adjustment of grid system to integrate the renewable energy sources, thus, discusses a greener energy future [25].

Generation, transformation and transmission are three fundamental features of electricity systems. Every one of these parts has an essential part, to do with the spontaneous operation for all intents and purposes of the whole cycle. All of this ties in with each other because it has to do with the movement of electrical energy at high voltages (along with auxiliary functions). This sophisticated network is, at its core, designed to transport significant quantities of electric power across long distances, so you can expect a good amount of losses due to transmission. These challenges require new and crucial controls to mitigate these losses effectively. These controls include key attributes such as voltage regulation and angle regulation, which are essential to the stability and efficiency of an entire system[26].

Furthermore, the fundamental factor of redistribution is closely linked with the operations of high voltage, which plays a pivotal role in ensuring the effective distribution of electrical energy to a vast array of substations. These substations kick-start the distribution circuit, adjusting the voltage to ethically safe levels for end use,

making electrical power more widely accessible [26]. So the whole system is synchronous, the one big orchestra under a semi-responsible operational model. Operators, which can be anyone who observes the system and users who use the energy can have specific requests for controls and execute necessary operations as they arise in the said model.

At the centre of this complex operational model is a well-defined hierarchy of tasking, where dominant resource control is the main concern. Amid such acute interplay, the system operator emerges as the foremost coordinating agent, directing and overseeing the holistic endeavor to ensure that energy flows are sustained and accountable for all grid stakeholders. Such collaboration between operators and users is essential for maintaining balance and ensuring reliability in the electric power system.

2.2.2. Challenges Posed by Integration on Grid Stability

Determined to counter the absolute and urgent threat of growing reliance within the fossil fuel system, countries across the globe have increasingly sought to endorse, subsidize, and publicize a range of investment incentives for the sustainable generation of renewable energy. Most of this developing and hopeful renewable electricity generating capability is due to the installation of wind generators, and the installation of photovoltaic techniques that correctly seize sun electricity. On the other hand, the energy collected from these renewable resources is by its very nature transient. Such intermittency and variability are highly unpredictable most of the time, thereby making the integration of such energy sources in the existing power system a major, multidimensional, and complex technological challenge that needs sufficient planning and execution. The increasing installation and widespread penetration of generation from diverse renewable energy sources may create significant operational challenges for the proper functioning of the overall power system. Among the most significant challenges in this context is integrating these renewable resources into the electric grid securely, reliably, and stable [27]. Moreover, there is a crucial need for the maintenance and conservation of the relevant reliability indices without any interference or disruptions to the normal functioning of the grid itself, ensuring that customers receive a consistently stable electricity supply. Addressing these challenges

will require innovative solutions and collaborative efforts from various stakeholders in the energy sector.

The rapid charging of electric vehicles at fast charging stations that are seamlessly integrated into the local distribution system results in noteworthy and substantial variable load changes for specific areas, which can lead to voltage unbalance, decreased power quality, and, in some cases, even overloading of the feeder or distribution transformer [28]. Despite these challenges, it is important to observe that the cutting-edge technologies provided by circuit breakers and analogous devices do indeed offer certain assurances to the transmission or distribution of power networks. The alterations proposed within these systems are relatively straightforward and less demanding than the extensive and deep modifications suggested to effectively coordinate and protect grid-scaled battery energy storage systems and flexible AC transmission systems. The data collection process and service assessments related to the energy charging provided by these highly efficient fast charging stations are significantly less complicated and can typically be executed with a much lower level of investment needed for data collection and associated work. Nonetheless, crucial aspects such as power quality, overall stability, and potential service interruptions must be thoroughly evaluated and assessed. Fast charging stations can also be likened to the shifts in demand that manifest when many households invest in and subsequently purchase battery electric vehicles. It is essential that this surge in demand does not occur simultaneously, as doing so could lead to voltage unbalance and immediate load fluctuations, which negatively impact the quality and stability of the overall power system. However, the growing installation of fast charging stations designated for electric vehicles within a distribution system presents a distinct array of challenges because it signifies a major technological transition [28]. The implications of this transition warrant careful consideration to ensure the continued reliability of the electrical infrastructure.

2.3. FAULT TYPES AND THEIR CAUSES

With substantial miles of overhead transmission lines being operated throughout the United States, transmission lines are deliberately exposed to a variety of potential

external events. These events can create abnormal or faulted condition on these active transmission lines. With such large distances of exposure to transmission lines, it is inevitable that electrical transmission line faults are going to occur, and it is just a matter of when these faults are going to occur. These faults can originate from many sources including weather, natural disasters events, animals, or from human interaction to name a few. Table 2.1, published by The North American Electric Reliability Corporation (NERC), defines the categories of different causes of transmission line faults and how frequent these electrical faults have occurred between 2012 and 2016 [29]. Many of these electrical faults have been linked to weather-related events, primarily associated with lightning strikes or unknown causes. In electrical power systems, an electric transmission fault is defined as an unusual circumstance that disrupts the normal flow of electrical current. These faults can be classified into two principal types: temporary and permanent. Temporary faults occur within the transmission system for a limited time, often resolving independently. For instance, when a branch falls onto a transmission line, it triggers a fault condition; however, once the branch is removed from the line, the fault is automatically cleared [30].

Table 2. 1. Transmission line fault cause codes and outage frequency [29]

TADS Transmission Line Fault Cause Code and Outage Frequency						
Initiating Cause Code	2012	2013	2014	2016	2016	2012-2016
Lightning	852	813	709	783	733	3890
Unknown	710	712	779	830	773	3804
Weather Excluding Lightning	446	433	441	498	638	2465
Failed AC Circuit Equipment California	261	248	224	255	362	1350
Miss Operation	321	281	314	165	249	1330
Failed AC Substation Equipment	248	191	223	221	214	1097
Foreign Interference	170	181	226	274	285	1109
Contamination	160	151	149	154	289	903
Human Error	212	191	149	132	153	837
Power System Condition	77	109	83	96	81	446
Fire	106	130	44	65	72	471
Other	104	64	77	77	78	400

Combined Smaller ICC Groups Study 1-3	57	53	49	37	47	243
Vegetation	43	36	39	32	34	184
Vandalism, Terrorism, or Malicious Acts	10	9	8	1	7	35
Environmental	4	8	2	4	6	24
All with ICC Assigned	3724	3557	3467	3587	3947	18282
All TADS Events	3753	3557	3477	3587	3947	18321

Permanent faults are abnormal conditions within the transmission system that cannot self-correct. A common permanent fault occurs when a current-carrying conductor breaks between two transmission towers and comes into contact with a grounded structure, resulting in a continual line-to-ground fault. This situation requires manual intervention to separate the conductor from the transmission tower. During these fault conditions, electrical currents can flow erratically between conductors and the ground or among various conductors. The behavior of electrical current in these circumstances is critical for classifying faults, often termed fault types. As shown in Figure 2.2, Different fault classifications in the electrical utility sector include single line-to-ground faults, double-line (line-to-line) faults, double-line (line-to-line) to-ground faults, and three-phase faults. The three-phase fault is unique as the only symmetrical fault, while single line-to-ground, double line, and double line-to-ground faults are classified as asymmetrical faults. Notably, a significant share of transmission system faults, particularly for voltages of 115 kV or more, is attributed to lightning, which causes flashover events on insulators. Research indicates that single line-to-ground faults account for approximately 70% to 80% of faults in transmission lines, whereas around 5% involve all three phases [31].

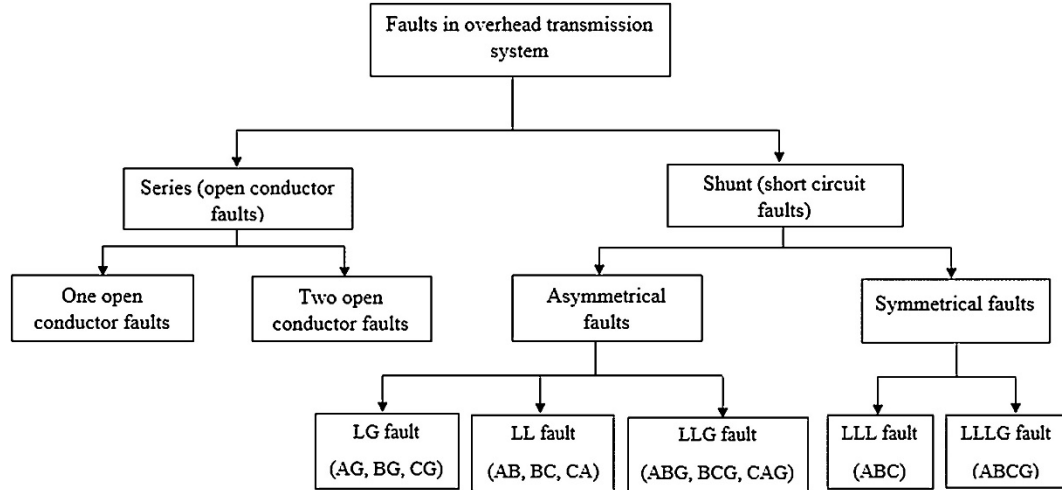


Figure 2. 2. Classification of transmission line faults [32].

2.3.1. Line to Ground Faults

The line-to-ground fault is the most widespread electrical malfunction encountered in transmission systems. Each transmission line consists of three conductors responsible for transporting electric current and a static or ground conductor wire, usually grounded at nearly every transmission structure. This configuration is known as a multi-grounded system. The three conductors carrying current are typically classified into phases labeled phase A, B, and C. The precise assignment of phases A, B, or C is mainly a matter of established convention, provided the labeling is consistently applied across the system. A line-to-ground fault is atypical in which one of the three current-carrying conductors touches a part of the transmission system with zero-voltage potential. To exemplify this occurrence, Figure 2.3 visually represents a hypothetical location along a transmission line where a fault from phase A to the ground occurs.

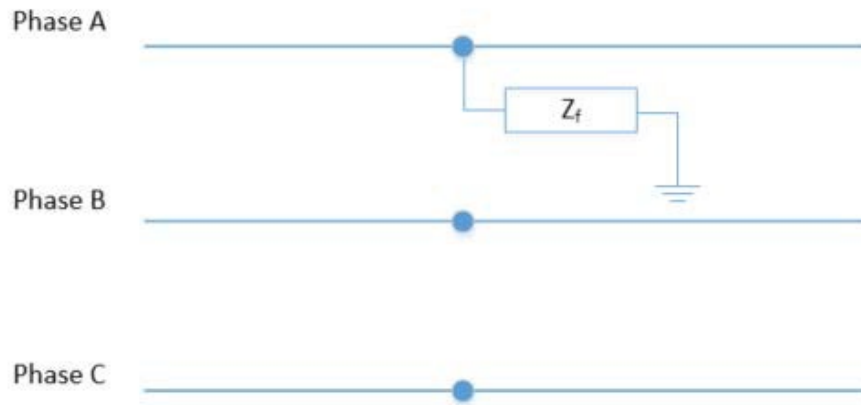


Figure 2. 3. The line to ground fault with Z_f fault impedance [33].

The Z_f fault impedance represents the impedance associated with the fault as it traverses through the current-carrying conductor toward the grounded equipment. The value of Z_f impedance is variable and influenced by the specific physical conditions that initiate the fault. To effectively comprehend this derivation or to evaluate any asymmetrical fault, it is crucial to have a fundamental understanding of symmetrical components [32].

2.3.2. Line-to-Line Faults

Line to line faults are faulted conditions that encompass connections between two of the current carrying conductors of the transmission line. The possible faulted classifications for these types of faults would consist of abnormal conditions between any two of the three current carrying conductors: phase A to phase B, phase A to phase C, or phase B to phase C. These line fault classifications are considered and analyzed within this research dissertation. Figure 2.4 provides a visual representation of a hypothetical point on a transmission line in which a phase B to phase C line to line fault has occurred. The Z_f fault impedance represents the impedance of the line-to-line contact. This Z_f impedance value can vary depending on the physical condition causing the fault [34].

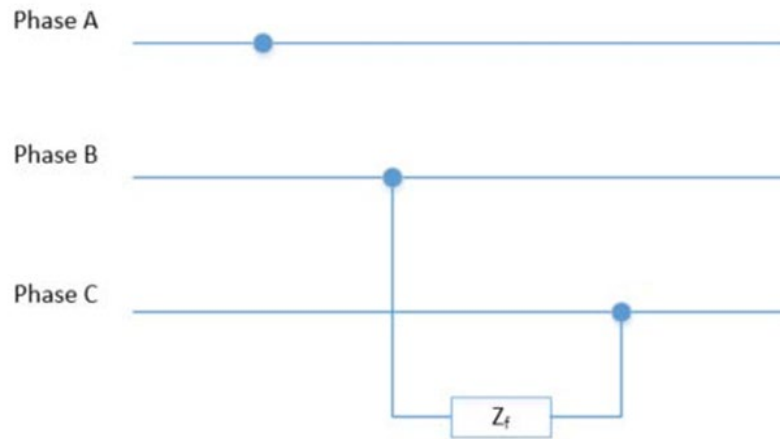


Figure 2. 4. Line to line fault with Z_f fault impedance [33].

2.3.3. Double Line to Ground (Earth) Faults

Double line-to-ground faults are defined based on the connections formed during the fault between the current-carrying conductors and the point of grounding within the transmission system. These defects can fall into one of three separates:

- Phase A – Phase B – Ground
- Phase A – Phase C – Ground
- Phase B – Phase C – Ground

In the present research dissertation, these line fault classifications are been analyzed to consider. An example of a transmission line when a double line to a ground fault (from phase B to phase C to ground) has occurred is shown in Figure 2.5. Whereas the Z_f fault impedance denotes the impedance associated with this double line-to-ground interaction. Here, with each fault condition, the fault current has passed through the current carrying conductors and returned to earth, so that the overall faulted impedance can be represented at the grounding connection. The Z_f impedance value can vary depending on how physical conditions cause the fault. 11 line-to-line to ground fault of healthy phases, please refer to [35].

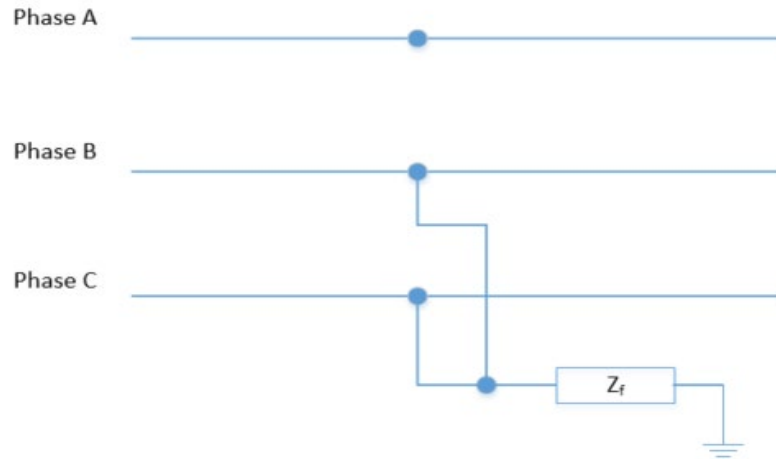


Figure 2. 5. Line to line to ground fault with z_f fault impedance [33].

2.3.4. Symmetrical Three-Phase Faults

The final type of fault examined in this dissertation is the three-phase fault, which is noted for being the least frequently encountered among the array of potential fault conditions. This scenario materializes when all three current-carrying conductors make contact with one another. In a three-phase electrical system, the sole configuration that can lead to a fault is one where phases A, B, and C simultaneously engage. Much like the other three fault types discussed earlier, three-phase faults exhibit a specific level of fault impedance. This fault impedance, denoted as Z_f , is structured so the fault current flow navigates through the fault impedance within each phase. Figure 2.6 illustrates the nature of three-phase faults [36].

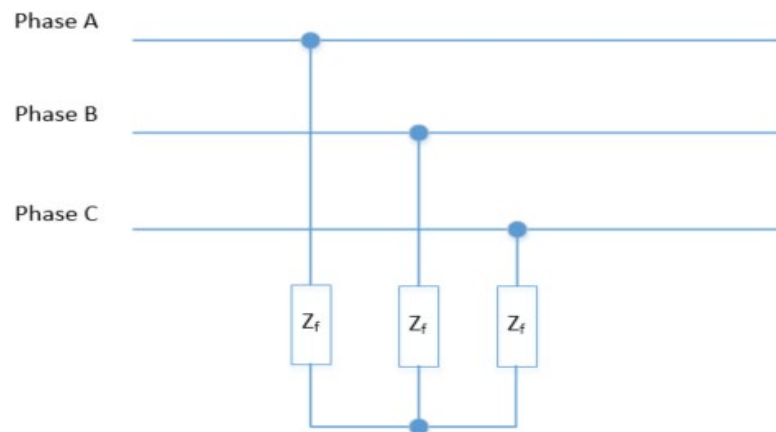


Figure 2. 6. Three-phase fault with z_f fault impedance [33].

2.3.5. Open Conductor Faults

There is one other type of fault that can occur which is the open conductor fault. The past four fault classifications that have been discussed are shunt fault types. The open conductor fault is a series fault type. During this faulted condition, an open circuit occurs in one or more phases of the transmission circuit. This type of fault classification was not studied within this research. Figure 2.7 provides a diagram of an open conductor fault example. This example shows that phase A has become an open circuit while phases B and C remain intact. This diagram indicates that phase A has established an open circuit, while phases B and C continue functioning normally [36].



Figure 2. 7. Open conductor fault [33].

2.4. FAULT DETECTION TECHNIQUES

This section addresses fault detection techniques in electric transmission power systems, often composed of various renewable energy sources and fast charging stations. The objective is to thoroughly discuss the main features, challenges, and practical applications of these techniques relating to electric transmission power systems. The fault detection techniques must operate at high speeds; therefore, considerable precautions must be diligently adopted to avoid inadvertent tripping critical elements within the system. This approach ensures the entire system can function safely and reliably while maintaining the required quality standards. Moreover, selecting the most appropriate scheme to be implemented in electric

transmission power systems represents a careful compromise between achieving rapid fault detection and attaining a satisfactory level of discrimination in identifying fault conditions [37].

The state-of-the-art monitoring techniques currently being developed are specifically aimed at performing efficient fault detection and enabling quick recovery during normal operational conditions. To attain this critical objective, the timely interventions of protective devices, including relays, circuit breakers, and various terminal units, are necessary. Any lack of an appropriate intervention can lead to severe consequences for the electrical system's overall performance, resulting in greatly increased costs due to the system's inherent fragility. Additionally, this can lead to significant losses in profit due to restricted operational capacity or reductions in overall power generation output. Numerous fault detection techniques have been developed for application in electric power transmission systems. However, various factors can influence their implementation, including size, complexity, cost, and the techniques employed. Furthermore, the specific characteristics of the system structure, the type of voltage and power it handles, and its associated peripheral elements can profoundly impact the adoption and efficacy of these techniques [38].

2.4.1. Conventional Methods

According to the fundamental principles utilized for its operation, the protection system functioning is categorized into several specific sections: protection against short circuits, protection against earth faults, and protection against potentially dangerous over-voltages and overloads, among others. With a handful of exceptions, most power transmission lines today are classified as overhead lines. Given this important context, earth faults are particularly interesting, primarily due to the direct relationship with the initial conditions observed in the installation zone, specifically within a transmission or transformer substation [36]. It follows logically that the model implemented for simulating the power system under the analyzed circumstances is that of transmission lines outfitted with various types of parameters, taking into account a medium length of the line and also the potential for interconnecting the power system

with distributed energy sources, which adds another layer of complexity to the analysis.

The paper also presents an original and innovative concept of power transmission lines that can be effectively connected to renewable energy sources and fast charging stations specifically designed for electric vehicles. The fundamental approach taken from the article highlights the transposition to the classical protection equipment utilized in the power transmission system of the recommended fault diagnostic algorithms. These algorithms are an integral part of the new strategy encompassing embedded control, embedded diagnosis, and fault compensation, and all applied to power systems with a high level of renewable energy sources. This integration enhances the reliability and efficiency of the power transmission systems and significantly supports the transition towards a more sustainable energy infrastructure [39].

2.4.2. Advanced Techniques

The theoretical background utilized to formulate the experiments presented in this work is grounded primarily in an extensive study, along with numerous additional references that have emerged in recent years. The focus of these references is to apply several tools whose theoretical basis converges towards a good fault diagnosis that is both efficient and fault tolerant. This is an integrated approach with other disciplines and industries. In conclusion, most of the studies related to the research themes used in the present paper mainly focus on topics related to energy generation and utilization, energy distribution, modern electrical engineering issues, integration of RES, establishment and management of fast charging stations for EVs, devices in power electronic systems, and implementation of protection and monitoring systems. Moreover, these research topics encompass critical subjects for data processing, pattern analysis, and technical implementation of anomaly detection for power systems. Finally, these themes also encompass optimization strategies, management systems, and issues regarding critical infrastructures in contemporary society [40].

The advanced techniques are demanded of being an essential part of an effectively and efficiently filtering unit that approaches the nonstationary faults in three-phase electric high-voltage transmission power systems. This method is capable of addressing the complicated and usually difficult nonstationary faults existing in the defined three-phase electric transmission power systems where the computational costs throughout this dimension of treatment can be effective and reasonable as well as can be measured. These advanced techniques help in the realization of multiple real-time applications in different domains [41]. The minimum average run length and the sequential probability ratio test are just a few examples of optimized approaches which, by their implementation, considerably enhance and improve the classification of the protective devices' connection state in relation to the faulted line. Systems can now differentiate whether the connection is held or open in order to ensure better operation of the entire system.

2.5. FAULT CLASSIFICATION METHODS

DFCM extensively leverages the differences in energy crosstalk spectra from a suite of different classifiers, including variants of the artificial neural network family, to identify and classify faults in a power transmission line model. The method is based on one or more of the following aspects: the modelling of the power transmission line, the handling of the test vector sets specific for the classifiers, the fault spotting, and the accurate fault type and position classification. To effectively analyze these components, a 3-phase power transmission line is modeled in the time domain, as the analyses and categorization pertinent to the power industry heavily depend on the power spectrums and power spectral density of the state variables of line devices [18]. Typically, electromagnetic transient simulations that necessitate time-based methodologies are utilized. Simulation results have confirmed the method's validity, showcasing its promise as an effective tool for detecting and classifying faults and accurately locating faults in electric power transmission lines, along with a spectrum of other use cases. Looking ahead, future research will incorporate the effects of measurement inaccuracies into the simulations and aim to evolve the technique into a more robust, industry-standard system. This evolution will include applying

measurement data gathered in real-time and expanding the range of faulty line conditions analyzed, moving beyond just phase-to-phase and phase-to-ground faults.

2.5.1. Rule-Based Systems

As previously noted and thoroughly examined, rule-based systems have played a significant role in electrical engineering. These systems boast a strong mathematical foundation, bolstered by intelligent improvements to algorithms that efficiently and automatically detect, classify, and locate faults in electrical transmission lines. The primary aim of implementing rule-based systems in power systems is to address the limitations of traditional power protection mechanisms, which include prolonged fault detection times, the transmission of disturbances throughout the electrical network, and the lack of effective communication regarding recorded sensor data. A distinctive advantage of the employed rule-based systems lies in their capability to meet the increasing demands for future electrical energy dispatch, seamlessly integrating into the new framework introduced in this chapter. Furthermore, integrating sophisticated methodologies within smart models designed for online fault location of electrical failures is highly sought after. The deployment of rule-based systems should be executed similarly to robust systems; two effective approaches to accomplish this involve the implementation of advanced technologies that provide Anti-Windup Compensators within Input-Output Reflection (IR) systems or alternatives that circumvent the challenges associated with the BMI (Bilinear Matrix Inequalities) approach on high-speed, optimized CPU platforms [42]. This strategy aims to yield a resilient model with minimal coding requirements and, consequently, rapid response times, ultimately steering towards an optimal software solution. According to the literature surrounding rule-based systems, the inference mechanism operates based on rules articulated through conjunctive expressions presented in disjunctive normal form. In these rules, the antecedents typically align with the input linguistic variables, while the consequent specifies the output linguistic variables derived from the rule.

2.5.2. Machine Learning Approaches

Much work has been done recently using machine learning and pattern recognition as an alternative to classic methods for detecting and classifying power system faults, particularly in electric transmission power systems. The methods mentioned earlier are primary information, such as voltage signals, current, or other information obtained in the electric transmission power system. The time zone is very important. The advantages of artificial intelligence are that it saves much time and effort for manual analysis, achieves objective decisions, increases capacities with time, can deal with large data sets, and can analyze many data sources. Machine learning consists of techniques that form the central area of artificial intelligence [18]. The main machine learning models used with transmission lines are the Bayesian network, fuzzy neural network, Kohonen self-organized map, radial basis function neural network, support vector machine, evolving fuzzy neural network, and k-means algorithm.

Support vector machines are a set of supervised learning methods used mostly for classification, regression, and outlier detection. Their advantage is that they allow for the calculation of the shortest distance while considering all signals in the training. These features distinguish them from other classifiers. Neural networks have interconnections organized in fixed patterns, which help self-learning, thus making them adaptable. In these networks, one layer works as input, while another does the classification. There are also references to this layer of radial neurons. These two methods, support vector machines and neural networks, are also known as static, offline, and non-adaptive. It can be said that neural networks, compared with SVM, require more iterations for convergence than support vector machines. A disadvantage of SVM is that it does not offer a direct probability explanation; only local probability estimates exist for each class. A significant advantage of SVM is that it focuses on the separation error [32].

2.6. INTEGRATION OF RENEWABLE ENERGY SOURCES

With many renewable energy source technologies that require electric energy conversion for their coupling to the power system, this system has lower security. They

are often uncontrolled power sources with significant dynamics, and they also depend on unpredictable atmospheric conditions. This implies that electric disturbances and failures will tend to occur more often. It is foreseen that the power system operation mode will have to undergo significant transformations to maintain the required quality of the supplied energy continuously. It can be said that combined electricity transmission systems will have to be developed. The objective of these new designs will be to combine several sets of principles, ensuring the quality of electricity supply at acceptable and competitive costs, such as the safety and performance of the equipment. These combined systems could, for example, include HVDC and HVAC technologies, with and without overhead lines, and the option of incorporating energy storage systems [23].

In addition to these technologies, data integration and management systems are required to assist power system operators in the decision-making process. This type of system should also make it possible to consider intermittent production characteristics, using the hierarchical structure within the power system, to improve security. The need to treat volatile renewable electricity sources, such as wind, solar, and others, is a factor weighing increasingly heavily over the power system. However, this is not the only change that has become more common in an integrated electric power supply system. Electric vehicles have gained greater relevance and deserve to be recognized. Due to the current performance of accumulators for these vehicles, there is an interest in increasing the number of charging stations. In this area, bi-directional systems are emerging that can interface with the general power system, assisting both in stabilizing its operation and, in a certain way, compensating for voltage sags [43].

2.6.1. Solar Power

Small photovoltaic sources are rapidly becoming a significant and expanding segment of investments in renewable energy globally. These smaller generation sources primarily connect to the power grid at points that are essentially considered load connections. While small solar power sources represent merely a small portion of overall energy generation compared to larger solar facilities, they present unique challenges in management within transmission and distribution systems. This is

especially true in operational distribution networks, which can be complicated by introducing numerous smaller inputs [43]. While many researchers focus their efforts on ensuring the successful and efficient connectivity of large solar power plants to high-capacity transmission networks, a substantial body of work is dedicated to optimizing the integration of smaller photovoltaic systems into low-voltage distribution networks. In various countries, distributed generation technologies, including photovoltaic systems, are not typically required to compensate for their usage of the distribution power grid. This situation often leads to new financial pressures whenever there is a need for further investment in the expansion or upkeep of the distribution infrastructure. Integrating collective solar systems, like those found on buildings or carport structures in commercial centers and community residences, involves myriad consumers that collectively increase the risk of faults within the power grid. If these collective systems were to connect as individual consumer units, the likelihood of grid failures could rise dramatically, particularly if the distribution network's necessary operation, maintenance, or expansion efforts are not adequately executed. This interconnectedness underscores the importance of addressing the complexities and responsibilities of integrating small renewable energy systems [43].

2.6.2. Wind Power

Registered power in wind sources refers to the maximum power generated by the wind power defined by its manufacturer and is defined as per its design parameters for the power generated at its respective wind speed. The generator operates continuously so that the mechanical part of the wind turbine can feed the generator the wind-generated power, which is generally indicated by the mechanical power maximization characteristic curve known as the P-w characteristics curve. This registered wind power is almost constant for a much wider applied torque, and the wind speed variation does not significantly impact the generator's mechanical power. If the wind speed level increases beyond a particular value, then a stall occurs, and the generator should be safely shut off to remain within its design specifications. With magnetic saturation, emerging anomalies develop into faults under the aileron mode, caused by severe mechanical wear, and they are ultimately torn into shrapnel. This could cause heavy damage to the gearbox if ejected at high speed[43].

Protective measures must be taken into account if both extreme winds occur as forecasted and the fan speed curve exceeds the power registered in the mechanical power maximization curve. The main purpose of the presented early wind generator limit detection model is to automatically recognize the highest generator power value allowed by the actuator/multi-pitch controller combination for each wind speed value. The wind power may then be continuously maximized for a wide performance curve range without the risk of damage to the wind turbine and the associated gearbox. Minimum tip speed control is employed for the maximum power point tracking designed to keep the turbine on the mechanical characteristic power curve. With a Kalman filter on the real and reactive power, the cascaded control can update the proposed control steering for fast and precise control action. Using a Kalman filter method, a conspicuous nameplate power of $P_m = 1 \text{ MW}$ is obtained [44].

An effective control strategy utilizes pitch control and tip speed ratio as key parameters in the context of variable-speed wind turbines. Enhancements in control actions and fault tolerance performance can improve efficacy in the bagging scheme through additional features like model predictive selection. Pitch control actions are validated via simulations and experiments using a proportional-integral-derivative controller. Applying a finite fault-tolerant control approach significantly enhances the performance of wind turbines when exposed to actuator faults. There is a demand for robust backstepping sliding mode techniques and certification processes to implement and refine this control strategy. Parallel-series compensators, along with PPF and PSC, are introduced. In drivetrain torque faults, the emphasis of adaptive robust fault-tolerant control lies primarily on robustness. The proposed fault-tolerant control's feasibility is integrated into the operational system, where backstepping sliding mode methods are applied to regulate wind turbine speed.

2.7. FAST CHARGING STATIONS

New electrification business models have emerged with the increased share of electric vehicles in road transport and on the horizon of a fast energy transition towards the sustainability of stationary energy supply with distributed generation based on

renewable sources. Due to the urban lifestyle, the growth in the number of electric vehicle owners, and the development of batteries, especially energy storage systems integrated into the vehicle, the charging time and quality aspects are of interest, especially in the case of business mobility over long distances. Advanced vehicle technologies equipped with lithium-ion batteries and power electronic converters can charge transparently in 30 minutes up to 80% of the battery capacity, with load powers of about 150 kW. The charging time mainly depends on the communication and control infrastructure of smart charging stations. By integrating quick energy storage systems based on supercapacitors or batteries with long life and high-power storage systems of larger dimensions based on flywheels, the time can be reduced to 10[44].

2.7.1. Technologies and Standards

The reduced dependency on oil and the need to reduce greenhouse gases have led many nations to apply steady policies to foster renewable energies to diversify their energy matrix and promote economic development. In many places in the world, the strong growth of the wind and solar industries has been leading entrepreneurs of these sources to sell this energy to electrical systems, which can offer higher reliability than other sources. Therefore, in electrical systems with wind and solar generation, fault detection and classification must be time-efficient and should be integrated with renewable energy sources installed. The protection speed in electrical systems has been long described in standards; whereas standards have been established in some countries to help the utilities of these countries increase the speed, accuracy, and integrity of their data and thereby achieve a higher level of transmission system performance [44].

Most of the world's international electric energy trade happens in the High Voltage Direct Current transmission system. The world's interconnected electric grid uses HV lines and substations to send and receive power from/to the synchronous electrical systems to allow the required energy and power support to the loads and the safe operation of the generators. The integrated application of transmission lines and Voltage Source Converter into a single Transmission System is used to mitigate issues of Flexible AC Transmission Systems [45].

2.7.2. Impact on Power System Operations

Reduced costs and simplicity of use of renewable energy sources have led to small centralized renewable units and large permanent units, increased LV level connections, and distributed storage units. Traditionally, the LV grid is a passive deposit with only unidirectional interaction between the source and the load. However, massive renewable energy sources connected at that level have transformed the LV grid into a more dynamic form with two-way power flows. Traditional grids have always been finely calibrated for the loads measured yearly and have always accommodated backward power flow. However, with the massive connection of renewable energies, the number of sources with power flow has increased both at MV and HV levels and at LV levels [46].

The injection of these new sources with power flow capability means that although load remains the most sensitive factor to protect, renewable sources also become critical elements that make it necessary to rethink the protection architecture of the current network and, in this context, with the help of new technology. The power system usually operates unattended for long periods of time, and under normal operating conditions, very few acts by the operating center need to be performed. Information is collected, decisions are made, and orders are executed to act on the devices that execute the required actions. However, when a fault occurs in the power system, a range of events dictate the levels of action required of the system. As the large amount of renewable units and the increasing number of devices connected to the distribution grid connection has been carried forward, malfunctions and protections have also increased [47].

2.8. FAULT MODELING IN RENEWABLES AND FCS

The renewable energy source, referred to as distributed generation, is accurately represented by a controlled voltage source with the essential ability to regulate and restrict the total active power exchanged between the renewable source and the electrical network. This source is managed through a complex gain system, encompassing the power and voltage specifications required to satisfy the load's

demands. In this scenario, voltage-source converters display both quadratic and derivative components in the correlation between the complex power produced by the converter under the active-reactive power, or PQ, control strategy and the line currents along with their corresponding line voltages. An initial straightforward analysis of the relationship between the odd harmonics present in the voltage-source converter line current spectrum and the fault-related injected currents introduces a new angle-reflected power algorithm. This algorithm operates by recording the harmonic currents for each grid phase. Following this step, a calculation is executed to assess the reactive power exchanged by the renewable energy source [46]. Ultimately, the angle and reflected power are precisely calculated to pinpoint the faulted phase within the system.

This research paper examines the use of fuzzy logic decision systems in accurately determining the faulted phase in electrical systems. It focuses on calculating symmetrical components based on measured electric currents and voltages, which act as the essential input signals to the neural network integration. In each step, an auxiliary process is running: a fuzzy logic classifier, a fuzzy system responsible for phase identification, and a fault-specific current sequence, improving the classifier capabilities of the neural network. Next, the signals obtained from each of the scenarios are needed for the complex Gabor wavelet normalization. A complete set of simulations is carried out in this regard at different topographies of ground faults, at different distances between fault occurrences and data measurements points, and at various angles of fault occurrences. In particular, preliminary results from the simulations will be presented describing a distribution network found in an industrial park, which has been recognized for utilizing the aforementioned renewable energy sources [47].

2.8.1. Dynamic Modeling of PV, Wind, And FCS

Integrating renewable energy sources, including photovoltaic and wind energy, and various power system equipment is crucial in comprehensively analyzing electric transmission power systems. This analysis encompasses critical areas such as fault detection, power quality, and voltage level stability. In exploring the models linked to

these contributions for fault detection and classification, we configure them through meticulous fault behavior analysis that derives insights from voltage variations observed in the transmission partial systems. For photovoltaic systems, establishing a quadrangular characteristic curve is essential for both operational and modeling purposes. This characteristic curve plays a significant role in ensuring that the system functions optimally under varying conditions. The reverse equations covered in this work are often used equations in PV systems; however, only the direct equations are the focus of the research presented in this paper. Thus, the photovoltaic model is developed in a way that it can be extracted directly from experimental real-world data gathered from installed systems. When modeling the photovoltaic system, two incompleteness factors are the focus. The first of these factors pertains to the upper power packet, which is limited by the maximum power that can be commanded from the corresponding power electronics. This restriction is important to ensure system reliability and performance. Conversely, the output power packet's lower bound is fixed when solar irradiation is low, close to zero (an estimated 30% of the nominal power emission) [48]. This above fact facilitates the fulfilment of appropriate operational parameters during time frames of low solar production, ultimately serving to create a complete framework under which the analysis of PV system behaviors within electric transmission networks may be considered.

The wind turbine studied in this work [49] is a fixed-speed synchronous generator with DIST conversion, and it follows a steady state three-phase power model. It uses the kinematic wind equation, and the aerodynamic is optimized specifically based on the cost of the turbine characteristics. The model consists of five wind phenomena that develop throughout the change of wind. Each of these stations is a modern steel structure with high resolution chargers and inverters inside, and great care has been taken to ensure the inverters can safely and efficiently minimize charge times. Moreover, the FCS model is a complex model of loads transmitted as it acts in on-grid and off-grid situation by using the power converter. The Battery Energy Storage System (BESS) incorporated within this framework serves as an essential component, providing the functionality of electric vehicle charging energy exchange, which is increasingly vital in today's energy landscape. Considering that this work presents an innovative fast charging station model, it is significant to note that no prior pre-

modeled fast charging stations are included or referenced under this study, particularly in system dynamics. Consequently, this research not only enhances current understanding but also paves the way for future advancements in renewable energy and electric vehicle charging infrastructure [49].

2.8.2. Load Fluctuations and Renewable Intermittency

One of the most complex aspects of the integrated bidirectional fast charging points is the respective imposed regulation for guaranteeing the reliability of the electric transmission power system. The popularity of electric vehicles is steadily growing, propelled by the increasing global attention on the significant issues surrounding global warming. Consequently, there is a trend where more stringent regulations are being implemented to effectively supply and control the fast-charging power points to ensure their efficiency and reliability. The large charging powers combined with the brief fast charging windows result in highly distorted load current patterns, making it quite challenging to control and regulate the detected peak power levels effectively. Due to the integration of decentralized renewable energy sources, integrating bidirectional fast charging points represents a promising alternative with notable technical and economic advantages, which do not hinder the widespread deployment and utilization of these charging points across various regions. Furthermore, the towing vehicles' bidirectional power capabilities could be harnessed to develop innovative new services and products in conjunction with the distribution of fast charging points, thereby creating new opportunities for enhancing the overall user experience while promoting sustainability in the transportation sector [50].

The output fluctuations of solar panels and wind turbines can change swiftly from peak production to the lowest minimal output levels, primarily due to factors such as cloud sharing and wind gust enhancement, among other environmental influences. Understanding and managing these variations is crucial for ensuring optimal energy production from these renewable sources. The maximum available power is vital in maximizing the total output power that can be harnessed from these sources. However, achieving this requires significant additional investment, increasing the stored energy cost. Importantly, during the hours when the variable renewable energy generation

sources are not producing, the consumption value of this maximum available power diminishes at a consistent rate. This aspect highlights the interconnected nature of renewable energy generation and consumption patterns. Meanwhile, the minimum available power can be defined as the unacceptable threshold value for electric systems to maintain stability and reliability. This threshold will be a key cost-driving parameter generated by the integrated bidirectional fast charging point that is used as the generator in vehicle-to-everything mode, bringing additional complexities to energy management and distribution systems [51].



PART 3

MACHINE LEARNING AND DEEP LEARNING OVERVIEW

The emergence of artificial intelligence has positioned deep learning as a central theme in this new era of AI technology and research. The primary focus of this field is on "learning" and "building," where the process of understanding the machine's essential layers contributes to the model's learning framework. This collection of learning layers defines the "depth" of the model, leading to the terminology of deep modeling technology, known as "deep learning." From a computational perspective, the term "deep" refers to manually disentangling the machine signal representations, which is carried out through various learning strategies such as supervised, self-supervised, unsupervised, or reinforcement learning. Initially, the role and potential of the intuitive connectionist model, specifically artificial neural networks, is recognized as the foundational element of deep learning. In deep learning models that utilize artificial neural networks, surrogate low-dimensional mappings are processed through essential activation functions or efficient non-linear mappings inherent in natural neural networks. The chaining capabilities that arise from combining multiple layers with artificial neural networks are fundamental to the current advancements in "deep learning" technologies. The connectionist frameworks of these neural networks closely resemble the biological neural networks found in the human brain. Recent research has further corroborated the neural basis underlying the dynamic nature of the brain, where the functional dynamics of neurons and the modeled geometric structures of these "natural neural networks" are leveraged to comprehend the brain's non-linear functions and provide advanced medical interventions for individuals with existing "brain faults" [52].

3.1. OVERVIEW OF MACHINE LEARNING AND AI TECHNIQUES

One of the most rapidly developing technological advancements of the past few years has been the rise of artificial intelligence (AI). Figure 3.1 shows a brief guide to AI and Machine Learning Methods, and as a subset of artificial intelligence, machine learning, or the application of statistical learning methods to construct models that can make predictions or classify, is also becoming increasingly popular. Artificial neural networks, a computational system that mirrors biological and human brains, are pillars of deep learning and machine learning. Despite artificial intelligence and other subsets of it, such as deep learning and machine learning, seemingly associated with their problem-solving skills, these subsets are also crucial in solving problems such as credit card and transportation fraud, analyzing data, building chatbots, and visual and voice recognition.

Other notable machine-learning techniques include Linear Regression, Nearest Neighbors, Decision Trees, Random Forests, Gradient Boosting Machines, and others [53]. Courses, training content, and educational institutions have also made strides in including machine learning topics to stay competitive with the times. Various training content that is suitable for students, business professionals, IT professionals, people with no programming experience, or people who have basic Python programming knowledge has made it easier to learn machine learning using mediums such as videos, textbooks, assignments, problem-solving projects, quizzes, solutions, and verified certifications. This saves a lot of time and money for institutions and students alike, with easy-to-use interfaces, forums, and direct interaction capabilities between problems and solutions, thus facilitating a more interactive and immersive learning experience.

Due to the increasing application of machine learning in various branches of science and technology in the last several years, among which are computer vision, speech recognition, and data analysis, to name but a few, it has gradually become a technological requirement. Deep learning is a hot term, and many military organizations and various technological companies are investing significant money in developing such systems. In such large investments, it is hard not to mention the role

of deep learning in these AI systems. The remarkable success of deep learning, especially in tasks such as object recognition, has led to enthusiasm and the use of similar methods in many critical areas with much practical application. With these works, we describe the historical consideration of these systems in detail [54].

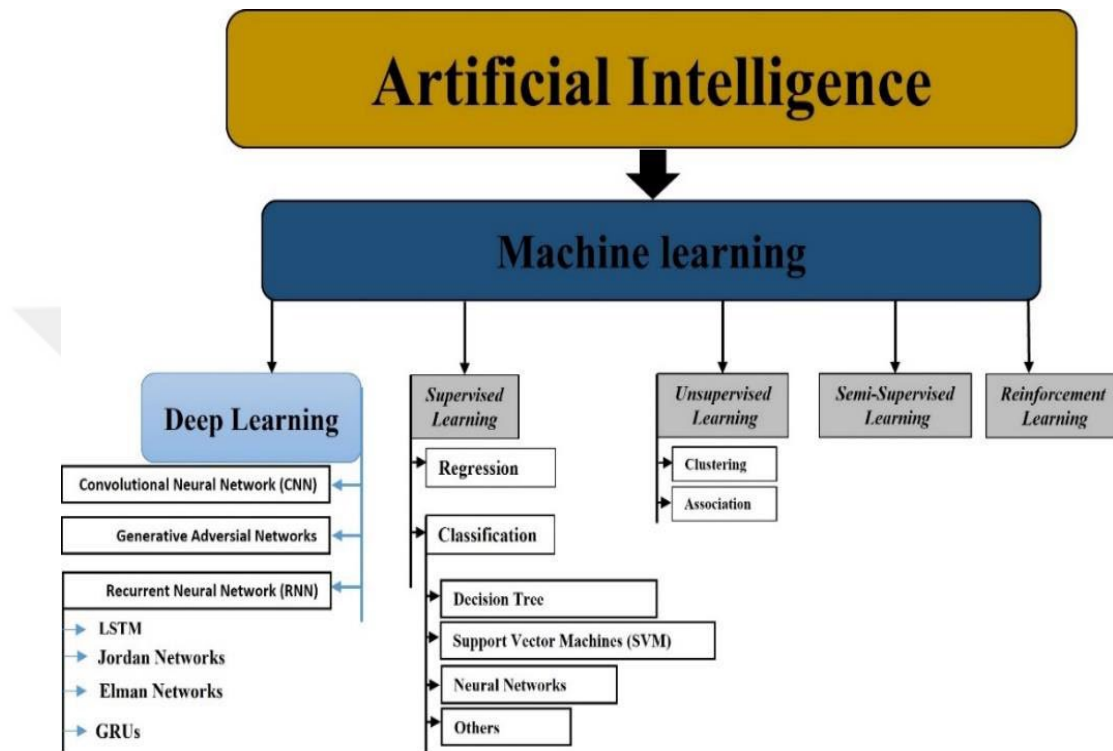


Figure 3. 1. A brief guide to AI and machine learning methods [51].

3.1.1. Advances In Artificial Neural Networks

Artificial Intelligence (AI) has been achieving major breakthroughs in recent years. The great advancements that have been made are largely due to Deep Learning (DL) techniques, particularly deep Artificial Neural Networks (ANNs). Deep learning has successfully solved many challenges that seemed infeasible only a few years ago. Advances in deep learning have redefined how we explore and exploit technologies to address existing and new business-related challenges. Horizon scanning studies, as well as those that examine the global policy issues of various forms of AI, continue to advance. Examples of these complex systems include those that involve topics such as versatile and capable artificial intelligence and emergent artificial intelligence [51].

Neuroscience has had a considerable impact on human society. Over the past three hundred years, research has focused on simulating the brain to achieve advanced computational outcomes while using limited resources. The breadth of these investigations is expanding, with the potential to significantly influence artificial intelligence (AI). As brain-simulation technologies increasingly overlap with the broader AI domain, researchers from diverse fields, each offering distinct perspectives on data representation, have started collaborating to advance AI research centered on artificial neural networks. These cross-disciplinary initiatives have resulted in noteworthy progress in AI by integrating various computational models. Consequently, simulated neural elements have developed into Artificial Neural Networks (ANNs) [55].

The development of ANNs presents both engineering and challenging theoretical issues. This research discipline, which promises to be one of the cornerstones for the future of AI, is currently experiencing a growing resurgence. As engineers do not fully understand how ANNs parallel the human brain function at this time, a variety of discipline-oriented target concepts have been used with a focus on providing practical solutions. As a practical matter, focusing on the capabilities of the ANN on a particular problem area and mapping high-level representation to a specific network architecture allows for the development of successful working ANN models. Several ANNs are used to work and solve daily problems across many fields [56]. In this context, research in both hardware and software is important. The mathematical foundations that underpin these types of learning are an essential prerequisite for AI algorithms, techniques, and systems to become soundly reliable.

3.1.2. Evolution Of Deep Learning Techniques

The field of artificial intelligence, in general, has caught the attention of the scientific community and the industrial capitalists. Artificial intelligence products have mass-migrated into the general financial markets in the past decade. Innovations in artificial intelligence have made major contributions to the fields of robotics, computer vision, natural language processing, and so on. Technological improvements have allowed the

creation of simulation environments capable of operating similar to and in a manner comparable to human functions. Thus, technology has blazed the trail, making possible the aspirations of developing a peripheral field that spans artificial intelligence, deep learning, and operations research to attain the ultimate goal in the future [54].

With the idea of deep reinforcement learning being comprehended by looking at genetic connections, this section will encompass the mode of operation of artificial neural networks from their inception. Furthermore, the progression of novel learning and the development of the rebirth of legitimate artificial intelligence techniques and algorithms is scrutinized. Additionally, how various forms and stages of neural networks provide learning done on a deeper level is revealed. Following this, the convoluted method of properly harnessing the hidden potential of deep reinforcement learning is delineated. Lastly, an overview of the progression of artificial neural networks secondary to other methods is highlighted [55].

3.1.3. Motivation For Deep Reinforcement Learning

In recent years, significant progress has been made in artificial intelligence. Notable advancements can be found in natural language processing, predictive analysis, and computer vision. Generally, these AI sectors are what allow chatbots to be trained to have natural, human-like conversations, and they contextualize a series of images into a story in one word, respectively. A major reason behind the rapid progression of these fields is the study of artificial neural networks, especially deep learning architectures. However, one key sector that had not seen progress until recently was physical systems' simulation, prediction, and control [56]. This challenging subdomain directly impacts embodied systems, like robotics, that require constant decision-making. Although traditional methods existed for specific problems, the development of general methods has only recently broadened. This is due to the development of deep reinforcement learning, which utilized deep learning in conjunction with reinforcement learning to solve problems in this sector.

These problems were traditionally addressed using optimization-based techniques. However, significant progress was made in this field only after the publication of a

series of successful results. Major innovations were made in this decade-long field with two main motivations. Firstly, it has been observed that neural networks can approximate optimal functions with arbitrary accuracy. This has led to the understanding that neural networks are approximators of value functions central to reinforcement learning algorithms. Secondly, training deep learning systems become more efficient due to several advancements and factors spread over the years, including developing new hardware and acquiring and labeling large data sets. These factors play a fundamental role in making deep reinforcement learning a main focus of today's research scene [57].

3.2. FUNDAMENTALS OF ARTIFICIAL NEURAL NETWORKS

An artificial neuron serves as the fundamental building block of an artificial neural network. It represents a simplified abstraction of a biological neuron. Biological neurons perform a rich array of behavioral functions in the brain, encompassing diverse processes such as signal integration, noise filtering, spontaneous pattern generation, and even the crucial maintenance of rhythm. In contrast, a single artificial neuron is designed to emulate only specific computational functions akin to those of a biological neuron, ideally focusing on the aspects that are of particular relevance and practical interest in the context of the learning process. Consider an artificial neuron with p inputs, weight $w_0, w_1 \dots w_p$. This wiring allows the neuron to function properly within a neural network structure. The artificial neuron can also be more abstractly represented by it having some inputs $(x_0, x_1, \dots, x_p)$ that doesn't activate its input gates that is going to be applied. The neuron applies a weighted sum to each input, which is passed on to an activation function, to calculate an output. Since this activation enables the neuron to model a broad family of nonlinear input-output functions, it is a very powerful building block [53].

This linear combination of weights and inputs, net , is determined by the equation:

$$net = w_0 * x_0 + w_1 * x_1 + \dots + w_p * x_p$$

The output of the neuron, y , is determined by the equation:

$$y = f(net) = f(w_0 * x_0 + w_1 * x_1 + \dots + w_p * x_p)$$

3.2.1. Neuron Models and Activation Functions

Neural networks have been studied before (and still) around many tasks, but the models differ mainly in the number of layers. These patterns have helped to develop the neuron unit model for more flexible, larger, and more complex neural networks. The unit model consists of taking the weighted combination of the unit inputs, transforming this into output through the activation function which is also called the neuron model. As researchers strived to establish models that would exemplify the learning process, the understanding of the actual brain and the neurons and synapses that make it up led to the development of models closer to real simulation. The default reference neuron model is a model made of seven components that represents the biological neuron theory with various activation functions like Heaviside, Rectified Linear Unit as well, as Tanh , and one can apply it to mimic and characterize the behavior of different neurons, and the model is not required to have a biological basis [53].

The usage of more than one activation function can be particularly effective in enhancing the performance of neural networks, and there are both functional advantages as well as some disadvantages to using a variety of different activation functions. In the context of artificial neural networks, one significant problem that has seen an increase in focus is the construction of a uniform model that accurately simulates the learning processes observed in different brains, which is inherently reliant on its unique characteristics. In such complex cases, in addition to the chosen activation function, the processes of tuning, data normalization, learning parameters, and the selection of model optimization algorithms played an extremely important role. They constitute some of the key aspects of the processes that are vital for the concrete development of practical artificial neural network models that are meant to delineate and faithfully mimic the intricate mechanism of learning. Also, before getting to such critical processes it is even more important to have architecture and learn what data you need to learn that affects your ensemble learning procedure [58].

Hence, the proper examination of these components is crucial to attain the best performance and reliability of ANNs.

3.2.3. Feedforward and Backpropagation Algorithms

The feedforward algorithm and the backpropagation algorithm are the most frequently used algorithms for feedforward neural network initialization, training, and testing [52]. Often called the forward calculation or forward propagation of the activations through the network, the feedforward algorithm is used primarily to compute and find the output given fixed or already optimized network connections. Using this algorithm, we can know the predictions of the network for some input values. This is an essential step to making a working network operational. Conversely, the backpropagation algorithm is used to train the network and fine-tune the weight of links as well as the biases, by using the feedforward algorithm to feed-forward and calculate the error to be sent back via the existing connections of the network during the training runs iterations. This is absolutely necessary for the network to learn. To use the feedforward algorithm effectively, the weights of links connecting an input layer, hidden layers, and an output layer must all exist or be properly determined and stored. It is important to maintain and optimize these weights to ensure that the network works in a proper and accurate manner [59]. The very first layer (input layer) processes certain known input patterns and predicts output through the layers of a feedforward neural network model. During this phase, output activations have emerged from the operations of the input layer. The different inputs are then fed into the neurons of the first hidden layer, which apply a transfer function to the weighted sum of the applied inputs. This sequence of operations plays a crucial role in transforming the input data into meaningful outputs. The last hidden layer or output layer takes all the incoming signals from hidden layers, applies the transfer function on the weighted sums of those signals, and propagates the activations onward to the output nodes. It shall only be when all the transfers of signals have been conducted clearly and thoroughly through to the output layer that the final network response to any given set of inputs becomes apparent. The feedforward algorithm demonstrates the crucial interplay between the two algorithms in producing the output of the

artificial neural network associated with this flow of information and activations, which manifests ultimately as output.

3.3. ADVANCED DEEP LEARNING ARCHITECTURES

The previous sections discussed the role of ANNs and CNNs, showcasing how they have been used to solve problems concerning vision, speech, and image classification. In this chapter, we will take a closer look at some advanced deep learning techniques [52], including but not limited to artificial (and many types) recurrent neural networks, temporal (or time-based) convolutional neural networks, autoencoders, generative adversarial networks, and long short-term memory networks LSTM, a specific subtype of RNN. It explains and discusses the core characteristics of such advanced ANN architectures and describes their real-life usages. These databases and usages cover several generations, from series modeling through image generation to temporal analysis. RNNs are meant to solve problems related to sequences, commonly referred to as sequence prediction or generation problems, and have been established as the go-to solutions for such types of problems throughout the years [60]. So, this section aims to prepare the learner to employ recurrent neural networks for appropriate tasks. It learns about sequence problems in machine learning for LSTM networks, Wavenet, temporal convolutional neural networks, other advanced RNN models, and more complex sequence prediction problems. Readers coming to this chapter are expected to have a good basic knowledge of Python and deep learning. Overall, an RNN can be used for general prediction defined by types of prediction, such as next time step prediction, future time step prediction, future time step type prediction, and even comprehensive time-step predictions. Since RNNs can predict the next observation based on previous observations, this is useful for time series prediction. They may also perform prediction of future time steps when the temporal patterns are identified or detected. Moreover, RNNs are capable of predicting the expected type of item of a future time step given a previous sequence, a process known as time step type prediction. The last (i.e., the first last) class of prediction tasks of which RNNs is target to predict every time step of the sequence [60].

3.3.1. Convolutional Neural Networks (CNNs) and Applications

A limited but promising front in applied AI and a class of deep learning models with the name “convolutional neural networks,” contributed much needed insights into the computational structure, design and architecture of the human visual system. CNNs have proven to be revolutionary, particularly in image recognition, and they laid the foundation and offer inspiration for more complex neural networks in the future. Convolutional networks operate on image data by applying layers of trainable filters to detect increasingly complex features, making the concept suitable for other similar data types. It was observed that the human brain decomposes objects into distinct parts at multiple levels of complexity, ultimately forming an even more complex whole object recognition process. The inspiration behind the architecture of CNNs stems from this concept. A CNN consists of various types of layers including the input, convolutional, pooling, fully connected, and classification layers connected through neurons [52].

The convolutional layers play a critical role in CNNs, where they apply learnable filters during the training phase, with receptive field extents over the input data to construct feature maps that produce local representations. The challenge in deep learning comes from determining the oversized number of weights from the many connections, as the volume of neurons in each layer relative to each other expands across the network architecture. Researchers have alleviated and even eliminated the challenges encountered in deep learning via the introduction and training of deeper CNN architectures. These architectures act as informative feature extractors and discover abstract semantically composed features of image data across numerous convolutional layers that none of the previous attempts have hitherto attained [61].

3.3.2. Recurrent Neural Networks (RNNs) and Their Variants

Recurrent Neural Networks (RNNs) are a class of artificial neural networks that exhibit temporal dynamic behavior through cycles of connections given by the use of a mirrored hidden layer in the architecture [61]. RNNs are capable of handling sequences, and these sequences unfold the network in time, wherein they can

demonstrate iterative behavior, meaning they demonstrate dynamic temporal behavior where the gradient of the network is forwarded through time. This makes the architecture suitable for handling time series data. RNNs share similarities with feedforward neural networks, with input layers, hidden layers, and output layers in the architecture, with a major exception wherein connections can be made not only in the feedforward direction between input, hidden, and output layers, but can also be directed backwards, allowing feedback to be passed through the hidden layers, enabling recurrence in the architecture. Fundamentally, feedback connections are necessary to make the network invariant to time. Theoretically, they enable the network to deal with and memorize sequences beyond arbitrary constant length, which can be seen in purely feedforward neural networks.

3.3.3. Generative Adversarial Networks (GANs)

Generative Adversarial Networks are an extension of Generative Models, and they are generally comprised of two sub-networks. The first network, known as the generator network, is trained to generate new data that is similar to the original training set data. The discriminator network, the second network, after being trained, gives a judgment score from 0 to 1 if the sample is from the original training set (real sample) or is generated by the generator network (fake sample). The discriminator network aims to produce a score of 1 for real samples and 0 for generated samples. Thus, in GAN, the discriminator's goal is to minimise the objective function while the generator does the opposite. The goal is to reach the point where both networks can better imitate the original input in terms of distribution [62].

In GANs, the goal of the generator network is to deceive the discriminator network as much as possible by producing fake samples that are generated from a distribution similar to the true data distribution. That is, the generator tries to minimize the function that measures cross-entropy for the samples classified as real by the discriminator. As for the discriminator, its objective is to minimize the probability of misclassification of generated and real samples. The GAN that derives these objectives and definitions is called GAN.

3.4. TRAINING AND OPTIMIZATION IN DEEP LEARNING

Training is a critical step in a deep learning network's lifecycle, and several techniques have been developed to complete this process effectively. Most of these techniques revolve around optimizing the network capacity by minimizing the loss function over the entire dataset. Different optimization techniques and algorithms are applied to minimize the loss function. For an optimization algorithm to be effective in training deep neural networks, it should not only minimize the loss function quickly and efficiently but also not get trapped in local minima [52].

Batch optimization represents the application of the optimization algorithm using every single sample in a batch. Several optimizers apply batch optimization techniques to train the neural network, and stochastic gradient descent is one of the most popular. However, as the batch gets larger, the speed of batch optimization becomes much slower. This gives us a trade-off between the model's accuracy and the model's speed. In the case where the dataset is very large, reaching the minimum using batch optimization could become very slow, and therefore, deriving critical properties, such as age, is simple yet could lead to more errors in prediction. Mini-batch optimization addresses this issue by simulating batch optimization for the first few epochs and then performing batch operations. This provides fast and accurate results and efficiently uses available computing power[63].

3.4.1. Loss Functions and Regularization

Our networks want to learn some underlying structure or pattern within our available data. There are various techniques for making this happen: simpler models, more complex datasets, or well-chosen hyperparameters. Being quite fundamental in deep learning, loss functions and regularizations are paramount for correctly shaping the learned parameters and making our network generalize well enough to new, unseen data. Whenever we talk about something that our model is supposed to do, like regression, the loss function then determines how to estimate how close our outputs are to the expected targets [60].

Loss functions can also be viewed as numerical encodings of a model's performance at various training points. These metrics then imply whether the training process is going well and can even trigger some hypermodel optimization tricks. Because useful model design patterns include using previously proven loss functions as model errors, keeping them under fast review when applying deep neural networks to different real-world or industry-driven use cases is useful. The abuse of heavy loss functions just because better results are reported, without careful thought about what those loss functions do, is a typical mistake made by researchers and practitioners in the field of deep learning. There is no one-size-fits-all solution for loss functions, and careful thought about which real problem a certain loss function is meant to solve is sometimes beneficial for successfully applying deep learning in practice [61].

3.4.2. Optimization Techniques: SGD, Adam, RMSprop

3.4.2.1. Stochastic Gradient Descent (SGD)

SGD, or Stochastic Gradient Descent, has emerged as the most popular optimization method primarily due to its simplicity and robust theoretical guarantees. The foundation of SGD lies in its ability to find a minimum of any differentiable function, which is critical in various machine learning applications. Its design not only makes it a practical choice for a wide array of tasks but also particularly advantageous when dealing with large-scale datasets that are often encountered in real-world scenarios. SGD is made efficient by only storing and processing small random batches of training examples. This enables the algorithm to take steps towards a more optimal solution quicker whilst updating model parameters on a smaller subset of the data at each iteration. A variant closely related to SGD is mini-batch SGD and it has become one of the most commonly used optimization techniques in training deep networks efficiently. Conceptually, mini-batch SGD chooses a small batch size, denoted as 'a' which is the number of training examples picked randomly from the entire dataset. Once we pick this mini-batch, we now calculate the loss function and its derivative for this batch. After we compute the gradient with respect to the parameters, we will update the model parameters. This process is represented by the equation:

$$\theta(t+1) = \theta(t) - (\text{learning rate}) g(\theta(t)).$$

Here, $\theta(t)$ is the value of the parameter in the t -th iteration, and (learning rate) is a small positive constant representing the size of the step we will take in the direction of the inverse of the gradient of the loss (the $-g(\theta(t))$). This iterative approach of gradient descent provides a quicker means of navigating the loss landscape and backtracking towards optimal parameter convergence [64].

3.4.2.2. Adam Optimizer

Yet another stochastic optimization algorithm that has received a remarkable amount of research interest in light of its application to machine learning and deep learning is the Adam optimization algorithm. An adaptive learning rate optimization algorithm that aims to take advantage of the strengths of both AdaGrad and RMSProp. It allows deep learning algorithms to achieve impressive performance through the novel use of learning rates that dynamically change with every pass, and that are individually tuned per parameter that must be optimized. The name of the method itself can be problematic, as it is named largely after its key sense of operation, which is where momentum-based methods and RMSProp methods are combined, rather than either method being exclusively adopted. It is specifically designed based on properties related to the use of adaptive learning rates per parameter, as well as momentum parameters. This is somewhat similar to the RMSProp algorithm, but the key difference is that Adam takes into account the momentum term to adjust the bias of the parameters. This correction takes place similarly to bias correction methods used in RMSProp [62]. Because of this advanced alteration, Adam always outperforms RMSProp by a wide margin in diverse deep-learning tasks. The algorithm successfully handles the decay average of previous gradients and keeps the second decay average of past squared gradients. Such averages, in turn, are typical components of momentum terms and give rise to the diagonal matrix which is used to normalize the gradients in the optimization, thereby improving the convergence rates and stability. And this capacity to tune the hyperparameters while deep in training is why Adam is such a touchstone for practitioners using large-scale deep-learning models in varied contexts.

3.4.2.3. RMSProp

This is an advanced optimization readers trying to find the minimum of a well-differentiable function. Although this technique is particularly useful in the large domain of deep learning, it is most widely applied to training recurrent neural networks. When it comes to in-text classification tasks, RMSProp is significant as it is used to predict the relevant subdocument in a larger document. The differentiating factor in RMSProp is that it keeps a running average of the squares of the gradient. This, therefore, acts as a weighted average, which it can use to divide that particular parameters learning rate by the average of those accumulations [61]. Thanks to the core methodology of exponentiated moving averages, RMSProp achieves a dynamic evolution of the learning rate using the decaying average of past squared gradients. This average is computed with exponential decay and has a crucial impact, as it is directly used in normalizing the step-size term of the optimization. RMS (or Root Mean Square) indicates its mathematical roots. RMSProp is the adaptive learning algorithm, which subtly adjusts appropriate learning rates for individual weights at every training iteration.

3.4.3. Hyperparameter Tuning and Cross-Validation

In this section, we discuss how we select the best-performing hyperparameters for our models. The process of finding the best-performing model is called hyperparameter tuning, hyperparameter optimization, or model selection. Cross-validation is essential in this process.

3.4.3.1. Hyperparameter Tuning

In the realm of machine learning models, as well as deep learning models, a crucial aspect that must be managed is hyperparameters, which are specific configurations that the programmer must set before the training phase begins. These hyperparameters can significantly affect the performance of the model. Consequently, it falls upon the analyst to explore a multitude of combinations of these hyperparameters, ultimately

selecting the combination that yields the best performance results. A good choice for optimization when hyperparameters to be considered are few, specifically two or three hyperparameters, is grid search. This process basically consists of analysing each possible combination between hyperparameters within the specified intervals for best performance results. While grid search is a popular method, it does come with some downsides; it can be inefficient and computing-heavy, especially when there have been a large amount of hyperparameters to test. So, to overcome this drawback, another approach known as random search is introduced [65]. Roaming over various combinations of hyperparameters is a different way to overcome the curse of dimensionality, addressing the challenge introduced by more hyperparameters! Random search is believed to be much more efficient from a computational perspective, especially in cases where a number of hyperparameters is much larger; hence, an efficient approach for hyperparameter optimization for complex models is easier to achieve.

3.4.3.2. Cross-Validation

When we evaluate machine learning models, utilising some data as a test set that the model has not encountered or seen before is ideal. This type of evaluation can be fair and unbiased, as relying on only a single data split might lead to some degree of overfitting, where the model performs well on the training data but poorly on unseen data. Instead, we further split the initial training set into multiple training and testing sets, often called "folds." Each fold allows for a more robust evaluation of the model's capabilities. It is essential that the model must never be trained on the specific training set that is used for validation during the K models' cross-validation process to avoid data leakage. If managed and implemented properly, the cross-validation assessment will provide insights into the model's performance on the test set, thus serving as a dependable measure of whether this particular model has generalized effectively [60]. After completing hyperparameter tuning and selecting the most suitable model, we then proceed to build the final version of the model utilizing the entire training set without the validation step. Following this, we apply our finalized model to the test set to verify how well it performs on data it has never been exposed to before.

3.5. DEEP REINFORCEMENT LEARNING

Reinforcement learning (RL) is an agent's ability to act to maximize some notion of cumulative reward in an environment. It is the training of machine learning models to make a sequence of decisions. The agent learns to achieve a goal in an uncertain, potentially complex environment. The agent decides what to do, given that its goal is to maximize some sort of numerical value. The whole point of reinforcement learning is to learn from the environment, to act long-term, and maximize the rewards. Specifically, the development of truly "deep" representations that serve to ground reinforcement learning remains a challenge for the reinforcement learning community, and the interactions between representation learning, experience replay, and the structure of RL agents is an important direction for future work currently at the cutting edge. There are two primary objectives in reinforcement learning algorithms [62]. The first is to learn a value function that estimates how good a particular state or action will be for the agent to be in. Value-based methods are about learning a value function, allowing the agent to act optimally. The second objective is to learn a policy that tells the agent the best action to perform. Policy-based methods deal with learning the probability distribution of the agent's options. They directly learn what to do without a value function, and of the two methods, policy-based methods might be the hardest. Regarding RL, Q-learning might be the most important since it is the best-known algorithm for solving the Markov decision process problem. We will utilize it in the development of deep reinforcement learning.

3.5.1. Introduction to Reinforcement Learning

Reinforcement learning (RL) is an essential subfield of artificial intelligence (AI), and it complements areas such as supervised and unsupervised learning. Reinforcement learning deals with how an intelligent agent may take action in a given environment to accomplish its objectives in the best possible way. This way of learning behaves like a boomerang. In RL, the agent learns to define which actions are best suited to perceive, assimilate, or explore the environment. The environment changes as a result of the reactions produced by the agent. The agent receives feedback about the actions

it performs through specific measures or signals of reward or compensation guided to maximize results or produce a utility function[62].

Reinforcement learning has already been successfully applied in many different and challenging areas of artificial intelligence, such as learning to play chess, checkers, and problem-solving in general. It has also been used in robotic control and learning to control elevators, and it has applications in computational finance, intelligent adaptive tutoring systems, and organic chemistry. At the same time, reinforcement learning research is growing to be incorporated into the development process in order to make more efficient commercial applications. One of the most important recent advances in the field was the use of reinforcement learning to train agents to play various video games, attaining superhuman performance levels.

3.5.2. Q-Learning and Policy Gradient Methods

In this section, we briefly introduce and analyse Q-Learning and policy gradient methods—two of the most popular deep reinforcement learning methods for agent learning processes based on predicted future rewards. Q-learning is a typical method using tables to store "state-action-value." Because of its simplicity, Q-learning has become one of the most popular reinforcement learning methods. Q-learning often has a huge amount of computational complexity, and the fact that it can be used to handle high-dimensional state and action spaces is also an attractive feature.

As for the policy gradient method, the term policy gradient refers to all methods that produce a difference measure for reinforcement learning concerning the policy: it improves the evaluation by comparing better and worse actions. A policy gradient [65], is a reinforcement learning method that examines policies' performance with respect to their current parameter settings and then updates the policy to enhance performance. The method focuses on state and action paths rather than current state-action pairs. In recent years, it has been increasingly attracting the attention of researchers in the deep reinforcement learning field. These methods are proven to be efficient by a growing number of empirical applications. Meanwhile, policy gradient methods also make

efficient the choice of the initial parameters of the Q-network and greatly improve the learning performance of the agent.

3.5.3. Deep Q-Networks (DQN) and Their Variants

Deep Q-Networks (DQN) and Their Variants If there is one milestone in the field of AI that attracted unprecedented industrial and popular attention towards deep learning, it would have to be the demonstration that a deep convolutional neural network could learn a high-dimensional environment like the video game console of Atari 2600. The seminal approach was coined the Deep Q-Networks (DQN). This first work has since been followed by a string of innovations in both the supervised and reinforcement learning fields. The DQN is essentially a standard CNN. In DQN, Q-learning, one of the most popular RL algorithms, is used to allow the agent to map input states to the Q values corresponding to various actions. This allows the design of an end-to-end deep learning framework where training is performed on mini-batches generated from replaying the agent's experiences by storing and sampling approximately 1 million memorable states [61].

The target Q values for training involve a "frozen" separate copy of the DQN, which is fixed and also employed by the agent to decide the current action. This "target Q" is used to update the Q function. The "fixing" is required as the normal Q function might, at some point, strongly influence and corrupt the process of training the Q values. Deep Q-learning must cope with two of the most prominent instability sources: (1) correlation between samples in the replay memory and (2) the loss distribution of the Q-values. The combination of these two sources results in excessively large feedback loops.

3.6. DUAL-LAYER AUTOENCODER (DLA)

In this type of autoencoder, an extra layer is added above the normal autoencoder's hidden layer. The additional layer helps prevent overfitting and makes it possible to increase the capacity of the encoder and, therefore, to improve the coding at the output layer. However, the extra complexity of such a network involves a greater task both in

the learning and execution stages. Furthermore, with a DLA [66], the lower layer would try to preserve information about having multiple targets in a supervised learning problem. The second layer is focused on learning to code supervised targets into a compact output layer. Some researchers believe that using a DLA should dramatically improve performance on poorly matched data. Finally, we have a prefix-based binary tree, each leaf of which is a supervisory signal. With a DPC approach, we would have a DLA and make the system two-fold to increase overall performance.

3.6.1. DLA Architecture

As described, the architecture of the Deep Learning Agent (DLA) consists of various modules: the input layer, the convolutional exploration layers, the long short-term memory layer, the fully connected reward layer, and finally, the output layer. The input layer receives the visual input observations. The seven expanded modules are designed to concatenate, at the end of the layer's activation function, the attached scalar output values from the policy network. As can be noticed, the assumption that there are seven exploration module structures is based on the $1 + 1$ convolution.

The output layer accepts two different inputs. The first input comes from the output data of the final output layer module. The second input consists of the probability distribution values, which come from the two navigator outputs, pointing to the eight directions around the DLA [67]. The design of the DLA agent uses the Navigator CNN as well as the Navigator CNN. Each CNN is primarily designed for the purposes of the Q-values and the policy probability distributions, respectively. The Q-value and policy probability results of the agents' actions are then used in conjunction with the DLA's intelligent handling. With an inner layer shared structure that ties the implicit fingers and decision-making roadmaps directly into the general bottleneck, navigational strategies get directly connected to higher and lower human cognitive process levels. The integration of sophisticated convolutional and pooling structures regulates substantial amounts of navigable image content data descriptors. Upon finalizing the Navigator CNN training, the DLA learns multiple object target localization modeling best strategies.

3.6.2. Efficiency of DLA in Handling Large Datasets

Modern businesses collect an enormous amount of data to garner insights that improve their products and services. Large datasets are hard to handle, involve a high cost, and can be subject to different sorts of computational activities. However, in real-world scenarios, one may not want to make predictions from the entire dataset. It's simply infeasible as large datasets demand large memory and CPU benchmarks just to run searching or dimensionality reduction algorithms on them. In the data preprocessing step, it is natural to work on a small fraction of data, and modern big data frameworks are designed to keep as much data as possible transformable in the available memory. Traditionally, these tasks are accomplished by using different platforms designed for performing these computations. However, with the emerging technology in DLA [68], which generates both the input and the network weights, searching and learning can be combined into a single precision operation to simultaneously serve the task of machine learning methods while operating on a large set of data very efficiently. This one-shot learning framework is very advantageous in the fact that with it, a deep learning iteration may be completed in a fraction of the time and one-thousandth of the data, giving the minimum batch size needed to yield relevant results and minimizing the computation cost. The simple replacement of the network learning weights with read-only fixed-point data arrays optimized for inference saves a substantial amount of energy, which in turn provides improvements in the computation costs. The importance of data reduction is also maximized as it minimizes the cost of the communication overhead typically associated with distributed deep learning systems.

3.6.3. Feature Extraction and Dimensionality Reduction in DLA

As stated before, DLA models have millions of parameters. Performance linearly goes up with the number of parameters, but a model that has significantly more parameters is an obvious mismatch. In addition, with more parameters, computation needs increase, and training and prediction times rise. In practice, to decrease the number of parameters, MLP models are deployed that increase the number of neurons in the input only. For such models, the input needs to be meaningful. The two most used methods

are SIFT and PCA. SIFT detects local extremum points in 3D and finds descriptors by using histograms to detect more representative features. To accelerate computation, networks are designed in layers from panoramic images, and the outcomes of the lower layers are again combined and used in a sophisticated manner to decrease the size of 3D volumes [66]. In addition, feature maps are subject to pooling, which provides a relaxed nature of the model. A MobileNet architecture introduces depth-wise separable convolutions. In other networks, a complete depth separation is contributed to by those two steps. Reduction in the aggregate of convolutions through a single Conv-Deconv operation in U-NET architecture is also applied.

3.7. LONG SHORT-TERM MEMORY (LSTM)

This section continues with the deep learning techniques. Firstly, Long Short-Term Memory is described, and then the LSTMs are used. Also, the advantages of the LSTMs over the other deep neural networks in time series classification problems are presented. To show the effectiveness of the LSTMs, a time series classification problem, Parkinson's disease detection, is chosen as an application. For this purpose, the data is trained by using both the standard LSTMs and the smart LSTMs, and then their classification performances are compared. Results demonstrate that LSTMs are successful in detecting PD. The increase in the classification performance in favor of the smart LSTMs concludes the text. The main point of the section is that although the LSTM models are successful in detecting diseases, they can be improved further by smart LSTMs. It is noticeable that the work increases the accuracy rate of classifying PD by approximately 11% utilizing smart LSTMs in comparison with the classical LSTMs.

LSTM, which can learn long-term dependencies and supports the possibility of passing through information, hence addressing vanishing gradient problems, can be expressed by introducing a memory cell to an internal context to the LSTM [62]. An input gate, a forget gate, an output gate, and a memory cell are included in each LSTM unit. The input gate processes the input information from the previous output and outputs a new candidate memory cell. The forget gate decides whether to carry the output information of the input gate forward to the candidate memory cell.

3.7.1. LSTM's Architecture

Long Short-Term Memory Networks, usually just called LSTMs, are a special kind of RNN, capable of learning long-term dependencies. LSTM is well-suited to classify, process, and make predictions based on sequences of data. They have an architecture that enables them to recognize and learn from data that has long-term dependencies. The complex architecture of an LSTM enables it to resolve the problems associated with a series of times when variables are very far apart. In fact, an LSTM is designed to remember past data and related outcomes for an extended period so that when new input data arrives, it will use that memory to make sense of the new data. Their expert ability to remember important data for a long period means they are expected to handle complex input data as input data both near and far apart in relevance and of different weights [62].

LSTMs do not require stationary data; they were developed to reconcile errors from both Long Short-Term Memory networks and traditional time-series methods by pooling their forecasts. The ability of LSTMs to detect long-term memory and to learn from data that has a very long gap from the present time is quite useful for us. As such, it is well-suited to work on problems with time input data and can be used to solve other complex problems that are not real-time sequences. Smoothing out the instability of LSTMs and Granger-wise pooling of proposed forecasts from LSTMs in individual calendars has been proposed for long-term future forecasting across all available contract months.

3.7.2. LSTM's Ability to Capture Temporal Dependencies

Standard recurrent neural networks (RNNs) exhibit a limitation that restricts their ability to capture long-range dependencies effectively. This limitation arises from the scalar gradient, which negatively impacts the propagation of these long-range dependencies over time. If you repeatedly multiply the weights in RNNs, the gradient vanishes (when weights < 1). It is these two phenomena that lead to the peripheral error issue where weights less than 1 end up getting multiplied, meaning that it relatively

approaches 0, and increases the error rate. This specifically is related to the activation function, and this can be tamed by optimizing the zig-zag between the hardest dimensions of the model with a small weight, such that the weights can remain near 1 for a very long time-- yielding larger derivatives which with the greatest work to navigate the difficult area. The Long Short-Term Memory (LSTM) architecture incorporates two memory components in every unit and uses update, forget, and output gates to control the information in and out of the network. This design enables LSTM networks to provide more accurate memory retention compared to traditional RNN architectures, which tend to forget certain memories relatively quickly, thereby bypassing the long-range dependency challenges faced by RNNs. Additionally, Long Short-Term Memory networks, as a subtype of gated recurrent neural networks, retain the full capabilities of RNNs while incorporating an enhanced gated long-term memory feature for processing text or voice, thus allowing automatic learning of the context-weighted significance of the diverse words within a statement [68].

Furthermore, deep, long, short-term memory networks can be employed in training complex models, which enable the use of a large dataset in one pass over the RNN and have been created and run successfully on large high-performance systems, possessing a high likelihood of parallel and actual performance on the newest machines. Gated units are capable of spotting and capitalizing on long-range dependencies while also sensing various time scales, which is a crucial feature of language. Each neuron within the long short-term memory networks has connections to the previous neurons in the next layer, ignoring the others, and thus possesses the capacity to capture dependencies across words as the words are processed incrementally by the LSTM, which are dependent or related. Long short-term memory networks effectively solve the gradient problem, which is implicit in feedforward networks through the use of error linear unit feedback, removing the constraints that the utmost probability objective imposes naturally [62].

3.7.3. LSTM's Application to Sequential Fault Classification

The classification of extracted fault signatures produced by a sophisticated damage diagnostic methodology will undoubtedly incorporate multiple fault classes. These

classifications will also be made on sequential data that is specifically considered time series data, and numerous advanced techniques borrowed from the time series data domain have been employed to apply supervised learning utilizing a feedforward neural network. The use of a feedforward neural network necessitates that the data be time equivalent, which means that it will contain data values that are equidistant in time. Although this requirement would not pose a significant challenge for certain diagnostic signals, such as the envelope acceleration of a process motor within a gear system, sequential fault classification using a variety of classification algorithms is particularly useful for enhancing overall engine performance evaluation and improving failure prediction reliability. This methodology ensures a comprehensive analysis of fault signatures, thus enabling more accurate assessments and predictions in various operational contexts[68].

The use of classification algorithms on irregularly spaced lines with tachometer data from engines and electrical generator motors has been documented, but the use in sequence modelling proved to be a fruitful collaboration. The sequential classification task used two electric motors and drove experimental rigs to carry out stator winding damage and rotor bar damage tests. A classifier was developed using a classifier that used 60 minutes of line frequency to test this approach.

3.8. EXTREME GRADIENT BOOSTING (XGBoost)

XGBoost is short for eXtreme Gradient Boosting. It is an ensemble machine learning algorithm based on the Gradient Boosting framework [69]. It is both distributed and efficient and scales across multiple processors. It is an optimized distributed gradient boosting library that is used in a variety of applications. It has been designed and optimized, especially for boosting tree algorithms. The boosting trees are rich variables when presented in the form of numerical input or output; hence, it has become a popular algorithm in the competitive machine-learning community. The best feature of XGBoost is that it is fast, efficient, and has accurate modeling. The technique maximizes the use of memory and executes efficiently, which has led to several successful and widely recognized competition results from its users.

XGBoost is a supervised learning function that takes the set of features and the target to learn from them a non-linear function, and then it makes predictions from new data using a gradient boosting framework [69]. It is used for both regression and classification problems. XGBoost can be used to optimize the tasks that are performed in parallel, and the inbuilt model from XGBoost is quite fast. The biggest advantage of this tool is that it has a sense of overfitting and better accuracy. The learning rate is another powerful tool that is available in the XGBoost library. The key features that are characteristic of XGBoost use a “strength” to successfully model using conformation in LAD and Huber loss functions, and it also allows for parallel computing on single machines and distributed computing for clusters. The downside of XGBoost is that it is limited by machine usage because of its high usage and classified model training time.

3.8.1. XGBoost Architecture

XGBoost follows the greedy algorithm model [70]. First, it calculates the reduction in the loss when a tree is added to the model. It then selects the best split across all the variables. It specifies the variable that achieves the maximum reduction in the loss. In this stage, we use a second-order Taylor expansion to determine the best-split point. Next, it calculates the improvement of the loss with the new split. It calculates the resulting score of the new tree. After that, tweak or regularisation is added. After constructing the tree, the nodes are pruned according to the scoring. Finally, it adds a scaled learning rate to control the shrinkage of the contribution of each tree.

The main ideas of the XGBoost model are described next [71]. We start by establishing two different sets, $L = \{(x_1, y_1), \dots, (x_N, y_N)\}$ and $\Omega = \{\omega_1, \dots, \omega_M\}$. The set L contains the samples of the training data with x_i being the i -th sample, and y_i the corresponding label. The set Ω contains the different possible labels. In the case of binary classification, Ω is the set of 2 numbers $\{-1, 1\}$. Each sample has the form $x_i = \{x_{i1}, \dots, x_{in}\}$, where the first coordinate x_{i1} represents the response for the continuous or categorical variable x_i that must be classified. The set Ω is divided for each attribute based on the values of the samples of the training set, and it is randomly divided by

the XGBoost algorithm to obtain a tree. The method is run recursively multiple times in order to obtain different trees.

3.8.2. Xgboost's Advantages for Multi-Class Classification

Besides the excellent cost function, XGBoost has several advantages for multi-class classification . One of the advantages is that it can create a parallel tree structure, and the computational complexity is largely determined by the depth of the tree rather than the number of classes. This increases the training efficiency of XGBoost.

Another advantage of XGBoost is that when we perform sampling, the sampling method of XGBoost is random sampling. The training set is directly used as the sampling set, and the relevant trees are built on the sampling set according to the sampling requirements, so this process provides different data for the structures of different trees, which also enhances the diversity of the overall model.

XGBoost can naturally output the probability for each type as the score of a classifier. The result is generated by normalizing the score. The score value is interpretable in terms of confidence because it is standardized to a range. In the competition or in the game, due to practical reasons, in many cases, we would like to evaluate the confidence of our classification result more than the classification result itself. XGBoost is very suitable as the last layer of the output of a model [69].

3.8.3. Xgboost's Suitability for Refining Fault Detection Outputs

The suitability of the XGBoost algorithm for refining the fault detection output becomes apparent when its methodology is considered. XGBoost takes a weak learner or numerous such weak learners who possess a significant edge over random guessing and makes them work in a team against a problem at hand to solve it efficiently. That team is like a cricket team, with all 11 players possessing some capacity to influence the outcome of the match. Few players possess a much larger capacity to influence the match from their previous performances, like the all-rounder, who also happens to be the vice-captain of the team. Other players might have had a rough patch in the past

and might have lost their commanding touch on the game. However, the experienced player, the captain, has observed these changes in the team statistics and is formulating a strategy by providing some instruction to a few players during training and a few to different players. When a player follows these instructions, penalties or incentives are given. As time progresses, the team's performance improves, and everyone becomes dedicated, and each player learns from the other. This technique has proven to have superior performance over traditional decision trees with bagging, as it introduces important regularization parameters to control model complexity[70,72].

The various regularization parameters are employed during the model training. The importance of the features is calculated at the end of the training. This technique has provided a description of the importance with a reduction in the model's complexity without sacrificing accuracy. The nature of interpretation is at the fundamental level. The feature importance plot is not the exact contribution towards a model outcome. The exact individual feature contribution can be a computationally expensive task, which trivially can be a good data visualization tool. XGBoost can provide feature importance, feature interactions, growth statistics, frequencies for second point-wise output, and a basic measure of permutation importance that is not sensitive to correlation, with the option to provide the inverse of various measures related to feature importance values. Typically, the varying costs are avoided if parameter-dependent lines are attached. It may be worth focusing on interpreting the multi-class and/or larger datasets and their function. As the name indicates, simple, unoptimized XGBoost is a relatively straightforward implementation for XGBoost learners. The model interprets the XGBoost output based on additive assumptions. The ensemble of a set of sub-models, such as regression and decision trees (with boosting), are combined with a regularized additive loss function. This function acts as a generalizing approximation of the full set of weak learners. As ensemble models use non-differentiable methods with respect to the input features, XGBoost is not a differentiable model, unlike linear models. While linear models such as generalized linear models, linear regression, ridge, or lasso regression predict the individual variable's effect from the coefficients and intercepts, ensemble methods predict the output with the combination of individual weak learners [71].

PART 4

LITERATURE REVIEW

4.1. INTRODUCTION

The reliability of supplying power to end users has always been one of the main indices that reflect the interconnected power systems during the last century. Identifying and detecting faults and failures in a timely way can significantly increase the efficiency of these operational systems and improve the power supply situation for end users. Power fault detection significantly reduces the time needed to locate faults, isolate them, and restore power. It can also reduce the threat of connected equipment damage, to say nothing of the threat to human life. The transmission line is one of the most important branches of any power system, and data on society's increasing need for energy is clear evidence.

However, identifying a fault in this extensive and widespread network is complex, wide, and multifaceted. Several researchers have addressed or presented their methods and technologies for detecting and classifying faults on power transmission lines in recent decades. Given the increasing energy demand, identifying a fault on a network, such as a transmission line, using only a traditional technique or nature is inefficient. In recent decades, the detection system for high-impedance faults on the transmission line has grown significantly in terms of technique and method control. Many innovative methods, techniques, and devices have been proposed to identify, classify, locate the fault, and send the proper commands to these protection relays. The major loss and limitation of existing protection and classifying fault systems is its responsibility to repair and develop the reliability of the new digital tools and vigilance methods[72].

The transmission power lines have a crucial challenge in their safe and secure operation. The proper functioning of transmission power lines will directly affect the reliability and safety of all the power systems. Therefore, developing advanced strategies for fast detection and localization of faults, including signal processing and machine learning techniques, is the main interest of various researchers. The primary targets of this literature review are: a) to provide an overview of the technologies and methodologies for the detection and classification of faults in transmission power lines; b) to identify the critical points related to these processes on electric lines; c) to provide a solid base for further investigations, analysis, and optimization of existing technologies for quick fault detection and classification. The review presented here is a comprehensive source of information about strategies related to fast fault detection and classification in longitudinal faults in power lines.

Energy demand is increasing daily; hence, new power generation plants must be connected to the grid, where power systems are interconnected. Furthermore, more emphasis is placed on renewable energy sources to cope with greenhouse emissions and energy security. Therefore, more demand is needed for new high-voltage transmission power line capacities. In addition, the existing transmission power line networks face various challenges primarily due to their continuous operation for decades, environmental weather conditions, and loadings. Most of the population lives in urban areas; therefore, the right of way space to install new power lines is limited, as it is expensive. One of the major assets in the utility business is transmission power lines, and continuous and reliable power transmission is only possible if there are no complete outages or long-duration power cuts in the system. This reliance on transmission lines can significantly impact power systems and economic conditions and may affect the standard of living in a region. Therefore, a sufficient and secure transmission line network must be maintained. The need for fast and effective strategies for detecting and classifying shunt faults in power lines is more critical, given aggressive maintenance strategies for early fault removal to minimize disruption and maintain system reliability, as well as increasing network complexity [72]. Additionally, electricity companies incur significant losses in the event of wrong isolation of fault sections. In case of any voltage or current fluctuation from the

primary targets, it is mandatory to detect the type of fault and isolate the faulty section from the rest of the system.

4.2. IMPORTANCE OF FAULT DETECTION AND CLASSIFICATION IN TRANSMISSION POWER LINES

The electric power grid is considered a critical infrastructure due to its constant need to maintain delivery services. The application of fault interruption and location systems can prevent or minimize power outages following a problem with the transmission line. Therefore, the application of the fault detection and classification technique is urgently needed. The cost of these power failures to utilities and affected consumers has been estimated to be billions of dollars, which perhaps does not include indirect costs, such as a reduction in the productivity of factories and, worst of all, potential threats to public safety. Sudden distribution network failures have severe economic and environmental consequences. Power outages are accompanied by pollution increases between restarting generators or connecting distribution transformers and industries while straining already insulated weak distribution segments. In addition, the long-term negative financial impact of power outages on utilities has led to a growing interest in transmission system reliability and enhanced provisioning, characterized by a global surge in revenues accompanied by intensive efforts to upgrade transmission infrastructure with technologies that can deliver output properties from waveform formulas. Yet, the success of these efforts is heavily dependent on the practical application of efficient fault detection and classification methods, as the development of quantitative model treatments can be based on estimates. The power grid consists of multiple transmission lines and substations. Each substation has high-kV transformers and circuit breakers. It is necessary to ensure the reliability of transmission lines. However, higher power transmission lines cannot completely avoid the occurrence of faults. In some cases, the transmission line voltage surges occur transiently or steadily. This may cause serious operational stability problems in the main transformer [73].

This paper discusses the methods and technologies for fault detection and classification in power posts in a systematic way. In the context of electric power

systems, fault detection and classification are precise detection and identification of faults and their location in the substations. Any disturbance in power lines or equipment changes the system parameters, and if this disturbance leads to the loss of the power line or a piece of equipment, it should be called a fault. Faults are undesirable, lead to damage to equipment, and contribute to the loss of life. Moreover, any disturbance in sensitive equipment frequently triggers a primary and secondary sleep system. This inadvertently accumulates larger amounts of energy, which prematurely affects the financial aspects, so prompt repair systems and equipment are fundamental. Consequently, fast detection of the fault increases the effectiveness of the operation of the utility company.

Research on various networks has shown that advanced faulty management systems have intensified in some of today's power systems. Moreover, the penetration rate of advanced devices is exponentially increasing due to the deregulation of the power market. Regarding sensitivity, the presence of advanced devices impedes the severity of the fault. Indeed, the current trend of increasing the capacity of sensitive devices is to increase the strength of the systems, which is a reflection of the randomized response of each company's managerial and industrial documentation. Faults adversely affect power reliability and quality; they damage the transmission and distribution equipment and cause a reduction in the outlet power of the system. In addition, faults frequently cause some safety hazards in the environment near the power equipment and systems. Regulators in the fault location section are asking that the presence of electricity posts be detected at a speed of a few milliseconds. The fault detection waits between 20 milliseconds for such a claim, and one sent 5 milliseconds are specified [19].

4.3. HISTORICAL OVERVIEW OF FAULT DETECTION TECHNOLOGIES

An efficient and effective secondary protection system for power transmission networks is vital to minimize the hazardous consequences of a power line fault. Over the years, various fault detection technologies have been developed. Most of the early fault detection technologies used manual supervision or offline inspection. Over time,

these methods gradually got replaced by automated systems integrated with communication and control.

Early fault detection technologies started with a manual inspection of the lines. The knowledge of a line fault was initially gained through call-ins. With the advent of telephones, automatic alarms were developed. This system could detect, within the whole system, two faults, whether they persisted simultaneously or alternately to ground on private or telephone lines. An alarm system installed on the Switch Room's ground floor at a Load Dispatch Centre had alarms based on telemetry from relays developed by a manufacturer. These relays protected ten 400 kV transmission line differential characteristics. Telemetry from relays, which provided differential protection for 400 kV lines, was communicated to the central alarm system of the respective railway. The central alarm system was interfaced with the control center through data acquisition. The alarm system had 48 signals and provided audible and visual indications[74].

In the 1980s, technology was developed to transmit differential-based protection system data to a remote control center. This was then used to identify the presence of a fault in the line. This was commonly used for EHV lines. Tests were conducted, standards were developed, and a database of typical vibration patterns led to the development of automatic monitoring systems. The experience of other industries, which have significantly reduced unreliability by adopting predictive reliability-focused maintenance and remotely operated communications-based monitoring systems, suggests that these systems could benefit transmission power lines. An understanding of the development of technologies used for fault or defect detection and the inventions that have been tested and led to them can be used as guidance to shape the future of monitoring.

This review identifies and discusses the main methods and techniques used to detect transmission line faults from the early stages to date. It starts by briefly discussing fault detection technologies in the early days of electrical power engineering. It then presents more recent methods and their advantages compared to earlier technologies. Trends for the reviewed publications are analyzed, and recommendations for future

research are presented based on them. The presented breakdown summarises the major developments in the field over the decades. It reflects a paradigm shift from the first manual techniques to the advanced state-of-the-art solutions based on computational intelligence approaches, which require minimum human intervention. Further, a discussion on the major problems faced by the reviewed approaches is also presented, which paves the way for current research in the field.

Fault detection technologies have rapidly evolved since the field of electrical transmission and distribution of electricity started. Customers have now come to expect high-quality electrical power from their power utilities. Currently, some technologies in transmission line fault detection include sequence components, wavelet packet transformation, support vector machines, wireless sensor networks with powerful data processing, wavelet transformation, fault component oscillation, artificial neural networks, power spectral density, adaptive neuro-fuzzy inference systems, deep learning, cog vector, and synchro-control [75]. These methods are a product of a lengthy training operation. The field has moved towards using sensor networks for fault identification, necessitating large-scale readability. Current research appears to be moving towards distributed automatic fault localization and classification using data and sequence components by computational intelligence approaches.

4.4. SIGNAL PROCESSING-BASED FAULT DETECTION TECHNIQUES

Signal processing methods have played a critical role in detecting faults in power transmission lines for a considerable duration. These techniques examine voltage and current waveforms to identify anomalies that arise from faults. Their capacity to extract characteristics from power system signals establishes a robust basis for both fault detection and classification. The primary signal-processing techniques consist of:

- **Fourier Transform (FT):** The Fourier Transform is a widely utilized technique for frequency-domain analysis within power systems. It aids in the detection of steady-state harmonics and spectral elements, which is beneficial for identifying irregularities such as harmonic distortions and resonance

phenomena. However, FT has some disadvantages in dealing with transient signals because it cannot determine the time localization of signal components [76].

- Wavelet Transform (WT): The Wavelet Transform addresses the weakness of the Fourier Transform as it allows procedure of time-frequency analysis and is appropriate for locating transient fault such as lightning strike, switching surges, and short circuit. By adjusting the resolution where needed, it can effectively play a part in capturing events of curiosity at a multitude of scales, particularly in detecting transmission line faults in dynamic cases [77].
- Short-Time Fourier Transform (STFT): The STFT method allows both time and frequency domain analysis by breaking the signal into segments of fixed duration (windows). While the Short-Time Fourier Transform does temporally localize better than the Fourier Transform, the fixed window size can limit resolution when the signal of interest has rapidly changing frequency content [78].
- Empirical Mode Decomposition (EMD): A data-driven methodology based on analytical decomposition of non-linear and non-stationary signals through their integral multiples of intrinsic mode functions (IMFs). The IMFs capture the intrinsic fault features of the complex signal so that the fault detection of the renewable energy integrated system or the system with dynamics load can be effectively achieved [79].
- The Hilbert-Huang Transform (HHT): is an extension of the Empirical Mode Decomposition (EMD) in the aspect of instantaneous frequency analysis of signal components. This feature is important to identify initial stage faults and the subsequent transient phenomena at the power lines. This property of the HHT enables it to adapt to the features of non-linear signals, thus making it more useful for application in modern power grids Hilbert-Huang Transform in Fault Detection[80].

Signal processing methods are widely recognized for their efficiency in computation and are commonly implemented in conventional power systems for various applications. Nevertheless, their inherent reliance on manual feature extraction, coupled with significant challenges when managing overlapping or intricate fault

situations, limits their effectiveness and applicability in modern, highly interconnected power networks. As these systems evolve and become complex, alternative approaches may be necessary to address these emerging challenges effectively.

Since the wavelet and discrete wavelet transforms are mostly used in signal processing-based fault detection techniques, they will be expressed more in the following two subsections.

4.4.1. Wavelet Transform

The primary objective of the wavelet transformation is to analyze a signal in terms of its time and frequency properties [77]. High-frequency components are assumed to occupy shorter time intervals, while low-frequency components are associated with longer durations. Thus, the wavelet transform achieves high resolution for transient high-frequency signals and periodic low-frequency components. A continuous wavelet transform (CWT) mathematically decomposes a continuous time function into wavelets. In contrast to the Fourier transform, the CWT provides time-frequency representations of a signal with excellent localization in both time and frequency domains. The continuous wavelet transforms of a function $x(t)$ at a certain scale ($a > 0$) and a corresponding translation value $b \in \mathbb{R}$ is expressed (as in Equation 4.1):

$$X_{a,b}(t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.1)$$

where ψ is a continuous function referred to as the mother wavelet, and a is the scaling parameter indicating the degree of compression for the mother wavelet. B is the translation parameter determining the wavelet's time location. The primary purpose of the mother wavelet is to generate daughter wavelets, scaled and translated versions of itself.

4.4.2. Discrete Wavelet Transform

Composing a signal aims to analyze it at varying resolution levels, with the resolution representing a descriptor of its frequency components [78]. The discrete wavelet transform is implemented to analyze time signals $x(t)$, allowing for complete decomposition into approximations employing a dyadic scaling function known as the father wavelet. The decomposition yields detailed coefficients through the mother wavelet, as illustrated in the associated Figure 2.1, which can be expressed in the following equations (4.2 & 4.3):

$$\phi_{jk}(t) = 2^{-\frac{j}{2}}\phi(2^{-j}t - n) \quad (2.2)$$

$$\psi_{jk}(t) = 2^{-\frac{j}{2}}\psi(2^{-j}t - n) \quad (2.3)$$

Where $n \in \mathbb{Z}$, j and k are integers.

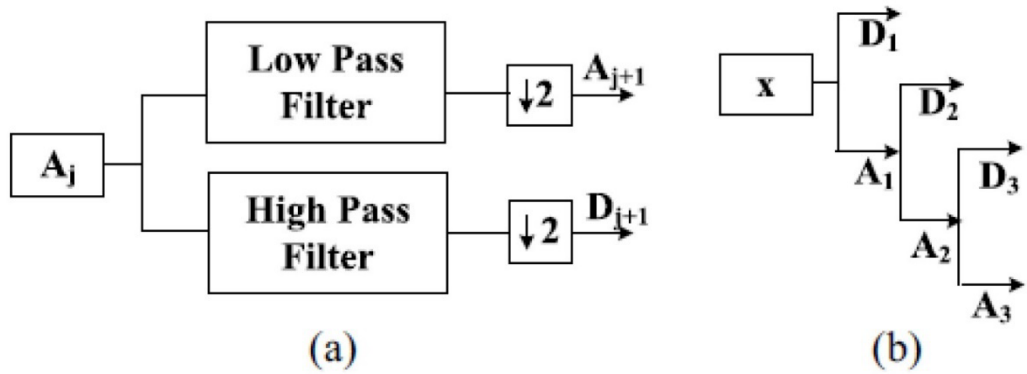


Figure 4. 1. Discrete Wavelet Transform.

Given that j and k are integers, the basis function is scaled by a factor of 2^{-j} and translated n units in time. The scaling function is linked to low-pass filters with coefficients $H = \{h_n\}$, and the wavelet function is related to high-pass filter coefficients $G = \{g_n\}$. It is essential to recognize that G is the counterpart of H interspersed with sign alterations. Filters are produced as a result of the two-scaling function (equations 4.4 & 4.5):

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t - n)$$

$$\psi(t) = \sum_n g(n)\sqrt{2}\psi(2t - n)$$

There are a maximum of "M" decomposition layers for a signal $x(t)$ with a length of 2^M . Here is the signal expression using wavelet transform:

$$x(t) = \sum_{k=0}^{2^{M-j}-1} a_{jk}\phi_{jk}(t) + \sum_j \sum_{k=0}^{2^{M-j}-1} d_{jk}\psi_{jk}(t)$$

The signal's approximation coefficient at level j is a_{jk} , while its detail coefficient is d_{jk} . Both high-pass and low-pass filters are used to pass the signal. The second step is a 2-fold down sampling of the two filters' outputs. At the first level, we get the approximate and detailed coefficients ($A1$ and $D1$). This procedure will be repeated till the degree of breakdown is attained that was previously determined.

4.5. AI-BASED FAULT DETECTION TECHNIQUES

AI-driven techniques have transformed fault detection by automating feature extraction and utilizing data-driven models for precise fault classification. These methods are especially beneficial in tackling the complexities presented by contemporary transmission systems, which face varied fault conditions and the need for real-time responses, necessitating adaptive and scalable solutions. Key AI techniques encompass:

- **Artificial Neural Networks (ANNs):** Another well-known and pioneered technique of artificial intelligence technique for Fault detection about which ANN. These models are capable of finding out the patterns and relationships inside historical data, allowing reliable fault classification. While ANN has been proven beneficial in the multi-class fault regime, the performance is

highly dependent on both the accuracy and size of the employed training datasets [81].

- **Support Vector Machines (SVMs):** Support Vector Machines (SVM) are trustworthy classification instruments that distinguish line between fault and non-fault conditions by maximizing the margin amongst various groups. Their application is seen in complex fault scenarios because of their efficiency in non-linear data space with the kernel functions. Decision Trees and Random Forest- Decision trees and random forest are techniques that produce interpretable models for fault detection. Decision trees have the same computing efficiency with random forests improve accuracy by averaging prediction from multiple trees. Such procedures lend themselves to data suffering from noise, common to high-voltage systems [82].
- **Deep Learning Models:** Advanced architectures like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs) have significantly enhanced the efficacy of fault detection. Long short-term memory (LSTM) networks are well-suited to investigate sequential data such as transient fault waveforms, while convolutional neural networks (CNN) are capable of investigating spatial features obtained from signal time-frequency representations [83,84].
- **Extreme Gradient Boosting (XGBoost):** A scalable and efficient machine learning algorithm that improves fault classification via the iterative improvement of decision trees. It is especially appropriate for a large-scale transmission network which is subjected to the various types of fault conditions [85].
- **Reinforcement Learning (RL):** Reinforcement Learning (RL) is a promising artificial intelligence approach to learn the best control policies by interacting with the environment. It has shown promise in identifying and making the right decisions for adaptive faults in diverse operational conditions [86].

AI-based techniques are very accurate and tunable; they become necessary for modern power systems. Their computational burden and reliance on massive training datasets make careful execution essential.

4.6. HYBRID APPROACHES

Hybrid methods combine the best signal processing and artificial intelligence methods to alleviate the drawbacks of stand-alone techniques. These frameworks leverage the feature extraction power of signal processing along with the rigorous classification accuracy of ML / DL models to provide adequate solutions for fault detection in power transmission lines. Hybrids consists of examples that are a combination of:

- **Wavelet Transform with ANNs:** WT extracts the time-frequency characteristics of the transient signals and uses artificial neural networks (ANN) for classification. It improved fault detection in high voltage systems with renewable energy integration [87].
- **EMD with SVM or Deep Learning:** Empirical Mode Decomposition (EMD) decomposes signals to the Intrinsic Mode Functions (IMFs) and use these as the input features for SVMs or deep-learning. This combination of methods is capable of detecting failures in nonlinear and dynamic operational conditions [88].
- **HHT with LSTM Models:** The combination of HHT's effective frequency analysis with the sequential data processing capability of LSTM has been shown to improve transient fault detection in transmission lines, particularly those embedded with time-varying renewable energy sources[82].

Such hybrid approaches are promising not only in enhancing the detection accuracy but also ensuring the computational efficiency as well as adaptability, especially vital for today grids that are exposed to dynamic load condition and high penetration of renewable energy sources.

4.7. CHALLENGES AND LIMITATIONS IN FAULT DETECTION AND CLASSIFICATION

There are a number of associated and intervening variables that challenge fault detection and classification methods, technologies, and practices. These alterations can lower detection accuracy, causing false positives or false negatives. Additionally, the

reliability of detecting faults decreases with the uncertainty of environmental conditions. New technological solutions, therefore, require a scalable and reconfigurable human-computer interface that is able to provide appropriate assistance in recognizing new and varied faults.

At the moment, transmission and distribution fault detection and classification systems rely exclusively on uniquely performing centralized databases, a solution not necessarily capable of detecting and recognizing new and emerging fault types. The training and performance of all HV electrical network incident operational staff, as well as the spectrum of protection settings, directly influence the effectiveness of any new fault detection and classification solution. The main barrier to the adoption of advanced automatic fault detection is not the primary performance of the technology but the economic one, which must balance investment and performance benefits. The cost and capacity of harsh environmental testing is another limitation, and there is no guarantee of overall system performance when tested based on severe real-world malfunctions. Further research and development in fault detection technologies will be able to improve reliability and costs thus far [30].

Transmission power line systems can be considered the least observable part of power systems. Accordingly, any mistake in fault detection and classification overloads the subsequent systems, which entails disastrous results. The complexity and diversity of possible system faults constitute two main challenges in detecting and classifying power line faults. Among the main factors that contribute to the high variability in system faults are the permanence of changing environmental factors, the different responses of various intrinsic system components to these environmental factors, more than 20 types of system faults, and the possible sequence and connection of different fault poles to form compound faults. Most manual fault detection and fault classification methods are either pristine algorithms with moderate performance or state-of-the-art algorithms that suffer from reliability and accuracy issues, especially in adverse weather conditions. The main reason for the low detection performance is a set of significant obstacles that restrict the progress of different fault detection systems and methods. In terms of hardware constraints, long transmission lines require increased processing speeds and sufficient storage capacities for transmission data. In

terms of data constraints, the complexity of high-value feature extraction and a suitable learning classification model is often increased except under certain conditions. Most of the available data are also not labeled, which is required to train machine learning-based systems. The accumulation of these obstacles is cost-effective for practicing solutions, but sophisticated proposals are required to fulfill the work's demands satisfactorily. A shortage of labeled data also limits the design of a universal practical system. Additionally, given the adverse impact of line faults and the dangers of their misclassification, the existing fault detection and classification methods and systems are not yet effective, and continuous research and development are required to avoid new problems [76].

4.8. COMPARATIVE ANALYSIS OF DETECTION TECHNIQUES

Signal processing methods remain a dependable option for identifying faults in conventional transmission lines, providing both computational efficiency and distinct understanding of fault attributes. Still, their limits on addressing overlapping or complex fault scenarios render them less significant in today's grid environments.

Due to their high accuracy and scalability, AI-based methods, operating in the non-linear and dynamic fault space, stand out in the fault detection and diagnosis field. Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) have displayed remarkable performance in fault detection under multiple grid scenarios. However, their requirement for immense computational resources and dependence on large training datasets can be hindrances to deploying in resource-constrained environments.

Hybrid methods combine the computational efficiency of signal processing with the flexibility of machine learning. These frameworks would be beneficial for modern transmission systems, as modern generation mixes with renewable energy, and rapid load-changing scenarios require advanced fault protection techniques that utilize the novel features of the frameworks.

4.9. ML/DL ADVANCES IN FAULT DETECTION

Transmission systems, especially low-voltage networks, are particularly vulnerable to a range of faults, such as short circuits, insulator breakdowns, and environmental disturbances like thunderstorms [3]. Beyond conventional shunt faults, modern power networks face progressive issues such as voltage sag (VS) and high-impedance faults (HIF), which are known to degrade power quality and disrupt system operations [4]. HIF, for instance, often results from tree branches contacting electrical conductors or damaged cables and is challenging to detect due to minimal fault currents, rendering traditional protection schemes inadequate [5].

The growing penetration of FCS in power networks exacerbates these challenges by introducing dynamic load fluctuations and additional complexity to system operations. FCS often operate in conjunction with renewable energy sources such as photovoltaic (PV) systems, which, while promoting sustainability, also increase system instability due to their intermittent nature [89], [90]. Voltage sag, one of the most prevalent power quality issues, accounts for up to 90% of reported disturbances in modern grids and is frequently caused by sudden load changes associated with FCS charging cycles or renewable energy intermittency [91], [92]. These disruptions have significant implications, particularly for industrial and commercial operations reliant on sensitive devices such as programmable logic controllers (PLCs) and variable frequency drives (VFDs) [93]. Similarly, low-impedance faults (LIF) contribute to these disruptions, making robust and scalable fault detection frameworks essential for modern power systems. Therefore, to address these problems, it is crucial to develop a protective system that takes into account high impedance failures (HIF), low impedance faults (LIF), and voltage sags. This will ensure that appropriate actions are implemented to maintain the safety and security of the system.

Numerous studies have proposed techniques for detecting and mitigating faults and voltage sags in power networks. Traditional methods such as Fourier Transform (FT) and Wavelet Transform (WT) have been widely used for fault detection and classification, leveraging their ability to process time-varying signals [94], [95]. These techniques have also been integrated with machine learning approaches such as

Artificial Neural Networks (ANN) and Support Vector Machines (SVM), providing improved accuracy in fault diagnosis [96]. For instance, SVM models coupled with Discrete Wavelet Transform (DWT) have demonstrated effective HIF detection in medium-voltage distribution networks, leveraging statistical features such as standard deviation to enhance classification accuracy [97]. In another approach, fine-tuned LSTM networks, optimized with the improved arithmetic optimization algorithm (IAOA), and db6 wavelet transform enhance fault feature extraction, as shown in [98]. Similarly, decision tree-based methods utilizing harmonic features have been used for HIF detection in medium-voltage network [99]. Despite these advancements, many of these approaches face challenges in real-time implementation due to their high computational requirements, sensitivity to noise, and limited adaptability to new fault conditions [100]. In addition, studies examining the probabilistic forecasting and identification of voltage sags in power networks [101], [102] often fail to address the underlying causes of these disturbances.

Table 4.1 provides a comparative summary of existing fault detection methodologies, outlining their feature extraction techniques, classification algorithms, and fault types addressed. While these methods have laid the groundwork for fault detection in power systems, they often lack the robustness, scalability, and efficiency needed for real-time applications in environments with high penetration of FCS and renewable energy sources.

Table 4. 1. Literature Review.

Reference	Feature Extraction Method					Detection and Classification Method			Types of faults		
	PM U Based	Signal Processing Based	Supervised Learning	Unsupervised Learning	Other Methods	Supervised Learning	Unsupervised Learning	Computational Method	HIF	LIF	V S
[2]						✓			✓		
[3]					✓ (TKEO)	✓			✓		
[91]		✓				✓ (SVM)			✓		
[93]		✓				✓ (ANN)				✓	
[94]		✓				✓ (Fuzzy Logic)				✓	
[95]		✓				✓ (PNN)				✓	

[96]	✓	✓			✓		✓
[97]		✓			✓		✓
[98]		✓			✓ LSTM with (IAOA)		✓
[99]		✓			✓ (Decision Tree)		✓
[101]				✓ (HOS)	✓ (SVM)		✓
[102]				✓ (TKEO)		✓	✓
[103]		✓			✓		✓
[104]		✓			✓ (SVM)		✓
[105]		✓			✓ (Wavelet Transform)		✓
[106]		✓				✓	✓
[107]						✓	✓
[108]	✓	✓	✓	✓ (Convolutional Autoencoder)		✓	✓
[109]				✓ Stacked Convolutional Sparse Extractor)		✓	✓
[110]				✓		✓	✓
[111]			✓	✓ (Discriminative Extractor)			✓
[112]			✓	✓ (Denoising Sparse Extractor)			✓
[113]	✓	✓	✓		✓		✓
[114]			✓		✓ (XGBoost)		✓

To overcome these challenges and rather than emphasize convolutional autoencoders for fault detection, particularly in high-impedance fault scenarios as in [108], this study proposes a hierarchical protection framework tailored to the IEEE-14 bus power system integrated with eight concurrently operating FCS and PV systems. The framework leverages advanced machine learning techniques, including a two-layer autoencoder for efficient feature extraction, Long Short-Term Memory (LSTM) networks for sequential fault classification, and XGBoost for refining fault categorization. This approach enables accurate classification of symmetrical and

asymmetrical faults, such as single-line-to-ground (LG), line-to-line-to-ground (LLG), three-phase (LLL), and three-phase-to-ground (LLLG) faults, under varying operational conditions.



PART 5

SYSTEM DESCRIPTION

In this section, the fundamental laws and in-depth analysis of the test system and methods that will be utilized in the proposed process are included. The first subsection outlines the test system. Then, the second and third subsections briefly describe the autoencoder and its specific experience in extracting latent features and the LSTM method that will be used to detect and classify faults.

5.1. POWER SYSTEM MODEL

An analysis of a power system has been performed using an IEEE-14 bus power transmission network. The network is illustrated in Figure 5.1 and built using the PSCAD [115] framework. The system consists of an 11 kV electric power grid that includes 14 buses, 5 synchronous machines, 20 transmission lines, and 3 transformers. The model parameters are thoroughly scrutinized. The inclusion of PV penetration in this system serves to strengthen the authenticity of the system in line with current trends. The system is supported by two 250kW photovoltaic (PV) systems [116]. The solar farm consists of a power plant controller, a photovoltaic (PV) array, a boost converter, and a DC-AC inverter. The main purpose of the power plant controller is to adjust the desired power levels (both active and reactive) of the inverter in the solar farm, using the measured values of voltage, active power, and reactive power at the point-of-common-coupling (PCC). The maximum power output of each photovoltaic (PV) array is 250 kilowatts (kW) under the conditions of a temperature of 28°C and an irradiation level of 1000 watts per square meter (W/m²). The boost converter is employed to control the direct current (DC) voltage to achieve optimal power point tracking. The DC-AC inverter regulates the functioning of the solar farm. Table 5.1 explains PV system parameters.

Table 5. 1. Description of PV system

Input Parameters for PV Array	Values
Series Connection (No. of modules per array)	35
Parallel Connection (No. of modules per array)	130
Series Connection (No. of cell strings per module)	36
Parallel Connection (No. of cell strings per module)	1
Reference irradiation (Γ_r)	1000
Reference cell temperature (γ_r)	28
Short Circuit current at reference conditions per cell	0.0025 kA
Open Circuit voltage at reference conditions per cell	0.000843 kV
Current at maximum power point	0.300 kA
Series Connection (No. of modules per array)	8.33 kV

In general, fast charging stations for electric cars (EVs) necessitate substantial power to swiftly recharge batteries. These charging stations require a power requirement of hundreds of kilowatts (kW) per charger, which is far more than the average power consumption of households. Therefore, the high demand or sudden increase in load created an added burden on the power system, resulting in a significant decrease in voltage known as a dip or sag. To examine this phenomenon, we incorporated eight FCSs that were connected in parallel to the system. These fast-charging stations (FCS) are connected to the transmission test system at the point of common coupling (PCC) [116]. Every FCS is linked to the system using a distribution transformer T_i which is thereafter affixed to the corresponding FCS_i via a breaker called S_i , where i represents the number of the FCS. The output power of FCS_1 , FCS_2 , FCS_3 , and FCS_4 is equal to the output power of FCS_5 , FCS_6 , FCS_7 , and FCS_8 accordingly. The rated kVA of transformers T_1 , T_2 , T_3 , and T_4 are equal to the kVA rating of transformers T_5 , T_6 , T_7 , and T_8 . Each distribution transformer, T_i , has a primary voltage range of 4.16 kilovolts (kV) to 0.480 kilovolts (kV), which corresponds to the voltage level specified in the FCS_i . Each circuit breaker, S_i , has two unique states: it can either be in the "on" state or the "off" state. The state 'on' signifies that the breaker S_i is in a closed position and is actively conducting FCS_i , thereby occupying it. When the breaker S_i is in the 'off' position, the FCS_i does not flow through the circuit. While two electrical currents can be proportional or equal, one breaker may have a certain state, represented as " S_i ,"

which does not impact the status of the other breaker. The description of fault analysis is given in the next section.

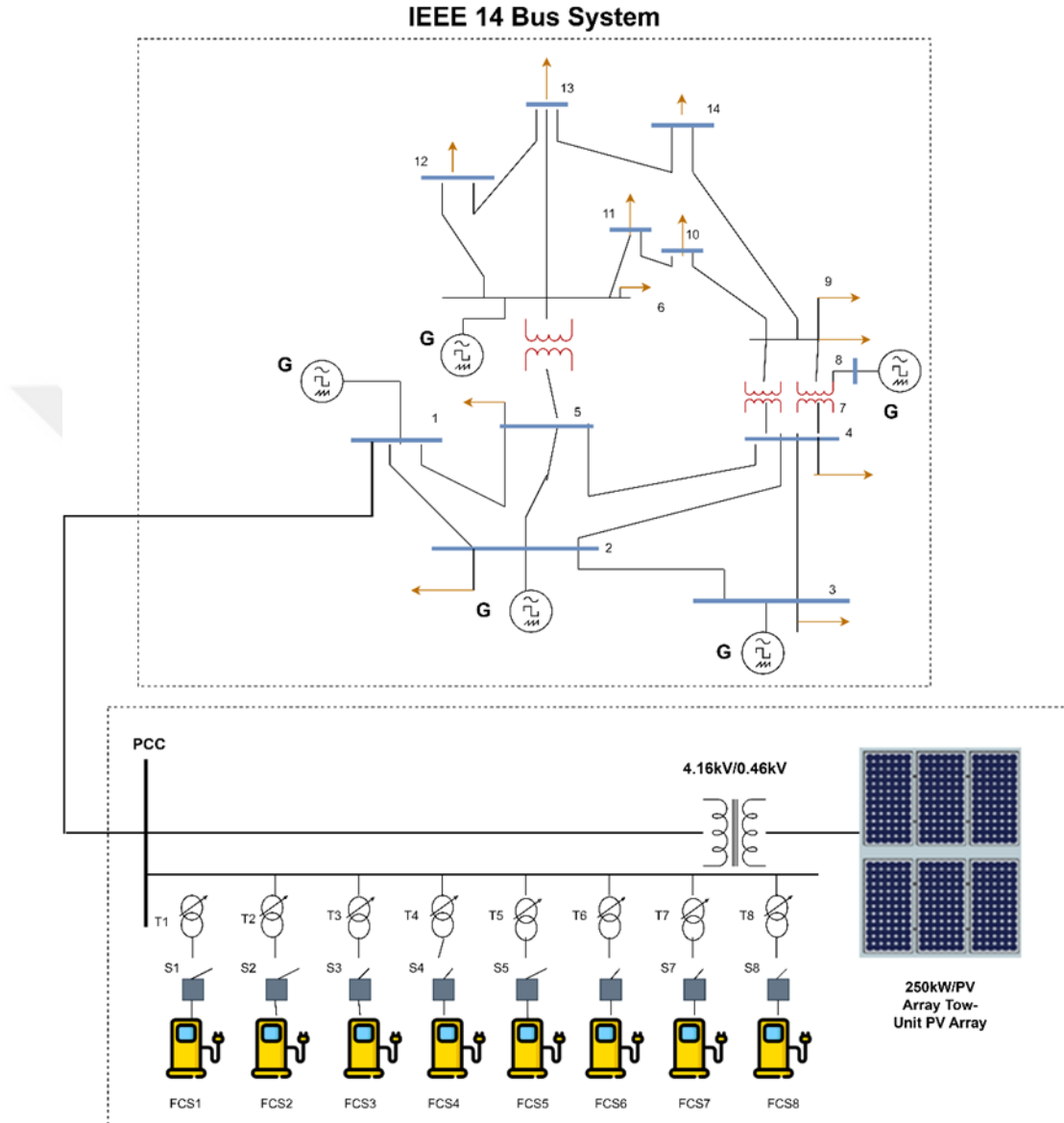


Figure 5. 1. Modified test system.

5.2. DESCRIPTION OF FAULT CASES

Three operation modes have been explored for the test of 14 bus system. In these modes, the FCS with a PV system can be adjusted via a static switch (SW1), as seen in Figure 5.1. The switch's closed state is represented by the number '1', while its open state is represented by the number '0'. The first scenario considers a 14-bus system

with a PV system but without any faults or FCS connection. The data is collected for cases where no fault is involved. In the second scenario, the HIF and LIF data are collected without attaching the FCS, but the PV is connected to the test system. During this process, the data for LIF and HIF is obtained from buses 5 and 6, where the fault is introduced. During the third operational mode, the FCS and the Photovoltaic (PV) system are connected to the PCC through the Low Voltage (LV) transmission system. This connection allows for the collection of data on VS from the PCC.

The eight FCSs are linked to the system with the help of a distribution transformer, which is then attached to the relevant circuit breaker, as shown in Figure 5.1. To acquire voltage sag data, several combinations of circuit breaker opening and closing are utilized. The voltage sag is observable with the other faults, as shown in Figure 5.2, for the voltage of phase A.

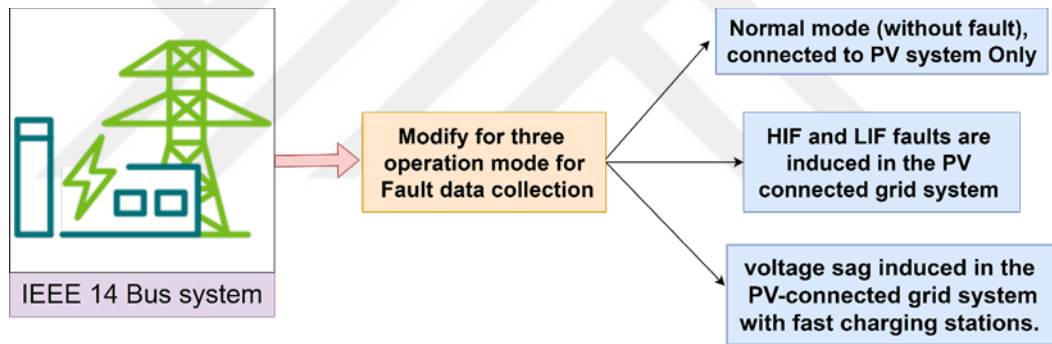


Figure 5. 2. Description of fault cases.

A voltage sag is observed between the time interval of 2 to 2.5 seconds for the entire 7-second simulation session. A range of fault resistances (R_f) from $10\ \Omega$ to $100\ \Omega$ is selected to encompass a wide spectrum of resistances, ranging from low to high. The test cases are analysed throughout a range of fault inception angles (FIA), spanning from 0° to 180° with a 15° gap. The study extensively examines all eight types of fault instances, including symmetrical faults LLLG_ABCG and unsymmetrical faults LG_AG, LG_BG, LG_CG, LLG_ABG, LLG_ACG, LLG_BCG, LLL_ABC. The study also investigates faults with differing fault inception angles (FIA) and fault resistances. Additionally, the study analyses voltage sag at PCC which can be shown in the test system diagram, Figure 5.1. However, to analyse the random behavior of

HIF and LIF, a model for HIF and LIF is created and thoroughly examined by simulating it at various points inside the test system. A comprehensive study was conducted on a total of 72 fault situations involving HIF, LIF, and VS. The fault resistance ranged from 10Ω to 100Ω , and the fault inception angle varied from 0° to 180° . The description of the fault parameters is given in Table 5.2.

Table 5. 2. Description of fault parameter.

Variations in Parameters	Dataset for Training	Dataset for Testing
Series Connection (No. of modules per array)	Variation in Resistance = 10Ω to 100Ω	Variation in Resistance = 10Ω to 100Ω Step Size= 10Ω
Parallel Connection (No. of modules per array)	Step size = 10Ω	Variation in Resistance = 0Ω to 10Ω Step Size = 5Ω
Series Connection (No. of cell strings per module)	Variation in Resistance = 0Ω to 10Ω	Random Angles = 0° to 180°
Series Connection (No. of modules per array)	Step Size = 5Ω	$\pm 5\%$

5.3. AUTOENCODER FOR FEATURE EXTRACTION

An autoencoder is a neural network architecture designed to learn how to generate its input as its output through training. An autoencoder, in contrast to a conventional neural network, seeks to generate an output that closely mimics its input [117]. This approach allows for the use of unsupervised learning [109], which can be applied to create a feature extractor. An autoencoder network consists of two main components: an encoding function that transforms the input (voltage and current values) (x) into a hidden layer, The latent vector, also known as the hidden layer, is represented by $t = \sigma(x)$ where t is the latent vector and σ is the activation function applied to (x). As the second component, the decoding function attempts to recreate the input, based on the latent representation of the input, which is expressed as $\hat{x} = \sigma(t)$. Autoencoders are frequently used for data compression, data denoising and feature extraction [118], [119], [120], [121].

An autoencoder performs with success only by achieving an appropriate size of latent vector (t). The latent vector can be smaller than the input vector x to encode the original input, which is why an autoencoder in which the dimension of its latent vector

is less than the dimension of its input is said to be undercomplete. An incomplete autoencoder can very well memorize the principal properties of the input space. Let x represent the input. We can express the encoder and decoder as follows:

$$t = \sigma(W_e x + b_e) \quad (1)$$

$$\hat{x} = \hat{\sigma}(W_d t + b_d) \quad (2)$$

The output of the encoder, denoted as t , represents the desired features that need to be obtained. The encoder function's weight and bias are represented as " W_e " and " b_e ", respectively. The variables W_d and b_d , respectively reflect the weight and bias of the decoder function. The symbols $\sigma(*)$ and $\hat{\sigma}(*)$ denote activation functions, which are crucial for a neural network to capture nonlinear information.

To achieve this optimization, the loss function is minimized, as denoted by the equation [111]:

$$L(x, \sigma(\hat{\sigma}(x))) \quad (3)$$

The loss function L penalizes the difference between the output $\sigma(\hat{\sigma}(x))$ and the input x .

Regarding feature extraction, a conventional autoencoder solely employs a single-layer encoder. Extracting deep features poses a significant challenge for this basic model [122]. In this study, the authors utilize a two-layer autoencoder to more effectively extract the profound characteristics of the unsupervised data. This autoencoder is utilized to extract the important features of the three-phase current and voltage signals.

5.4. BASELINE LSTM FOR CLASSIFICATION AND DETECTION

The LSTM is a type of recurrent neural network that is constructed based on the concepts of deep learning. Sequential data classification has been widely employed in several applications, including forecasts, voice recognition, fault identification and categorization, and islanding detection. Unlike normal RNNs, LSTM networks avoid the vanishing gradient descents problem. Hence, it has the ability to fetch and retain data for a long time [123]. In addition, it operates well with bigger networks of connections.

The essential components of the Long Short-Term Memory (LSTM) model include the input gate (i_t), output gate (o_t), forget gate (f_t), and memory cell state (c_t) [124]. The input gate decides what new information is stored in the memory and the output gate defines what information to pass to the next hidden state. Thus, the forget gate handles whether information obtained from the previous stage should be discarded or retained for the relationship of input patterns versus classes in the next state. A refined categorization technique is utilized that preserves the fundamental characteristics while rejecting extraneous information. Assuming (x_i) as the input for an individual unit at a time (t), and (h_t) is the hidden layer. The weights (w_{ix}), (w_{ox}), (w_{fx}), (w_{cx}) correspond to the connections between input (x) and the input gate (i_t), output gate (o_t), forget gate (f_t), and memory cell (c_t) respectively. Similarly, the weights (w_{ih}), (w_{fh}), (w_{oh}), (w_{ch}) represent the corresponding connections between the output (h) and the input gate (i_t), output gate (o_t), forget gate (f_t), and memory cell (c_t). The variables pertaining to different layers for a particular input (x_i) are represented as:

$$i_t = \sigma(w_{ix}x_t + w_{ih}h_{t-1} + b_i) \quad (4)$$

$$f_t = \sigma(w_{fx}x_t + w_{fh}h_{t-1} + b_f) \quad (5)$$

$$o_t = \sigma(w_{ox}x_t + w_{oh}h_{t-1} + b_o) \quad (6)$$

$$c_t^* = \tanh(w_{cx}x_t + w_{ch}h_{t-1} + b_c) \quad (7)$$

$$c_t = c_t^* \Theta i_t + c_{t-1} \Theta f_t \quad (8)$$

$$h_t = \tanh(c_t) \Theta o_t \quad (9)$$

The bias vectors b_i , b_o , b_f and b_c represent all the elements of an LSTM unit. The classifying model is trained through an iterative process of adjusting the weight values to establish a robust association between patterns and classes.

A neural network consists of a series of successive layers. Furthermore, a standard deep learning model consists of additional layers such as the activation layer, the drop-out layer, the batch normalization (BN) layer, and the flatten layer. When employing classification methods, it is common practice to utilize dense layers and the SoftMax layer during the final stage. Each of these layers aims to:

Activation Function: The use of a non-linear function allows the model to capture complex patterns in the data. It takes as input the output signal from the previous cell and is converted into form that can be passed on (as the input) to the next cell. Regular activation functions such as Rectified Linear Unit (ReLU), sigmoid, and hyperbolic tangent are generally used.

$$\alpha_{Relu}(x) = \max(0, x) \quad (10)$$

$$\alpha_{sigmoid}(x) = \frac{1}{1+e^{-x}} \quad (11)$$

$$\alpha_{tanh}(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (12)$$

Inside the bustle of the city, all kinds of sounds fill the air, which interfere with each other like vying musicians. The brilliant neon lights of colossally tall buildings pierce the night sky. This bizarre glow blankets the city below them. A numbing crowd for pedestrian movement. In firecracker like succession, the smell of grilling food on street side mixes with the fumes of gaseous states. This all is a sort of overdosage from aroma

saffron that lingers plenty long afterwards. For all this mayhem, there's an unmistakable vitality that flows through the city. It's a relentless heartbeat that pushes things into night.

$$y_i(x) = \gamma \frac{x_i - \mu_x}{\sqrt{\sigma^2 + \varepsilon}} + \beta \quad (13)$$

Flatten layer: It reduces the dimensionality of the features to a one-dimensional array which is later used to feed into the dense layer. For this reason, it is commonly used prior to the dense layer.

Fully Connected Layer: This layer stands out with all neurons being wired with all of the neurons from the preceding layer. This layer is commonly used in neural networks to operate as the last few layers to project the features in the space where they can be classified.

SoftMax Layer The fully connected layer also uses SoftMax activation, and it is often used as the final activation function in a neural network. The number of neurons in the layer equals the final classification layer.

$$S_i = \frac{e^{x_i}}{\sum_{i=1}^J e^{x_i}} \quad (14)$$

Loss Function: In order to evaluate the attractiveness of the model, its predicted value is compared to the intended target value, using the notion of a penalty value. Hence, the optimizer adjusts the model parameters by reducing the loss function to enhance the precision of the model in making predictions and classifications [125].

$$\sum_{c=1}^N y_c \log P_c \quad (15)$$

PART 6

PROPOSED METHODOLOGY

To address the difficulties of fault identification and categorisation in contemporary power systems with high FCS and PV penetration, the suggested methodology employs a hybrid strategy integrating state-of-the-art machine learning algorithms. Scalability, accuracy, and applicability in real-time are guaranteed by this methodology, which combines efficient fault classification algorithms with strong feature extraction and sequential pattern recognition. In Figure 6.1, we see the process flow diagram showing how the suggested approach incorporates feature extraction, sequential fault classification, and fault categorisation procedures. Next, we'll dive into the approach, covering everything from the system architecture and data preparation to how the detection and classification models were implemented.

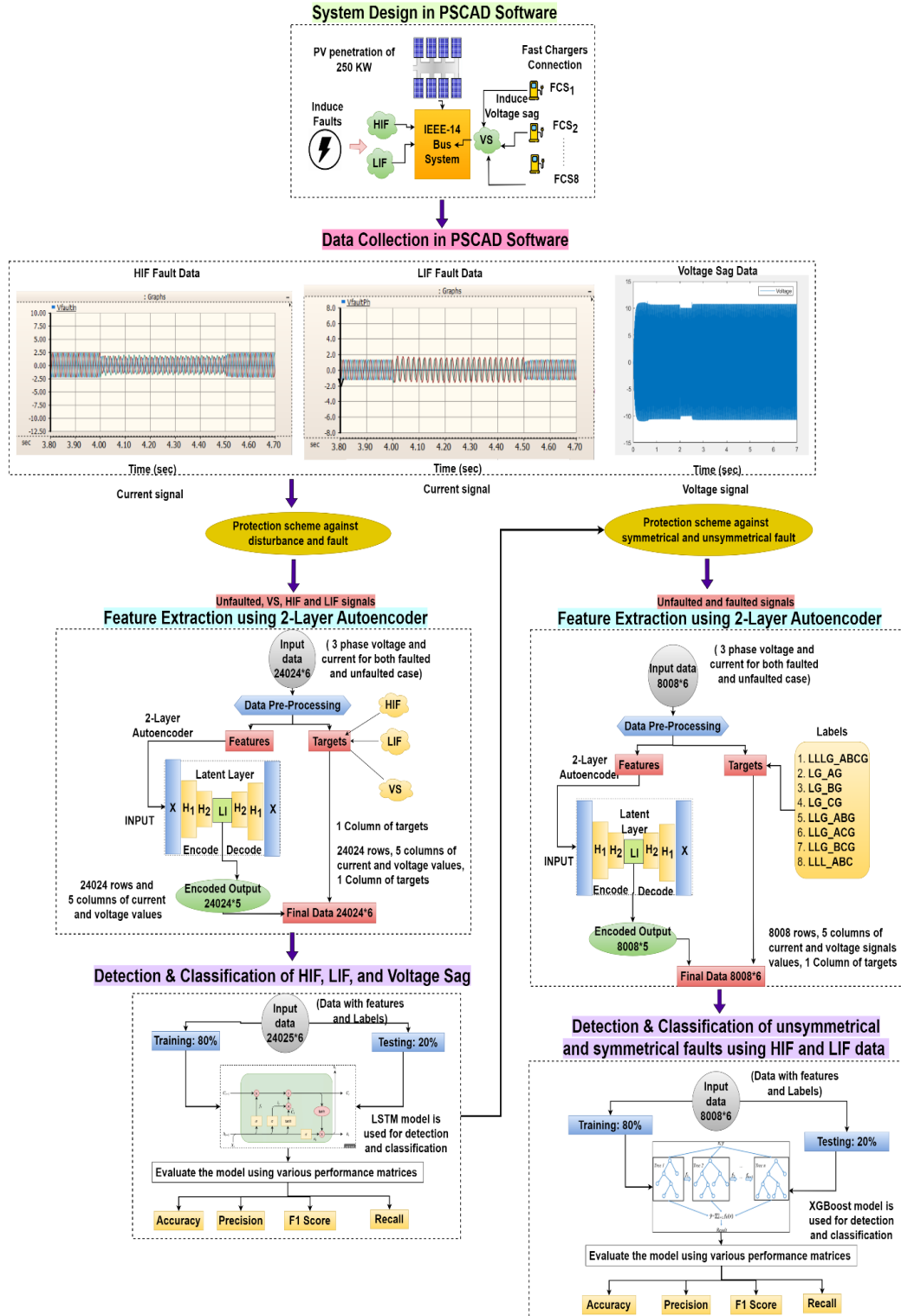


Figure 6. 1. Proposed methodology.

6.1. DESCRIPTION OF PROPOSED METHODOLOGY

As discussed in the introduction section, reliability and stability of the power system are the main concerns on which the power industry should focus. The methodology in the research paper has been developed keeping this challenge in mind. The methodology starts with designing the model in PSCAD. The FCSs with PV systems are integrated into the IEEE 14 bus system for analysing the effects of FCS on voltage sag. The HIF and LIF faults are also induced in the system with different values of resistance for analysing the fault effects in the system. Hence, three modes of operation are designed using HIF, LIF, and VS as described in Section 2.2, for the collection of data. The data is in the form of voltage and current signals. Once the data was gathered in PSCAD, we proceeded to carry out the necessary steps like data normalization using step 2. for detection and classification using Python. We first propose the feature extraction technique after conducting a thorough analysis of the data as define in step 3 in algorithm 1. Due to the noise robustness and dimensionality reduction properties of autoencoders, a two-layer autoencoder is proposed for feature extraction. A dataset consisting of 24024 instances and 6 features is inputted into TLA. The dataset is then transformed into 5 columns by feature reduction. To prepare the data for classification, an additional column of labels (HIF, LIF, and VS) is added, resulting in 24024 instances with 6 columns (x_1). Therefore, the final dataset contains 24024 instances with 6 columns as can be seen in third block of Figure 6.1. As this data is time series-based, an LSTM model is proposed for the detection and classification of faults using step 12 of algorithm 1, and can be seen in the fourth block of Figure 6.1. To evaluate the model, accuracy, precision, recall, and F1 score performance matrices are proposed. After the initial classification of broad categories of faults, classification and detection are suggested for protection against symmetrical and unsymmetrical faults for additional protection. This subsequent classification utilizes 8008*6 HIF and 8008*6 LIF input data. The feature extraction technique employs the same TLA. We extract the output from the encoded latent layer, which has five dimensions, resulting in a reduced feature data set of 8008*5. We propose XGBoost for subsequent fault detection and classification, providing one column with eight labels for LG, LLG, LLL, and LLLG faults as labels, as shown in the last block of Figure 6.1. The separate

evaluation of XGBoost has been done for HIF and LIF faults, and the results will be explained in the coming sections. The pseudocode can be seen in algorithm 1.

Algorithm 1. Pseudocode for Fault detection and prediction

Input: HIF, LIF, and VS data ($X = x_1, x_2, x_3$), model parameters as in Table 2. N number of samples in t , x_{1n} number of samples in x_1 , x_{2n} is number of samples in x_2 , x_{3n} is number of samples in x_3

- Initialize the process using HIF, LIF, and VS data
- Carry out the signal data preprocessing (normalising and encoding) on X
- For $epoch$ in 0: 100:
- Train the Autoencoder for signals feature extraction (section 2.3), as per equation (1-3).
- Generate t as per equation 1.
- End
- Split data t and X into train (80%) and test (20%)
- For $epoch$ in 0: 100 #training loop
- For N in t do:
- Train the LSTM classifier model (as per section (2.4)), evaluating the model over cross entropy loss function, as per equation (15) and classify the labels HIF, LIF, and VS (3 labels).
- End
- Split data t and x_2 and x_3 into train (80%) and test (20%)
- For $epoch$ in 0: 100 #training loop 2 for XGB
- For N in t do:
- Train the XGB classifier model (as per section (3.5) and evaluating the model over cross entropy loss function and classify the labels (LG, LLG, LLL, LLLG) of 8 labels from x_2 and x_3 as seen in Figure 6.1.
- End

6.2. STEPS IN METHODOLOGY

This section describes how unsupervised learning and deep learning are utilized to discern and detect faults in FCS integrated transmission systems. The processes involved in the classification process are illustrated in detail, along with an explanation, in Figure 2. The subsequent steps outline the process:

Data Collection: The 14-bus system with FCS and PV was simulated in PSCAD, as seen in the early stages in Figure 2. The LIF and HIF are implemented among bus 5 and bus 6. Data is subsequently gathered between 3 to 3.5 sec for eight distinct categories of short-circuit faults: LG_AG, LG_BG, LG_CG, LLG_ABG, LLG_ACG, LLG_BCG, LLL_ABC, and LLLG_ABCG. We have detected all eight defects with varying resistances as listed in Table 3. In addition, the voltage sag data is gathered by manipulating the switches attached to FCSs at PCC, as described in section 3.1. The data is collected for 2 to 2.5 sec in which voltage sag occurs. To extract features, we collect the current and voltage values of each phase under these conditions using a two-layer autoencoder. We gathered the data under various circumstances to enhance the realism of the testing system.

Feature Extraction: Post preprocessing, the two-layer autoencoder is employed to extract characteristics of the three-phase current and voltages. A similar two-layer autoencoder is employed both for the first classification (VS, LIF, and HIF) with a label of 3, and for the subsequent classification (Eight classes of LIF and HIF) with a label of 8, serving as a feature extraction technique. The introduction section discusses the benefits of utilizing autoencoders compared to other approaches. The purpose of this was to extract important features for accurate training of the LSTM model as well as the XGBoost model, which are used for the classifications and detection of faults.

Training and testing Phase: The LSTM classifier, which is based on deep learning, is trained to perform initial classification using the current and voltage values collected from the autoencoder in the three phases. The LSTM model is trained using three labeled datasets (VS, LIF, and HIF) to classify and detect disturbances or anomalies. Initially, we have a total of $24024 * 6$ data points, with 8008 rows in each class (LIF,

HIF, and VS). Following the use of autoencoder for feature extraction, we obtained a dataset consisting of 24,024 instances, each containing 5 features, with one feature being reduced. This enables the data to be trained efficiently, leading to enhanced accuracy. Following the initial classification, XGBoost is employed for future identification and detection of LIF and HIF, which are assigned the label 8. At this stage, we have a total of $8008 * 6$ data points for each LIF and HIF. We utilized a total of 8008 input data points, comprising 6 columns for each phase current of bus 5 and 6, due to the presence of 8 labels. The data is partitioned into groups with a ratio of 80:20 to optimize the efficiency of the training and testing procedure using steps 13 and 18 in Algorithm 1.

Results of Detection and Classification: In this particular instance, a portion representing 20% of the data samples was used to assess the efficacy of deep learning-based LSTM classifiers and XGBoost classifiers in detecting various events occurring in the established transmission network. Afterward, the various assessment elements are evaluated, including memory, accuracy, F1 Score, and precision.

6.3. FEATURES EXTRACTION USING TWO-LAYER AUTOENCODER

As we discussed in the introduction section, there are lots of methods available for feature extraction, but we have used a two-layer autoencoder for its superiority over other methods. The description of a two-layer autoencoder is given in section 2.1. An autoencoder is composed of an encoder and a decoder. The encoder's role is to convert input data with a high number of dimensions into a latent space with fewer dimensions. The encoder decreases the number of dimensions in the input data to a latent space, while the decoder endeavours to recreate the initial input based on this latent representation. We have taken the output from the encoder only; thus, the shape of the data will be the shape of the latent space. We have given latent dimension as 5 so the input data $24024 * 6$ will become $24024 * 5$. The autoencoder is trained on the normalized current and voltage data from transmission line sensors using a mean squared error loss function. The optimization is performed with the Adam optimizer, with a learning rate of 0.001. After training, the encoder part of the autoencoder is utilized to extract compact feature vectors from the input signals. The same type of

two-layer autoencoder will be used for both the classification. For generating the training and testing dataset, Voltage Sag, different types of LIF and HIF with varying parameters like FIA and line parameter variation, is considered. Table 3 provides the specifics of the parameter adjustments. The hyperparameters utilized for the two-layer autoencoder are shown in Table 4. The procedure of extracting features for fault detection and classification is illustrated in Figure 2.

6.4. VOLTAGE SAG, LIF AND HIF DETECTION AND CLASSIFICATION USING LSTM

The protection scheme has been developed to establish protection against faults and other disturbances for low-voltage transmission lines connected with fast charging stations using the LSTM classifier. The functionality of LSTM is elucidated in section 2.2. This method utilizes sequential data, specifically post-fault current samples, to accurately identify and classify faults. Following data collection from the PSCAD model, the data is fed into an autoencoder to extract features. The signals are then used as consecutive input for the categorization model. The size of the input data in the cell array is 24024*5 after feature extraction. The LSTM model is trained using these data. The LSTM model comprises an Input layer, an LSTM layer, a fully connected layer, a fully connected layer, and a SoftMax layer. These layers are used for the detection and classification of LIF, HIF, and VS, followed by the output layer. The categorical array consists of targeting labels that correspond to various fault case conditions. The various performance matrices are used to assess the performance of the LSTM model.

6.5. SYMMETRICAL AND UNSYMMETRICAL FAULTS DETECTION AND CLASSIFICATION USING XGboost

Once the faults and disturbances have been identified and classified using LSTM, we proposed XGboost to categorize symmetrical and unsymmetrical faults using LIF and HIF data. Figure 6.1. provides a visual representation of this. This protection strategy is designed to provide accurate protection against specific faults such as line-to-line (LL), line-to-line-to-ground (LLG), line-to-line-to-line-to-ground (LLLG), etc. by utilizing HIF and LIF data. This approach comprises two separate stages: the training

stage and the testing stage. Before the training and testing split, the data's features are retrieved using a two-layer autoencoder. The initial data size of the three phase current inputs is 8008×6 . After feature extraction, the number of columns is reduced to 5. Therefore, once the feature extraction and labelling process is completed, the data size is again 8008×6 .

The maximum depth of each tree is denoted by the term "maximum tree depth". When focusing on enhancing performance, using a deeper tree can lead to greater complexity and a higher danger of overfitting. The value is always positive and has a default value of 6. The defect detection module is most effectively implemented with a tree depth of 3, using the trial-and-error method. The model initiates the optimization process by progressively increasing its learning rate at each step toward its aim. It necessitates further iterations with a small learning rate. However, the likelihood of achieving the optimal outcome is also enhanced. After conducting multiple experiments, the optimal learning rate is determined to be 0.1. The quantity of trees in an ensemble is indicated by the number of estimators. Performing a solitary boost cycle is tantamount to a single tree. After conducting multiple trials, it has been determined that the optimal number of trees is 100. The problem of overfitting is regulated by the minimal kid weight. The greater the value, the less likely it is for a model to learn associations using very explicit samples to construct a tree. The regularization parameters utilized are lambda (λ), alpha (α), and gamma (γ). A value of 1 for lambda decreases overfitting. When Alpha is set to 0, the model's performance improves as the number of features increases. When Gamma is set to 0, it functions as a pseudo regularization parameter.

PART 7

RESULTS AND DISCUSSION

This section discusses the various fault cases like HIF, LIF, and VS with variable resistances and FIAs. We induced the faults at buses 5 and 6 and at different resistance configurations and collected the data, as specified in Table 7.1. To analyse the increased demand effect of a fast-charging station on the grid, we have induced the VS using different switching strategies[116]at various instances and collected the data. This section discussed the voltage and current signal patterns at different faults. This will allow us to explain why we use this variational data and show the non-linearities presented in the signals. The three-phase current and voltage waveforms for each fault can be shown in Figures 7.1, 7.2, and 7.3.

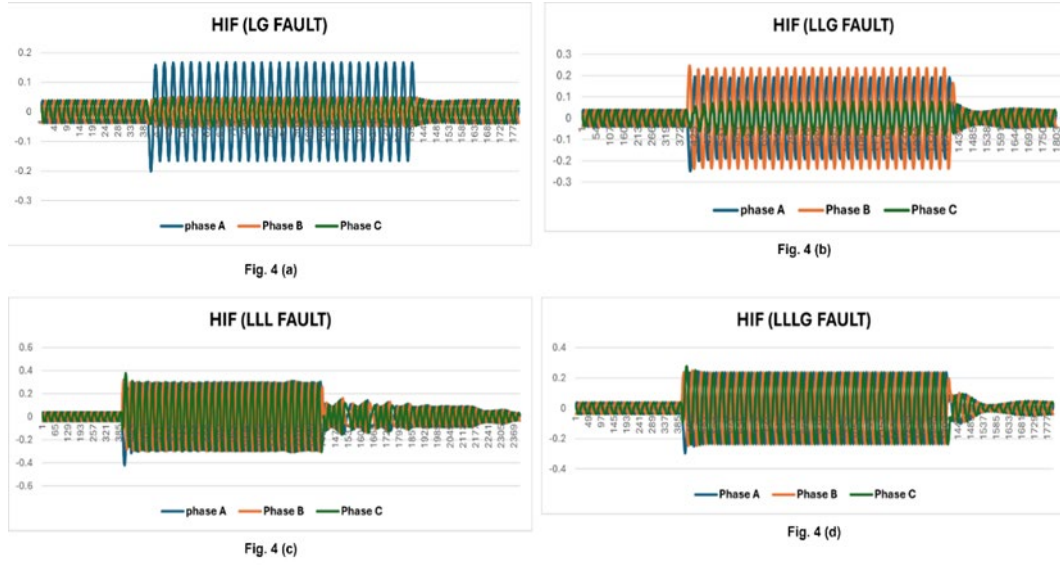
7.1. ANALYSIS OF VARIOUS FAULT CASES

This section discusses the various fault cases like HIF, LIF, and VS with variable resistances and FIAs. We induced the faults at buses 5 and 6 and at different resistance configurations and collected the data, as specified in Table 7.1. To analyse the increased demand effect of a fast-charging station on the grid, we have induced the VS using different switching strategies [116] at different instances and collected the data. In this section, we discussed the voltage and current signal patterns at different faults. This will allow us to explain why we are using this variational data and show the non-linearities presented in the signals. The three-phase current and voltage waveforms for each fault can be shown in Figures 7.1, 7.2, and 7.3.

7.1.1. Impact Assessment of High Impedance Faults

This section examines the effects of HIF on the current waveforms of a three-phase system, primarily focusing on LG, LLG, LLL, and LLLG faults. The LG (AG) fault

can be shown in Figure 7.1 (a); as can be seen in Figure 7.1, phase A experiences notable disruptions in comparison to phases B and C. The initial increase in the amplitude of Phase A's current waveform indicates that the fault primarily impacts this phase. The continuing presence of the fault causes the waveform to exhibit oscillations, indicating the effect of the fault on the current in Phase A. The oscillation seen can be ascribed to the intrinsic properties of high-impedance faults, which result in intermittent contact with the ground, causing periodic interruptions. Figure 7.1 (b) illustrates the LLG fault occurring between phases A and B and the ground. In these types of problems, the impedance is considerably greater than in conventional fault conditions, leading to smaller fault currents that are difficult to detect. Figure 7.1 (b) shows that the current value of phase B is greater than that of phase A, while the current of phase C remains relatively constant. Figure 7.1 (c) depicts the LLL fault. During the fault onset, there is a brief occurrence of strong current spikes throughout all phases. Over time, a significant amount of electric current is consistently maintained. However, it is worth noting that there are oscillatory patterns with different amplitudes, which indicate the intricate interaction among the three phases during the fault. Figure 7.1 (d) displays the current waveforms for phases A, B, and C when there is a high-impedance LLLG fault. This fault type occurs when all three phases of an electrical system make contact with the ground simultaneously. It creates a complex fault situation where the fault currents are lower because of the high impedance nature of the problem. The graph illustrates the present behavior over a period of time, reflecting the system's reaction to this failure state. The fault's low current magnitude and intermittent nature may not be enough to trigger typical overcurrent protection devices, thereby requiring the development of more sensitive and advanced detection systems. Hence, the advanced pattern recognition method is required.



(a) LG (AG) fault current signal at the Bus-5, (b) LLG (ABG) fault current signal at the Bus-5, (c) LLL (ABC) fault current signal at the Bus-5, (d) LLLG (ABCG) fault current signal at the Bus-5.

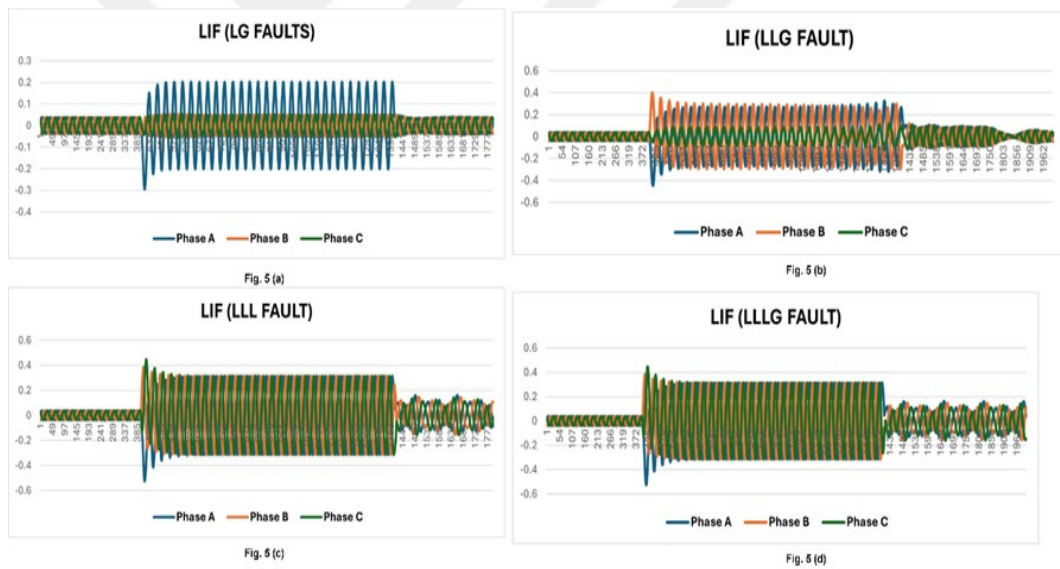
Figure 7. 1. HIFs at 3 to 3.5 seconds with 0° inception angle and $100 \, \Omega$ resistance.

7.1.2. Impact Assessment of Low Impedance Faults

This section presents the findings of the simulation conducted on the IEEE14 bus model. The experiment involved inducing low-impedance faults at a specific time interval of 3 to 3.5 seconds. The recorded results focus on the faults categorized as LG, LLG, LLL, and LLLG. Current waveforms for three phases (A, B, and C) under different low-impedance faults are shown in Figure 7.2. In the LG fault scenario (Figure 7.2(a)), Phase A's current is much higher than Phases B and C's, indicating a straight path to the ground. All phases have stable currents cycling about the baseline before the fault; however, Phase A surges sharply while Phases B and C encounter slight disturbances. The large rise in faulty phase current makes this fault easier to identify. Phases A and B have higher currents than Phase C in the LLG fault scenario (Figure 7.2(b)), indicating a fault involving two phases and ground. Due to coupling effects, the affected phases oscillate during fault initiation, while Phase C stays stable. LLG faults are detected by monitoring phase currents and zero-sequence components.

In the LLL fault scenario (Figure 7.2(c)), all three phases have high and oscillatory currents. A symmetrical fault is indicated by phase-balanced magnitudes and

oscillations. This defect causes a considerable current increase across all phases with similar waveforms. Differential protection and overcurrent relays are needed to identify severe LLL faults with huge fault currents. In the LLLG fault scenario (Figure 7.2(d)), all three phases have higher currents like the LLL fault, but the ground complicates things. This causes more complex oscillations. Due to the ground component, the waveforms interact differently from the LLL fault. Comprehensive protection solutions must include phase and ground fault detection for LLLG fault detection. Figure 7.2 shows how low-impedance fault current waveforms differ. Unique waveform patterns and magnitudes for each fault type emphasize the need for specific protection solutions for accurate detection and mitigation. Maintaining system reliability and safety requires advanced protection systems that monitor and analyse these waveforms.



(a) LG (AG) fault current signal at the Bus-5, (b) LLG (ABG) fault current signal at the Bus-5, (c) LLL (ABC) fault current signal at the Bus-5, (d) LLLG (ABCG) fault current signal at the Bus-5.

Figure 7. 2. LIFs at 3 to 3.5 seconds with 0° inception angle and 10Ω resistance.

7.1.3. Impact Assessment of Voltage Sag Due To Fcs

Figure 7.3 provides a clear representation of how fast charging stations affect voltage levels in an electrical system, demonstrating a significant reduction in voltage across all phases. The voltage at the PCC is compared to the voltage sag threshold. The

voltage sag scenario that arises during a charging event at any of the fast chargers (depicted in Figure 1) results in the voltage at the point of PCC remaining less than 90% of its designated voltage for a specific number of cycles [116]. Figure 7.3 depicts the calculated voltage waveforms for the three-phase system. Figure 7.3 shows that the breaker S6 is activated at time t , indicating the occurrence of a charging event at FCS3. Furthermore, the activation of S5 occurred prior to this, suggesting that a charging process is currently taking place at FCS5. Upon activating breaker S6, the current being drawn at the PCC experienced a rise, causing the voltage to drop below the specified threshold. The period of voltage sag and voltage drop can be observed in Figure 7.3. The droop observed is indicative of the substantial power demand exerted by these stations, leading to transient voltage reductions. Understanding these effects is essential for developing strategies to mitigate power quality problems and ensure the reliable operation of electrical systems as the adoption of fast-charging infrastructure increases.

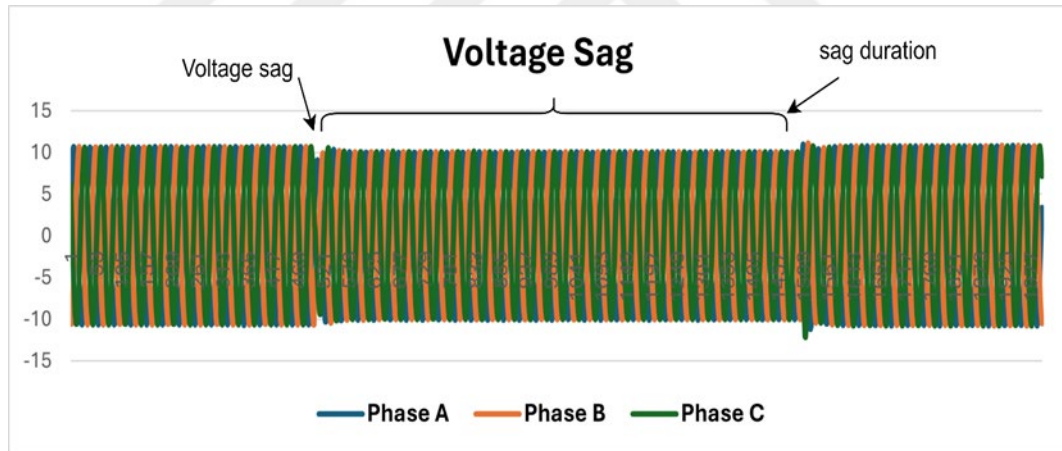


Figure 7. 3. Voltage sag at 2 to 2.5 seconds due to FCSs.

7.2. PERFORMANCE ANALYSIS OF DEEP LEARNING-BASED MODEL FOR DETECTION AND CLASSIFICATION OF FAULTS

The efficacy of the proposed methodology for safeguarding transmission lines has been proved in this section. We have developed three operating fault scenarios: normal mode, HIF mode, and LIF mode, as well as a mode for connecting to a fast-charging station. In the first scenario, no fault has been detected. In the second scenario, both

high-impedance and low-impedance faults have been detected. In the third scenario, fast-charging stations have been attached to measure the voltage drop caused by the fast-charging process. The fault is caused by variations in resistances and the angle of fault impedance. Occurrences of the second and third scenarios might result in a decrease in voltage at the PCC, also known as voltage sag. The third scenario is primarily designed to differentiate between the sources of the voltage sag in order to ascertain whether it is caused by system failures or the activation of a Fault Current Limiter (FCS). The features chosen by the autoencoder are utilized for training and evaluating the LSTM model. The fault signals are measured at different resistances and fault inception angles to incorporate the non-linear signals, which are well managed by TLA. The deep learning model is implemented in Python after collecting the data from PSCAD. Table 7.1 displays the specific operating scenario details that were taken into account throughout the training and validation of the LSTM modules. The details provided in Table 7.1 list the hyperparameters utilized for LSTM.

Table 7. 1. Hyperparameters of autoencoder, LSTM, and XGBoost.

AUTOENCODER	
Latent Dimension	5
Epochs	100
Learning Rate	0.001
Loss Function	MSE
Optimizer	Adam
LSTM MODEL	
Hidden Layers	50
Epochs	100
Learning Rate	0.001
Batch Size	64
Loss Function	Cross Entropy Loss
Optimizer	Adam
XGBOOST	
Learning Rate	0.3
Number of estimators	100
Maximum depth	6

The data was collected in order to identify and classify the fault, which is labeled as HIF, LIF, and VS. A LSTM model has been used to address a multi-class classification problem with three classes. The fault type classification model is trained using 100 epochs and the learning rate is initialized to 0.001. The LSTM model has attained a precision of 94%, signifying that 94% of the instances classified as HIF were accurately recognized as genuine HIF cases. The F1-score of 0.95 achieved by HIF demonstrates a well-balanced performance in detecting high impedance issues, demonstrating equal emphasis on precision and recall. The HIF has a recall rate of 0.96, indicating that it correctly detected 96% of the HIF instances in the dataset.

A HIF classification accuracy of 0.96 shows that the model accurately categorized 96% of the events. The model's accuracy is in its ability to distinguish high impedance difficulties from other flaws and typical scenarios. Figure 7.4 displays the performance metrics of the LSTM model, organized by class.

The LSTM model attained a precision of 0.96 in the LIF category, indicating that 96% of the LIF events were correctly detected. By ensuring a high level of accuracy, the occurrence of false positives is minimized, hence ensuring the reliability of LIF detection. The model's recall of 0.94 signifies that it accurately recognized 94% of the LIF instances in the dataset. The recall rate demonstrates the model's ability to detect most low-impedance defects, while a small number may remain undetected. The model demonstrates its efficacy in addressing low impedance difficulties by accurately detecting the presence or absence of LIF in 96% of cases. The F1 score for LIF was 0.95, indicating that the model achieves a strong trade-off between precision and recall. It will decrease the occurrence of both false negatives (cases of LIF that are missed) and false positives (cases that are incorrectly classified as LIF when they are not).

The LSTM model demonstrated outstanding performance in accurately predicting Voltage Sag caused by Fast Charging Stations, with a perfect precision score of 1.00. In essence, the model has the ability to precisely identify all VS results and correctly categorize them as authentic without any erroneous classifications. The recall rate was 100%, indicating that the model correctly identified every occurrence of voltage sag.

The F1-score achieved a perfect value of 1.00, indicating exceptional precision in correctly recognizing VS instances and avoiding any misclassifications.

The accuracy for VS classification was slightly lower at 0.99, yet this still shows a very high level of accuracy, as the model accurately detected VS cases with almost perfect precision and recall. This demonstrates the model's exceptional capacity to effectively address the specific and crucial issue of voltage sag, which is commonly associated with fast charging stations. As a result, it guarantees reliability and resilience in real-world scenarios.

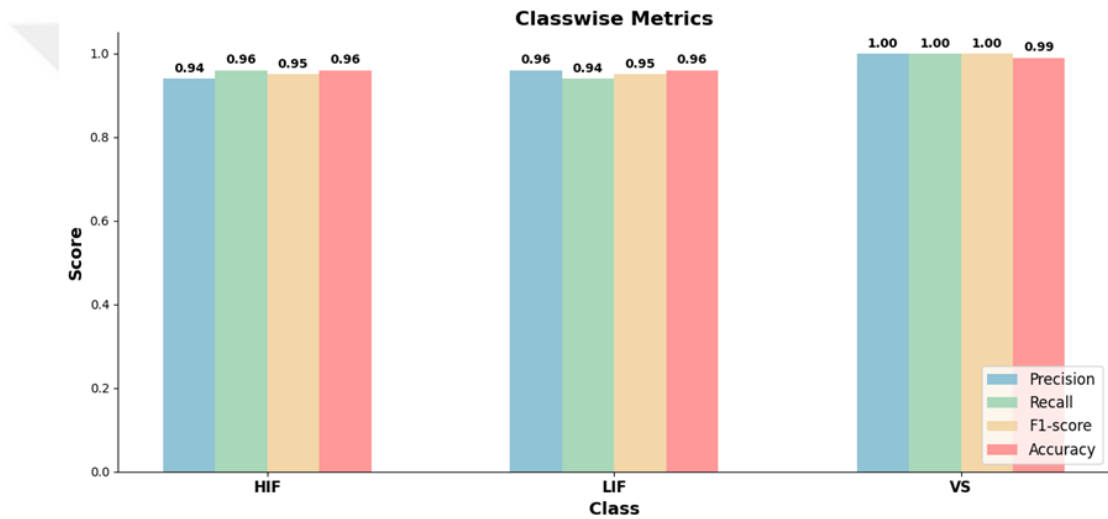
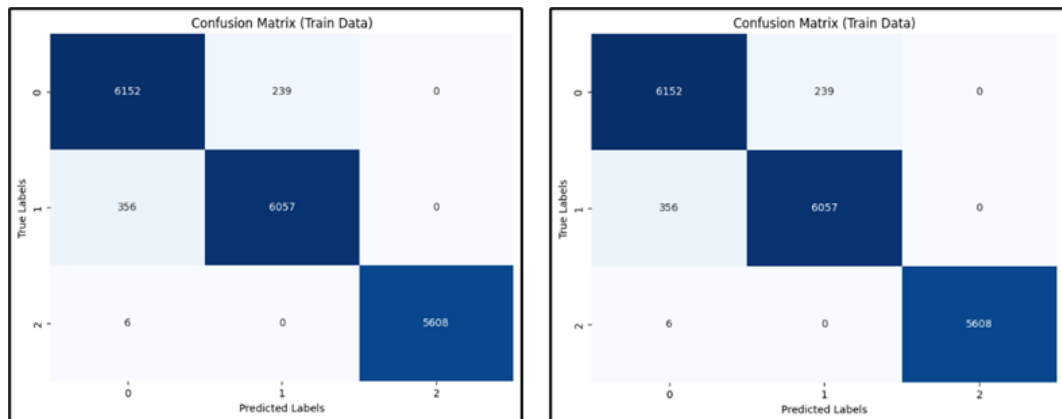


Figure 7. 4. Performance metrics of LSTM.



7.5. (a)

7.5. (b)

Figure 7. 5. Confusion matrix of LSTM. 7.5 (a) Train data 7.5 (b) Test data.

The confusion matrix for the training and testing of the LSTM model is depicted in Figures 7.5 (a) and 7.5 (b). The confusion matrices offer a comprehensive and comparative evaluation of the model's predictive precision, highlighting its strengths and areas for enhancement in the system model. The confusion matrices for High Impedance Fault (HIF) demonstrate that the model accurately recognized 6,152 instances in both the training and testing data, suggesting a consistent capacity to detect HIF situations with precision. The training data revealed that the model exhibited 6 instances of false positives and 239 instances of false negatives. This indicates the model's general competence but also highlights its difficulty in accurately differentiating HIF from other sorts of faults. During the testing phase, the model maintained the same level of accuracy in correctly identifying instances. However, it showed a modest increase in the number of erroneous negatives, namely 356, while the number of false positives remained unchanged. The confusion matrices for the training and testing phases of the LSTM model are sequentially displayed in Figure 7.5 (a) and Figure 7.5 (b). These matrices offer a comprehensive assessment of the model's prediction precision, highlighting its very advantages along with the areas that need enhancement.

The model successfully identifies 6,152 occurrences of High Impedance Faults (HIF) in both the training and testing data, indicating its consistent ability to fully detect HIF cases. Nevertheless, training data also uncovered six instances where the model incorrectly identified a problem as positive and 239 instances where it failed to identify a fault that was actually there, suggesting difficulties in accurately identifying HIF from other sorts of faults. During the testing phase, the model exhibits a slightly elevated count of false negatives (356), while maintaining the same count of false positives (6). This indicated that the model had some challenges in effectively applying its acquired knowledge to unfamiliar data. The model demonstrated strong and reliable performance for Low Impedance Faults (LIF) in both datasets. During the training phase, it accurately detects 6,057 instances of LIF cases, with only 239 incorrect positive identifications and one incorrect negative identification. The testing data demonstrates consistent accuracy through 6,057 correct identifications, a low number of false positives, and a limited number of false negatives. Voltage Sag, LIF, and HIF frequently display pronounced temporal patterns and interdependencies. LSTM, as a

variant of Recurrent Neural Network (RNN), is specifically intended to efficiently capture sequential data patterns. Hence, LSTM is well-suited for identifying the changing attributes of these types of faults as time progresses.

7.3. PERFORMANCE ANALYSIS OF SYMMETRICAL AND UNSYMMETRICAL FAULT DETECTION AND CLASSIFICATION MODEL

Once the voltage sag, LIF, and HIF faults have been successfully classified, it is necessary to classify the subcategory faults (such as LLG, LL, etc.) that come under HIF and LIF in order to obtain accurate information about these faults. This will enhance the system's protection. There are a variety of factors that might cause LIFs, which are the most prevalent type of fault in the transmission system. One of the contributing reasons is the occurrence of a tree branch falling into the power cables. They provide a pathway with low impedance for current flow, therefore, the fault current typically has a quite large magnitude. As a result, they appear to be particularly hazardous for the system. This has the potential to produce a fire, as well as a potential danger to human safety and the equipment that is attached. This study examined the effects of several LIF cases with resistances from 0 to 40 ohms.

The XGBoost model is proposed for this subsequent classification. The classification of faults such as LG, LL, and LLL (e.g., LG_AG, LG_BG, LLG_ABG, etc.) is often determined using structured data that includes many attributes generated from LIF and HIF data. XGBoost does exceptionally well in managing structured data that has a large number of features, where the relationships between the features are more stable rather than changing over time. Hence, the XGBoost model has been used to address a multi-class classification problem with 8 classes. The fault type classification model is trained using 100 number of estimator, and the learning rate is initialized to 0.3, as can be seen in Table 4. The model has achieved a perfect level of precision, accuracy, F1 score, and recall, all of which are at 100%. The measured accuracy of the train data is 97 %. As depicted in Figure 7.6. The XGBoost model has exceptional performance across all investigated metrics over testing data in finding and classifying problems in the HIF category. The system thoroughly investigates each fault class, from one-phase

to three-phase complicated faults, achieving nearly 100% recall, precision, F1-score, and accuracy. It is evident that the model employed in power systems for identifying diseases produces consistent results for different types of faults. By doing this, we are providing a comprehensive solution for the management and diagnosis of issues in power systems.

Figure 7.7 displays the confusion matrix according to the test-train data of several situations. The XGBoost model has demonstrated high accuracy in both the training and testing phases of HIF fault detection, as well as fault classification. During the training phase, the model achieved almost 100% accuracy in identifying and categorizing all the fault categories, with very few misclassifications. Throughout the testing phase, the model has consistently demonstrated a comparable degree of accuracy to that observed during the training phase. Nevertheless, it exhibits some certain degree of misclassification, particularly when confronted with intricate fault detection. In general, the model performed well, with only minor errors in classifying defect categories.

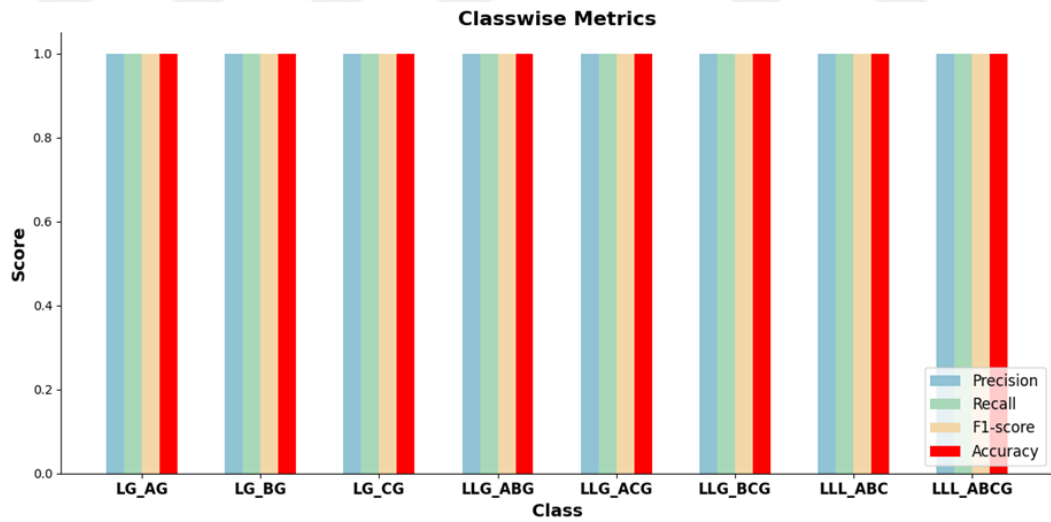


Figure 7. 6. Performance metrics of XGBoost for HIF.

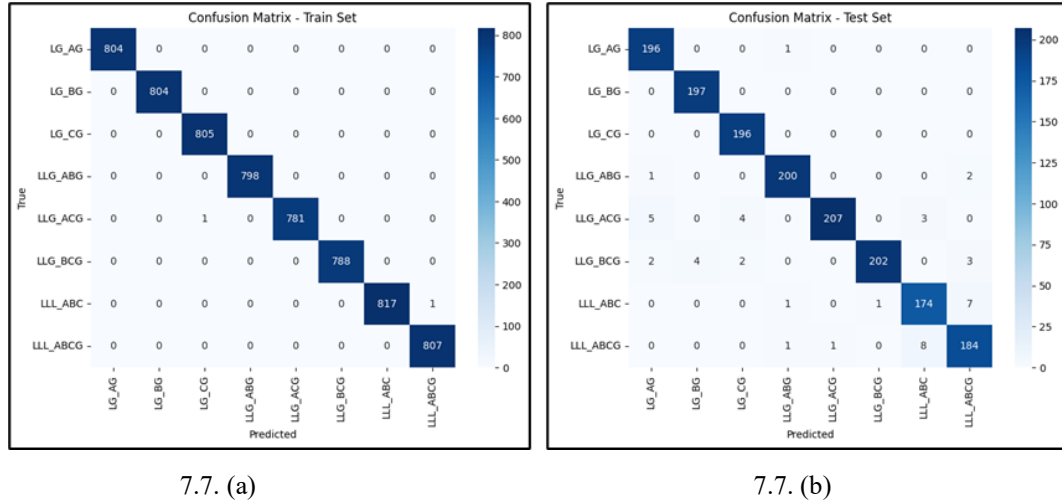


Figure 7. 7. Confusion matrix of XGBoost for HIF. 7.7 (a) Train data 7.7 (b) Test data.

The performance metrics of the XGBoost model for LIF are displayed in Figure 7.8. The model demonstrates exceptional performance in identifying and categorizing single-line-to-ground problems (LG_AG, LG_BG, LG_CG) with precision, recall, F1-scores, and accuracy all approaching or reaching 100%. These problems arise when one phase conductor comes into touch with the ground or a grounded item. Power systems commonly experience this form of fault more frequently than any other. These flaws have the potential to result in excessive electrical current and could result in damage to equipment or power outages. This suggests that the model exhibits a high level of dependability in these specific fault categories. The model also demonstrates outstanding performance for LLG failures. These failures occur when there is a failure in the insulation or physical contact between two phase conductors and the ground. The consequences of these faults are graver compared to single-line-to-ground faults, resulting in increased fault currents and more extensive damage. The model exhibits remarkable performance in detecting these errors, with high precision, recall, F1-score, and accuracy. This indicates that the model effectively classifies defects and has low misclassification. Nevertheless, the model demonstrates moderate performance when it comes to LLL errors, achieving a precision, recall, and F1-score of 0.65 each. The marginal decrease in precision and recall indicates a certain degree of misclassification, yet, on the whole, the model is still useful. These rare errors are exceedingly severe, resulting in significant electrical currents and potentially inflicting catastrophic damage to the system. Furthermore, Figure 7.8 demonstrates that the model attains equivalent performance for LLLG faults as it does for LLL faults. The

overall accuracy of the test data set is 0.77, whereas the accuracy of the train data set is 0.91. These findings indicate that although the model is generally successful, there is potential for enhancing its performance in dealing with more intricate defect types to attain even greater precision and recall.

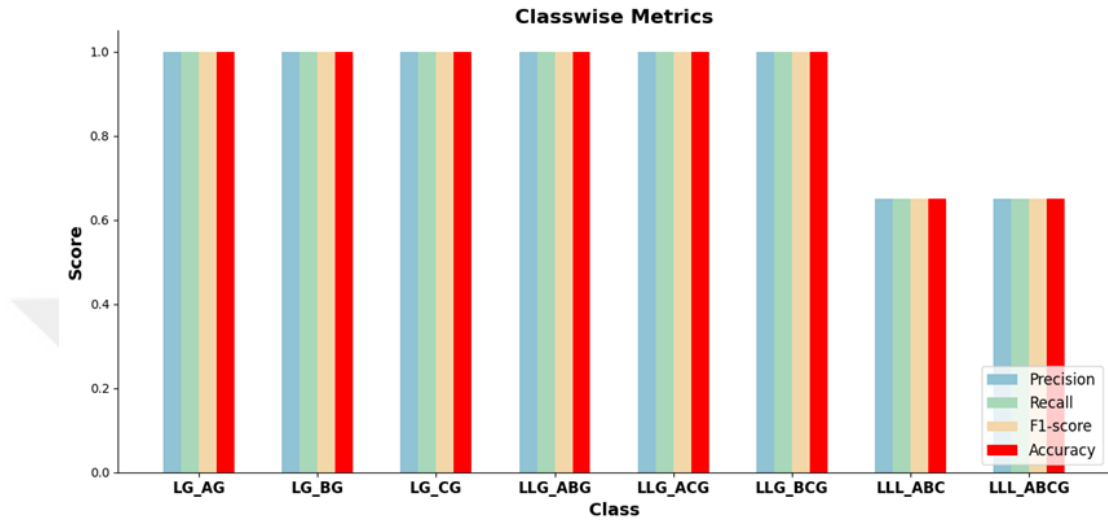
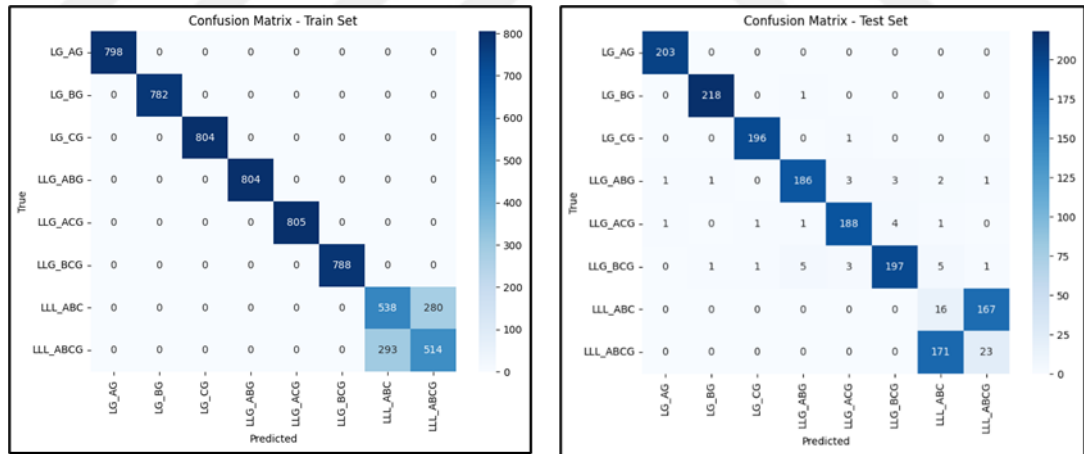


Figure 7. 8. Performance metrics of XGBoost for LIF.



7.9. (a)

7.9. (b)

Figure 7. 9. Confusion matrix of XGBoost for LIF. 12 (a) Train data 12 (b) Test data

The confusion matrices for the test set and train set illustrate the efficacy of the XGBoost model in precisely detecting and categorizing LIF into eight distinct categories. The confusion matrices of XGBoost can be seen in Figure 7.9. The model demonstrates a remarkable level of accuracy in identifying single-line-to-ground problems in the test set, with minimal occurrences of misclassification. The majority

of errors occur in the more intricate three-phase faults and three-phase-to-ground faults, suggesting a failure to differentiate between these types of problems. Similarly, the model exhibits outstanding performance in the train set for single-line-to-ground and line-to-line-to-ground defects, achieving virtually perfect classification. Nevertheless, there are notable misclassifications that occur between the LLL and LLLG groups. The findings indicate that the model exhibits robust performance in dealing with less complicated fault types, but faces difficulties when it comes to more intricate classifications. This problem could potentially be resolved by enhancing the model or incorporating supplementary distinctive characteristics.

7.4. STATE-OF-THE-ART COMPARISION

Table 7.2 displays the comparison findings of comparing the proposed scheme to various intelligent methods for fault classification and detection that have been reviewed in the literature for transmission line systems [3], [101], [104], [126], [127], [128], [129]. In the experiments mentioned in [3], TKEO and XGBoost were used to detect and classify faults. However, this literature did not consider the voltage sag in their study. Furthermore, the feature extraction technique they have used is incapable of learning complex and high dimensional featured data and unable to capture intricate patterns present in the data, whereas the two layer autoencoder that we used in our proposed method has all these capabilities. In addition, we have proposed a comprehensive approach for fault detection and classification including HIF, LIF, and VS. The study presented in [126] and [127] utilizes the SVM and LSTM models for detection and classification without using any feature extraction techniques. The accuracy they have achieved is commendable, but they did not consider the HIF fault as well as voltage sag, resulting in limited protection against other faults. In the research conducted in reference [128] for fault detection, the statistical machine learning framework was used. Although it demonstrated robustness for simpler networks in certain conditions, this study also considered limited fault types in comparison to ours. By contrast, our hybrid method identifies HIF, LIF, and VS faults with 99% accuracy, exhibiting exceptional precision and adaptability within a power system integrated with a PV system and fast charging stations. In this study [129], the author utilizes the wavelet transform for feature extraction, and for fault location, they

use a Generalized Neural Network (GNN). However, it encounters difficulties in precisely identifying LIF, and VS faults since overlapping of transient features may occur or are less distinct. Besides that, its high-frequency sampling (100 kHz) dependency and the wavelet analysis are computationally expensive, which will limit the real-time applications from scaling up.

Table 7. 2. Comparison of the proposed scheme with other methods.

Reference s	Signal recorde d	Data size	Feature Extraction	Classificatio n and detection technique	HIF Accurac y	LIF accurac y	VS accurac y
[3]	V & I	-	TKEO	XGBoost	100%	100 %	-
[101]	V	-	HOS	NN, SVM	-	-	99.56
[104]	V & I	-	-	SVM	-	-	99.17
[126]	V & I	27000	Not required	LSTM	-	99.45%	-
[127]	I	24000	Not required	LSTM autoencoder	-	97.6%	-
[128]	V	4950	Wavelet Transform	GridSearchC V optimized statistical ML	-	99.46 %	-
[129]	I	177	Wavelet Transform	GNN	94%	-	-
Proposed	V & I	24,02 4	Two-layer Autoencode r	LSTM, XGBoost	96%, 100%	96%, 91%	99.8%

Previous research studies[101], [104] have addressed the classification and detection of voltage sag but they have not considered the LIF and LIF faults. Additionally, the research does not provide information about the source of voltage sag. When compared to our proposed method, we have considered many types of faults, as well as the source of voltage sag, which is defined as FCS. The accuracy of our methodology is also higher when compared to [101] and [104]. Overall, in some cases, we have achieved less accuracy when compared to other methods, but we have covered many types of

faults and their subcategories as well; hence, our proposed methodology enhances the security of the power system by addressing a wide range of disturbances and faults.



PART 8

CONCLUSIONS

The proposed study introduces a hierarchical hybrid approach that significantly advances fault detection and classification in modern power systems integrating fast-charging stations (FCS) and photovoltaic (PV) systems. By combining the sequential data modeling capabilities of Long Short-Term Memory (LSTM) networks with the precise classification accuracy of XGBoost, the framework effectively addresses challenges such as voltage sag (VS), high-impedance faults (HIF), and low-impedance faults (LIF). The methodology achieves 96% accuracy for HIF detection, 91% for LIF detection, and 99.8% for voltage sag detection, outperforming existing methods in both accuracy and scalability. The framework's computational efficiency ensures real-time applicability and provides a reliable solution for mitigating critical power system challenges.

The results underscore the pivotal role of machine learning in intelligent power network management. The LSTM-XGBoost framework demonstrates near-perfect performance metrics, including precision, recall, and F1 scores, for detecting VS and classifying symmetrical and asymmetrical faults. The framework achieves an optimal balance of computational efficiency and adaptability by leveraging LSTM's ability to analyze sequential data and XGBoost's capacity to handle non-linear relationships.

Future research will explore integrating advanced machine learning techniques, such as attention mechanisms to improve the detection of subtle fault patterns like HIF [130], [131], and ensemble learning to enhance robustness in imbalanced datasets [132]. Graph Neural Networks (GNNs) could also aid in fault localization and cascading failure prediction [133]. Developing real-time hardware prototypes will further validate the framework's industrial applicability in microgrids and distributed energy systems. These advancements aim to enhance the framework's scalability,

generalizability, and reliability, ensuring it meets the evolving demands of modern power networks.



REFERENCES

1. Talab, O., Avci, İ., and Sultan, A. L., "MICROGRID ENERGY MANAGEMENT WITH DAY AHEAD AND NOVEL EGWO SOLUTION", (2023).
2. Phadke, A. G., Wall, P., Ding, L., and Terzija, V., "Improving the performance of power system protection using wide area monitoring systems", *Journal Of Modern Power Systems And Clean Energy*, 4 (3): 319–331 (2016).
3. Bhatnagar, M., Yadav, A., and Swetapadma, A., "Enhancing the resiliency of transmission lines using extreme gradient boosting against faults", *Electric Power Systems Research*, 207: (2022).
4. Gómez, J. C. and Morcos, M. M., "Voltage sag and recovery time in repetitive events", *IEEE Transactions On Power Delivery*, 17 (4): 1037–1043 (2002).
5. Davis George Beam Ctrlistopher J Melhom, T. E., Member, S., Member, I., Technology Center Pennsylvania Power, I., and Usa, P., "Voltage Sags: Their Impact on the Utility and Industria€ Customers", .
6. Bindi, M., Piccirilli, M. C., Luchetta, A., and Grasso, F., "A Comprehensive Review of Fault Diagnosis and Prognosis Techniques in High Voltage and Medium Voltage Electrical Power Lines", *Energies*, 16 (21): (2023).
7. Cebrian, J. C., Giacomini, J., Carneiro, C. A., Silva, G. B., and Morales-Paredes, H. K., "A Novel Methodology to Estimate Probability Density Function of Voltage Sag Duration and Failure Rates on Power Distribution Systems", *IEEE Access*, 11 (February): 16863–16874 (2023).
8. Uzair, M., Li, L., Eskandari, M., Hossain, J., and Zhu, J. G., "Challenges, advances and future trends in AC microgrid protection: With a focus on intelligent learning methods", *Renewable And Sustainable Energy Reviews*, 178 (February 2022): (2023).
9. Liu, Z., Wu, G., He, W., Fan, F., and Ye, X., "Key target and defect detection of high-voltage power transmission lines with deep learning", *International Journal Of Electrical Power And Energy Systems*, 142 (PA): 108277 (2022).
10. Tzelepis, D., Psaras, V., Tsotsopoulou, E., Mirsaeidi, S., Dyško, A., Hong, Q., Dong, X., Blair, S. M., Nikolaidis, V. C., Papaspiliotopoulos, V., Fusiek, G., Burt, G. M., Niewczas, P., and Booth, C. D., "Voltage and current measuring technologies for high voltage direct current supergrids: A technology review identifying the options for protection, fault location and automation applications", *IEEE Access*, 8: 203398–203428 (2020).

11. Mikulović, J., Šekara, T., and Forcan, M., "Power definitions for three-phase systems in terms of instantaneous symmetrical components", *International Journal Of Electrical Power And Energy Systems*, 147 (February 2022): (2023).
12. Sinsel, S. R., Riemke, R. L., and Hoffmann, V. H., "Challenges and solution technologies for the integration of variable renewable energy sources—a review", *Renewable Energy*, 145: 2271–2285 (2020).
13. shafiei, M. and Ghasemi-Marzbali, A., "Fast-charging station for electric vehicles, challenges and issues: A comprehensive review", *Journal Of Energy Storage*, 49 (February): 104136 (2022).
14. Smolenski, R., Szczesniak, P., Drozd, W., and Kasperski, L., "Advanced metering infrastructure and energy storage for location and mitigation of power quality disturbances in the utility grid with high penetration of renewables", *Renewable And Sustainable Energy Reviews*, 157 (December 2021): 111988 (2022).
15. Akhtar, S., Adeel, M., Iqbal, M., Namoun, A., Tufail, A., and Kim, K. H., "Deep learning methods utilization in electric power systems", *Energy Reports*, 10: 2138–2151 (2023).
16. Labrador Rivas, A. E. and Abrão, T., "Faults in smart grid systems: Monitoring, detection and classification", *Electric Power Systems Research*, 189 (August): 106602 (2020).
17. Stefanidou-Voziki, P., Sapountzoglou, N., Raison, B., and Dominguez-Garcia, J. L., .
18. Shair, J., Li, H., Hu, J., and Xie, X., .
19. Belagoune, S., Bali, N., Bakdi, A., Baadji, B., and Atif, K., "Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems", *Measurement: Journal Of The International Measurement Confederation*, 177: (2021).
20. "Electricity Grid Simple- North America - Electric Power Transmission - Wikipedia",
https://en.wikipedia.org/wiki/Electric_power_transmission#/media/File:Electricity_grid_simple-_North_America.svg (2025).
21. Bindi, M., Piccirilli, M. C., Luchetta, A., and Grasso, F., .
22. Ahmad, F., Iqbal, A., Ashraf, I., Marzband, M., and khan, I., .

23. Laninga, J., Nasr Esfahani, A., Ediriweera, G., Jacob, N., and Kordi, B., "Monitoring Technologies for HVDC Transmission Lines", *Energies*, 16 (13): (2023).
24. He, H., Du, E., Zhang, N., Kang, C., and Wang, X., "Enhancing the power grid flexibility with battery energy storage transportation and transmission switching", *Applied Energy*, 290: (2021).
25. Jiang, W., Zou, D., Zhou, X., Zuo, G., Ye, G. C., and Li, H. J., "Research on key technologies of multi-task-oriented live maintenance robots for Ultra High Voltage multi-split transmission lines", *Industrial Robot*, 48 (1): 17–28 (2021).
26. Pinheiro, M. G., Madeira, S. C., and Francisco, A. P., "Short-term electricity load forecasting—A systematic approach from system level to secondary substations", *Applied Energy*, 332: (2023).
27. Sánchez, A., Zhang, Q., Martín, M., and Vega, P., "Towards a new renewable power system using energy storage: An economic and social analysis", *Energy Conversion And Management*, 252: (2022).
28. Wang, L., Qin, Z., Slangen, T., Bauer, P., and Van Wijk, T., .
29. "Annual U.S. Transmission Data Review", (2018).
30. Dashti, R., Daisy, M., Mirshekali, H., Shaker, H. R., and Hosseini Aliabadi, M., "A survey of fault prediction and location methods in electrical energy distribution networks", *Measurement: Journal Of The International Measurement Confederation*, 184: (2021).
31. Wang, Y., Liu, J., Zhang, Z., Zheng, T., Ren, S., and Chen, J., "A faulty line detection method for single phase-to-ground fault in resonant grounding system with CTs reversely connected", *International Journal Of Electrical Power And Energy Systems*, 147: (2023).
32. Wong, S. Y., Choe, C. W. C., Goh, H. H., Low, Y. W., Cheah, D. Y. S., and Pang, C., .
33. "POWER SYSTEM ANALYSIS", .
34. Wang, X., Yan, Z., Zeng, Y., Liu, X., Peng, X., and Yuan, H., "Research on correlation factor analysis and prediction method of overhead transmission line defect state based on association rule mining and RBF-SVM", *Energy Reports*, 7: 359–368 (2021).
35. Teimourzadeh, H., Moradzadeh, A., Shoaran, M., Mohammadi-Ivatloo, B., and Razzaghi, R., "High Impedance Single-Phase Faults Diagnosis in Transmission Lines via Deep Reinforcement Learning of Transfer Functions", *IEEE Access*, 9: 15796–15809 (2021).

36. Yepes, A. G., Lopez, O., Gonzalez-Prieto, I., Duran, M. J., and Doval-Gandoy, J., "A Comprehensive Survey on Fault Tolerance in Multiphase AC Drives, Part 1: General Overview Considering Multiple Fault Types", *Machines*, 10 (3): (2022).
37. Jin, L., Mao, Y., Wang, X., Shi, P., Lu, L., and Wang, Z., "Optimization-Based Maximum-Torque Fault-Tolerant Control of Dual Three-Phase PMSM Drives Under Open-Phase Fault", *IEEE Transactions On Power Electronics*, 38 (3): 3653–3663 (2023).
38. Wang, X., Wang, Z., He, M., Zhou, Q., Liu, X., and Meng, X., "Fault-Tolerant Control of Dual Three-Phase PMSM Drives with Minimized Copper Loss", *IEEE Transactions On Power Electronics*, 36 (11): 12938–12953 (2021).
39. Zhang, S., Wang, R., Si, Y., and Wang, L., "An Improved Convolutional Neural Network for Three-Phase Inverter Fault Diagnosis", *IEEE Transactions On Instrumentation And Measurement*, 71: (2022).
40. Bauwens, T., Schraven, D., Drevling, E., Radtke, J., Holstenkamp, L., Gotchev, B., and Yildiz, Ö., .
41. Gao, Z. and Liu, X., .
42. Qin, H., Abulaiti, A., Maimaiti, A., Abulaiti, Z., Fan, G., Aili, Y., Ji, W., Wang, Z., and Wang, Y., "Integrated machine learning survival framework develops a prognostic model based on inter-crosstalk definition of mitochondrial function and cell death patterns in a large multicenter cohort for lower-grade glioma", *Journal Of Translational Medicine*, 21 (1): (2023).
43. Qian, H., Lee, B., Wu, Z., and Wang, G. P., "Research on DC arc fault detection in PV systems based on adjacent multi-segment spectral similarity and adaptive threshold model", *Solar Energy*, 264: (2023).
44. Song, Y., Wang, L., Sheng, L., Zhang, M., Liang, H., Ren, D., Cui, H., Zhang, H., Xu, H., and He, X., .
45. Xiao, H., Sun, K., Pan, J., Li, Y., and Liu, Y., .
46. Zhang, H., Xiang, W., Lin, W., and Wen, J., "Grid Forming Converters in Renewable Energy Sources Dominated Power Grid: Control Strategy, Stability, Application, and Challenges", *Journal Of Modern Power Systems And Clean Energy*, 9 (6): 1239–1256 (2021).
47. Nasser, A. R., Azar, A. T., Humaidi, A. J., Al-Mhdawi, A. K., and Ibraheem, I. K., "Intelligent fault detection and identification approach for analog electronic circuits based on fuzzy logic classifier", *Electronics (Switzerland)*, 10 (23): (2021).
48. Erdiwansyah, Mahidin, Husin, H., Nasaruddin, Zaki, M., and Muhibbuddin, .

49. Raouf, A., Tawfiq, K. B., Eldin, E. T., Youssef, H., and El-Kholy, E. E., .
50. Alanazi, F., "Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation", *Applied Sciences (Switzerland)*, 13 (10): (2023).
51. Nasser Mohamed, Y., "CONDITION MONITORING AND FAULT DETECTION FOR ELECTRICAL POWER SYSTEMS USING SIGNAL PROCESSING AND MACHINE LEARNING TECHNIQUES", (2024).
52. Sarker, I. H., .
53. Montesinos López, O. A., Montesinos López, A., and Crossa, J., "Multivariate Statistical Machine Learning Methods for Genomic Prediction", *Multivariate Statistical Machine Learning Methods for Genomic Prediction, Springer International Publishing*, (2022).
54. Saha, S., Gan, Z., Cheng, L., Gao, J., Kafka, O. L., Xie, X., Li, H., Tajdari, M., Kim, H. A., and Liu, W. K., "Hierarchical Deep Learning Neural Network (HiDeNN): An artificial intelligence (AI) framework for computational science and engineering", *Computer Methods In Applied Mechanics And Engineering*, 373: (2021).
55. Soori, M., Arezoo, B., and Dastres, R., .
56. Nti, I. K., Adekoya, A. F., Weyori, B. A., and Nyarko-Boateng, O., .
57. Perera, A. T. D. and Kamalaruban, P., .
58. Wang, C., Luo, D., Deng, Q., and Yang, G., "Dynamics analysis and FPGA implementation of discrete memristive cellular neural network with heterogeneous activation functions", *Chaos, Solitons And Fractals*, 187: (2024).
59. Narkhede, M. V., Bartakke, P. P., and Sutaone, M. S., "A review on weight initialization strategies for neural networks", *Artificial Intelligence Review*, 55 (1): 291–322 (2022).
60. Fang, W., Chen, Y., and Xue, Q., "Survey on Research of RNN-Based Spatio-Temporal Sequence Prediction Algorithms", *Journal On Big Data*, 3 (3): 97–110 (2021).
61. Mienye, I. D., Swart, T. G., and Obaido, G., "Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications", *Information*, 15 (9): 517 (2024).
62. Ashraf Zargar, S., "Introduction to Sequence Learning Models: RNN, LSTM, GRU", .
63. Heidari, A., Jafari Navimipour, N., Unal, M., and Toumaj, S., .

64. Yousuf, H., Lahzi, M., Salloum, S. A., and Shaalan, K., .
65. Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., and Poria, S., "A Review of Deep Learning Techniques for Speech Processing", (2023).
66. Michelucci, U., "An Introduction to Autoencoders", (2022).
67. Berahmand, K., Daneshfar, F., Salehi, E. S., Li, Y., and Xu, Y., "Autoencoders and their applications in machine learning: a survey", *Artificial Intelligence Review*, 57 (2): (2024).
68. Menghani, G., .
69. Athlur, S., Saran, N., Sivathanu, M., Ramjee, R., and Kwatra, N., "Varuna: Scalable, Low-cost Training of Massive Deep Learning Models", (2022).
70. Mora, E., Cifuentes, J., and Marulanda, G., "Short-term forecasting of wind energy: A comparison of deep learning frameworks", *Energies*, 14 (23): (2021).
71. Tan, K. and Wang, D., "Towards Model Compression for Deep Learning Based Speech Enhancement", *IEEE/ACM Transactions On Audio Speech And Language Processing*, 29: 1785–1794 (2021).
72. Abid, A., Khan, M. T., and Iqbal, J., "A review on fault detection and diagnosis techniques: basics and beyond", *Artificial Intelligence Review*, 54 (5): 3639–3664 (2021).
73. Omitaomu, O. A. and Niu, H., .
74. Nsaif, Y. M., Hossain Lipu, M. S., Ayob, A., Yusof, Y., and Hussain, A., "Fault Detection and Protection Schemes for Distributed Generation Integrated to Distribution Network: Challenges and Suggestions", *IEEE Access*, 9: 142693–142717 (2021).
75. Li, W., Huang, R., Li, J., Liao, Y., Chen, Z., He, G., Yan, R., and Gryllias, K., "A perspective survey on deep transfer learning for fault diagnosis in industrial scenarios: Theories, applications and challenges", *Mechanical Systems And Signal Processing*, 167: (2022).
76. Gu, Y. and Bollen, M. H. J., "Time-Frequency and Time-Scale Domain Analysis of Voltage Disturbances", (2000).
77. Mallat, S., "A Wavelet Tour of Signal Processing", (2008).
78. Mateo, C. and Talavera, J. A., "Short-time Fourier transform with the window size fixed in the frequency domain", *Digital Signal Processing: A Review Journal*, 77: 13–21 (2018).

79. L, S. R., Shih, H. H., i-C, N. Y., L, H. H., and Huang and, N. E., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", (1998).
80. German-Sallo, Z. and Grif, H. S., "Hilbert-Huang Transform in Fault Detection", (2019).
81. Mohd Amiruddin, A. A. A., Zabiri, H., Taqvi, S. A. A., and Tufa, L. D., .
82. Widodo, A. and Yang, B. S., .
83. Breiman, L., "Random Forests", (2001).
84. Chen, T. and Guestrin, C., "XGBoost: A scalable tree boosting system", (2016).
85. Nikanjam, A., Morovati, M. M., Khomh, F., and Ben Braiek, H., "Faults in deep reinforcement learning programs: a taxonomy and a detection approach", *Automated Software Engineering*, 29 (1): (2022).
86. Hadad, K., Pourahmadi, M., and Majidi-Maraghi, H., "Fault diagnosis and classification based on wavelet transform and neural network", *Progress In Nuclear Energy*, 53 (1): 41–47 (2011).
87. Ramesh Babu, N. and Jagan Mohan, B., "Fault classification in power systems using EMD and SVM", *Ain Shams Engineering Journal*, 8 (2): 103–111 (2017).
88. Feng, D., "Embedded Parallel Computing Platform for Real-Time Recognition of Power Quality Disturbance Based on Deep Network", *Tehnicki Vjesnik*, 30 (6): 1920–1928 (2023).
89. Bollen, M. H. J., "Voltage sags: effects, mitigation and prediction", .
90. Wang, Y., Li, S., and Xiao, X. Y., "Estimation Method of Voltage Sag Frequency Considering Transformer Energization", *IEEE Transactions On Power Delivery*, 36 (6): 3404–3413 (2021).
91. Chaitanya, B. K., Yadav, A., and Pazoki, M., "An Intelligent Detection of High-Impedance Faults for Distribution Lines Integrated with Distributed Generators", *IEEE Systems Journal*, 14 (1): 870–879 (2020).
92. Ghaderi, A., Ginn, H. L., and Mohammadpour, H. A., .
93. Baqui, I., Zamora, I., Mazón, J., and Buigues, G., "High impedance fault detection methodology using wavelet transform and artificial neural networks", *Electric Power Systems Research*, 81 (7): 1325–1333 (2011).
94. Tonelli-Neto, M. S., Decanini, J. G. M. S., Lotufo, A. D. P., and Minussi, C. R., "Fuzzy based methodologies comparison for high-impedance fault diagnosis in

- radial distribution feeders", *IET Generation, Transmission And Distribution*, 11 (6): 1557–1565 (2017).
95. Moravej, Z., Mortazavi, S. H., and Shahrtash, S. M., "DT-CWT based event feature extraction for high impedance faults detection in distribution system", *International Transactions On Electrical Energy Systems*, 25 (12): 3288–3303 (2015).
 96. Wang, S. and Dehghanian, P., "On the Use of Artificial Intelligence for High Impedance Fault Detection and Electrical Safety", *IEEE Transactions On Industry Applications*, 56 (6): 7208–7216 (2020).
 97. Veerasamy, V., Abdul Wahab, N. I., Ramachandran, R., Thirumeni, M., Subramanian, C., Othman, M. L., and Hizam, H., "High-impedance fault detection in medium-voltage distribution network using computational intelligence-based classifiers", *Neural Computing And Applications*, 31 (12): 9127–9143 (2019).
 98. Al Sultan, M., Avci, İ., and Talab, O., "ENHANCED FAULT DETECTION AND CLASSIFICATION IN TRANSMISSION LINES USING FINE-TUNED LSTM MODEL AND DBN TRANSFORM-BASED FEATURE SELECTION", .
 99. Sheng, Y. and Rovnyak, S. M., "Decision tree-based methodology for high impedance fault detection", *IEEE Transactions On Power Delivery*, 19 (2): 533–536 (2004).
 100. L'Heureux, A., Grolinger, K., Elyamany, H. F., and Capretz, M. A. M., "Machine Learning with Big Data: Challenges and Approaches", *IEEE Access*, 5: 7776–7797 (2017).
 101. Nagata, E. A., Ferreira, D. D., Bollen, M. H. J., Barbosa, B. H. G., Ribeiro, E. G., Duque, C. A., and Ribeiro, P. F., "Real-time voltage sag detection and classification for power quality diagnostics", *Measurement: Journal Of The International Measurement Confederation*, 164: (2020).
 102. Qader, M. R., Bollen, M. H. J., and Allan, R. N., "Stochastic Prediction of Voltage Sags in a Large Transmission System", .
 103. Naidoo, R. and Pillay, P., "A new method of voltage sag and swell detection", *IEEE Transactions On Power Delivery*, 22 (2): 1056–1063 (2007).
 104. Mohammadi, Y., Miraftabzadeh, S. M., Bollen, M. H. J., and Longo, M., "Voltage-sag source detection: Developing supervised methods and proposing a new unsupervised learning", *Sustainable Energy, Grids And Networks*, 32: (2022).
 105. Gencer, Ö., Öztürk, S., and Erfidan, T., "A new approach to voltage sag detection based on wavelet transform", *International Journal Of Electrical Power And Energy Systems*, 32 (2): 133–140 (2010).

106. Karlsson, R. C. L. and D., "Voltage Sag Source Location Based on Voltage Measurements Only", *Elect. Power Quality Utilisation J.*, 14 (Jun.): 25–30 (2008).
107. C. Becker, W. Braun, K. Carrick, T. Diliberti, C. Grigg, J. Groesch, B. Hazen, T. Imel, D. Koval, D. Mueller, T. S. John, and L. E. C., "Proposed Chapter 9 for Predicting Voltage Sags (Dips) in Revision to IEEE Std 493, the Gold Book", *IEEE Trans. Ind. Appl.*, 30 (3): 805–821 .
108. Rai, K., Hojatpanah, F., Badrkhani Ajaei, F., and Grolinger, K., "Deep learning for high-impedance fault detection: Convolutional autoencoders", *Energies*, 14 (12): (2021).
109. Wang, S., Chen, H., Wu, L., and Wang, J., "A novel smart meter data compression method via stacked convolutional sparse auto-encoder", *International Journal Of Electrical Power And Energy Systems*, 118: (2020).
110. Perera, L. P. and Mo, B., "Ship performance and navigation data compression and communication under autoencoder system architecture", *Journal Of Ocean Engineering And Science*, 3 (2): 133–143 (2018).
111. Gogna, A. and Majumdar, A., "Discriminative Autoencoder for Feature Extraction: Application to Character Recognition", *Neural Processing Letters*, 49 (3): 1723–1735 (2019).
112. Qiu, Y., Zhou, W., Yu, N., and Du, P., "Denoising Sparse Autoencoder-Based Ictal EEG Classification", *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, 26 (9): 1717–1726 (2018).
113. Shadi, M. R., Ameli, M. T., and Azad, S., "A real-time hierarchical framework for fault detection, classification, and location in power systems using PMUs data and deep learning", *International Journal Of Electrical Power And Energy Systems*, 134: (2022).
114. Trizoglou, P., Liu, X., and Lin, Z., "Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines", *Renewable Energy*, 179: 945–962 (2021).
115. "IEEE 14 Bus System | PSCAD", <https://www.pscad.com/knowledge-base/article/26> (2024).
116. Alshareef, S. M., "Voltage Sag Assessment, Detection, and Classification in Distribution Systems Embedded with Fast Charging Stations", *IEEE Access*, 11: 89864–89880 (2023).
117. Tschannen, M., Bachem, O., and Lucic, M., "Recent Advances in Autoencoder-Based Representation Learning", (2018).

118. Perera, L. P. and Mo, B., "Ship performance and navigation data compression and communication under autoencoder system architecture", *Journal Of Ocean Engineering And Science*, 3 (2): 133–143 (2018).
119. Gogna, A. and Majumdar, A., "Discriminative Autoencoder for Feature Extraction: Application to Character Recognition", *Neural Processing Letters*, 49 (3): 1723–1735 (2019).
120. Qiu, Y., Zhou, W., Yu, N., and Du, P., "Denoising Sparse Autoencoder-Based Ictal EEG Classification", *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, 26 (9): 1717–1726 (2018).
121. Qian, J., Song, Z., Yao, Y., Zhu, Z., and Zhang, X., .
122. Qian, J., Song, Z., Yao, Y., Zhu, Z., and Zhang, X., .
123. Al-Selwi, S. M., Hassan, M. F., Abdulkadir, S. J., Muneer, A., Sumiea, E. H., Alqushaibi, A., and Ragab, M. G., .
124. Sherstinsky, A., "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network", *Physica D: Nonlinear Phenomena*, 404: (2020).
125. Trizoglou, P., Liu, X., and Lin, Z., "Fault detection by an ensemble framework of Extreme Gradient Boosting (XGBoost) in the operation of offshore wind turbines", *Renewable Energy*, 179: 945–962 (2021).
126. Rafique, F., Fu, L., and Mai, R., "End to end machine learning for fault detection and classification in power transmission lines", *Electric Power Systems Research*, 199: (2021).
127. Rafique, F., Fu, L., and Mai, R., "LSTM autoencoders based unsupervised machine learning for transmission line protection", *Electric Power Systems Research*, 221: (2023).
128. Uddin, M. S., Hossain, M. Z., Fahim, S. R., Sarker, S. K., Bhuiyan, E. A., Mueen, S. M., and Das, S. K., "On the protection of power system: Transmission line fault analysis based on an optimal machine learning approach", *Energy Reports*, 8: 10168–10182 (2022).
129. Jamil, M., Kalam, A., Ansari, A. Q., and Rizwan, M., "Generalized neural network and wavelet transform based approach for fault location estimation of a transmission line", *Applied Soft Computing Journal*, 19: 322–332 (2014).
130. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I., "Attention is all you need", *Advances In Neural Information Processing Systems*, 2017-Decem (Nips): 5999–6009 (2017).

131. Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., and Zhang, Y., "Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network", *IEEE Transactions On Smart Grid*, 10 (1): 841–851 (2019).
132. Chen, T. and Guestrin, C., "XGBoost: A scalable tree boosting system", *Proceedings Of The ACM SIGKDD International Conference On Knowledge Discovery And Data Mining*, 13-17-Aug: 785–794 (2016).
133. Kipf, T. N. and Welling, M., "Semi-supervised classification with graph convolutional networks", *5th International Conference On Learning Representations, ICLR 2017 - Conference Track Proceedings*, 1–14 (2017).



RESUME

Muhamed G. Mehmud AL-SULTAN is a computer engineer with a strong academic background and extensive industry experience. He earned his B.Sc. in Computer Engineering from Mosul Technical College, Iraq, in 2004, followed by an M.Sc. in Computer Engineering from Çankaya University, Turkey, in 2015. In January 2025, he obtained his Ph.D. in Computer Engineering from Karabük University, Turkey. Since 2006, he has been a key contributor to the Ministry of Electricity in Iraq, where he has played a pivotal role in advancing power systems and smart grid technologies. His research and technical expertise span a wide range of domains, including. Wireless Communication & Wireless Sensor Networks (WSNs), Internet of Things (IoT) & Smart Cities, Artificial Intelligence (AI), Machine Learning, & Deep Learning, Artificial Neural Networks (ANNs), Fault Detection & Classification in Power Transmission Systems & Microgrids. As an active IEEE member, he remains dedicated to cutting-edge research, technological innovation, and the development of intelligent systems for modern energy infrastructures.