Computing&AI Connect
ISSN: 3104-4719
2025, Article ID. 2024.0007, Cite as: https://doi.org/10.69709/CAIC.2024.103229

SCIFINITI

Systematic Review

OPEN ACCESS

A Systematic Review on the Integrating Artificial Intelligence for Enhanced Fault Detection in Power Transmission Systems: A Smart Grid Approach

Nijat Taghizad[1,*] **Pietro Lio**[2,3,4]

[1] Acami Nakchivani Street, Binaqadi District, Baku 1000, Azerbaijan
[2] Complex Systems and Nonlinear Dynamics-University of Firenze, 50122 Florence, Italy
[3] University of Pavia, 27100 Pavia, Lombardy, Italy
[4] Cambridge University, Cambridge CB2 3AP, England, UK

Article History
Submitted: August 17, 2024 Accepted: November 21, 2024 Published: February 07, 2025

Abstract
Modern electrical systems rely on sensors and relays for fault detection in three-phase transmission lines and distribution transformers, but these devices often face time complexity issues and false alarms. In this study, the fault detection accuracy is compared in models studied in 2023 and 2024 following PRISMA guidelines. The objectives were to identify fault types, utilize machine learning models to assess their predictive efficacy, and establish accuracy levels. To explore this further, a systematic literature review was performed based on AI accuracy in fault detection from scholarly articles published in reputable journals. The inclusion criteria required journals published between 2023 and 2024 that tested AI systems for three-phase transmission lines and distribution transformers, while sources older than 2023 were excluded. The selected journals used both simulated and real databases to assess AI-based fault detection accuracy. A total of 12 sources were searched, with two selected for comparative analysis based on their relevance to the study's objectives. The risk of bias was assessed using the Robvis method. Findings were presented using narratives, graphs, and tables, and the results were synthesized through comparative data analysis. The Novel Glass Box-Based Model was ranked as the most accurate fault detection model (99% accuracy), followed by Convolutional Neural Network (98%), Gated Recurrent Unit (92%), Random Forest (90%), Logistic Regression (74%), and Support Vector Classifier (63%). Both selected studies compared older fault detection systems with AI-based models, demonstrating the superior accuracy of modern AI approaches. However, the study was limited by reliance on only two sources, which introduced potential bias in the exclusion criteria. The findings suggest that AI developers should aim for 99% accuracy in fault detection systems to meet industry requirements.

Keywords:
integrating artificial intelligence; fault detection; power transmission systems; smart grid approach; machine learning; deep learning; models; accuracy; receiver operating characteristic curve

1. Introduction

1.1. Rationale

Power grids are a crucial part of people's daily lives. Power grids are required to supply electricity to important places daily, including homes, offices, and industries. The disadvantages include frequent faults caused by insulation failure, natural disasters, and harmonic disturbances. [1]. Additionally, errors occur due to failures within power transformers. Because of these issues, electric supply companies have been collecting data from transformers and power lines in real life.

* **Corresponding Author:**
Nijat Taghizad, Baku, Acami Nakchivani Street, Binaqadi District, Baku – Azerbaijan, tagizadenijat@gmail.com

The power transition lines' and systems' data collected are simulated to create information that can be used for machine learning [2]. Modern electrical systems utilize relays and sensors to detect faults. Sensors and relays have helped improve fault detection accuracy by collecting more data from transmission lines [3]. Simulated data include Voltage value in Phase A (VL1), Voltage value in Phase B (VL2), Voltage Value in Phase C (VL3), Current Value in Phase A (IL1), Current Value in Phase B (IL2), Current Value in Phase C (IL3), Oil temperature Alarm (OTA), Oil Temperature Indicator (OTI), Ambient Temperature Indicator (ATI), Winding Temperature Indicator (WTI), and finally, Magnetic Oil Gauge Indicator (MOGI) [4].

In testing faults in power transmission lines and systems, the data collected by the sensors and relays are simulated. Using conditioners such as the min-max scaling reduces the dimensionality and complexity of the selected datasets by giving the datasets a smaller range of values suitable for training and testing [5]. Another conditioner is the Principal Component Analysis (PCA), which ensures that every variable contributes equally to the analysis by normalizing the data. The PCA is a dimensionality-reducing and machine-learning tool that simplifies an extensive data set into a smaller set without significantly changing the patterns and trends. PAC data normalization involves giving each feature a zero mean and a unit variance. It also computes a covariance matrix to assess the correlation between each and all other variables [6].

Over the years, fault detection has evolved from traditional to modern machine learning approaches. Due to the demand for efficiency, new models have been developed to conduct accurate fault detections in electric power transmission lines and systems. Due to the outlined demand for the best fault detectors, this study was undertaken to review comparisons of different machine learning models and conclude from among them the best fault detectors between 2023 and 2024.

1.2. Objectives

1.2.1. General Objective

To compare the levels of accuracy of fault detection presented by different models studied between 2023 and 2024

1.2.2. Specific Objectives

1. To determine types of faults in power transmission systems;
2. To examine fault detection machine-learning models in power transmission systems and
3. To assess levels of accuracy of fault detection techniques in power transmission systems.

2. Literature Review

2.1. Types of Faults in Power Transmission Systems

Faults are driven from two datasets: the distribution transformer fault prediction model and the general transmission line error classification. Datasets are found from a classic transmission line simulation under various situations. The main faults in power transmission lines include short-circuit and symmetrical faults. Short-circuit faults include Line to Ground (L-G) and Line-to-Line (L-L). On the other hand, symmetrical faults include the L-L-L fault (Between A, B, and C phases), the L-L-G fault (Between two phases and the ground), and the L-L-L-G fault (Three phase symmetrical fault) [7].

Chakraborty, De, and Nama studied the effect of the length of the line on fault current using the L-G transmission line simulated fault data on MATLAB Simulink. The study explains that the L-G, the origin of the fault in the transmission line, recorded lower fault levels than points close to the source. L-L short circuit fault diagnosis in various micro-grid systems. According to the study, the L-L fault happens every time two transmitting lines are short-circuited [8]. The L-L fault's leading cause is the wind-swinging transmission lines that touch together, causing a short circuit. L-L faults occur between 15% and 20% [9].

The L-L-L fault in an electrical system happens when all three phases are short-circuited. The primary cause of this fault is equipment failure. An example of an equipment failure is an insulation breakdown that causes direct contact (short circuit) among all three phases. The L-L-L fault's occurrence is rare (below 15%). The L-L-G fault in an electrical system happens when two phases are short-circuited with a ground connection. The major causes of this type of fault are insulation breakdown and equipment failure. The fault's occurrence is relatively uncommon. Finally, the L-L-L-G fault in an electrical system happens when all three phases are short-circuited with a ground connection. The leading causes of the fault are severe equipment failure and insulation breakdown [10].

2.2. Fault Detection Machine-Learning Models in Power Transmission Systems

2.2.1. Traditional Machine Learning Models

Logistic regression is a standard statistical modeling method for binary classification tasks where the objective is to classify data into one of two groups. The model applies a

2025, Vol. 2, Article ID. 2025.0007
https://doi.org/10.69709/CAIC.2024.103229

logistic (sigmoid) function to a linear equation of the input features. This function converts the linear equation's result into a probability score. The logistic function's S-shaped curve makes it a good fit for models in which probability is the output. Through training, the coefficients of the linear equation are modified. Determining which model best fits the data is to find coefficients that minimize the discrepancy between the model's predictions and the observed results. Logistic regression is a popular statistical technique because of its ease of use and interpretation of results [10].

Another traditional model is the Decision Tree Classifier. The Decision Tree Classifier is a machine-learning model that divides data input into subset values. Every node symbolizes a characteristic, and each branch denotes a decision rule that makes additional results. These are the final choices and classes. The Decision Tree Classifier is a simplistic and straightforward model for interpreting and visualizing faults [11].

The AdaBoost Classifier is another traditional machine-learning model that uses boost strategies to classify fault detection accuracy. It combines several weak classifiers into one robust classifier. Every classifier in the series corrects the weights of the cases by concentrating on those that the preceding ones incorrectly classified. This procedure enhances the performance of the model in challenging situations. AdaBoost is effective because it adapts to ensemble errors, enhancing accuracy across various scenarios. Adaptive Boosting, as an ensemble technique, leverages machine learning to achieve improved performance [12].

2.2.2. Modern Machine Learning Models

Machine learning involves identifying a trained model using input data to generate a preferred output. The model is subjected to learning how to predict data by providing an output. A deep learning model is a simple three-layer neural network. Each layer of the model uses activation functions to capture complex relationships and then produce nonlinear relationships. Fault detection is measured based on accuracies, F1 scores, and other performance metrics when selecting the suitability of a machine-learning model.

The Perceptron-based neural networks is the first model suitable for detecting and classifying faults within electrical systems and distribution transformers. The perceptron-based neural networks use a supervised deep learning approach. Unlike other models, the Perceptron-based neural networks utilize a single-layer network for input data classification into two categories. The model adjusts weights produced by errors and, therefore, learns the input patterns to increase the accuracy of its output. The iterative process continues until the model detects a suitable fault [13].

Another model is the Hidden Markov Model (HMM). It uses an algorithm to handle electrical data formatted as a multivariate time series. HMM makes probabilistic observations of the signal received to predict anomalies. HMM is trained on the data it receives. The model applies two key variables, namely states and observations. Each unique state calculated by the HMM represents a disparate fault or normal operating conditions within the transmission line system. Therefore, the electrical data collected is measured over time with considerations based on the influence of the type of fault. HMM is trained using the Baum-Welch Algorithm. The Baum–Welch algorithm is a famous electrical engineering statistical computing and bioinformatics algorithm used to determine unknown HMM parameters. The Baum–Welch algorithm approximates the values of the expressions above according to the fault type and the electrical signal. After calculating the value, it adjusts them for an accurate measure. The decoding process identifies the most probable hidden sequence and then Backpropagation to find the highest value in the data generated by the electrical systems. The Stochastic gradient descent optimizer is an example of a deep learning model that conducts Backpropagation. The Backpropagation manipulates values to generate output that represents the actual value. Stochastic find partial derivatives of the weights and biases, creating a gradient vector to increase accuracy and F1 scores. Further, Backpropagation trains the neural network's ways of computing updated parameters [14].

Another model is the Capsule network. The Capsule network filters data through deep learning to detect and classify transmission line faults. It can preserve data in spatial hierarchical relationships. The model transforms three-phase current and voltage signals into a wavelet energy matrix for input, making the network highly effective and flexible for fault scenarios and line conditions [15].

2.3. Levels of Accuracy of Fault Detection in Power Transmission Systems

Output accuracy (fault detection) can be measured using the Confusion matrices. A Confusion matrix is a table that visually lays out algorithm results. It help to find areas where models easily confuse one class with the other. The Confusion matrix is structured like a square, and it most commonly shows the frequencies of the predicted variables relative to the actual labels, allowing for a clear representation of the model's performance. It consists of four different entries, namely the True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN). The True Positives (TP) are cases where the model correctly predicts a positive outcome. The False Positives (FP) are cases where the model incorrectly predicts a positive outcome. The True Negatives (TN) are cases where the model correctly predicts the negative outcome. Finally, False Negatives (FN) are cases where the model incorrectly predicts a negative outcome [4].

In addition to being a square-shaped performance-displaying tool, confusion matrices also calculate evaluation metrics, which include accuracy, precision, recall, f1-score, and specificity. Accuracy measures the frequency with which the model produces correct predictions. Precision refers to when the model predicts a positive outcome and when it is accurate. Recall refers to how frequently the model correctly predicts the positive class. F1-Score refers to the average precision and recall, which gives us a good idea of the model's overall performance. Finally, specificity refers to how often the model correctly classifies negative classes [4].

The level of accuracy is also measured using the Receiver Operating Characteristic Curves (ROC). ROC are graphical plots and a contemporary statistical technique used to illustrate the predictive capabilities of a machine-learning model when the decision threshold is changed. They involve combining several confusion matrices under different decision thresholds and evaluating the result according to the area under the graph. This gives an unbiased point of view of the performance of this model, even if the dataset is imbalanced [16].

2.4. Gaps in Current Literature & Studies

The study aimed to analyze the integration of artificial intelligence in enhancing fault detection in power transmission systems. Based on information presented in the literature, there were several gaps that previous studies left. First, the sources used in the literature section discussed traditional machine learning models that detect faults in transmission lines and systems, including Logistic Regression, Decision Tree Classifier, and AdaBoost Classifier. This study, therefore, aimed at studying the Random Forest (RF), the Logistic Regression (LR), the Support Vector Classifier (SVC), and the Gated Recurrent Unit (GRU) as traditional models used to detect faults in transmission lines and system. Additionally, the literature exhibited bias in focusing on perceptron-based neural networks, the Hidden Markov Model (HMM), and the Capsule Network as modern machine learning approaches for fault detection in transmission lines and systems. To overcome this gap, this study introduced the systematic literature review of the Convolutional Neural Network (CNN) and the Novel glassbox-based proposed EB approach in fault detection in transmission lines and systems. Filling these gaps enabled this study to compare traditional and modern fault detection models in transmission lines and systems.

3. Materials and Methods

The PRISMA guidelines guided this section of the study (Chart 1) below.

3.1. Eligibility Criteria

The inclusion criteria were that the study used journals published between 2023 and 2024. The study also included journals that tested Artificial Intelligence systems' accuracy in fault detection in three-phase transmission lines and distribution transformers. The study excluded sources that were older than 2023.

3.2. Information Sources

The study used an extraction form to identify the sources that are suitable for the study. The extraction form was as presented below as [17] recommended. The extraction form presented as Table 1 is as an illustration of the data collection form. After the data collection, the empty extraction form (Table 1) was filled latter on and presented as shown in Table 2.

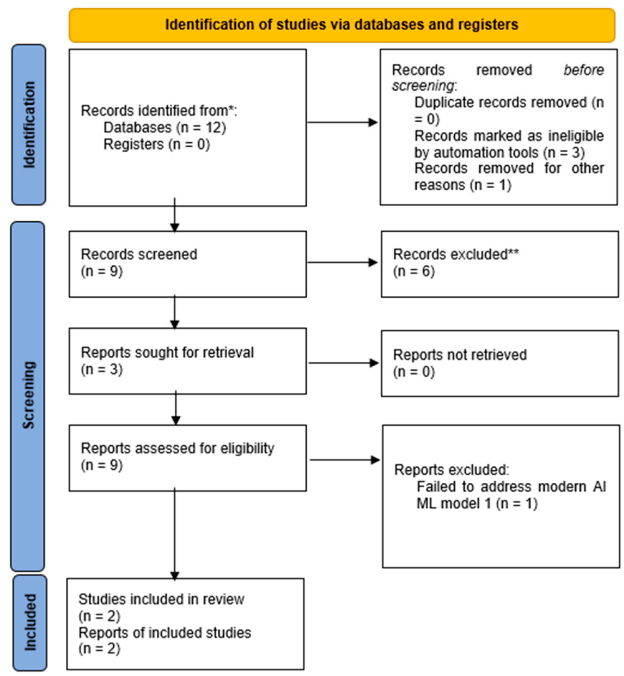
The sources of information were journals written by professionals and scholars in the field of Artificial intelligence and engineering. The sourced journals came from Google search engine that helped the researcher establish a total of 12 credible journals for the systematic review.

3.3. Search Strategy

The study used a qualitative research design. Systematic literature review involved identifying scholarly articles from respectable international journal websites. In this case, this study conducted a systematic literature review to determine types of faults in power transmission systems, examine fault detection machine-learning models in power transmission systems, and assess levels of accuracy of fault detection techniques in power transmission systems.

The study applied a direct search strategy since the articles reviewed were readily provided in the internet (i.e., Google and Google scholar). There filtering of the sources was limited to the extraction form (Table 1).

2025, Vol. 2, Article ID. 2025.0007
https://doi.org/10.69709/CAIC.2024.103229



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).
**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.
Source: Page MJ, et al. BMJ 2021;372:n71. Doi: 10.1136/bmj.n71.

Chart 1: PRISMA 2020 flow diagram for new systematic reviews which included searches of databases and registers only.

3.4. Selection Process

The selection process involved assessing the sources based on the extraction form. The researcher identified the sources according to the presence of the protocols section presented in the extraction form. All the two sources met the inclusion criteria of the review. Four reviewers reviewed the sources, each screening and retrieving the records. The supervisor provided the sources: a professor at Cambridge University. Each reviewer worked independently in the absence of the author. No automation tools were used in this process.

3.5. Data Collection Process

The research used an extraction form to collect qualitative and quantitative data from the two journals. The study also used comparative analysis to present the findings from the two sources. Further, two reviewers were responsible for independently collecting data from each report. The other three reviewers confirmed the authenticity and accuracy of the data from the study's investigators. The process was manual, and thus, no automation tools were used.

3.6. Data Items

List and define all outcomes for which data were sought. Specify whether all results compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect.

List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.

3.7. Study Risk of Bias Assessment

The study had a limitation in that it reported results based on only two sources, which introduced potential bias in the exclusion criteria. To assess the risk of bias in the included studies, the researcher used the Robvis methods that visualized risk-of-bias during the inclusion criteria for sources used in the systematic review. Robvis refers to an open-source Shiny web app and R package that is used to assess the risk-of-bias figures, thus providing an evidence synthesis workflow inside R. To back up the process, the researcher further used a manual approach whereby it em-

Table 1: Extraction form.

Study ID	Protocol	Was the Author Contacted for Additional Outcome Data	Contact Record	Date Last Searched/ Consulted.
	Published 2023–2024 Journal (type) Used simulated data Addressed types of faults Addressed fault detection ML models Addressed levels of accuracy of fault detection techniques	Yes, positive response Yes, but no response No It was N/A		

Source: Author (2024).

ployed a peer review approach to check and classify the sources. The study employed two reviewers to certify and select the two sources independently.

3.8. Effect Measures

To determine the effect measures, the study used the Odds Ratio (OR) to synthesize or present results. Odds Ratio (OR) is a type of risk estimate whereby the Odds of Event 1 are divided by the Odds of Event 2. In this case, ORs close to 1 mean the estimated effects are likely the same for groups 1 and 2. In reverse, ORs far from 1 mean a lower likelihood that the two events will be identical. In the case of this study, the OR equaled 1.00. This implied that the sources adequately complemented each other and were reliable for systematic review.

3.9. Synthesis Methods

The study used the extraction form provided in Table 1. According to Table 1, the two sources reviewed were eligible. Data was collected from secondary sources, namely scholarly articles dated between 2023 and 2024. The collected data was analyzed using qualitative analysis techniques. Qualitative analysis dealt with the thematic coding of data to determine the common themes and trends provided by the sources selected to inform the study. Comparative analysis was conducted to elaborate on the differences in the findings and come to a proper conclusion about the issue under study. Case studies demonstrated the practical implementation of the phenomena the study proposed to analyze.

Performance metrics that determine the accuracy of fault detection and classification models apply the receiver operating characteristic curve (ROC). The ROC curve was also chosen to demonstrate the binary classifier's capability.

In addition, machine learning, being a binary classification, led to this study using true positive (TP), false positive (FP), true negative (TN), and false negative (FN) to determine the accuracy of models in fault detection.

Accuracy is defined as a ratio that sums up correct predictions from a total of several dataset samples. It is calculated as:

$$\mathrm{Acc} = [(\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})] \times 100\% \quad (1)$$

Precision is the ratio of correct predictions of all positive classes from all optimistic predictions. It is calculated as:

$$\mathrm{P} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FP}) \quad (2)$$

Recall is a sensitivity ratio summarizing all correct predictions from a positive class and all incorrect predictions from a negative class.

$$\mathrm{R} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN}) \quad (3)$$

The F1 score is defined as a weighted average of recall and precision. The F1 score is valuable if the dataset has an uneven class distribution. It is calculated as:

$$\mathrm{F1-score} = 2 \times (\mathrm{Precision} \times \mathrm{Recall})/(\mathrm{Precision} + \mathrm{Recall}) \quad (4)$$

The study, therefore, provided results from other existing studies that presented a confusion matrix to summarize the accuracy of their machine-learning classification models.

3.10. Reporting Bias Assessment

To assess the risk of bias in the included studies, the researcher used the Robvis methods that visualized risk-of-bias during the inclusion criteria for sources used in the systematic review.

2025, Vol. 2, Article ID. 2025.0007
https://doi.org/10.69709/CAIC.2024.103229

Table 2: Study selection.

No.	Study ID	Protocol	Was the Author Contacted for Additional Outcome Data	Included/ Excluded	Date Last Searched/ Consulted.
1	Akhtar, Atiq, Shahid, Raza, Samee, Alabdul-hafith, [10]	[✓] Published 2023–2024 [✓] Journal (type) [✓] Used simulated data [✓] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [✓] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Included	25 November 2024
2	Turanl, Ben-teşen, [4]	[✓] Published 2023–2024 [✓] Journal (type) [✓] Used simulated data [✓] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [✓] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Included	25 November 2024
3	Maria, Michael [18]	[✓] Published 2023–2024 [✓] Journal (type) [×] Used simulated data [o] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [×] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024
4	Omitaomu, Niu. [19]	[×] Published 2023–2024 [✓] Journal (type) [×] Used simulated data [o] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [0] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024
5	Arévalo, Jurado. [20]	[✓] Published 2023–2024 [✓] Journal (type) [✓] Used simulated data [o] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [o] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024

Table 2: *Cont.*

No.	Study ID	Protocol	Was the Author Contacted for Additional Outcome Data	Included/ Excluded	Date Last Searched/ Consulted.
6	Wen, Shen, Zheng, Zhang [21]	[✓] Published 2023–2024 [✓] Journal (type) [×] Used simulated data [o] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [o] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024
7	Wang, Wang, Bhandari. [22]	[✓] Published 2023–2024 [✓] Journal (type) [✓] Used simulated data [o] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [o] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024
8	Stock, Babazadeh, Becker. [23]	[×] Published 2023–2024 [✓] Journal (type) [✓] Used simulated data [✓] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [o] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024
9	Zerguit, Youness, Derrhi. [24]	[✓] Published 2023–2024 [✓] Journal (type) [×] Used simulated data [✓] Addressed types of faults (Outcome A) [✓] Addressed fault detection ML models (Outcome B) [✓] Addressed levels of accuracy of fault detection techniques (Outcome C)	[ ] Yes, and a positive response [ ] Yes, but no response [✓] No [✓] It was N/A	Excluded	25 November 2024

Source: Author (2024).

3.11. Certainty Assessment

To assess the level of confidence of the sources used in the research, the study used the Cronbach alpha calculation, referred to as the reliability test. Before using the findings, the study generated a reliability of 0.76, sufficient to proceed with data analysis.

4. Results

4.1. Study Selection

The study collected data from two recommended studies by the supervisor at Cambridge University and reviewers of the list provided by the researcher, as presented in Table 2.

In Table 2, 9 studies were selected by the researcher to serve in the systematic review. Out of the nine studies, only two qualified to be used in the systematic review. The other studies failed to qualify, given they either presented the expected protocols partially or could not show them entirely.

4.2. Study Characteristics

This systematic review used the studies [10] and [4] which were characterized as having been published between 2023 and 2024; they were academic journals; they used simulated data; they addressed types of faults (Outcome A); they addressed fault detection ML models (Outcome B); and finally, they addressed levels of accuracy of fault detection techniques (Outcome C).

4.3. Risk of Bias in Studies

The risk of bias assessment follows Cochrane's review criteria, focusing on selection bias, detection bias, and reporting bias. Selection bias involved the data source being questionable by more than 50% of the reviewers (its adequacy is below 50%). Detection bias involved the source of data being questionable by 50% of reviewers and at the same time being approved by the other 50% of reviewers (its adequacy is at 50%). Finally, the reporting bias involved the data source approved by more than 50% of reviewers (its adequacy is almost 100%). Each paper was evaluated as "Low Risk" (+), "High Risk" (−), or "Unclear Risk" (?). To address the risk of bias in studies, the researcher presents that risk of bias assessment in Table 3.

Based on the information in Table 3, most of the studies were of unclear risk because of their bias in providing the required information. Only two studies were adequate for the study.

4.4. Results of Individual Studies

The effect estimate is measured using standardized mean difference, odd ratio, and correlation coefficient in statistics. In this case, the researcher used the odd ratio and presented the results in Section 3.8.

4.4.1. Summary Statistics for Each Group

Turanl and Benteşen studied the Convolution Neural Network (CNN) model and established that the CNN's Confusion matrix of public dataset (Acc = 0.98); Simulink dataset with (Acc = 1.00); real dataset (Acc = 0.99); and the CNN's ROC from public dataset (AUC = 1.00); Simulink dataset (AUC = 1.00); real dataset (AUC = 1.00) [4].

Akhtar et al studied a Novel glass-box-based proposed EB approach. They established that the EB approach's Confusion matrix of the real dataset (Acc = 0.99); the EB approach's ROC from the actual dataset (AUC = 1.00); the RF's, LR's, SVC's, and GRU's Confusion matrix of the real dataset (Acc = 0.99); and finally, the RF's and LR's ROC from a real dataset (AUC = 0.97 and 0.58) [10].

Effect estimate is precise (e.g., confidence/credible interval), ideally using structured tables or plots.

4.5. Results of Syntheses

4.5.1. Results of All Statistical Syntheses Conducted

4.5.1.1. Convolution Neural Network (CNN) model

Turanl and Benteşen examined the classification of faults in power transmission lines using deep learning of different datasets. Regarding the types of faults in power transmission systems, datasets were generated using a simulated power system [4]. The categorization of the various faults is presented in Tables 4 and 5.

Table 4: Symmetrical faults.

L-L-L Fault	L-L-L-G Fault
ABC	ABCG

Source: [4].

Table 5: Asymmetrical faults.

L-G Fault	L-L Fault	L-L-G Fault
AG	AB	ABG
BG	BC	BCG
CG	AC	ACG

Source: [4].

According to Tables 4 and 5, the data encompassed ten short-circuit fault categories: AB, BC, AC, ABG, BCG, ACG, AG, BG, CG, and ABC. The study acquired the L-G (AG, BG, CG), L-L (AB, BC, AC), L-L-G (ABG, BCG, ACG) (asymmetrical), L-L-L-G (ABCG), and L-L-L (ABC) (symmetrical). The fault detection machine-learning model in power

Table 3: Risk of bias in studies.

Code	References	Protocol Results	Selection Bias	Detection Bias	Reporting Bias	Risk of Bias
1	Akhtar et al. [10]	100%	+	+	+	Low risk
2	Turanl, Benteşen. [4]	100%	+	+	+	Low risk
3	Maria, Michael. [18]	58%	?	-	-	High risk
4	Omitaomu, Niu. [19]	60%	?	?	-	Unclear Risk
5	Arévalo, Jurado. [20]	83%	+	+	?	Unclear Risk
6	Wen et al. [21]	67%	?	?	-	Unclear Risk
7	Wang et al. [22]	83%	+	+	?	Unclear Risk
8	Stock, et al. [23]	75%	+	?	?	Unclear Risk
9	Zerguit, et al. [24]	83%	+	+	?	Unclear Risk

Source: Taghizad (2024).

transmission systems was the deep network model, namely the one-dimensional Convolution Neural Network (CNN), with the input being the MATLAB-Simulink software-generated datasets. The accuracy levels of fault detection techniques in power transmission systems are presented in Figure 1. The ROC's AUC value for all classes is presented in Figure 2.

The accuracy levels of fault detection techniques in power transmission systems were an average of 0.98 (98%). The ROC presented an AUC value of 1.0 (100%).

4.5.1.2. Novel Glass-Box-Based Proposed EB Approach

Akhtar et al. examined novel glass-box machine's fault detection in power transmission systems. The study used the benchmark dataset, which entailed a collection of line currents and voltage fault conditions on types of faults in power transmission systems. The data with faults was modeled using the MATLAB simulator. The study analyzed signal data for faults in the transmission line. In the current and voltage analysis over time, the Va voltage values (ranging between -0.6 and 0.6) indicate no fault, while the Va values near 0.0 indicate fault. The current signals over time, such as Ia current near 0.0, indicated no fault, while Ia values (between -750 and 750) showed the existence of a fault [10]. The study presented the accuracy (Acc) levels of fault detection techniques in power transmission systems in Figure 3. The ROC's AUC value for all classes is presented in Figure 4.

Akhtar, Atiq, Shahid, Raza, Samee, and Alabdulhafith also reviewed the accuracy of the Random Forest (RF), the Logistic Regression (LR), the Support Vector Classifier (SVC), and the Gated Recurrent Unit (GRU) [10]. The study presented the accuracy (Acc) levels of fault detection techniques in power transmission systems in Figure 5. The ROC's AUC value for all classes is presented in Figure 6.

4.5.2. Results of All Investigations of Possible Causes of Heterogeneity Among Study Results

Both studies used in the review utilized the same methodology. For example, they tested simulation data to determine the accuracy of the models they built to detect faults in transmission lines. Both studies' approaches were similar, including reporting the findings based on ROC and Confusion matrix.

4.5.3. Results of All Sensitivity Analyses Conducted to Assess the Robustness of the Synthesized Results

The robustness of the results was reported based on the objectives of this paper. The study aimed to test the hypothesis: There is no significant relationship between fault detection machine learning models and the accuracy levels of fault detection techniques in power transmission systems. The study established a substantial relationship between fault detection machine-learning models and the accuracy of fault detection techniques in power transmission systems. All the models showed acceptable accuracy in fault detection. Therefore, no sensitivity analyses affected the robustness of the synthesized results.

4.6. Certainty of Evidence

By using two of the nine selected journals, the research achieved a reliability of over 0.7. This implied that the

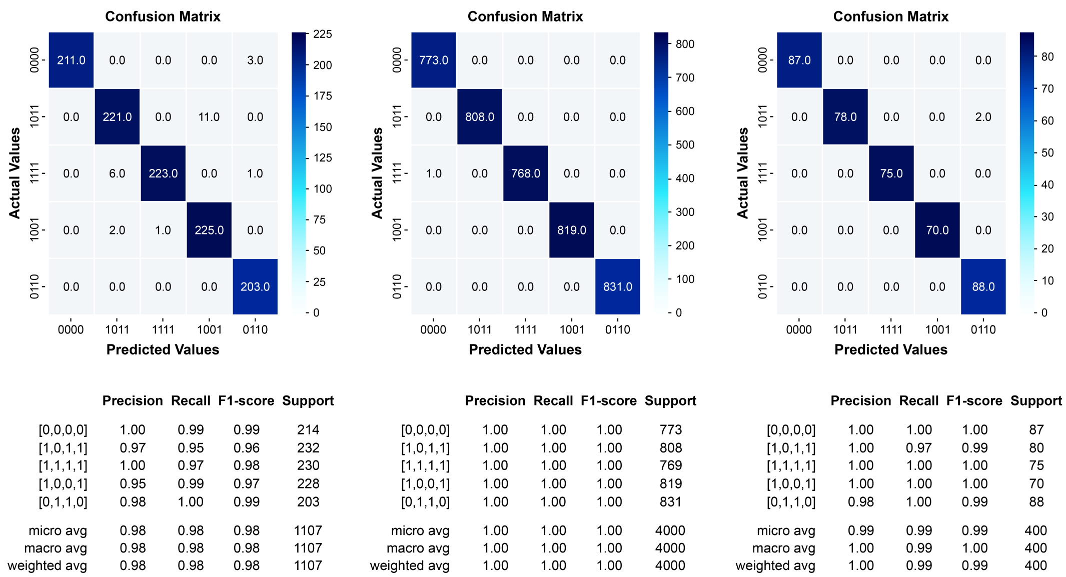


Figure 1: CNN's Confusion matrix of public dataset (Acc = 0.98); Simulink dataset with (Acc = 1.00); real dataset (Acc = 0.99). Source: [4].

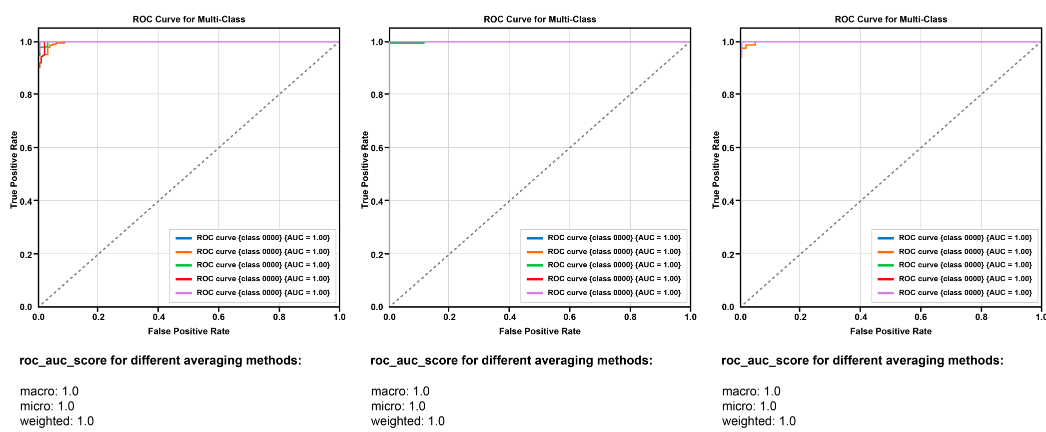


Figure 2: CNN's ROC from public dataset (AUC = 1.00); Simulink dataset (AUC = 1.00); real dataset (AUC = 1.00). Source: [4].

study reviewed sources that adequately informed the study, and thus, the certainty of evidence was regarded as high.

5. Discussion

5.1. Provide a General Interpretation of the Results in the Context of Other Evidence

For simulated data, the results of the RF model were an accuracy of 0.89, a combination of an average precision of 0.91, an average recall of 0.89, and finally, an average FI-score of 0.89. The LR model had an accuracy of 0.63, a combination of an average precision of 0.63, an average recall of 0.63, and an average F1-score of 0.63. The SVC model presented an accuracy of 0.68, which entailed an average precision of 0.80, an average recall of 0.68, and an average F1-score of 0.64. CNN's Confusion matrix from the simulated dataset was found to have an accuracy of 1.00. Using actual data, the results of the RF model were an accuracy of 0.90, a combination of an average precision of 0.92, an average recall of 0.90, and finally, an average FI-score of 0.88. The LR model had an accuracy of 0.74, a combination of an average precision of 1, an average recall of 0.42, and an average F1-score of 0.59. The SVC model presented an accuracy of 0.63, which entailed an average precision of 0.64, an average recall of 0.63, and

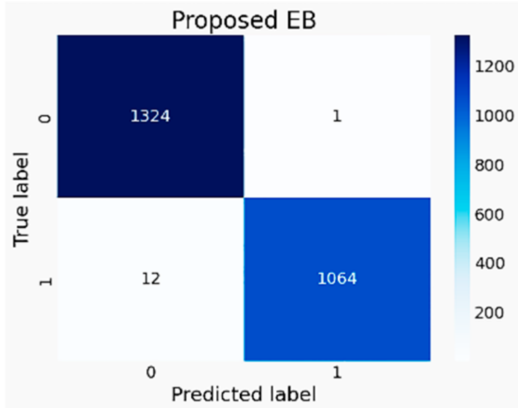


Figure 3: EB approach's Confusion matrix of the real dataset (Acc = 0.99). Source: [10].

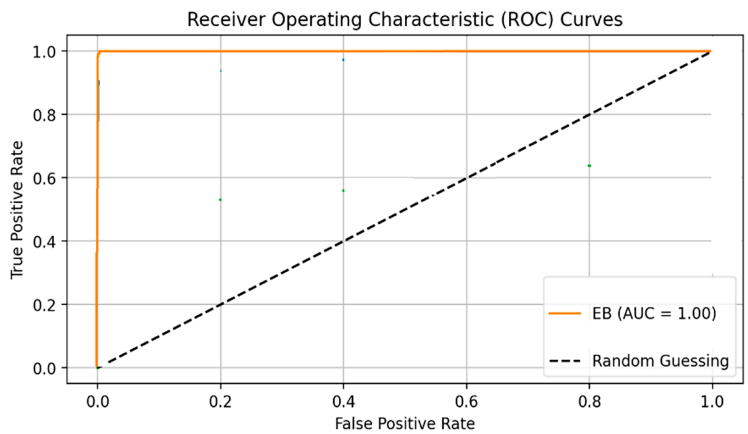


Figure 4: EB approach's ROC from the actual dataset (AUC = 1.00). Source: [10].

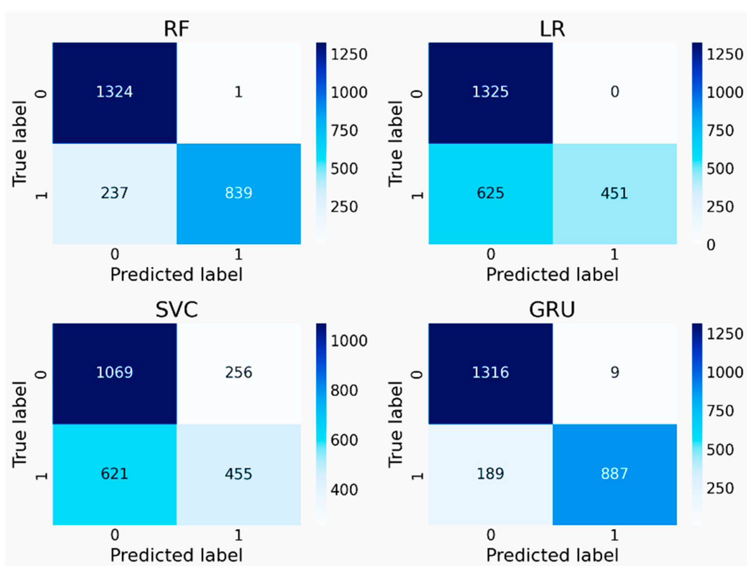


Figure 5: RF's, LR's, SVC's, and GRU's Confusion matrix from real dataset. Source: [10].

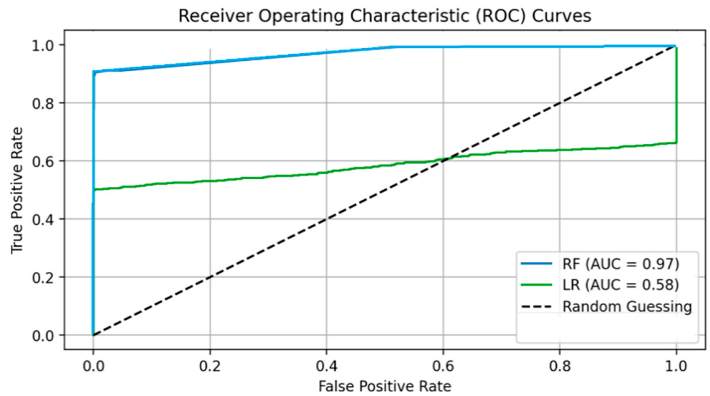


Figure 6: RF's and LR's ROC from real dataset (AUC = 0.97 and 0.58). Source: [10].

an average F1-score of 0.62. The GRU model had an accuracy of 0.92, a combination of an average precision of 0.93, an average recall of 0.92, and an average F1-score of 0.92. Finally, the EB approach had an accuracy of 0.99, a combination of an average precision of 0.99, an average recall of 1.00, and an average F1-score of 1.00. CNN's Confusion matrix from the real dataset was found to have an accuracy of 0.99. For ROC, the study established that RF's AUC = 0.97, LR's AUC = 0.58, EB's AUC = 1.00, and finally, CNN's AUC = 1.00.

The study's findings align with the findings of Zerguit, Youness, and Derrhi, who studied how integrating AI-enhanced fault detection industrial systems supported the findings of this study. The author revealed that advanced neural networks and hybrid models had a higher accuracy detection than traditional models. Traditional models include SVM, DT, KNN, and ANN. Regarding accuracy in detecting faults, the SVM model's configuration presented moderate accuracy. The DT model presented a relatively lower accuracy. The KNN model showed moderate accuracy. Finally, the ANN model displayed high accuracy. Advanced Neural Networks included the CNN model, which showed high accuracy. Unfortunately, the study was not in line with this study's findings, given it considered the GRU as having a high level of accuracy [24]. In conclusion, the study confirmed that advanced and hybrid AI-driven machine learning models demonstrated superior accuracy in defect detection compared to traditional models and approaches, aligning with the findings of this research.

5.2. Discuss Any Limitations of the Evidence Included in the Review

The main limitation of the two sources was the lack of adequate comparison of different models developed to detect fault. Because of this limitation, this systematic review was conducted to investigate the accuracy of fault detection ML models used in electric power transmission lines.

5.3. Discuss any Limitations of the Review Processes Used

The main limitation presented by the review process is the insistence on using the most recent studies between 2023 and 2024. Because of this, the findings were restricted from assessing older studies that could have shaded proper comparative analysis findings on the differences between the new and old or traditional ML models.

5.4. Discuss the Implications of the Results for Practice, Policy, and Future Research

The study recommends policymakers use the study's findings to make laws that will guide the development of ML models with ethical perspectives in their operations. This will ensure transparency measures to establish fault detection accuracy since the lesser the accuracy, the higher the tax cost for the citizens. In addition, the more accurate the fault detection, the more the government saves as opposed to traditional models that lead to loss and wastage of electric power and also as a result of power outages that affect the economy.

Practically, the study recommends engineers apply the suggested ML model in fault detection as this will increase work efficiency and reduce wastage of resources and working hours.

For future researchers, the study recommends applying the study to address the development of new ML models for fault detection. By doing so, future researchers can fill the gaps this study leaves and other studies in the review.

6. Conclusions

In conclusion, the fault detection machine-learning models in power transmission systems included the Random Forest (RF), the Logistic Regression (LR), the Support Vector Classifier (SVC), the Gated Recurrent Unit (GRU), the Convolutional Neural Network (CNN) and the Novel glassbox-based proposed EB approach whose measurements are presented in Appendix A Table A1. Despite the findings being adequate, it was established that there existed a gap caused by the lack of ROC data for the SVC and GRU models. The study also failed to extract data regarding the accuracy of fault detection linked to simulated data for the GRU model. Finally, the study also failed to review the TP, TN, FP, and FN data for the CNN model. Despite these missing elements, the study was adequate to conclude that the EB approach was the most advanced and most suitable model for detecting faults in electric power transmission lines.

List of Abbreviations

Acc	Accuracy
ANN	Artificial Neural Network
ATI	Ambient Temperature Indicator
AUC	Area Under the Curve
CNN	Convolutional Neural Network
DT	Decision Tree
EB	Explainable Boosting
F1-Score	Formula One Score
FN	False Negatives
FP	False Positives
GRU	Gated Recurrent Unit
HMM	Hidden Markov Model
IL1	Current Value in Phase A
IL2	Current Value in Phase B
IL3	Current Value in Phase C
KNN	K-Nearest Neighbors; k Means the Number of Nearest Neighbors Considered in the Analysis
L-G	Line to Ground
L-L	Line-to-Line
L-L-G	Between Two Phases and the Ground
L-L-L	Between A, B, and C Phases
L-L-L-G	Three-Phase Symmetrical Fault
LR	Logistic Regression
MATLAB	Matrix Laboratory
MOGI	Magnetic Oil Gauge Indicator
OTA	Oil temperature Alarm
OTI	Oil Temperature Indicator
PCA	Principal Component Analysis
RF	Random Forest
SVC	Support Vector Classifier
SVM	Support Vector Machine
TN	True Negatives
TP	True Positives
VL1	Voltage Value in Phase A
VL2	Voltage Value in Phase B
VL3	Voltage Value in Phase C
WTI	Winding Temperature Indicator

Author Contributions

Conceptualization and supervision, P.L.; methodology, N.T.; validation, P.L.; formal analysis, N.T.; investigation, N.T.; resources, N.T.; data curation N.T.; writing---original draft preparation, N.T.; writing---review and editing, N.T. and P.L.; project administration, N.T.; funding acquisition, N.T. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials

Further, the study used a systematic literature approach to collect and analyze data. The method used journals published between 2023 and 2024. The sources of data were open access and online articles, which are considered public information and therefore accessible to the public, including the author who used them to address the topic of research.

Consent for Publication

Not applicable.

Conflicts of Interest

The study was developed based on the author's practical experience in the field of artificial intelligence, which may have influenced the selection and interpretation of sources. As a result, perspectives from related fields such as electrical engineering may not have been fully incorporated. During the peer review process, the reviewers recommended focusing on more recent literature published between 2023 and 2024. This suggestion was taken into account, and the manuscript was revised accordingly. The author declares no financial or personal conflicts of interest that could have influenced the outcomes of this research.

Funding

The authors did not receive any external funding for this article. The source of funds was the author’s family contributions and budget. The author received no sponsorship or donations towards the completion of the research.

Acknowledgments

The authors would like to thank the reviewers for their valuable feedback in shaping this paper and recognizing its contribution to the academic community. We also extend our sincere gratitude to our families and friends for their support and for funding the completion of this project.

Standards of Reporting

PRISMA guidelines and methodology were followed.

Appendix A

Table A1: Summary of Results.

Models	AUC	Acc. Simulated Data	Acc. Real Data	True Positive	True Negative	False Positive	False Negative
RF	0.97	0.89	0.90	1324	839	237	1
LR	0.58	0.63	0.74	1325	451	625	0
SVU	-	0.68	0.63	1069	455	621	256
GRU	-	-	0.92	1316	887	189	9
CNN	1.00	0.98	0.99	-	-	-	-
EB	1.00		0.99	1324	1064	12	1

References

[1] A. Algamluoli, X. Wu, H. Abdulhadi, K. Gautier, and M. Khan, “Ultra-high voltage gain non-isolated DC-DC converter based on new interleaved triple boosting technique,” *Front. Energy Res.*, vol. 11, 2023. [CrossRef]

[2] A. G. Migisha, J. M. Ntayi, M. S. Adaramola, F. Buyinza, L. Senyonga, and J. Abaliwano, “Effects of technical and security factors on grid electricity reliability: Evidence from Uganda national electricity grid network,” *Technol. Sustain.*, vol. 3, pp. 41–67, 2024. [CrossRef]

[3] M. Kunicki, J. Fulneček, and P. Rozga, “Sensors and fault diagnostics in power system,” *Sensors*, vol. 24, p. 5999, 2024. [CrossRef]

[4] O. Turanl and Y. Benteşen, “Classification of faults in power system transmission lines using deep learning methods with real, synthetic, and public datasets,” *Appl. Sci.*, vol. 14, p. 9590, 2024. [CrossRef]

[5] A. Demircioğlu, “The effect of feature normalization methods in radio mics,” *Insights Imaging*, vol. 15, pp. 2–11, 2024. [CrossRef]

[6] G. Li and Y. Qin, “An exploration of the application of principal component analysis in big data processing,” *Appl. Math. Nonlinear Sci.*, vol. 9, pp. 1–13, 2024. [CrossRef]

[7] P. L. Brito, “Hybrid approach for detection and diagnosis of short-circuit faults in power transmission lines,” *Energies*, vol. 17, p. 2169, 2024. [CrossRef]

[8] K. Chakraborty, S. De, and P. Nama, “The effect of line length on the fault current due to a line-to-ground fault in the transmission line using MATLAB Simulink,” *World J. Adv. Eng. Technol. Sci.*, vol. 11, pp. 413–417, 2024. [CrossRef]

[9] J. A. c-Malanche, E. Ramirez-Velasco, F. J. Villalobos-Pina, and S. K. Gadi, “Short-circuit fault diagnosis on the windings of three-phase induction motors through phasor analysis and fuzzy logic,” *Energies*, vol. 17, p. 4197, 2024. [CrossRef]

[10] I. Akhtar, S. Atiq, M. U. Shahid, A. Raza, N. A. Samee, and M. Alabdulhafith, “Novel glass-box-based explainable boosting machine for fault detection in electrical power transmission system,” *PLoS ONE*, vol. 19, pp. 309–459, 2024. [CrossRef]

[11] M. Iwagami, R. Inokuchi, E. Kawakami, T. Yamada, A. Goto, and T. Kuno, “Comparison of machine-learning and logistic regression models for predicting 30-day unplanned readmission in electronic health records: A development and validation study,” *PLOS Digit Health*, vol. 3, p. 578, 2024. [CrossRef]

[12] K. Mittal, K. Gill, R. Chauhan, and H. Pokhariya, “Exploring AdaBoost’s classifier’s power in electrical fault detection: A comprehensive analysis,” *Res. Gate*, vol. 1, pp. 1–5, 2024. [CrossRef]

[13] A. B. Bhadra and H. R. Jalilzadeh, “*A multi-layer perceptron neural network for fault type identification for transmission lines*,” Orlando, FL: Southeast Con, Press, pp. 198–203, 2023. [CrossRef]

[14] Y. Qifeng, C. Longsheng, M. T. Naeem, “Hidden Markov Models based intelligent health assessment and fault diagnosis of rolling element bearings,” *PLoS ONE*, vol. 19, pp. 1–12, 2024. [CrossRef]

[15] E. Arboleda, J. Daang, A. Omas-as, “Advancements in fault detection techniques for transmission lines:

A literature review," *Res. Gate*, vol. 1, pp. 16–24, 2024. [CrossRef]

[16] A. Indrayan, R. K. Malhotra, M. Pawar, "Using ROC curve analysis for prediction gives fallacious results: Use predictivity-based indices," *J. Postgrad. Med.*, vol. 70, pp. 91–96, 2024. [CrossRef]

[17] K. Dwan, D. G. Altman, J.A. Arnaiz, J. Bloom, A. W. Chan, E. Cronin, "Systematic review of the empirical evidence of study publication bias and outcome reporting bias," *PLoS ONE*, vol. 3, 2008. [CrossRef]

[18] M. J. M. Louis, "AI integration for enhanced smart grids and electric vehicle efficiency," *Res. Gate*, vol. 1, pp. 1–12, 2024. Available: https://www.researchgate.net/publication/380184015_AI_Integration_for_Enhanced_Smart_Grids_and_Electric_Vehicle_Efficiency.

[19] O. A. Omitaomu, H. Niu, "Artificial intelligence techniques in smart grid: A survey," *Smart Cities*, vol. 4, pp. 548–568, 2021. [CrossRef]

[20] P. Arévalo, F. Jurado, "Impact of artificial intelligence on the planning and operation of distributed energy systems in smart grids," *Energies*, vol. 17, p. 4501, 2024. [CrossRef]

[21] Wen, X.; Shen, Q.; Zheng, W.; Zhang, H. "AI-driven solar energy generation and smart grid integration: A holistic approach to enhancing renewable energy efficiency". *Int. J. Innov. Res. Eng. Manag.* **2024**, *11*, 55–66. [CrossRef]

[22] X. Wang, H. Wang, B. Bhandari, "AI-empowered methods for smart energy consumption: A review of load forecasting, anomaly detection, and demand response," *Int. J. Precis. Eng.-Manuf.-Green Technol.*, vol. 11, pp. 963–993, 2024. [CrossRef]

[23] S. Stock, D. Babazadeh, C. Becker, "Applications of artificial intelligence in distribution power system operation," *IEEE Access*, vol. 1, pp. 1–23, 2021. [CrossRef]

[24] Y. Zerguit, H. Youness, M. Derrhi, "Integrating AI for enhanced fault detection in industrial systems: Evaluating machine and deep learning approaches in the Industry 4.0 and IIOT era," *J. Theor. Appl. Inf. Technol.*, vol. 102, pp. 3795–3816, 2024. Available: https://www.jatit.org/volumes/Vol102No9/8Vol102No9.pdf.