IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

Machine Learning Approaches for Fault Detection and Classification in Electrical Power Systems

Mohasin Shafiahamad Bijali[1] and Dr. Shyam Sunder Kaushik[2]
[1]Research Scholar, [2]Professor
Department of Electrical Engineering
Sunrise University, Alwar, Rajasthan

Abstract: *Electrical power systems are exposed to different types of faults that can adversely affect system stability, reliability, power quality, and continuity of electricity supply. Conventional protection systems primarily rely on predefined thresholds, impedance characteristics, and deterministic relay settings, which may face limitations under changing operating conditions, renewable-energy penetration, nonlinear system behavior, and complex fault scenarios. Machine learning (ML) has emerged as a promising approach for intelligent fault detection and classification because it can learn complex relationships from voltage, current, frequency, phasor, and other electrical measurements.*

This research paper presents a comprehensive review of machine learning approaches used for fault detection and classification in electrical power systems. Supervised learning techniques such as decision trees, support vector machines, k-nearest neighbors, random forests, artificial neural networks, and ensemble methods are discussed along with emerging deep learning approaches including convolutional neural networks, recurrent neural networks, long short-term memory networks, and hybrid models.

Keywords: Machine Learning, Artificial Neural Networks, Support Vector Machine, Deep Learning, Power System Protection.

I. INTRODUCTION

Electrical power systems consist of interconnected generation units, transmission networks, distribution systems, transformers, circuit breakers, protection devices, and loads. Reliable operation of these components is essential for maintaining a stable and continuous electricity supply. Faults such as single-line-to-ground (LG), line-to-line (LL), double-line-to-ground (LLG), and three-phase (LLL) faults can produce abnormal currents and voltages and may cause equipment damage, voltage instability, cascading outages, or large-scale blackouts if they are not detected and cleared rapidly. Consequently, fault detection and classification are fundamental functions of modern power-system protection.

Traditional protection schemes generally employ overcurrent, distance, differential, impedance, and directional relays. These techniques have demonstrated high reliability over decades, but modern power networks are becoming increasingly complex. The integration of distributed generation, photovoltaic systems, wind farms, battery energy storage, flexible AC transmission systems, and power-electronic converters changes fault-current characteristics and may reduce the effectiveness of protection settings developed for conventional networks.

Machine learning provides an alternative approach by learning patterns from historical and simulated electrical measurements. Instead of relying exclusively on fixed thresholds, an ML model can identify relationships among multiple variables and determine whether a disturbance corresponds to a particular fault. Measurements such as three-phase voltages, currents, symmetrical components, frequency, phase angles, rate of change of frequency, and synchronized phasor measurements can be used as model inputs.

The application of ML to power-system protection has expanded from conventional artificial neural networks to support vector machines, decision trees, random forests, ensemble learning, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks. These approaches can perform binary fault/no-fault detection, multiclass fault classification, fault-location estimation, and coordinated diagnosis.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

The primary objective of this paper is to examine the role of ML techniques in detecting and classifying electrical power-system faults and to compare their suitability for practical protection applications.

Table 1. Major Power-System Fault Categories

Fault Type	Description	Typical Electrical Signature	ML Classification Difficulty
LG	One phase connected to ground	High phase/ground current, voltage depression	Moderate
LL	Two phases shorted	Increased current between phases	Moderate
LLG	Two phases connected to ground	Strong unbalanced current and voltage	High
LLL	Three phases shorted	Large balanced fault current	Relatively low
LLLG	Three phases connected to ground	Severe three-phase disturbance	Moderate
High-impedance fault	Fault through relatively high impedance	Low/irregular fault current	High

II. MACHINE LEARNING FRAMEWORK FOR FAULT DETECTION

An ML-based fault-detection system generally contains five major stages:

Data acquisition → Preprocessing → Feature extraction → ML classification → Protection decision

Measurements may be obtained from conventional instrument transformers, digital relays, phasor measurement units (PMUs), intelligent electronic devices, or simulated power-system models.

The data are first cleaned and normalized. Noise filtering and synchronization may then be applied. Relevant features are extracted from the signals using statistical, frequency-domain, time-frequency, or transformation-based methods. Finally, the ML classifier determines whether a fault exists and identifies its type.

III. DATA ACQUISITION AND PREPROCESSING

The performance of an ML classifier depends strongly on the quality of training data. Data can be generated through power-system simulation or collected from real electrical networks.

Common simulation environments include MATLAB/Simulink, PSCAD, DIgSILENT PowerFactory, and other electromagnetic-transient or power-system simulation platforms.

Typical input variables include:

1. Three-phase voltage signals
2. Three-phase current signals
3. Positive-, negative-, and zero-sequence components
4. Voltage and current phasor angles
5. Frequency
6. Active and reactive power
7. Rate of change of current
8. Rate of change of voltage
9. Harmonic components
10. Wavelet coefficients

Preprocessing may involve normalization, noise removal, missing-value treatment, segmentation, feature scaling, and dimensionality reduction.

Table 2. Common Input Features for ML-Based Fault Diagnosis

Feature Category	Examples	Application
Voltage	RMS voltage, phase voltage, voltage angle	Fault detection

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

Current	RMS current, phase current, current angle	Fault detection/classification
Symmetrical components	Positive, negative, zero sequence	Fault-type identification
Frequency	System frequency, frequency deviation	Disturbance analysis
Power	Active and reactive power	Operating-state analysis
Harmonics	THD, harmonic magnitudes	Abnormal-condition detection
Wavelet features	Detail and approximation coefficients	Transient fault detection
PMU features	Magnitude, angle, frequency	Wide-area diagnosis

IV. FEATURE EXTRACTION TECHNIQUES

Feature extraction converts raw electrical signals into informative numerical representations. Selecting appropriate features can improve classification accuracy and reduce computational requirements.

1. Fourier Transform

The Fourier transform represents signals in the frequency domain and can identify harmonic components associated with disturbances.

2. Wavelet Transform

Wavelet transform is particularly useful for transient power-system faults because it provides simultaneous time and frequency localization. Wavelet coefficients can capture sudden changes in current and voltage.

3. Symmetrical Components

Positive-, negative-, and zero-sequence components provide physically meaningful information about unbalanced faults and are therefore frequently used in classification models.

4. Principal Component Analysis

Principal Component Analysis (PCA) reduces dimensionality by transforming correlated features into a smaller set of principal components. It can decrease computational complexity while retaining much of the information contained in the original dataset.

Table 3. Comparison of Major ML Techniques

Algorithm	Main Strength	Main Limitation	Interpretability	Real-Time Potential
Decision Tree	Simple decision rules	Overfitting	High	High
KNN	Simple implementation	Expensive prediction for large datasets	High	Moderate
SVM	Strong nonlinear classification	Kernel selection	Moderate	High
Random Forest	Robust and accurate	Larger model size	Moderate–High	High
ANN	Learns nonlinear relationships	Training requirements	Low	High
CNN	Automatic feature learning	Requires substantial data	Low	High
RNN	Sequential-data processing	Training complexity	Low	Moderate–High
LSTM	Long-term temporal learning	Computational complexity	Low	Moderate–High
Hybrid Models	High flexibility	Higher complexity	Low–Moderate	Depends on design

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

V. FAULT DETECTION VERSUS FAULT CLASSIFICATION

Fault detection determines whether an abnormal electrical event has occurred, whereas fault classification determines the specific category of the fault.

For example:

Detection:

Normal → Fault

Classification:

Fault → LG / LL / LLG / LLL

A third stage may estimate the fault location:

Detection → Classification → Localization

This hierarchical structure can reduce classification complexity and provide more useful information to protection engineers.

VI. PERFORMANCE EVALUATION

The performance of ML-based fault classifiers should be evaluated using multiple metrics rather than accuracy alone.

Table 4. Evaluation Metrics

Metric	Formula/Concept	Purpose
Accuracy	Correct predictions / Total predictions	Overall performance
Precision	TP/(TP+FP)	Reliability of positive predictions
Recall	TP/(TP+FN)	Ability to detect actual faults
F1-score	Harmonic mean of precision and recall	Balanced performance
Specificity	TN/(TN+FP)	Ability to identify normal states
False alarm rate	FP relative to total normal cases	Protection reliability
Detection time	Time required to identify fault	Real-time suitability
Classification latency	Decision time after data acquisition	Operational performance

For protection applications, detection speed and false-alarm rate can be as important as classification accuracy.

VII. CLASS IMBALANCE PROBLEM

Real electrical systems do not experience every fault type with equal frequency. Some faults occur much more frequently than others. Consequently, an ML dataset may become imbalanced.

For example, an algorithm trained predominantly on LG faults may classify uncommon LLL or high-impedance faults poorly.

Possible solutions include:

1. Oversampling minority classes.
2. Undersampling dominant classes.
3. Synthetic data generation.
4. Class-weighted learning.
5. Focal loss for deep-learning models.
6. Collection of additional real-world fault records.

Evaluation should therefore include class-wise precision, recall, and F1-score rather than relying exclusively on overall accuracy.

VIII. INFLUENCE OF RENEWABLE ENERGY ON FAULT CLASSIFICATION

Modern power systems increasingly incorporate solar photovoltaic and wind generation. Power-electronic interfaces can substantially alter fault-current magnitude, duration, phase relationships, and harmonic characteristics.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

This creates an important challenge for ML-based protection. A classifier trained using a conventional power-system model may not generalize effectively to networks with high inverter-based generation.
Therefore, training datasets should include multiple operating scenarios and different levels of renewable penetration.

IX. NOISE AND MEASUREMENT UNCERTAINTY

Electrical measurements can contain noise due to instrument transformers, communication systems, sensors, sampling limitations, and electromagnetic interference. ML models trained only on ideal simulation data may perform poorly when exposed to noisy real-world measurements.
Robustness can be improved by:

1. Adding noise during training
2. Using signal filtering
3. Applying wavelet denoising
4. Using robust feature extraction
5. Testing across different signal-to-noise ratios
6. Combining data from multiple measurement sources

X. EXPLAINABLE ARTIFICIAL INTELLIGENCE IN POWER-SYSTEM PROTECTION

One important limitation of deep learning is its black-box nature. Protection engineers may be reluctant to rely on a model that produces a classification without explaining why it reached that decision.
Explainable AI (XAI) methods can provide information about influential features, decision pathways, or model behavior.
For example, feature-importance analysis may indicate whether negative-sequence current, voltage magnitude, phase-angle variation, or wavelet coefficients contributed strongly to a particular classification.
Explainability is especially important for safety-critical protection systems where incorrect decisions can have substantial consequences.

XI. CYBERSECURITY CONSIDERATIONS

As ML-based protection systems become connected to digital substations, PMUs, communication networks, and supervisory control systems, cybersecurity becomes increasingly important.
An attacker could potentially manipulate measurements or introduce adversarial perturbations designed to cause an ML model to misclassify an electrical condition.
Therefore, intelligent protection systems should incorporate:

1. Secure data acquisition
2. Authentication
3. Communication security
4. Anomaly detection
5. Robust ML training
6. Adversarial testing
7. Continuous model monitoring

Cybersecurity should be treated as an integral component of intelligent power-system protection rather than as a separate concern.

X. COMPARATIVE RESEARCH FINDINGS

Table 5. Overall Assessment of ML Approaches

Approach	Accuracy Potential	Data Requirement	Computational Cost	Suitability for Protection
Decision Tree	Moderate–High	Low–Moderate	Low	Good

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 2, June 2026 Impact Factor: 8.2

SVM	High	Moderate	Moderate	Very Good
Random Forest	High	Moderate–High	Moderate	Very Good
ANN	High	High	Moderate	Very Good
CNN	Very High	High	High	Excellent with suitable hardware
LSTM	High–Very High	High	High	Excellent for sequential data
Hybrid CNN-LSTM	Very High	Very High	Very High	Promising
Conventional relay	High for designed conditions	Very Low	Very Low	Mature and highly reliable
ML + conventional protection	Very High potential	High	Moderate–High	Highly promising

The comparison indicates that no single ML method is universally optimal. Traditional algorithms may be preferable where computational resources and training data are limited, whereas deep learning can provide stronger automatic feature-learning capabilities when large and representative datasets are available.

XI. CONCLUSION

Machine learning represents a significant opportunity for improving fault detection and classification in electrical power systems. Conventional protection techniques remain essential because of their maturity, speed, deterministic behavior, and extensive field validation. Nevertheless, the increasing complexity of modern grids creates situations in which fixed protection rules may not capture all possible operating conditions. ML approaches can learn complex relationships from voltage, current, phasor, frequency, and transient measurements and can therefore provide additional intelligence for fault diagnosis.

Among conventional ML techniques, SVM, random forest, decision trees, and ANN have demonstrated strong potential for classification problems. Deep-learning approaches such as CNN, RNN, and LSTM can further improve automatic feature extraction and temporal-pattern recognition, particularly when large datasets are available. Hybrid models offer another promising direction by combining signal-processing techniques with machine-learning or deep-learning architectures.

However, high laboratory or simulation accuracy does not automatically guarantee reliable field performance. Dataset diversity, model generalization, measurement noise, class imbalance, computational latency, interpretability, cybersecurity, and changing network configurations must all be considered before practical deployment. The most promising future direction is therefore not necessarily the complete replacement of conventional protection but the development of hybrid intelligent protection systems, in which ML algorithms complement established protection mechanisms.

Overall, ML-based fault detection and classification can contribute to faster and more adaptive power-system protection. Continued research involving real-world datasets, hardware-in-the-loop testing, explainable AI, robust learning, and renewable-dominated networks will be essential for translating research results into dependable practical applications.

REFERENCES

1. Rai, P., Londhe, N. D., & Raj, R. (2021). Fault classification in power system distribution network integrated with distributed generators using CNN. *Electric Power Systems Research, 194*, 106914.
2. Belagoune, S., Bali, N., Bakdi, A., Baadji, B., & Atif, K. (2021). Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems. *Measurement, 177*, 109330.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, June 2026
Impact Factor: 8.2

3. Moradzadeh, A., Teimourzadeh, H., Mohammadi-Ivatloo, B., & Pourhossein, K. (2022). Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. *International Journal of Electrical Power & Energy Systems, 135*, 107563.
4. Londhe, N. D., et al. (2024). Representational and sequential feature learning model for fault classification in power system. *Australian Journal of Electrical and Electronics Engineering, 22*(1).
5. Alhanaf, A. A., Farsadi, M., & Balik, H. H. (2024). Fault detection and classification in ring power system with DG penetration using hybrid CNN-LSTM. *IEEE Access, 12*, 59953–59975.
6. Sahoo, A. K., & Samal, S. K. (2024). Deep learning based detection, classification, and location of power system faults. *Bulletin of Electrical Engineering and Informatics, 13*(6).
7. Joshi, A., Khathoon, R., Devikrishna, Angel Peter, P. V., & Vinod, V. (2023). A very fast and easily implementable support vector machine based relay algorithm to classify the fault and non-fault disturbance in VSC HVDC terminals. *Electric Power Systems Research, 220*, 109380.
8. Yang, Y., Tu, F., Huang, S., Tu, Y., & Liu, T. (2023). Research on CNN-LSTM DC power system fault diagnosis and differential protection strategy based on reinforcement learning. *Frontiers in Energy Research, 11*, 1258549.
9. Moradzadeh, A., et al. (2024). Fault classification and localization in power transmission line based on machine learning and combined CNN-LSTM models. *Energy Reports, 12*, 5610–5622.
10. Abubakar, M. K., Monsurat, B. O., Adesina, L. M., Musa, A.-W., Olalekan, O., & Jimada-Ojuolape, B. (2026). A hybridized artificial neural network and support vector machine model in power transmission fault detection. *Communication in Physical Sciences, 13*(7), 1056–1072.
11. Bidhuri, D. (2026). Deep learning-based fault detection and classification in smart grid systems using hybrid CNN-BiLSTM architecture with attention mechanism. *International Journal of Creative and Open Research in Engineering and Management, 2*(5).
12. Yenagimath, M., Ankaliki, S., & Girish, V. (2024). Three-tier-deep-learning model for fault classification in power system using PMU data: A hybrid meta-heuristic optimization. *International Journal of Intelligent Systems and Applications in Engineering*.
13. Sahoo, A. K., & Samal, S. K. (2024). Deep learning based detection, classification, and location of power system faults. *Bulletin of Electrical Engineering and Informatics, 13*(6).
14. Moradzadeh, A., Teimourzadeh, H., Mohammadi-Ivatloo, B., & Pourhossein, K. (2022). Hybrid CNN-LSTM approaches for identification of type and locations of transmission line faults. *International Journal of Electrical Power & Energy Systems, 135*, 107563.
15. Rai, P., Londhe, N. D., & Raj, R. (2021). Fault classification in power system distribution network integrated with distributed generators using CNN. *Electric Power Systems Research, 194*, 106914.