Transmission line faults in power system and the different algorithms for identification, classification and localization

Akshith BT23CSE001 | **Dhruva Narayan** BT23CSE016 | **Saketh Vasudev** BT23CSE038

Abstract:

Transmission line faults are among the most critical issues affecting the stability and reliability of modern power systems. Rapid **identification**, accurate **classification**, and precise **localization** of these faults are essential for minimizing power outages, ensuring equipment safety, and maintaining overall system performance. This paper presents a comprehensive review and comparative analysis of various algorithms developed for fault detection in transmission lines. Traditional techniques such as **impedance-based methods** and **traveling wave methods** are examined alongside modern approaches involving **machine learning**, **support vector machines (SVM)**, **artificial neural networks (ANN)**, and **wavelet transform-based techniques**. The study evaluates these algorithms based on key parameters such as detection accuracy, computational efficiency, speed, noise tolerance, and scalability. Through simulation results and existing literature, the paper highlights the strengths and limitations of each method, offering insights into their practical applicability in real-world power systems. The findings aim to guide researchers and engineers in selecting optimal algorithms for robust and intelligent fault management in evolving power grid infrastructures.

Introduction:

The protection of transmission lines has always been a critical concern in the field of electrical engineering, as these lines are the backbone of any power system. They are responsible for transferring large amounts of electrical energy from generating stations to load centers, often across vast geographical distances. Being continuously exposed to environmental conditions, these lines are susceptible to faults that can result in serious disruptions to the power supply. A fault on a transmission line not only threatens the reliability of power delivery but also poses a risk to equipment and public safety. With rising demand for uninterrupted and high-quality power, it is essential to detect, classify, and locate faults accurately and promptly to ensure minimal downtime and prevent cascading failures in the grid.

Faults in transmission lines are an inevitable aspect of power systems. These faults may arise from a variety of causes such as lightning strikes, tree contact, conductor sagging, insulation failure, equipment malfunction, or human interference. Some of the most devastating power outages in history, like the **2012 India blackout affecting over 400 million people**, or the **2003 North American blackout which disrupted life for nearly 50 million residents**, were attributed to transmission faults that went undetected or uncontained. These events underscore the pressing need for reliable and intelligent fault management systems in modern power grids. Traditional systems using protective relays and distance relays can isolate faults, but their accuracy can be compromised by varying operating conditions, fault resistances, and transient disturbances.

From a technical perspective, electrical faults in transmission lines are generally classified into ***symmetrical and unsymmetrical faults***. *Symmetrical faults*, though rare, involve all three phases equally—such as **three-phase-to-ground (LLL-G) or three-phase faults (LLL)**—and result in balanced system disturbance. These are usually severe but easier to analyze. On the other hand, *unsymmetrical faults*—including **single line-to-ground, line-to-line, and double line-to-ground faults**—are more common and complex to analyze due to the imbalance they introduce in the system. These faults can lead to considerable electrical and mechanical stress, endangering both equipment and human lives. Accurate fault classification is essential for proper relay operation and fault isolation

.

.

Advancements in signal processing have introduced powerful techniques such as the wavelet transform, which excels in detecting transient events like faults by analyzing both the time and frequency components of voltage and current signals. This allows for more precise detection and localization, particularly during the brief and dynamic disturbance window. Complementing these methods, fuzzy logic has emerged as a flexible decision-making tool that can handle uncertainties in input signals. The integration of wavelet transforms with fuzzy logic systems has led to intelligent fault classifiers capable of operating under diverse and unpredictable conditions.

The advent of communication technologies like Global System for Mobile Communication (GSM) and Wireless Sensor Networks (WSNs) has revolutionized fault detection strategies. These technologies enable real-time transmission of system health data and geographic fault location through GPS. WSNs use distributed sensors along transmission corridors to track electrical anomalies, and the time differences between detection events are used to triangulate the fault location accurately. This is especially helpful in challenging terrains or remote regions where manual inspection is time-consuming or impractical.

In recent years, artificial intelligence (AI) and machine learning (ML) have played an increasingly vital role in power system monitoring and automation. These technologies offer advanced capabilities such as pattern recognition, anomaly detection, and predictive maintenance. AI models can learn from historical fault data and improve fault detection accuracy over time. Applications span from demand forecasting and energy optimization to smart microgrids and real-time grid management. Traditional methods often fall short in complex, high-impedance fault scenarios or evolving grid topologies. ML-based methods, by contrast, offer faster response times, adaptability, and the ability to work under noisy or uncertain data environments, making them indispensable tools in the smart grid era.

Another essential component in fault management is the fault locator. Fault locators are vital for quickly identifying the fault point, especially in systems where physical inspection is impractical or inefficient. Their importance is amplified in rugged terrains, vast transmission corridors, or systems involving multiple jurisdictions. Even when the system is automatically restored, fault locators help in pinpointing the origin—be it from bushfires, vandalism, or hardware failure—thus enabling maintenance teams to arrive at the site before the evidence is lost and take corrective measures efficiently.

In conclusion, ensuring the protection of transmission lines through early fault detection, accurate classification, and precise localization is fundamental to maintaining the safety, reliability, and economic operation of power systems. As the demand for smarter and more resilient power infrastructure grows, the integration of AI, advanced signal processing, and communication technologies will continue to transform fault analysis into a more proactive, data-driven, and intelligent process.

Literature:

Conventional Approach

Single-Terminal Impedance-Based Fault Location Methods

Single-terminal impedance-based methods are among the most widely adopted conventional techniques for fault location in transmission lines. These methods are preferred in practice due to their low cost, simple implementation, and independence from communication infrastructure. According to [3], these techniques estimate the fault location by calculating the apparent impedance seen from one end of the line and comparing it with the known impedance per unit length of the transmission line.

$$d = \frac{Z_f}{Z_{line}}$$

where Z_f is the measured impedance during the fault and Z_{line} is the known line impedance per unit length. Despite their simplicity, these methods are prone to several inaccuracies. One major source of error arises from the fault resistance, particularly in ground faults, which causes the measured impedance to deviate from the actual value, leading to significant location errors. Additionally, the load current flowing during the fault affects the accuracy of the estimate due to the "reactance effect." Other critical factors influencing the error include incorrect fault type identification, uncertainty in zero-sequence impedance values, and modeling inaccuracies such as neglecting mutual coupling or representing untransposed lines as transposed. For instance, a 20% deviation in the zero-sequence impedance can result in up to a 15% error in the calculated fault location. The estimation error is generally expressed as a percentage difference between the estimated and actual fault distances:

$$\text{Error}(\%) = \left| \frac{D_{estimated} - D_{actual}}{D_{actual}} \right| \times 100$$

Takagi method

The Takagi method is a single-ended impedance-based fault location technique that enhances accuracy by minimizing the effect of fault resistance, which commonly introduces errors in traditional methods. It uses both the pre-fault and fault voltage and current waveforms to distinguish between load current and fault current components. By doing so, it isolates the fault-generated component and computes a more accurate estimate of the fault distance. This approach is particularly effective for high-resistance faults and is widely implemented in modern numerical relays. However, its accuracy can still be influenced by line modeling inaccuracies and the quality of fault type identification

Travelling Wave-Based Method

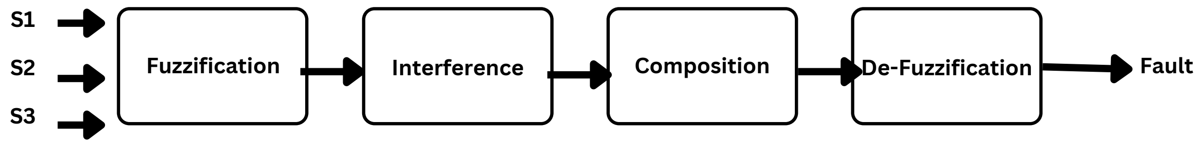
The travelling wave-based fault location method relies on detecting the high-frequency transient waves generated at the point of fault. These waves propagate along the transmission line and are captured at one or both ends of the line using high-speed recording devices. The time difference between the arrival of incident and reflected waves is used to estimate the location of the fault. This technique offers high precision and fast fault location, making it suitable for long-distance and extra-high voltage transmission lines. However, it requires fast data acquisition systems and precise time synchronization, such as GPS, which can increase implementation cost.

Fuzzy Logic-Based Fault Classification

FL is based on the ideas of fuzzy set theory and fuzzy set membership often found in natural (e.g., spoken) language. FL uses imprecision to provide robust solutions to problems. FL relies on the concept of a fuzzy set. The notation for fuzzy sets: for the member x, of a discrete set with membership μ, is μ/x. In other words, x is a member of the set to degree μ. Discrete sets are defined as:

$$A = \mu_1 / x_1 + \mu_2 / x_2 + \mu_3 / x_3 \ldots\ldots\ldots \mu_n / x_n$$

Fuzzy system



Where:

$$S1 = \frac{\mathrm{I_a} - \mathrm{I_b}}{\max\left(I_a, I_b, I_c\right)}$$

$$S2 = \frac{\mathrm{I_b} - \mathrm{I_c}}{\max\left(I_a, I_b, I_c\right)}$$

$$S3 = \frac{\mathrm{I_c} - \mathrm{I_a}}{\max\left(I_a, I_b, I_c\right)}$$

Rules to find nature of ground faults using values of S1, S2 and S3

- If (S1 is smallg) and (S2 is largeg) and (S3 is mediumg) then (F is AG)
- If (S1 is smallg) and (S2 is smallg) and (S3 is largeg) then (F is BG)
- If (S1 is mediumg) and (S2 is smallg) and (S3 is largeg) then (F is CG).
- If (S1 is Smallg) and (S2 is Largeg) and (S3 is Smallg) then (trip output is ABG)
- If (S1 is Smallg) and (S2 is Smallg) and (S3 is Largeg) then (trip output is BCG).
- If (S1 is Largeg) and (S2 is Smallg) and (S3 is Smallg) then (trip output is CAG).

Rules to find nature of phase faults.

- If (S1 is Smallph) and (S2 is Largeph) and (S3 is Smallph) then (trip output is AB)
- If (S1 is Smallph) and (S2 is Smallph) and (S3 is Largeph) then (trip output is BC)
- If (S1 is Largeph) and (S2 is Smallph) and (S3 is Smallph) then (trip output is CA)
- If (S1 is Mediumph) and (S2 is Mediumph) and (S3 is Smallph) then (trip output is ABC)
- If (S1 is Smallph) and (S2 is Mediumph) and (S3 is Mediumph) then (trip output is ABC)
- If (S1 is Mediumph) and (S2 is Smallph) and (S3 is Mediumph) then (trip output is ABC)
- If (S1 is Smallph) and (S2 is Smallph) and (S3 is Mediumph) then (trip output is ABC)

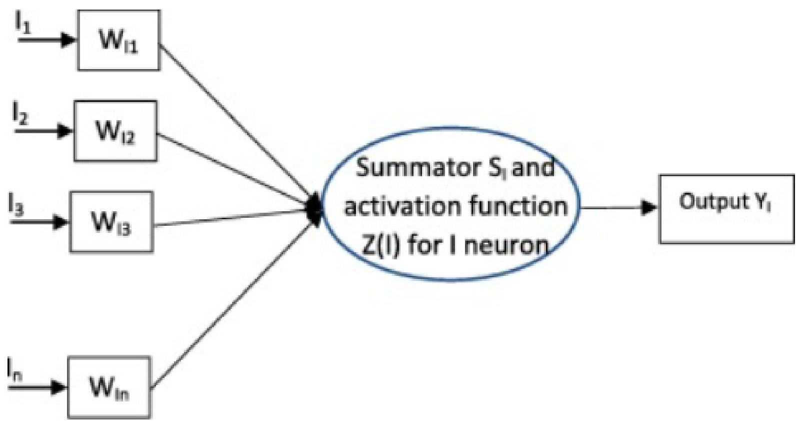
Deep Learning and MLTechniques

1. Artificial neural network

Artificial Neural Networks (ANNs) have emerged as a powerful tool in the domain of fault detection and localization in power systems due to their capability to model complex nonlinear relationships and learn from data. Inspired by the functioning of biological neurons, ANNs consist of interconnected processing units (neurons) that receive, process, and transmit information using weighted inputs and activation functions.

Each neuron processes the input vector $I = [I_1, I_2, \dots I_n]$ by multiplying it with corresponding weights $w_{I_1}, w_{I_2}, \dots$ *and* w_{I_n}, and computing a weighted sum S_I . This summation is then passed through a nonlinear activation function (AF) such as sigmoid, tanh, or ReLU, yielding the neuron's output Y_I

However, the primary challenge with ANN-based approaches lies in the selection of optimal network parameters and weights. Improper configuration can lead to issues like overfitting, underfitting, or poor generalization. Additionally, the requirement for substantial labeled data for training adds to the complexity of the method.



2. Long-short term memory

The system structure, known as the Recurrent NN (RNN), is applied to analyze time series. Its unique feature is that the preceding output information along with the current input information influence the present output. Although this network has certain benefits, gradient disappearance or explosion is more likely to happen when the data, needed by RNN, to learn is far from the present projected value. A variation of the RNN is LSTM. Through a device known as a gate unit, it is able to achieve the self-circulating weight change, successfully resolving the gradient disappearance and gradient explosion issues. As a result, it is more suited for handling time series data. LSTM is used for dynamically spatial forecasting problems. It learns through layers which helps to understand the pattern for better performance. It holds the information until required and then discards it when it is not required any more for prediction. There are three types of LSTM–

1. Frontward pass
2. Rearward pass
3. Bidirectional

Model Comparison and Performance Evaluation

To evaluate the effectiveness of various machine learning algorithms for transmission line fault detection and classification, we implemented and compared several models. The models tested include:

- Linear Regression
- Logistic Regression
- Polynomial Regression
- Naive Bayes
- Decision Tree
- Support Vector Machine (SVM)
- K-Nearest Neighbors (KNN)

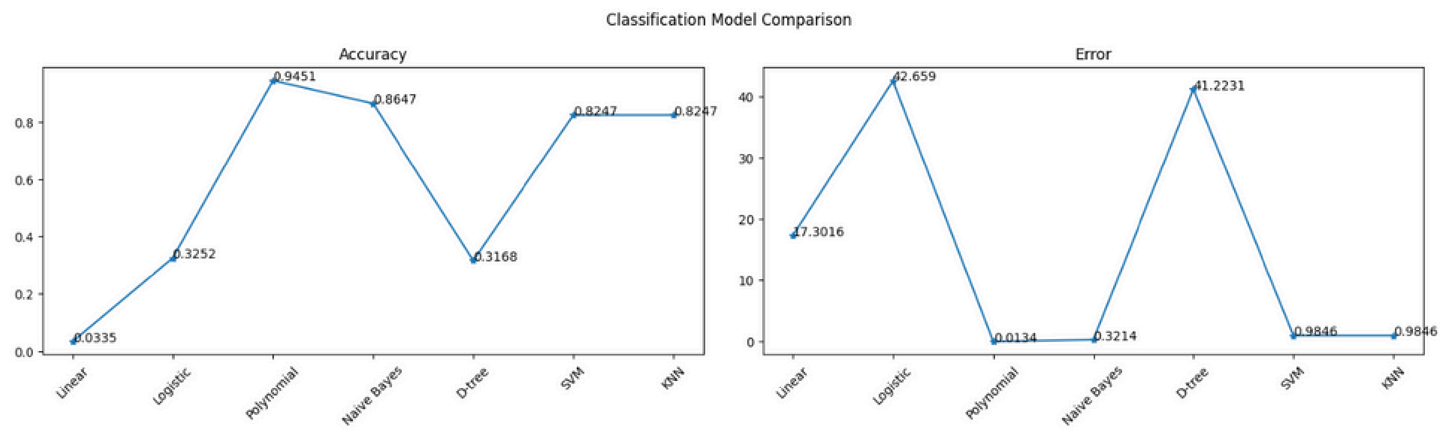
The objective here is to detect whether a fault has occurred. We trained each model using the preprocessed dataset and evaluated them on accuracy and error metrics.

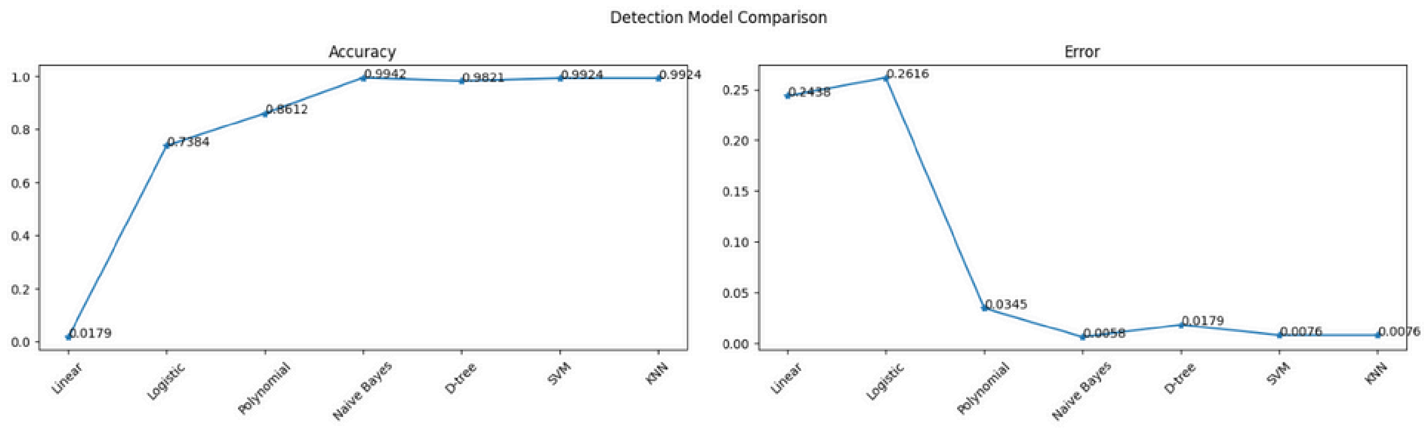
Detection Task:

Model	Accuracy(%)	Error (%)
Linear Regression	1.79	24.38
Logistic Regression	73.84	26.16
Polynomial Regression	86.12	3.45
Naive Bayes	98.06	1.94
Decision Tree	99.47	0.53
SVM	98.21	1.79
KNN	99.24	0.76

Classification Task:

Model	Accuracy(%)	Error (%)
Linear Regression	3.35	17.30
Logistic Regression	32.52	42.66
Polynomial Regression	94.51	0.01
Naive Bayes	79.69	2.11
Decision Tree	86.63	0.31
SVM	31.68	41.22
KNN	82.47	0.98





Observation

Among the various machine learning models tested for fault detection and classification in a three-phase transmission line system, the K-Nearest Neighbors (KNN) model demonstrated the highest performance in fault detection, achieving an accuracy of 99.24% with minimal error. For fault classification, the Polynomial Regression model outperformed others with an impressive accuracy of 94.51%. These results indicate that different models excel in different aspects of the fault analysis pipeline. The success of KNN in detection can be attributed to its ability to capture local patterns in the data, while Polynomial Regression effectively models the complex non-linear relationships required for accurate classification. These models are therefore well-suited for practical deployment in intelligent power system fault monitoring systems.

Proposed Solution

In this study, a comprehensive real time fault detection framework for transmission systems is proposed and implemented using Python. The simulation of the power network is carried out using the pandapower library, which offers a robust and flexible environment for modeling, analyzing, and simulating electrical power systems. To enable accurate fault identification, a machine learning based classifier using the K Nearest Neighbors (KNN) algorithm is integrated into the system. The classifier is trained on synthetically generated fault data by simulating various fault scenarios including single line to ground (LG), line to line (LL), double line to ground (LLG), and three phase (3P) faults using pandapower's shortcircuit module. During the training phase, critical electrical features such as line current magnitudes, voltage levels at both ends of each line, angle differences, their temporal derivatives, active power flow, and current imbalance are extracted from the system over a sliding window of time. These features are labeled based on the type of fault and used to train a binary fault detector and a multiclass fault type classifier.

The system initially operates in a training mode where faults are introduced randomly across various lines and categories, and the resulting data is collected. Once a sufficient number of labeled samples is acquired, the KNN model is trained using normalized feature vectors. The binary classifier is designed to distinguish between normal and faulty states while the auxiliary classifier is responsible for identifying the exact fault type. After training, the system transitions into a detection mode during which it continuously monitors the power system in real time. The power flows are updated every second to simulate realistic operational variations, and the extracted features are fed into the trained classifier to detect the presence of faults. When a fault is identified, the system also determines its type and provides a confidence score for the prediction.

The entire pipeline including network setup, feature extraction, training, and real time monitoring is implemented within a unified Python environment. Additional functionality for manual fault injection and clearing is also included to facilitate controlled testing. This solution demonstrates a seamless integration of data driven machine learning techniques with traditional power system simulations, offering a practical and modular approach to high frequency real time fault detection. It can serve as a foundation for more advanced protection, localization, and automated control strategies in future smart grid applications.

References:

- Khan, A. Q., Ullah, Q., Sarwar, M., Gul, S. T., & Iqbal, N. (2018). Transmission line fault detection and identification in an interconnected power network using phasor measurement units. IFAC-PapersOnLine, 51(24), 1356–1363. Your paragraph text
- Shukla, P. K., & Deepa, K. (2024). Deep learning techniques for transmission line fault classification – A comparative study. Ain Shams Engineering Journal, 15(2), 102427. https://doi.org/10.1016/j.asej.2023.102427
- Jadhavar, A., Gavali, S., Chavan, A., & Sayyad, A. P. (2023). Transmission line fault detection system. International Journal of Engineering Applied Sciences and Technology, 8(1), 289–292. http://www.ijeast.com
- Minh, N. Q., Khiem, N. T., & Giang, V. H. (2024). Fault classification and localization in power transmission line based on machine learning and combined CNN-LSTM models. Energy Reports, 12, 5610–5622. https://doi.org/10.1016/j.egyr.2024.11.061
- Hashim, M. M. I., Ping, H. W., & Ramachandaramurthy, V. K. (2009). Impedance-based fault location techniques for transmission lines. TENCON 2009 - 2009 IEEE Region 10 Conference. doi:10.1109/tencon.2009.5396253
- Choudhary, P., Bhaskar, M. K., & Parihar, M. (2023). Fuzzy logic-based fault classification for transmission line analysis. International Research Journal of Engineering and Technology (IRJET), 10(10), 287. https://www.irjet.net