Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

Nonlinear Optimization for Optimal Fault Detection and Classification in Power Systems Using Novel Semi-Supervised KNN Approach

Debshree Bhattacharya [1], Albert John Varghese [2,#], Rejo Roy [3]

[1] Assistant Professor, Electrical Engineering, Rungta College of Engineering & Technology, Bhilai

[2,3]Associate Professor, Electrical Engineering, Rungta College of Engineering & Technology, Bhilai

Email - [1]debshree.bhattacharya@gmail.com, [3]rejo.roy@rungta.ac.in

[#] Corresponding Author – [2]albert.varghese@rungta.ac.in

Article History:

Received: *19-08-2024*

Revised: *30-09-2024*

Accepted: *18-10-2024*

Abstract:
Power system reliability is paramount for ensuring uninterrupted energy supply. This paper addresses the critical challenge of accurate fault detection and classification in power systems, particularly in the context of increasing renewable energy integration. Traditional protection systems, while effective, often struggle to adapt to evolving grid conditions and emerging fault types. To overcome the limitations of traditional supervised learning methods, which require large amounts of labelled data, we propose a novel semi-supervised learning approach based on the K-Nearest Neighbor (KNN) algorithm for robust fault detection and classification in power systems. The proposed method leverages the strengths of semi-supervised learning, effectively utilizing both labelled and unlabelled data to improve classification accuracy of transmission line faults. By iteratively training the KNN model on an expanding dataset, the system can adapt to changing fault patterns and enhance its performance over time. To enhance the robustness and accuracy of fault classification, we formulate the fault detection and classification problem as a nonlinear variational inequality (NVI)-based problem. This NVI framework captures the inherent nonlinearity of power system dynamics and the variability of fault events, enabling the KNN classifier to accurately identify faults under diverse operating conditions. The proposed method is evaluated through extensive simulations in MATLAB. The effectiveness of the proposed approach is demonstrated through simulations on a 200 km transmission line under various fault conditions. The results indicate that the KNN classifier can accurately identify and categorize different fault types, contributing to improved system reliability and operational efficiency. This research provides a promising solution for enhancing power system protection, enabling timely fault detection and response, and ultimately minimizing system disruptions.

Keywords: K-Nearest Neighbour; Decision Tree; Support Vector Machine.

1. Introduction

The increasing global demand for reliable and sustainable energy has increasingly stressed power systems. Giving people access to enough, consistent power is the primary goal of every power system. Faults, often triggered by natural events like lightning or storms, disrupt the normal flow of electricity, potentially leading to widespread power outages and damage to infrastructure. To mitigate these risks, robust protection systems are essential.

Several types of gearbox line problems can occur like three-phase faults short circuits, line-to-line faults or line-to-ground faults. These faults can lead to damage to insulation and other critical components, maloperation of protective relays, cascading failures, affecting multiple components and systems.

Traditional protection systems, while effective, often rely on fixed thresholds and time delays. These methods may struggle to adapt to changing system conditions and emerging fault types. Machine learning offers a promising solution to address the limitations of traditional methods. It is critical to devise a coordinated control and protection systems capable of correctly and efficiently tracking any system-wide abnormalities and intervene in a timely manner to avoid further harm to the system and its components.

To optimize power quality, grid stability, and energy utilization, renewable energy sources require synchronization techniques. Phase-Locked Loops (PLLs) are a key technology for aligning the frequency and phase of PV-based systems with the grid, enabling uninterrupted power supply. [1]

Applications of Machine Learning in Power System Protection

Since current protection techniques accomplish their primary goal, researchers are concentrating on improving accuracy, decreasing operating time, and automating the protection process.

Fault Classification and Detection: Taking into account a wide range of challenges and operating circumstances an ANN-based Machine Learning approach can also be used for power system failure analysis [2]. Multi-label classification, specifically the binary relevance method, offers a powerful approach to detecting and classifying multiple faults concurrently. By employing individual binary classifiers for each label, this technique efficiently handles complex fault scenarios. [3]

Anomaly Detection: A semi-supervised learning method is becoming more popular since it can effectively handle both labelled and unlabelled data. Various types of transmission line problems can be detected and classified using the semi-supervised learning technique [4].

Real-time Monitoring and Control: This study compares the performance of K-Nearest Neighbors (KNN) and Decision Trees (DT) in semi-supervised classification tasks. Semi-supervised learning has found applications in various domains, such as innovative home load control [5] and HVAC system protection [6]. KNN is used to identify and categorise anomalies in malfunctioning events, so that safe operation of transmission lines is ensured by [7].

Hybrid Approach: For the purpose of analysing and predicting transmission line faults, a hybrid of KNN and fuzzy logic can also be employed [8]. To improve power quality and system efficiency, hybrid control methods, integrating Fuzzy Logic and Artificial Neural Networks, can be used to reduce harmonics in PV systems. [9] For robust transmission line protection, a combination of DWT and KNN algorithms is a promising solution. [10][11] Additionally, KNN-based techniques have been successfully applied to protect intricate six-phase transmission lines, enhancing power system reliability. [12]

2. K-Nearest Neighbors (KNN) algorithm

KNN is a semi-supervised learning algorithm that leverages a small labelled dataset to classify unlabelled data. This approach efficiently reduces the initial data requirements and eliminates the need for extensive classifier training. By incorporating new labelled data into the training set, KNN can adapt and improve its classification accuracy over time.

Nonlinear optimization techniques, particularly through semi-supervised learning approaches like K-Nearest Neighbors (KNN), have shown promise in enhancing fault detection and classification in power systems. These methods effectively utilize both labelled and unlabelled data, addressing the challenges posed by limited labelled datasets in complex power grid environments. The key aspects of Semi-Supervised Learning Framework are:

- Semi-supervised learning leverages a limited amount of labelled data in conjunction with a larger unlabelled dataset to enhance model generalization across diverse fault scenarios. [13]
- Techniques such as KNN can classify faults by calculating distances between labelled and unlabelled samples, achieving high accuracy rates (up to 97%) in fault detection [14]
- Advanced feature extraction methods, including wavelet packet analysis and frequency domain analysis, enhance the model's ability to identify faults [15][16]
- Feature selection techniques, including Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), can optimize feature subsets, enhancing the accuracy and reliability of fault classification models. [17] These techniques can also be used for forecasting applications to foresee occurrences of faults [18][19]

Case studies demonstrate the effectiveness of semi-supervised methods in real-world scenarios, with significant reductions in error rates even with minimal labelled data [15][16]. While semi-supervised learning offers substantial benefits, challenges remain in ensuring the quality of unlabelled data and the complexity of feature extraction, which can impact the overall performance of fault detection systems.

There are four critical steps that make up the KNN model. A key aspect of this KNN model is its use of a semi-supervised learning approach. This allows the classifier to learn from both labelled and unlabelled data. The process involves:

1. **Initial Training:** The classifier is trained on the small, initial labelled dataset.

2. **Labelling Unlabelled Data:** The classifier predicts the labels for unlabelled data points based on their similarity to labelled data points.

3. **Enhancing the Dataset:** The newly labelled data points are added to the training dataset, increasing its size and diversity.

4. **Retraining:** The classifier is retrained on the expanded dataset, improving its accuracy.

Two most important questions when implementing a KNN.

- How to choose the K-value: The minimum or cross-validation approach is employed for each test data point to calculate the value of K, which is frequently set to a constant and odd number. Numerous methods exist for determining how similar two sets of data are; some examples are the Minkowski distance, the Mahalanobis distance, and the Euclidean distance.
- Closeness measure between two data points: The nearby data points, or training samples, in the KNN model are the ones that have been correctly classified. Category classification is the foundational idea around which the technique is founded. The K-Nearest Neighbors (KNN) algorithm classifies data points based on the majority class of their k nearest neighbors.

The key advantages of the Semi-Supervised Approach:

- Requires fewer labelled data points.
- The model's performance can be continually enhanced through iterative learning.
- The iterative training process is computationally less intensive compared to traditional supervised learning.

The KNN classifier is trained on a relatively small dataset of labelled power system data. This initial dataset is used to establish the classifier's decision-making criteria. To evaluate the classifier's performance, a validation dataset consisting of 20 fault scenarios with varying fault resistances and locations is introduced. The classifier's predictions for these scenarios are compared to the known ground truth labels. The KNN classifier categorizes faults based on the fault current magnitude measured at bus B1. The classifier's decision-making process is guided by the patterns identified in the training data.

By effectively leveraging semi-supervised learning, the KNN classifier can provide reliable fault classification, even in scenarios with limited labelled data. The paper introduces the KNN algorithm, a classifier that can handle defect data points with and without labels, and it is based on a semi-supervised machine learning approach. Fault currents obtained at a single bus are sent into the classifier. After classifying unlabelled data, it is added to the labelled dataset. The KNN classifier then utilizes the k-nearest labelled samples to classify subsequent unlabelled data points. The system's performance is evaluated under various fault resistance levels on transmission lines.

3. TEST SYSTEM DESCRIPTION

Figure 1 illustrates a Three-Phase transmission line fault identification and categorization system. This system is simulated using MATLAB and configured as depicted in the figure.

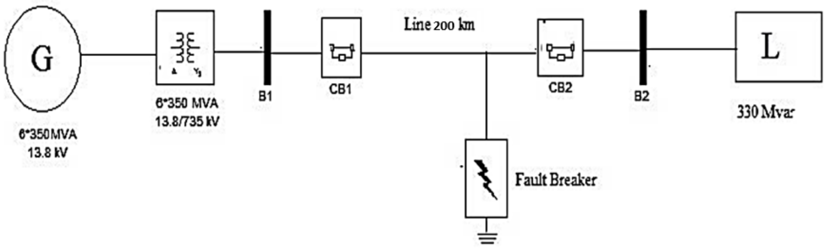

Fig 1: 3-Φ transmission line fault identification and categorization system [14]

System Configuration:

- **Power Transmission:** Power is transmitted from a power plant to a variable load network over a 200 km, 735 kV, 100 MW, three-phase transmission line operating at 60 Hz.
- **Generation:** Six 350 MVA generators are connected to bus B1.
- **Load:** The load is connected to bus B2.
- **Circuit Breakers:** Circuit breakers CB1 and CB2 are installed on both sides of the transmission line.

- **Fault Simulation:** A Fault Breaker block is used to simulate various fault types at specific locations and resistances along the line.

The Fault Breaker block in MATLAB provides access to data related to various short-circuit faults. This study focuses on ten common fault types in transmission lines: Three-Phase-to-Ground Fault (LLLG), Three-Phase Fault (LLL), Two-Phase-to-Ground Fault (LLG) and Line-to-Line Fault (LL). The fault current at bus B1 is utilized as input data for fault classification. During the training phase, each fault type is simulated on the test system, and the corresponding current data at bus B1 is extracted from the MATLAB workspace.

The classes of Faults are considered as follows : Line R-Y-B Fault, Line R- Line Y Fault, Line Y- Line B Fault, Line R- Line B Fault, Line R-Ground Fault, Line Y-Ground Fault, Line B-Ground Fault, Line R-Y-Ground Fault, Line Y-B-Ground Fault, Line R-B-Ground Fault is classified as Class A, Class B, Class C, Class D, Class E, Class F, Class G, Class H, Class I, Class J respectively. Because the types of defects are unique, the classifier may be trained using the supplied data. The classifier is trained using a dataset of fault simulations to accurately identify and categorize different fault types.

4. RESULTS

This research proposes a machine learning-based approach for detecting and classifying faults in a power transmission system. Using the IEEE 14-Bus Power Grid model as a simulation platform, a two-layer strategy is employed to accurately identify and categorize faults on transmission lines. A K-Nearest Neighbors (KNN) classifier is developed to identify fault types. The model is trained on a variety of fault scenarios to ensure accurate classification. The simulation is conducted using MATLAB. To evaluate the model's performance, a series of tests are performed, and the results, including identified fault classes and current amplitudes, are presented in Excel format.

Figures 2 and 3 illustrate the simulated responses for a 200 km transmission line with fault resistances of 0.002 ohms and 0.2 ohms, respectively, as determined by the KNN algorithm. KNN classification results for a 200 km transmission line with a fault resistance of 0.002 ohms is shown in Figure 2.

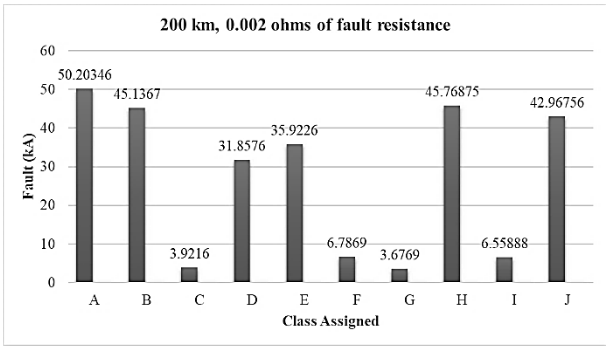

Fig 2: Simulated Response for 200 km, 0.002 Ohms of Fault Resistance based on KNN findings

KNN classification results for the same transmission line with a fault resistance of 0.2 ohms, using the updated class labels from the previous scenario are illustrated below in Figure 3.

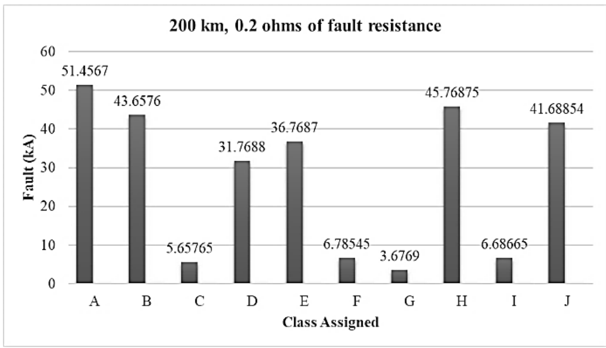

Fig 3: Simulated Response for 200 km, 0.2 Ohms of Fault Resistance based on KNN findings

By leveraging the KNN algorithm and a comprehensive training dataset, this study demonstrates the feasibility of accurately detecting and classifying faults in power transmission systems, contributing to improved system reliability and operational efficiency.

5. CONCLUSION

This paper presented a novel approach for fault detection and classification in power systems using a semi-supervised K-Nearest Neighbor (KNN) algorithm. The proposed method effectively utilizes both labelled and unlabelled data, overcoming limitations associated with traditional supervised learning methods that require large amounts of labelled data.

The KNN classifier model was trained on a diverse set of fault scenarios and it demonstrates high accuracy in identifying and categorizing various fault types, including those with varying fault resistances and locations. The KNN classifier integrated with a nonlinear variational inequality framework, identified fault types on a simulated 200 km transmission line under diverse fault conditions. The inherent nonlinearity of power system dynamics is robustly captured, enabling the KNN classifier to adapt to changing fault patterns and emerging fault types.

The results obtained through extensive simulations validate the effectiveness of the proposed approach in improving power system reliability and operational efficiency. By enabling timely fault detection and response, this method contributes to minimizing system disruptions and enhancing the overall resilience of power systems.

While the proposed method shows promising results, future research can explore the following avenues to further improve its performance and applicability:

- Developing real-time fault detection and classification systems that can quickly identify and respond to faults.
- Incorporating deep learning and ensemble methods to enhance the accuracy and robustness of the classifier.
- Integrating IoT and sensor data to improve the granularity of fault detection and enable predictive maintenance.

By continuously advancing the fault detection and classification, we can further enhance the reliability and resilience of power systems, ensuring a secure and sustainable energy future.

References

[1] Sourabh Kumar, Albert John Varghese, Rejo Roy and Debshree Bhattacharya, "Phase Locked Loop-based Synchronization of Solar PV System with Single-Phase Grid for Integrated Load Supply", International Journal of Intelligent Systems and Applications in Engineering, Vol. 12(4s), pp. 545–551 (2024)

[2] H. A. Tokel, R. A. Halaseh, G. Alirezaei and R. Mathar, "A new approach for machine learning-based fault detection and classification in power systems," 2018 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference, Washington, DC, USA, doi: 10.1109/ISGT.2018.8403343, pp. 1-5 (2018)

[3] Debshree Bhattacharya and Manoj Kumar Nigam, A Multilabel Approach for Fault Detection and Classification of Transmission Lines using Binary Relevance, International Journal on Recent and Innovation Trends in Computing and Communication Vol 11(7), pp. 261–269 (2023)

[4] Tamer S. Abdelgayed, Walid G. Morsi and Tarlochan S. Sidhu, "Fault detection and classification based on co-training of semi-supervised machine learning", 2019 International Conference on Innovative Trends and Advances in Engineering and Technology, doi: 10.1109/THE 2017.2726961 (2019)

[5] Robert A. Sowah, Nicholas A. Dzabeng, Abdul R. Ofoli, Amevi Acakpovi, Koudjo. M. Koumadi, Joshua Ocrah and Deborah Martin, "Design of Power Distribution Network Fault Data Collector for Fault Detection, Location and Classification using Machine Learning," 2018 IEEE 7th International Conference on Adaptive Science & Technology (ICAST), Accra, Ghana, doi: 10.1109/ICASTECH.2018.8506774, pp. 1-8 (2018)

[6] M. Dey, S. P. Rana and S. Dudley, "Semi-Supervised Learning Techniques for Automated Fault Detection and Diagnosis of HVAC Systems," 2018 IEEE 30th International Conference on Tools with Artificial Intelligence (ICTAI), Volos, Greece, doi: 10.1109/ICTAI.2018.00136, pp. 872-877 (2018)

[7] A. Yadav and A. Swetapadma, "Fault analysis in three phase transmission lines using k-nearest neighbor algorithm," 2014 International Conference on Advances in Electronics Computers and Communications, Bangalore, India, doi: 10.1109/ICAECC.2014.7002474, pp. 1-5 (2014)

[8] Y. Zhang, J. Chen, Q. Fang and Z. Ye, "Fault analysis and prediction of transmission line based on fuzzy K-Nearest Neighbor algorithm," 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Changsha, China, doi: 10.1109/FSKD.2016.7603296, pp. 894-899 (2016)

[9] Vinit Kumar, Rejo Roy, Albert John Varghese and Debshree Bhattacharya, Enhancing Power Quality in Solar-Fed Cascaded Multilevel Inverters: A Comparative Study of Fuzzy Logic and Neural Network Controllers for Output Voltage Regulation, International Journal of Intelligent Systems and Applications in Engineering: Vol. 12(8s), pp. 395–403 (2024)

[10] M.Pavani, S.Raghunath Sagar and G.Ganesh, "Fault classification in transmission lines using wavelet transform and k-nearest neighbors", International Journal of Education and Applied Research, Vol. 9 (2019)

[11] Akanksha Malhotra and Purva Sharma, "Fault Detection and Classification Using Discrete Wavelet Transform", International Journal of Research and Scientific Innovation, Vol. V, pp. 93-101 (2018)

[12] S. K. Shukla and E. Koley, "Fault detection and classification in six-phase transmission system using k-nearest neighbour algorithm," 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies, Kerala, India, doi: 10.1109/ICICICT1.2017.8342621, pp. 542-546 (2017)

[13] Weihua Li, Xiaoli Zhang and Ruqiang Yan, "Semi-supervised Learning Based Intelligent Fault Diagnosis Methods." Intelligent Fault Diagnosis and Health Assessment for Complex Electro-Mechanical Systems. Singapore: Springer Nature Singapore, pp. 95-191 (2023)

[14] P. P. Wasnik, N. J. Phadkule and K. D. Thakur, "Fault detection and classification based on semi-supervised machine learning using KNN," 2019 International Conference on Innovative Trends and Advances in Engineering and Technology, Shegoaon, India, doi: 10.1109/ICITAET47105.2019.9170220, pp. 79-83 (2019)

[15] Jiahao Zhang, Lan Cheng, Zhile Yang, Qinge Xiao, Sohail Khan, Rui Liang, Xinyu Wu and Yuanjun Guo, "An enhanced semi-supervised learning method with self-supervised and adaptive threshold for fault detection and classification in urban power grids", Energy and AI, Vol. 17 doi: 10.1016/j.egyai.2024.100377, (2024)

[16] Xinyang Li, Hongfa Meng, Xiaoling Peng, "Semi-supervised learning for fault identification in electricity distribution networks", Twelfth International Conference on Signal Processing Systems, doi: 10.1117/12.2589229, (2021)

[17] Homanga Bharadhwaj, Avinash Kumar, Abheejeet, Mohapatra, "A Synchro-phasor Assisted Optimal Features Based Scheme for Fault Detection and Classification", International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, doi: 10.1109/IJCNN.2019.8852371, pp. 1-8 (2019)

[18] Rejo Roy, Albert John Varghese and S,.R. Awasthi, "Compare and Evaluate One Hour Ahead Solar Power Forecasting with ANN Tuned Using Genetic Algorithm and Hybrid Genetics-Based Particle Swarm Optimization", Control Applications in Modern Power Systems. Lecture Notes in Electrical Engineering, Vol 870, Springer, Singapore, doi: 10.1007/978-981-19-0193-5_33, pp. 405-417 (2022)

[19] Rejo Roy, Albert John Varghese and S.R. Awasthi, "Short-Term Power Forecasting for Renewable Energy Sources Using Genetics-Based Harmony Search Algorithm". Smart and Intelligent Systems. Algorithms for Intelligent Systems. Springer, Singapore, doi: 10.1007/978-981-16-2109-3_34, pp. 357-367 (2022)