WILEY | Hindawi

Research Article

Optimal PMU Placement for Fault Classification and Localization Using Enhanced Feature Selection in Machine Learning Algorithms

Ayman Faza [ID],[1] **Amjed Al-Mousa** [ID],[2] and **Rajaa Alqudah** [ID][2]

[1]*Electrical Engineering Department, Princess Sumaya University for Technology, Amman 11941, Jordan*
[2]*Computer Engineering Department, Princess Sumaya University for Technology, Amman 11941, Jordan*

Correspondence should be addressed to Ayman Faza; a.faza@psut.edu.jo

Machine learning (ML) algorithms are increasingly used in power systems applications. One important application is the classification and localization of various types of transmission line faults. Using voltage and current measurements from phasor measurement units (PMUs), a number of useful features can be extracted, which can form the basis of a ML-based prediction of the fault type, line, and distance on the line. This paper proposes a technique to find the optimal number and placement of PMUs by performing thorough feature selection. The features are selected to maximize the accuracy of the ML classification and regression algorithms. The results show that for the IEEE 14 bus system, the use of only five PMUs is sufficient to obtain high levels of accuracy. For example, a testing accuracy of 99.0% and 97.1% can be achieved for the fault type and fault line location, respectively. As for the fault distance along the line, the testing MAE of 3.1% can be obtained along with an R^2 score of 94.4%. Adding more PMUs does not provide any additional value in terms of accuracy.

1. Introduction

Phasor measurement units (PMUs) have recently been employed in power grids with the goal of monitoring the state and operation of the grid with much higher accuracy and at a much higher data rate than its previous counterpart, the supervisory control and data acquisition (SCADA) systems. SCADA systems are still very widely used in monitoring today's grid, but the advancement in PMU technology and the high accuracy of synchrophasor measurement are promising to revolutionize the monitoring and control process of the power grid to provide higher accuracy and much faster action mechanisms that would ensure higher stability, reliability, and security for the grid.

In addition, with the recent advances in machine learning (ML) algorithms, the data provided by PMUs can be used to provide a faster and more accurate assessment of the system state, which can be used to detect abnormal conditions and take appropriate actions much

faster than before, which can provide a great benefit in regard to controlling the system operation, including the mitigation of different types of disturbances, line faults, or generator tripping events, as well as reacting to sudden changes in demand which can otherwise cause serious disruption to the grid operation if not addressed quickly enough.

In that respect, the use of PMUs in combination with ML algorithms can provide an accurate assessment of system operation, quickly detect abnormal situations, and help the control centers take quick action to mitigate any problems in the grid, or at least help isolate them to prevent further propagation of the problems.

In this paper, signals obtained from PMUs were used to classify and localize faults that occur in the power grid using ML algorithms. The IEEE 14 bus system is used as a test case, and various types of faults are simulated in some transmission lines in the network and at varying distances on each transmission line in question. PMUs collect positive

sequence voltages and currents at each bus, in addition to zero sequence components of the same voltages and currents, and store the results to be processed by the system.

By obtaining sufficient information from the grid, ML algorithms are able to provide an accurate classification of the fault type, determine at which transmission line the fault has occurred, and at what distance the fault has occurred as measured from the two buses between which the line is connected. In this paper, a support vector machine (SVM) classifier was used to determine the fault type and location, while random forest regression was used to predict the fault distance.

In addition, this paper shows that while we can theoretically place a PMU at each bus and obtain voltage and current information from every bus in the system, there is actually no need to place a PMU at all buses. Through correlation-based feature subset selection, we show that only a subset of the available signals can be used to accurately determine the type and location of the faults. By carefully selecting the most relevant signals, we show that rather than using signals from all 14 PMUs, it suffices to use the signals obtained from only 5 PMUs to accurately classify the type of fault and determine its exact location with a high degree of accuracy. In fact, our results show that adding more data to our ML algorithm does not necessarily improve the accuracy of our classification and regression any further. As such, it would be considered a waste of computing resources to continue processing data from all PMUs when it is sufficient to use only a few.

It is worth mentioning that the majority of the proposed approach workload is completed offline, including data preparation, optimal feature selection, training the machine learning model, and optimizing PMU placement. This offline setup ensures that real-time processing remains streamlined and does not require significant time or resources. This approach guarantees that the proposed technique is highly suitable and practical for this application.

Despite advancements in PMU placement studies, a notable research gap exists in addressing fault classification and localization as primary factors in the placement decision. Previous works primarily focused on system observability. The literature lacks comprehensive methodologies that specifically consider fault classification and localization in determining the optimal number and positioning of PMUs. This research is aimed at bridging this gap by proposing a novel machine learning-based approach that targets the optimization of PMU placement for improved fault classification and localization in power systems.

It follows that the main contribution of this work lies in proposing a machine learning-based methodology to determine the optimal number and positioning of PMUs for the classification and localization of faults in transmission lines.

The rest of the paper is organized as follows. Section 2 provides a thorough review of the related literature. Section 3 describes the details of the IEEE 14 bus system, the PMU placement, and the signals and relevant features obtained from the PMUs to develop the dataset used in this research. In Section 4, the methodology used to select the most important features in the dataset is outlined. In addition, the ML algorithms used were explained, and the methods used to evaluate their correctness are outlined. Section 5 presents the results of the ML algorithms and discusses their accuracy in terms of correctly classifying and localizing the transmission faults. In addition, in this section, we provide the rationale for choosing only a subset of the PMUs to provide the relevant signals in the system in light of the obtained results. Section 6 concludes this paper and provides directions for future research.

2. Literature Review

A summary of related literature is presented in this section. It includes a review of the most common ML applications in power systems, a survey of fault detection and classification research using ML methods, in addition to a summary of the most relevant research on the optimal placement of PMUs.

2.1. ML in Power Systems. ML techniques have many applications in power systems, and a number of recent papers have provided a good overview of the main areas of research that have benefited from various ML algorithms.

Miraftabzadeh et al. present several applications for ML algorithms in the areas of power flow solutions, load forecasting, forecasting for solar energy, and power quality considerations, including disturbance detection and classification [1, 2]. In addition, they list the main ML classification algorithms used in power system-related research, such as logistic regression, K-nearest neighbor (KNN), random forest, and support vector machines (SVM), all of which have applications in the aforementioned areas, mainly in the areas of detection and classification of power related disturbances and faults.

Regression techniques such as linear regression, regression trees, and deep neural networks have also been extensively used in various types of forecasting and control methods in power systems.

The work of Ozcanli et al. [3] mentions wind and solar energy forecasting as valuable applications of ML algorithms while also confirming that the detection and classification of different types of faults and/or disturbances can effectively be done using ML algorithms.

The work in [4] also presented ML applications in power systems, focusing on predicting power outages, decision-making, and system restoration, in addition to system stability assessment and stability control.

2.2. Machine Learning in Fault Detection. Several papers were found in the literature on the topic of using ML algorithms for fault detection. Examples of these algorithms are deep learning, artificial neural networks (ANNs), SVM, recurrent neural network (RNN), and long short-term memory (LSTM).

Deep learning is a widely used technique to detect faults in power systems. For example, Yu et al. [5] proposed a wavelet transform and deep learning-based technique to detect faults in microgrids. The proposed technique was applied to the CERTS microgrid and the IEEE 34 bus system. In [6], Wang et al. used deep learning neural networks

(DLNN) to diagnose faults in power systems. The data was collected from the power system dispatching department and fed into stacked autoencoders (SAEs) to train the DLNN. The number of hidden layers was examined to evaluate the model results. Markovic et al. [7] proposed a hybrid approach of two stages to detect the faults that may occur in a buck converter connected to a distribution network. The first stage used a model-based technique to estimate system parameters, while during the second phase, a deep neural network classifier was used to predict if the system was normal or faulty.

On the other hand, Barreto et al. [8] used the ANN technique for fault detection and identification in power systems. The study considered different types of faults like short circuits, line contingencies, and load contingencies. The study was carried out on the IEEE 39 bus system considering different numbers of PMUs.

The SVM technique was widely used in the scope of power system fault detection. Aleem et al. [9] presented a quarter-sphere SVM for fault detection and classification in power systems, while the author of [10] used two binary classification algorithms to detect and identify faults in smart grids. To enhance the performance of these classifiers, hyperparameter tuning and data balancing using the Synthetic Minority Oversampling Technique (SMOTE) were evaluated. Medeiros et al. [11] developed a field application system for detecting structural faults on anchor rods using frequency domain reflectometry analysis. The system utilizes ML techniques to classify the measured signals as normal or faulty, achieving an accuracy greater than 98%.

To compare SVM and ANN techniques in detecting and identifying faults in power systems, Manojna et al. [12] compared the effectiveness of these algorithms using a dataset of phase voltage, phase current, and their combinations. The paper concluded that the three models give high accuracy for fault detection while SVM gives the highest accuracy for fault classification.

Belagoune et al. [13] introduced three deep learning models based on long short-term memory (LSTM). These models are for fault region identification (FRI), fault type classification (FTC), and fault location prediction (FLP). The study was conducted on a two-area four-machine power system. The study shows that the LSTM-based models effectively identify faults in power systems. Tehrani and Levorato [14] extracted time domain and frequency domain features and used LSTM to detect faults and abnormalities in smart grids.

2.3. Fault Classification and Localization. Many papers address the problem of classifying and localizing faults in power systems. For example, to classify and locate frequency disturbance events (FDEs) such as generator trip (GT), line outage (LO), and load disconnection (LD) that occur in power systems, Shadi et al. [15] used a hierarchical methodology that employed two machine learning algorithms, namely, recurring neural networks (RNNs) and an LSTM. The study shows that RNN performs better than LSTM and converges faster with fewer epochs. Zhang et al. in [16] used an LSTM model to capture temporal features in the power grid, which performs fault detection. The LSTM is followed by an SVM classifier that can classify the fault type accurately.

On the other hand, Zhang et al. in [17] used RNN and bidirectional gated recurrent unit (Bi-GRU) for fault location on the IEEE 39 bus system. The model is used to maintain the temporal current signal characteristics. Chen et al. [18] used a graph convolutional network (GCN) to localize faults in power distribution networks. The model was tested using IEEE 123-bus and 37-bus benchmark systems. This model outperforms the SVM, random forest, and a three-layer fully connected neural network (FCNN) in terms of accuracy and robustness.

Ren et al. [19] used deep learning convolutional neural networks (CNNs) to classify and localize power system disturbances. A dataset was generated by conducting a dynamic simulation on the Polish 3120 bus system, and a hyperparameter search was done to optimize the CNN model, where the model accuracy was around 84% and 91% for the classification and localization, respectively.

SVM was widely used in classifying fault types of power systems. The authors of [20] used SVM to classify the fault type in a transmission system with a connected loop configuration. The model achieved an accuracy result of 92.13%. Mandava et al. [21] used SVM to identify if a system is working correctly and identify the type of fault if any.

Various research papers used ANN and DNN to classify and locate faults in hybrid power systems. The model proposed in [22] used feed-forward of three-phase voltages and currents while using backpropagation for fault classification. Resmi et al. [23] used ANN to locate the faults in transmission lines by identifying the zone in which the fault occurred. The paper focused only on unsymmetrical faults of types of line-to-ground faults, line-to-line faults, and double line-to-ground faults. On the other hand, Jain et al. [24] used ANN to classify and locate phase-to-phase faults in parallel transmission lines within one cycle after the fault happened. Nasrin et al. [25] used deep learning to locate ten types of faults that may occur in power systems.

To overcome the issue of data imbalance in the dataset, a data augmentation classifier (DAC) and multigenerator data augmentation classifier (MDAC), which are based on generative adversarial networks (GANs), were proposed in [26]. The paper shows the improvement of the accuracy results after balancing.

In this paper, the work presented uses a linear SVM classifier and random forest regressor to classify and localize faults in an IEEE 14 bus system using voltage and current data signals obtained from PMUs placed in the system. Furthermore, we show a method to determine the best location and optimal number of PMUs to be placed in the system for optimal fault classification and localization in the system based on the accuracy of the ML techniques used in the classification and localization.

2.4. PMU Placement. Various studies have focused on finding the optimal location for PMU placement in different power systems of various sizes. Most recent references on this topic focus on finding the best locations for PMUs from

the perspective of system observability for state estimation applications. However, very little work was done for placing the PMUs for the specific purpose of fault localization and classification. For example, Abbasy et al. present an approach for optimal PMU placement based on the observability of the network using a binary integer linear programming optimization method. The number of PMUs used was set to the minimum required to achieve the required observability [27]. Eladl et al. used binary integer programming as well for optimal allocation of PMUs by minimizing the cost of total PMUs given connectivity and observability constraints [28]. On the other hand, Taher et al. used the modified imperialist competitive algorithm for optimal allocation of PMUs and also for full observability with a minimal number of PMUs [29].

Sodhi et al. [30] presented a two-stage method for optimal placement of PMUs for system state estimation, where in stage 1, PMU placement was done using integer linear programming. In stage 2, a sequential elimination algorithm was used for numerical observability.

The work by Madani et al. [31] presents a methodology for the optimal placement of PMUs based on multiple criteria and not just state estimation. Their methodology includes a weighted average, whereby different criteria are given different weights in the decision-making process. In addition to state estimation, the criteria considered for optimal PMU placement include system robustness, critical network paths, and interarea oscillations.

The work in [32] by Li et al. considered PMU placement for fault line localization. They used convolutional neural networks (CNNs) to best place the PMUs for the best localization of faults in the system. They presented various accuracy and observability levels for different PMU placement algorithms.

Many studies have aimed to determine the optimal placement of PMUs across various IEEE networks, including the IEEE 14 bus system, employing diverse approaches. References [33, 34], for example, adopted an instruction-level parallelism (ILP) approach, while [35] utilized integer linear programming. The authors of [36] explored several methods, including depth-first search (DFS), graphic theoretic procedure using the merger method, original simulated annealing method, and recursive security N algorithm. Furthermore, [37] employed a hybrid algorithm, combining the graph-theoretic procedure and the recursive algorithm (RA), to address the optimal PMU placement problem. Reference [38] employed mathematical programming formulations, specifically mixed-integer linear programming (MILP) and nonlinear programming (NLP). The work in this paper used a machine learning approach not only to find the optimal placement of PMUs but also to optimize the number of needed PMUs for improved fault classification and localization in power systems.

It is noted that most of the previous studies have looked at the PMU placement problem from the perspective of system observability. Our work differs as it considers the problems of fault classification and localization as the main factors in placing PMUs in a system, which has, so far, received limited attention in the literature. Our work goes on to develop a methodology for placing the PMUs at the locations that will yield the most accurate results while also minimizing the number of PMUs required to obtain the required accuracy.

3. Dataset and Feature Generation

The dataset used in this paper was generated using MATLAB/Simulink. The IEEE 14 bus system was used as a test system and was fully implemented using Simulink, based on the transmission line, load, and generation information provided in [39]. MATLAB/Simulink is capable of providing accurate transient analysis for voltage and current signals in the system; therefore, it accurately captures the behavior of these parameters when a system disturbance such as a transmission line fault occurs. A PMU is connected to each bus in the system to measure the current and voltage parameters. Each PMU measures the voltage with respect to the reference at the bus to which it is connected and measures the current injected into the network at the same bus. Each PMU provides the following values based on the three-phase voltage and current measurements at each bus and at a sampling rate of 64 samples per cycle:

(i) The magnitude of the positive sequence voltage

(ii) The phase angle of the positive sequence voltage

(iii) The frequency of the positive sequence voltage

(iv) The magnitude of the zero sequence voltage

(v) The magnitude of the positive sequence current

(vi) The phase angle of the positive sequence current

(vii) The frequency of the positive sequence current

(viii) The magnitude of the zero sequence current

For example, Figure 1 shows the four voltage-related signals that the PMU provides for bus 3 when simulating a three-phase symmetrical fault on Line 2-3. Note that unsymmetrical faults produce currents and voltages that can be decomposed into their symmetrical components: the positive, negative, and zero sequence components. As such, including both the positive and zero sequence components in the analysis is useful in differentiating between symmetrical and unsymmetrical faults, in addition to distinguishing between the different types of unsymmetrical faults.

For the purposes of developing the dataset, several fault scenarios were simulated. Ten different fault types were simulated, as shown in Table 1, and 15 different transmission lines were considered, as shown in Table 2. This includes all transmission lines in the IEEE 14 bus system. Furthermore, the faults were placed at varying distances on the line itself in 5% increments of the total line length, starting from the bus of the lower number towards the bus of the higher number. The total number of simulated faults was, therefore, 2850 faults, each representing one unique data point in the dataset.
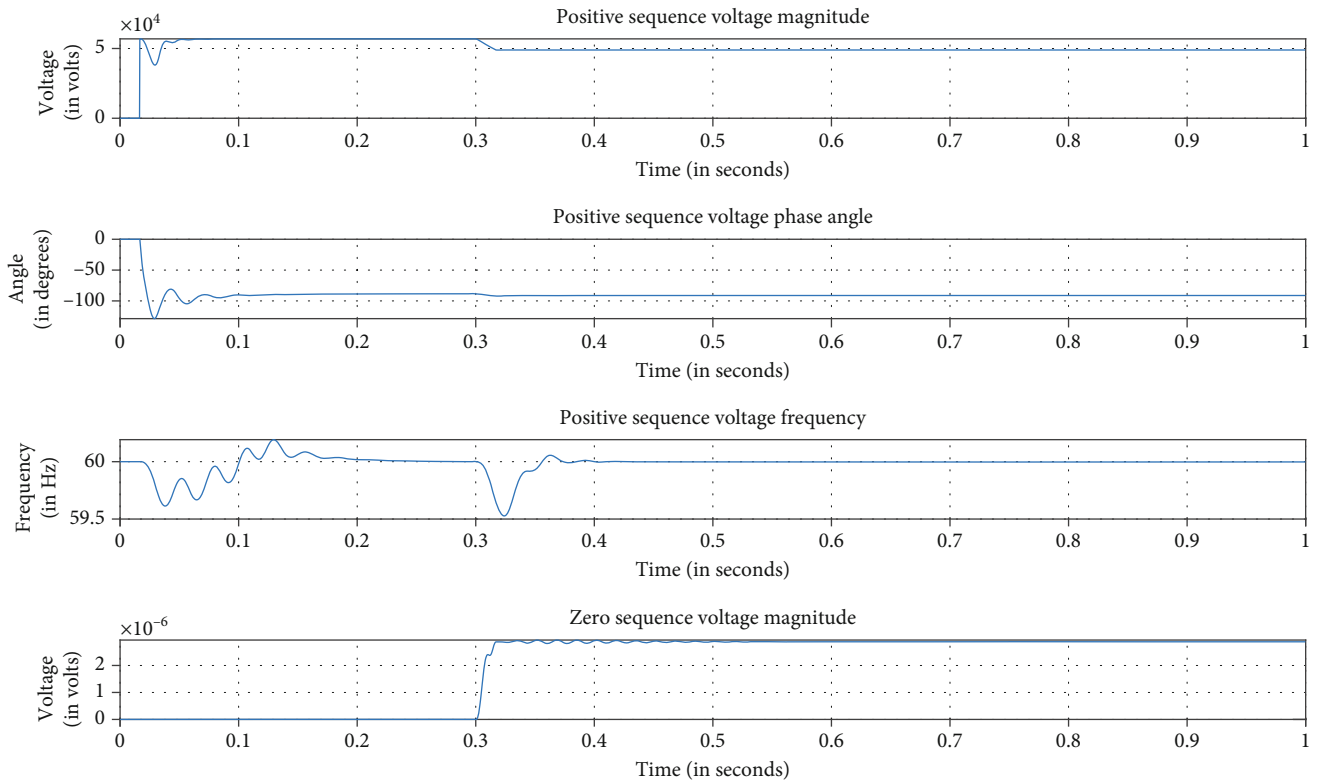
FIGURE 1: Screenshot for the PMU signals for voltage on bus 3.

TABLE 1: Summary of fault type simulated.

Fault type number	Fault type
1	Three-phase-to-ground ABC-G
2	Single-line-to-ground (A-G)
3	Single-line-to-ground (B-G)
4	Single-line-to-ground (C-G)
5	Line-to-line (AB)
6	Line-to-line (AC)
7	Line-to-line (BC)
8	Double-line-to-ground (AB-G)
9	Double-line-to-ground (AC-G)
10	Double-line-to-ground (BC-G)

TABLE 2: Summary of all possible fault lines in the IEEE 14 bus system.

Fault line number	Fault line
1	Line 1-2
2	Line 1-5
3	Line 2-3
4	Line 2-4
5	Line 2-5
6	Line 3-4
7	Line 4-5
8	Line 6-11
9	Line 6-12
10	Line 6-13
11	Line 9-10
12	Line 9-14
13	Line 10-11
14	Line 12-13
15	Line 13-14

For each fault type, line, and distance, a number of features were calculated using the eight PMU signals from all PMUs in the network. More specifically, the following features were directly taken from the PMU measurements:

(1) The voltage's positive sequence magnitude

(2) The voltage's positive sequence phase angle

(3) The voltage's zero sequence magnitude

(4) The current's positive sequence magnitude

(5) The current's positive sequence phase angle

(6) The current's zero sequence magnitude

These six values were obtained by averaging the first 100 signal samples right after the fault occurrence.

In addition, the frequency signals of the voltage and current right after the fault changes dynamically are used, and a
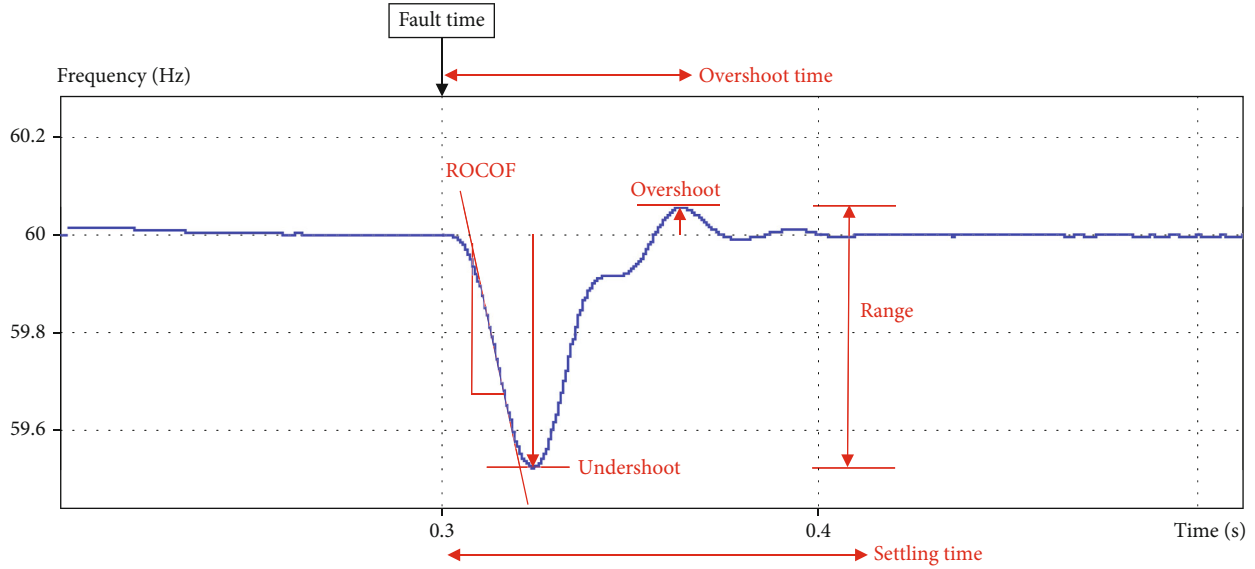
FIGURE 2: Description of frequency-related features.

number of features were extracted from each frequency signal that best describes the faults as follows:

(1) Frequency overshoot: the highest value the frequency reaches after the occurrence of a disturbance

(2) Frequency undershoot: the lowest value the frequency reaches after the occurrence of a disturbance

(3) Frequency overshoot time: the time from the moment fault occurs until the moment frequency reaches its maximum value

(4) Frequency settling time: the time from when the disturbance occurs until the frequency settles back to within less than 10% of the overshoot value

(5) Rate of change of frequency (ROCOF): the slope of the frequency directly after the disturbance occurs

(6) Frequency range: the range of the frequency change after the disturbance, which can be calculated as the difference between the overshoot and the undershoot values

Figure 2 depicts the features obtained from the analysis of each frequency signal. Note that the frequency signal has the ability to provide visual insight into the system dynamics when the fault occurs. The sudden change in the system configuration caused by the fault will lead to visible transients in the rotor dynamics of the generators in the system, which will lead to oscillatory changes in the frequencies of the voltage and current signals in the network.

Overall, this results in nine different features for every current signal and nine different features for every voltage signal. Therefore, each PMU in the network provides a total of 18 signals. Having connected a PMU to each bus, and given there are 14 buses in the system, there are a total of

252 features available for the classification of the fault. These 252 features will be used to classify the faults in the system in terms of their type, line, and distance on the line. A schematic diagram of the IEEE 14 bus system is shown in Figure 3 for reference.

The sampling rate of the PMUs used in this simulation is set at its highest value of 64 samples/cycle, as defined by the IEEE Standard for Synchrophasor Measurements for Power Systems [40]. This is a relatively high sampling rate and largely distinguishes PMU measurements from regular SCADA measurement tools. As such, the features obtained in this analysis are highly accurate. To demonstrate this accuracy, a similar analysis was conducted by reducing the sampling rate of the PMUs and observing the relative variation in the values obtained for the desired features.

Tables 3 and 4 show the variations (or errors) in the values of the features obtained by reducing the sampling rate of the PMU in factors of 2 starting at 32 samples/cycle down to 1 sample/cycle. Most of the results show very little variation in the obtained values even at the lowest sampling rate of 1 sample/cycle; therefore, we are only showing the results of simulation for the voltage magnitude at bus 1, considering all 10 types of faults, and assuming the fault occurs in the middle of the line.

From Table 3, we can note that there is a slight increase in the error as the sampling rate is reduced, which is expected. However, a small deviation of about 1 volt or less is very insignificant, given the typical voltage rating of several kilovolts. In that respect, the small increase in error is considered insignificant.

The more serious issue is in the reading of the ROCOF values displayed in Table 4, as it can be noted that there is a sharp increase in the error as the sampling rate is reduced. This is not surprising since the ROCOF rapidly changes during a fault and is considered an important indicator of the occurrence of the fault. Note that at high sampling rates
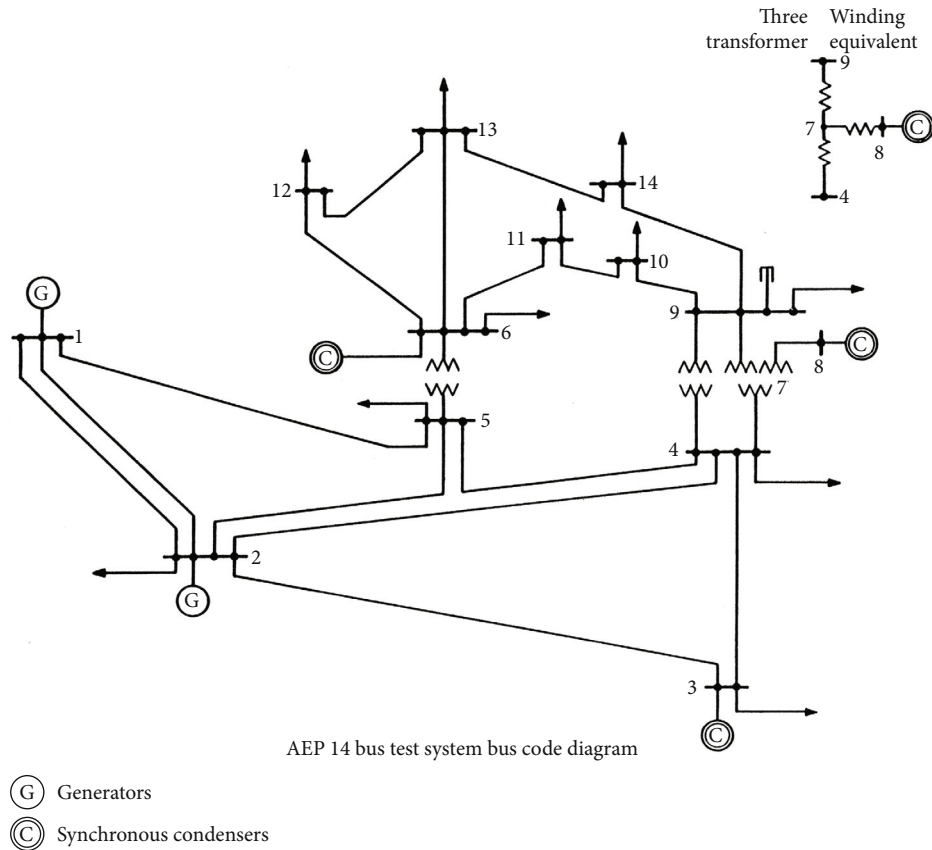
AEP 14 bus test system bus code diagram

(G) Generators

(C) Synchronous condensers

FIGURE 3: The IEEE 14 bus system (reprinted from [39]).

TABLE 3: Sampling error for voltage magnitude at bus 1.

V1 mag (V) Fault type	Samples/cycle Fault location	64	32	16	8	4	2	1
					Error (V)			
1		0	0.001	0.008	0.008	0.014	0.019	0.024
2		0	0.001	0.007	0.029	0.006	0.194	0.435
3		0	0.002	0.000	0.035	0.068	0.228	0.260
4		0	0.002	0.003	0.010	0.066	0.428	0.677
5	1	0	0.003	0.008	0.012	0.097	0.644	1.012
6		0	0.004	0.005	0.058	0.107	0.338	0.350
7		0	0.000	0.006	0.038	0.004	0.317	0.649
8		0	0.002	0.007	0.007	0.062	0.433	0.675
9		0	0.003	0.006	0.042	0.075	0.220	0.219
10		0	0.000	0.002	0.022	0.005	0.239	0.463

(16-64 samples/cycle), the error is relatively low. However, as the sampling rate values become lower, the error rises sharply. At a sampling rate of 2 or 1 sample/cycle, many of the ROCOF values cannot be properly read, thus emphasizing the importance of using a high sampling rate. A similar analysis is performed on all PMUs for both voltage and current readings, and given the low variation in the results at sampling rates between 16 and 64 samples/cycle, we can reasonably conclude that a sampling rate of 64 samples/cycle is sufficient for our purposes.

Moreover, a sampling rate of 64 samples/cycle is also sufficient for allowing timely protection for the system under fault. As mentioned in [41], the critical clearing times for a three-phase symmetrical fault is about 200-300 ms, which represents the worst-case scenario as three-phase faults are typically the most severe and require the fastest action for protection equipment. Since PMUs obtain 64 samples per cycle, it is sufficient to take the measurements from the first 2-3 cycles immediately after the occurrence of a fault to determine the type and

Table 4: Sampling error for ROCOF in voltage at bus 1.

V1 ROCOF Fault type	Samples/cycle Fault location	64	32	16	8	4	2	1
					Error (Hz/s)			
1		0	0.166	0.565	0.906	0.122	0.410	22.747
2		0	0.276	0.354	0.949	0.977	—	13.985
3		0	0.000	0.081	0.056	0.031	5.857	—
4		0	0.039	0.023	0.218	0.799	0.352	—
5	1	0	0.345	1.396	3.170	5.835	—	—
6		0	0.135	0.369	0.847	0.373	0.055	—
7		0	0.069	0.069	0.230	0.925	5.972	—
8		0	0.120	0.120	1.639	1.417	1.532	21.174
9		0	0.101	0.364	0.448	0.564	0.627	18.719
10		0	0.000	0.409	0.470	0.532	6.479	—

Table 5: List of features selected per output.

Fault type		Fault line		Fault distance	
I6_F_Sett_Time	10	I12_F_OS_Time	10	I12_Mag	10
V14_F_Sett_Time	10	I8_F_Sett_Time	10	I6_Mag	10
I6_F_OS_Time	8	V5_F_Sett_Time	10	I3_Mag	10
V2_F_Sett_Time	8	V3_F_OS_Time	10	I2_ROCOF	10
V10_F_Sett_Time	7	V3_F_Sett_Time	10	I2_Mag	10
V3_F_Sett_Time	6	V2_zero	10	I1_Mag	10
V1_F_Sett_Time	5	I1_F_OS_Time	10	I12_ROCOF	9
I8_F_Range	2	V1_F_OS_Time	10	I6_zero	9
I13_ROCOF	1	I9_F_Overshoot	9	I14_Mag	3
I9_F_Sett_Time	1	V8_F_OS_Time	9	I14_zero	1
		I5_F_OS_Time	9	I5_zero	1
		I10_F_Overshoot	1		
		V6_F_OS_Time	1		
		V5_F_OS_Time	1		

location of the fault, which would require no more than 50 ms, allowing sufficient time for the protection equipment to take the required breaking action.

4. Methodology

4.1. Feature Selection. As mentioned earlier in Section 3, 18 features are extracted from every PMU. In a system with 14 PMUs installed, this adds up to 252 features. This section focuses on extracting the most important features for each of the three outputs the system is trying to predict, which will be designated henceforth as the fault type, fault line, and fault distance.

For feature selection, the popular WEKA tool has been used [42]. WEKA is a data mining tool that can be used for multiple purposes; feature selection is one of them. The tool employs multiple algorithms and techniques for feature selection. The method used was the correlation-based feature subset selection (*CfsSubsetEval*). This method evaluates how good a set of features is by considering their

predictive ability as well as the degree of redundancy between the features [43]. The algorithm's ideal output is a set of highly correlated features—with the output/target—but with a low correlation among these features. In order to have better generalization ability, the tool also supports cross-validation when performing the feature selection.

The feature selection was performed three times independently, once for each output. Each time, 10-fold cross-validation was performed. The results of the feature selection runs are shown in Table 5. One can see the top features selected per each output and how many times each feature was selected during the 10-fold cross-validation runs. For example, when performing the feature selection for the fault type, the feature *I6_F_Sett_Time* (which represents the settling time of the current signal's frequency at bus 6) was selected ten times. This shows that this feature is consistently important and must be factored in the fault type classification. Based on the table, 10, 14, and 11 features were selected for the three outputs.

TABLE 6: Aggregated feature selection per PMU.

PMU	Sum of features selected
PMU_2	38
PMU_6	38
PMU_3	36
PMU_1	35
PMU_12	29
PMU_5	21
PMU_8	21
PMU_14	14
PMU_9	10
PMU_10	8
PMU_13	1

TABLE 7: Summary of all experiments.

Experiment	Number of PMUs	PMUs
0	All	All
1-4	1	{1}, {2}, {3}, {6}
5-10	2	{1,2}, {1,3}, {1,6}, {2,3}, {2,6}, {3,6}
11-14	3	{1,2,3}, {1,2,6}, {1,3,6}, {2,3,6}
15	4	{1,2,3,6}
16	5	{1,2,3,6,12}
17	6	{1,2,3,6,12,5}
18	7	{1,2,3,6,12,5,8}
19	8	{1,2,3,6,12,5,8,14}
20	9	{1,2,3,6,12,5,8,14,9}
21	10	{1,2,3,6,12,5,8,14,9,10}
22	11	{1,2,3,6,12,5,8,14,9,10,13}

In order to consider the practical aspect of this problem, if one has a PMU at a certain location, then one can fetch all the features available from that PMU. To be able to proceed with the analysis, the list of PMUs that generate the features in Table 5 are extracted and aggregated across all of the three outputs. The results are shown in Table 6, which clearly indicates that PMUs {2, 6, 3, and 1} are the most important PMUs. This list will be used as a priority list in devising the test cases and prioritizing what PMUs should be selected. It is worth noting that this analysis can be recomputed for any bus configuration.

4.2. Experimental Setup and Test Scenarios. As outlined in Section 4.1 and presented in Table 5, PMUs {1,2,3,6} were identified as the most critical PMUs within the power system under investigation. To construct the experiments for this study, various combinations of these PMUs were utilized. The details of the experiments and corresponding test scenarios can be found in Table 7.

The base case experiment (Experiment 0) involved the utilization of all 14 PMUs within the system. Subsequently, a series of experiments were conducted with individual PMUs out of the four critical PMUs {1,2,3,6} (Experiments 1-4) and combinations of two PMUs (Experiments 5-10). Additionally, four different experiments were carried out with three PMUs (Experiments 11-14).

Furthermore, an experiment was conducted involving all four PMUs (Experiment 15), followed by the incremental addition of PMUs {12, 5, 8, 14, 9, 10, 13} one at a time, resulting in seven additional experiments (Experiments 16-22).

In total, 23 experiments were executed for each label based on the selected features, as discussed in Section 4.1. All experiments were implemented using Python 3.0 and executed on Google Colab, a cloud computing service provided by Google [44].

For each experiment, all features of the included PMUs in that experiment will be fed to the machine learning model. For example, for Experiment 5 in Table 7, all features from PMU 1 and PMU 2 will be fetched from the dataset, which gives a list of 36 features (18 features per PMU) as discussed earlier in Section 3.

4.3. ML Algorithms and Metrics

4.3.1. ML Algorithms. Different machine learning algorithms were used to classify and localize faults that occur in the power grid. A linear support vector machine (SVM) classifier was used to classify the fault type and location, while a random forest regressor was used to predict the fault distance. SVM is one of the popular supervised ML algorithms that are used primarily for classification problems. In this work, linear SVM is used to classify the type of fault that occurred in the power grid as listed in Table 1 and to predict the line where the fault was found. Random forest regression is a supervised machine learning algorithm that uses ensemble learning and builds multiple decision trees while training. Random forest regression is used in this work to predict the distance where the fault occurred as measured starting from the bus of the lower number towards the bus with the higher number between which the fault line is connected.

These algorithms were optimized by conducting hyperparameter selection using the grid search technique. Grid search is used to exhaustively search from a list of predefined values for the hyperparameters to optimize these values. Below is a list of the hyperparameters used in the grid search for the classification and regression algorithms:

(i) max_iter for the linear SVM classifier: this value presents the maximum number of iterations needed for the algorithm to converge. The default value in the algorithm is 100, but after conducting a grid search over a list of values, the value of 1200000 was found to be the optimum and is sufficient for the model to converge

(ii) n_estimators for the random forest regressor: this value presents the number of trees in the random forest. The default value in the algorithm is 100, but with a grid search, it was found that setting this value to 50 gives better results

In addition, stratification and cross-validation were used to split the data between training and testing. The data was

split into ten stratified folds, preserving the percentage of data samples in each class type.

4.3.2. Evaluation Metrics.

4.3.2. Evaluation Metrics. Different metrics were used to evaluate the performance of the classification and regression algorithms. For the classification algorithm, these metrics are accuracy, precision, recall, and $F1$-score. The confusion matrix was also reported.

The accuracy was calculated by dividing the number of correctly predicted values by the total number of values. The correctly predicted values equal the sum of True Positives (TP) and True Negatives (TN), while the total number of values equals the sum of True Positives, True Negatives, False Positives (FP), and False Negatives (FN), as shown in the following:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \quad (1)$$

The precision was calculated by dividing the number of True Positives by the sum of the number of True Positives and False Positives, as shown in the following:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (2)$$

The recall was calculated by dividing the number of True Positives by the sum of True Positives and False Negatives, as shown in the following:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (3)$$

Finally, the $F1$-score was calculated using the values of precision and recall, as shown in the following:

$$F1 \text{ score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}}. \quad (4)$$

For the regression algorithm, the metrics used are R-squared score (R^2) and mean absolute error (MAE).

R^2, also known as the coefficient of determination, is used to evaluate how well a regression model fits a given dataset. R^2 usually ranges from 0 to 1, where 0 represents a poor fit, and 1 indicates an excellent fit of the provided dataset. Equation (5) shows the formula for calculating R^2.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(\text{actual values}_i - \text{predicted values}_i)^2}{\sum_{i=1}^{N}(\text{actual values}_i - \text{average of actual values})^2}. \quad (5)$$

MAE is defined as the average of the differences between actual and predicted values, as shown in Equation (6). In MAE, the absolute value is used so the negative errors are properly accounted, which makes MAE less sensitive to the outliers in the dataset.

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\text{actual values}_i - \text{predicted values}_i|. \quad (6)$$

In summary, Figure 4 presents the procedure outlined in Sections 3 and 4 by visually depicting the relationship between the three main procedures used in this paper, that is, the "Raw Data Preparation" outlined in Section 3, the "Optimal Feature Selection" addressed in Section 4.1, and the "ML Models Training and Evaluation" thoroughly explained in Sections 4.2 and 4.3.

As shown in the flowchart, this work combines the physical aspects of power system fault simulation with the prediction power of ML algorithms, through the optimal feature selection analysis to provide the optimal PMU list. This list would include the number and location of each PMU in the network, which would provide the optimal set of PMUs for correctly classifying and localizing the different fault types in the system.

This procedure can be easily extended and applied to any other similar power system by running the relevant line fault simulations and by using a similar procedure for feature selection and ML model training and evaluation. The following section presents the results obtained in this research.

5. Results and Analysis

5.1. Fault Type. The first set of results is concerned with predicting the fault type, which is a ten-class classification problem. Figure 5 shows both the average testing and training accuracy when running all the experiments reported in Table 7.

It can be seen that the accuracy starts around 87% for a single PMU and grows rapidly to around 94% with only two PMUs. However, it plateaus to above 99% after using 5 PMUs, where the return on investment for any additional PMUs will be limited. As expected, it can also be noticed that the testing accuracy is usually lower than the training accuracy, especially around the knee of the curve.

The detailed results of using only two PMUs are shown in Figure 6(a). While using any two of the four selected PMUs would yield an accuracy above 90%, it is clear that the PMU pair {1,6} produced the highest testing accuracy of 96.2%, while the pair {2,3} had the lowest accuracy of 92.0%. A similar analysis using three PMUs is demonstrated in Figure 6(b). The testing accuracy differences here are not as clear as in the two PMUs case; however, all of the four combinations' testing accuracy is now above 96.5%.

Since 10-fold cross-validation was used to produce the results, the standard deviation of the accuracy among these 10-fold is shown in Figure 7. It is clear that the training standard deviation starts with a high value of around 3.7% for a single PMU and drops significantly once four PMUs are used. A similar pattern appears for the testing accuracy. However, a residual value of around 0.5% does not diminish with increasing the number of PMUs.
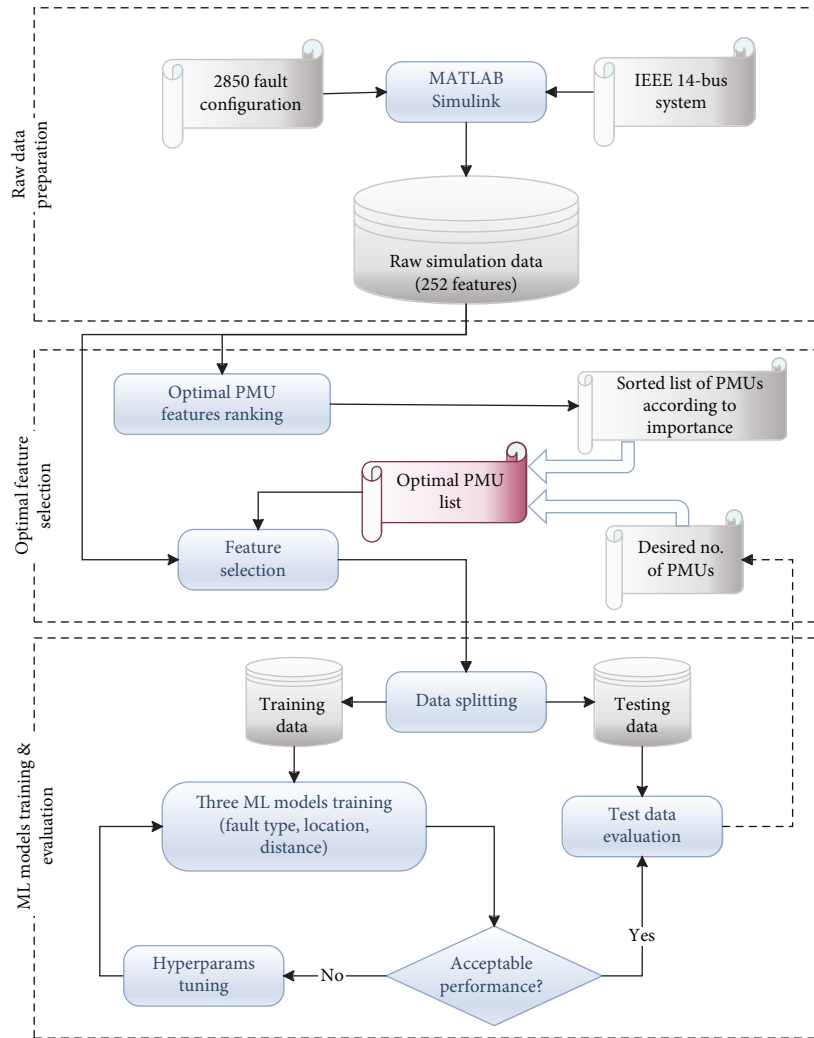
FIGURE 4: Overall technique block diagram.

Figure 8 compares other metrics such as precision, recall, and the F1-score for different fault types. A 3-phase to ground fault (ABC-G) is more likely to be detected correctly than other fault types. Meanwhile, the double-line-to-ground (AC-G) fault is the most difficult, with all its detection performance metrics around 93.5%. One can observe that, in general, for all fault types, all three metrics, precision, recall, and F1-score, are aligned. This can be attributed to the balanced number of test cases and the accuracy of the classifier, as well as the use of cross-validation.

Considering the results reported in Figure 5, it can be seen that using more than five PMUs does not significantly improve accuracy. Thus, an attempt to study the details of the case where five PMUs {1,2,3,6,12} are used is shown in Figure 9. The figure shows that most of the faults are detected correctly. However, it can be seen that there is an issue in the (A-G) and (B-G) faults, where 1.1% and 1.4% of the faults are being mixed up between these two types, respectively. Another significant problem is in the (BC-G) faults, where 1.1% and 0.7% are misclassified as (AB-G) and (AC-G).
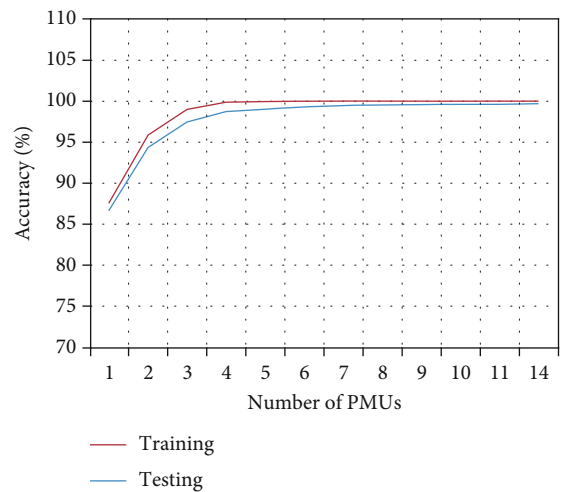
FIGURE 5: Fault type accuracy vs. the number of PMUs.

(a) Two PMUs
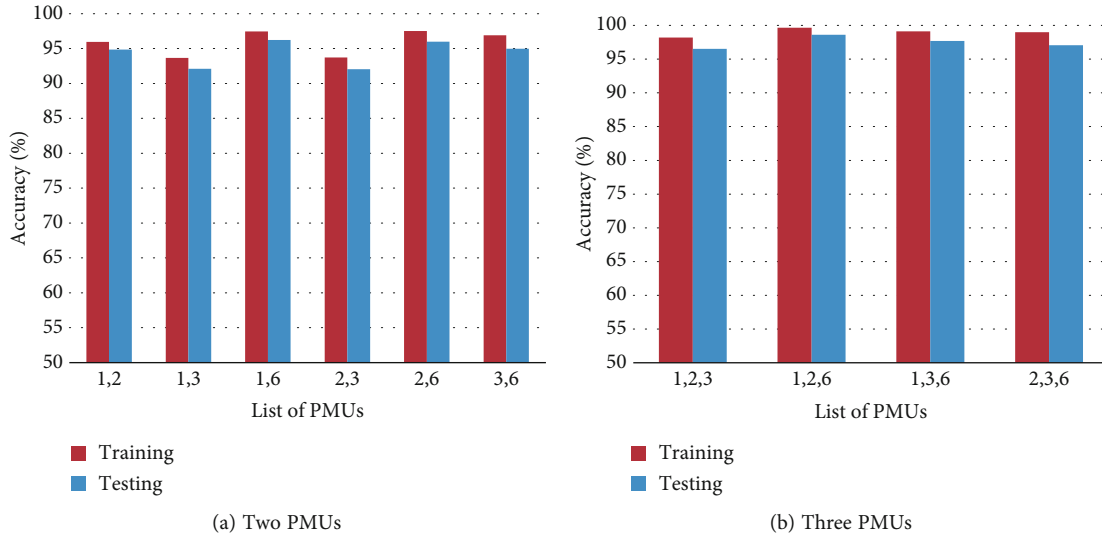
(b) Three PMUs

FIGURE 6: Fault type accuracy using only two or three PMUs.
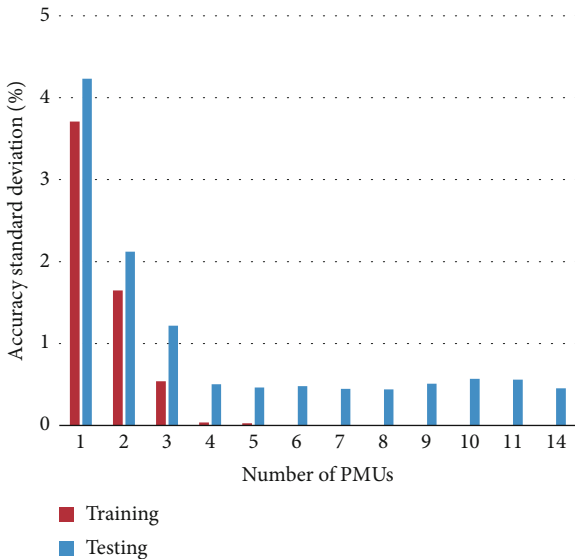
FIGURE 7: Fault type standard deviation in accuracy vs. the number of PMUs.

FIGURE 8: Fault type performance metrics vs. fault type.

5.2. Fault Line. A similar analysis to the results presented in the fault type section is discussed here for the fault line classifier. Figure 10 shows how the accuracy of the fault line classification improves as the number of PMUs increases. It is clear that the accuracy with a low number of PMUs is lower than the accuracy in the fault type classification. This is expected as the number of output classes (lines in this case) is fifteen, while it is ten for the fault type. Again, it is clear that having five PMUs guarantees roughly the highest possible accuracy with the least number of PMUs.

To compare the accuracy of multiple combinations of the cases where two and three PMUs are used, Figures 11(a) and 11(b) are presented. It can be seen that the use of PMUs {3,6} yields the highest accuracy of 83.1%. Note that bus 6 is central among the buses of the IEEE 14 bus system (see

Figure 3), with many lines connected to it. In addition, buses 3 and 6 have a synchronous condenser connected to them, dispatching reactive power immediately as a result of any disturbances, resulting in many of the features changing due to a fault. This makes the two PMUs connected to buses 3 and 6 among the most valuable ones in the system.

When considering the three PMU test cases, in Figure 11(b), the numbers generally improve, and a testing accuracy of 87.3% is achieved using PMUs {2,3,6}, confirming the status of the PMUs at buses 2, 3, and 6 as the most valuable ones, as presented earlier in Table 6.

Figure 12 shows the standard deviation of accuracy among the 10-fold runs for the fault line classification. The
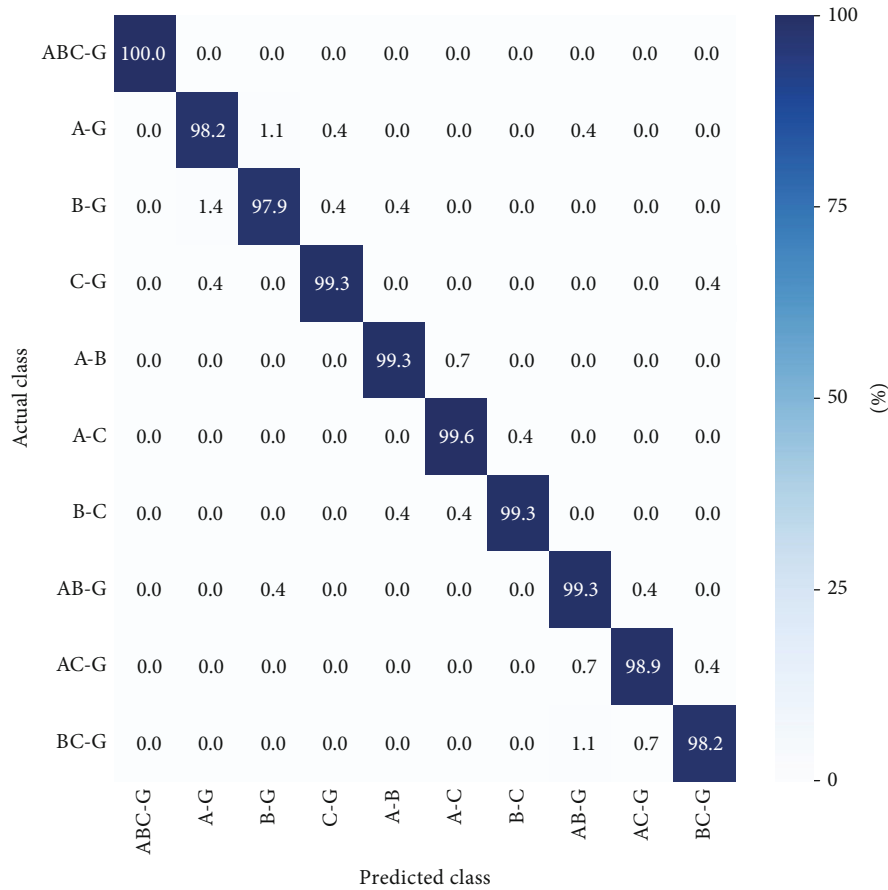
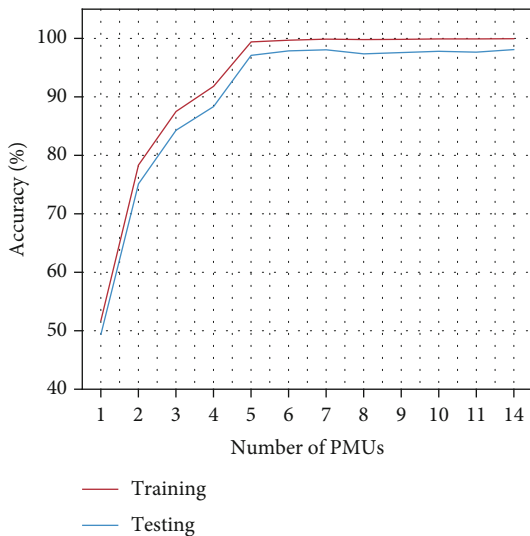FIGURE 9: Fault type confusion matrix (no. of PMUs = 5).

FIGURE 10: Fault line accuracy vs. number of PMUs.

results follow a similar pattern to that achieved with the fault-type classifier. However, the values in the case are slightly higher, meaning there is a higher variation in the fault line accuracy than in the fault type.

To study the variation of the performance metrics across different fault lines, Figure 13 is introduced. It can be seen

that the precision, recall, and $F1$-score are very high for Line 1-2, which connects the two generators in the system. Looking back at the IEEE 14 bus configuration, one can note that Line 1-2 is the line that carries the most amount of power under normal operating conditions. This makes any fault on this line easily detectable, as that would lead to a major change in voltages and currents in the system when this fault occurs, as a big redistribution of power flow in the system will be needed to resume normal operation. Another observation is the low recall value for faults on Lines 9-14 and 6-11. This can also be explained by the fact that these are low-power carrying lines, in addition to being away from generators and the presence of multiple alternative paths to the power flowing through them, making the identification of the fault line more difficult.

Finally, the confusion matrix of the five PMU cases is also introduced for the fault line classifier in Figure 14. Some of the major issues apparent in this confusion matrix are the 4% on Lines 6-11 and 9-10, which are being incorrectly classified as faults on the Line 10-11. This is expected since all three segments, Line 10-11, Line 6-11, and Line 9-10, are connected in series, with no branching on either line, which makes faults on these three lines have a similar effect on the system thus resulting in a few erroneously detected faults. Nonetheless, a 4% error is a very small one considering the
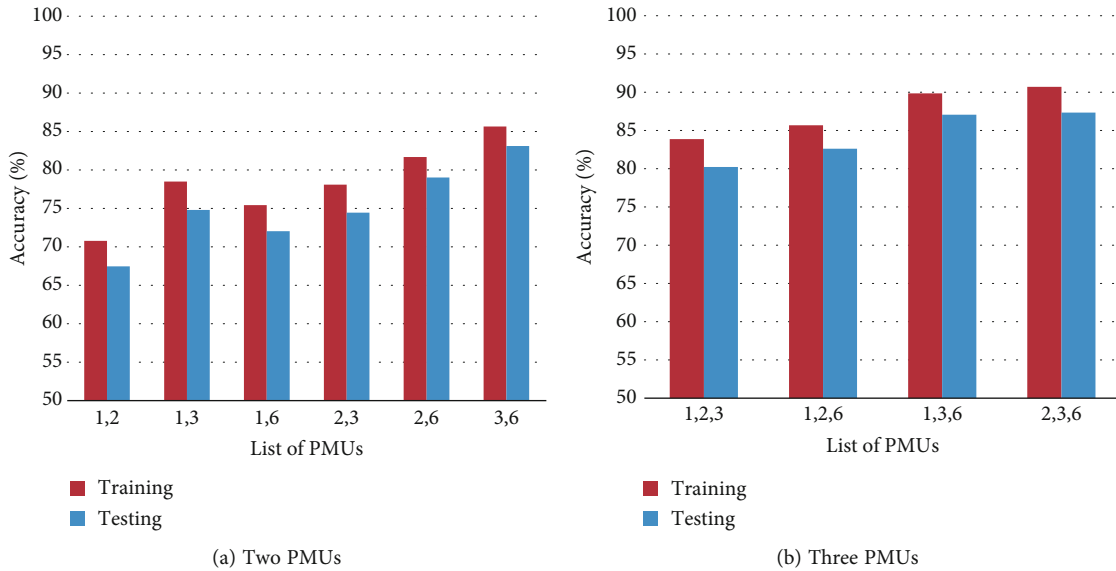
(a) Two PMUs

(b) Three PMUs

FIGURE 11: Fault line accuracy using only two or three PMUs.
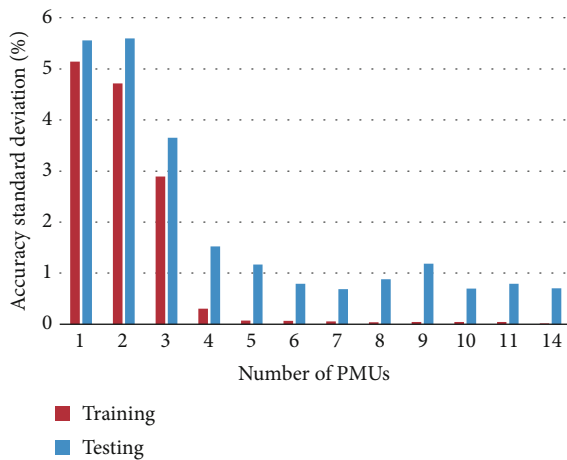
FIGURE 12: Fault line standard deviation in accuracy vs. the number of PMUs.

FIGURE 13: Fault line performance metrics vs. fault line.

similarity between the lines. Another issue to point out is the Line 2-4 faults, which are classified as either 2-3, 2-5, 4-5, or 9-10. Again, the close proximity of these lines means that sometimes, a fault on one of these lines will affect the system in a similar manner to faults on the surrounding lines. With only 5 PMUs used, however, these small errors in classification are acceptable.

5.3. Fault Distance. The last output of this classification is the fault distance, as measured on the transmission line starting from the bus with the lowest number. Recall that the data points used in the training and testing were simulated at distances of 5 percentage points of the total length along the line.

Therefore, since the fault distance is a numeric value, the metrics used to assess the classifier performance will differ. Figure 15 shows both the R^2 score and the MAE as the num-
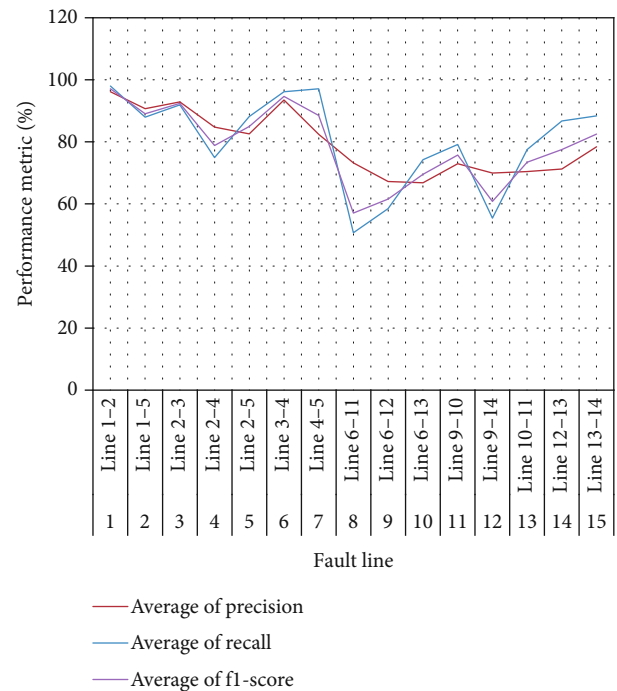
ber of PMUs increases. The figure shows a gradual drop in the testing MAE as the number of PMUs increased. However, this pattern stops once five PMUs are used, and the MAE starts fluctuating around an error of three percentage points. It can also be seen that the training MAE is consistently lower than the testing MAE. These results are expected as a 10-fold evaluation is used. Similarly, the R^2 score increases as the number of PMUs is increased. However, the knee of the curve is not as obvious in the training dataset as it is in the testing dataset.
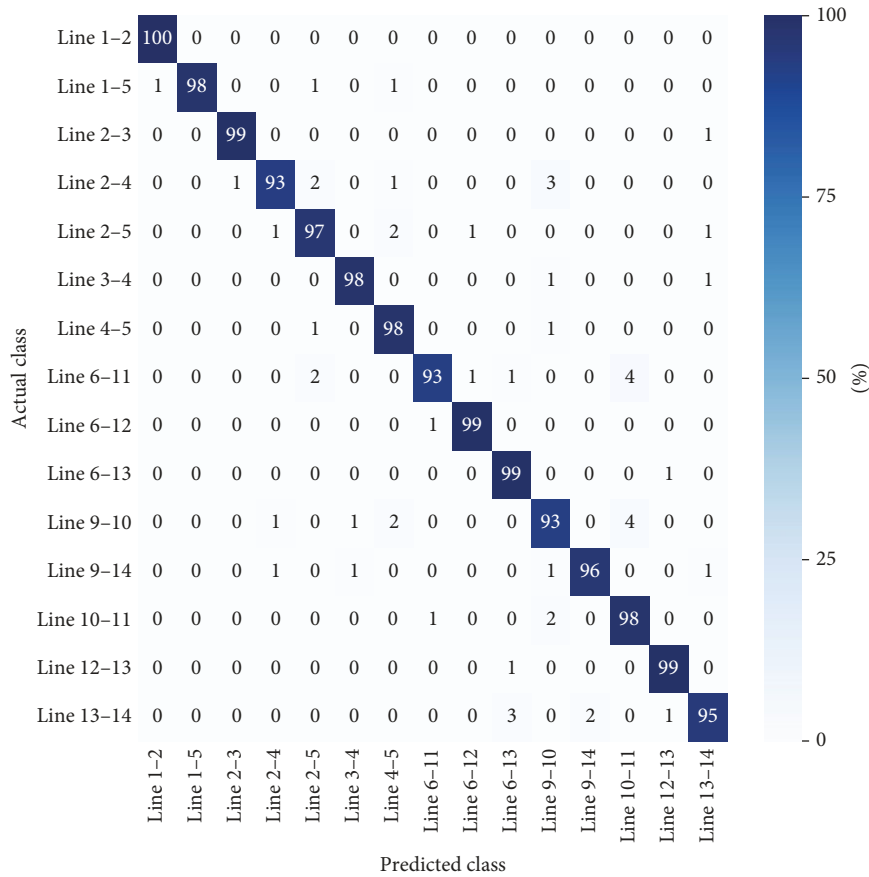
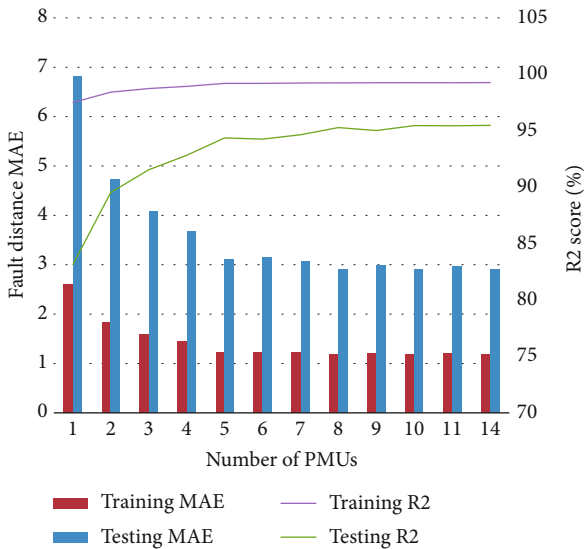FIGURE 14: Fault line confusion matrix (no. of PMUs = 5).

FIGURE 15: Fault distance error vs. number of PMUs.

One issue to remember when considering the fault distance MAE is that the data simulated was discrete in 5% intervals. This means there is up to 2.5% error resulting from discretization only. The average error due to using discrete values would be 1.25%, which would still exist even if an ideal classifier existed.

Figure 16 shows the MAE and R^2 score when only two or three PMUs are used. Figure 16(a) shows the MAE and the R^2 score when only two PMUs are used. It is clear that the combinations of {1,3} and {2,6} produced slightly worse results than the rest. Meanwhile, for the three PMU cases, the first three combinations produced roughly similar results. However, the {2,3,6} combination had slightly worse results. This could be explained by the lack of slack bus (bus 1) features, which can be very useful in determining the exact distance of the fault given its effect on the power, and hence, the current provided by the slack bus itself, and how much it changes when a disturbance or a fault occurs in the system.

Meanwhile, Figure 17 shows the MAE standard deviation for training and testing data. One can notice the sheer difference between the training and testing MAE standard deviation. In addition, one can notice that the standard deviation in testing MAE continues to fluctuate with the increasing number of PMUs. While the standard deviation of MAE tends to decrease when increasing the number of PMUs, a certain amount of error remains and cannot be reduced further as the values of the dataset are discrete. A smaller step in the distance between percentage points would reduce the error; however, the authors believe that a step of 5 percentage points is sufficient to obtain reasonable error results. However, this can be changed in practice to accommodate any desired level of accuracy.
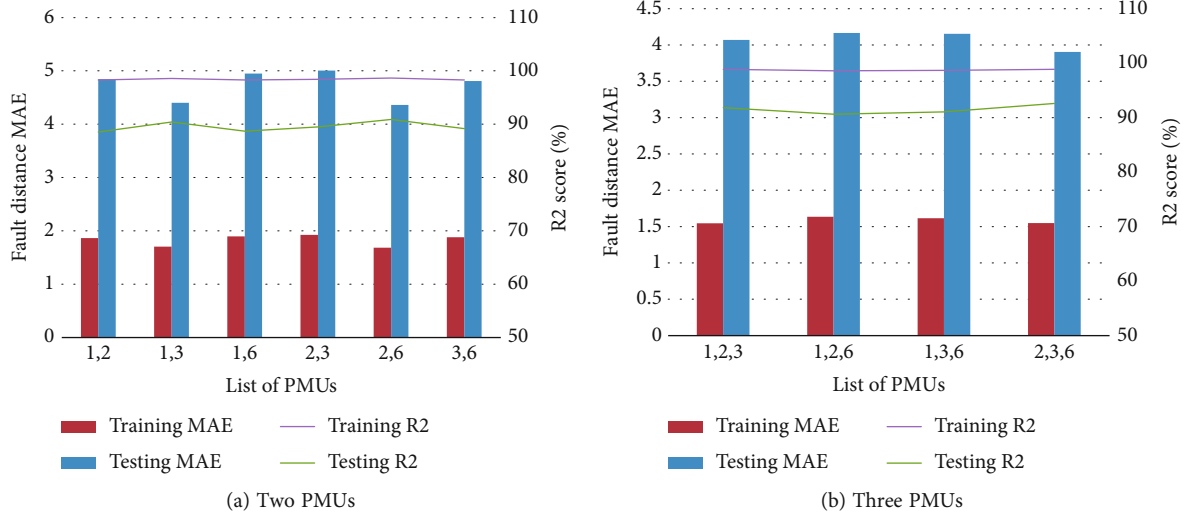
(a) Two PMUs

(b) Three PMUs

FIGURE 16: Fault distance MAE and R^2 score using only two or three PMUs.

In an attempt to further study the impact of the fault distance on the accuracy of the fault, the analysis in Figure 18 was performed for the case where 5 PMUs are used {1, 2, 3, 6, and 12}. Figure 18(a) shows how the MAE changes with the fault distance. The data in this graph is aggregated across all fault lines and across all fault types. It can be clearly seen that the error on the sides of the line is higher than on the center. Faults closer to the bus can lead to higher currents than faults in the center of the line, due to the lower impedance the fault current has to go through. This, in turn, can lead to higher variation in the measurements obtained by the PMU and result in higher errors in determining the distance of the fault. If one considers a single fault type at a time but aggregates it for all lines, similar results can be obtained. However, when only a single line is considered, the error will vary significantly based on where the line is located within the grid.

Figure 18(b) shows how the fault distance MAE changes with the fault type. It is clear that the 3-phase to GND fault produces the highest error among all fault types. This is expected as the 3-phase to GND fault results in the highest fault current, which in turn produces more variation in measurement. This large variation would translate to a higher error in the fault distance prediction.

Finally, the impact of the line on the fault distance error is shown in Figure 18(c). It can be seen that certain lines suffer from higher MAE in the fault distance than other lines. For example, Line 2-5 have an MAE of about 1.5%, while lines like 9-10 have an MAE of 6%. This discrepancy is more difficult to explain and can vary from one system to another, as the fault current in each line is attributed to multiple factors, including the line impedance and the power flowing in the line before the fault, which can, at times, have conflicting effects on the prediction error.

5.4. Additional Discussion. It is clear that the tools used in this work and the intelligent placement of the PMUs based on the results of the machine learning algorithms and the
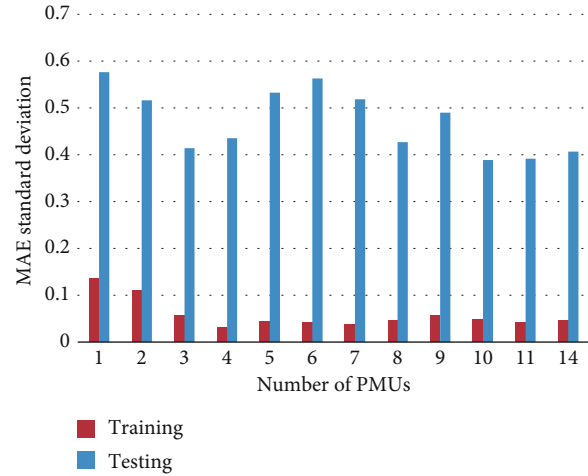

FIGURE 17: MAE standard deviation in the fault distance vs. the number of PMUs.

feature selection were sufficient to correctly classify the type of fault and correctly identify the line and distance at which the fault occurs. It is important here to state the main limitations of this work, which can be addressed in subsequent research efforts.

First, this work was done on the IEEE 14 bus system, and while this system is a typical power system, and many of the conclusions derived can be extended to other systems of different sizes, a similar analysis can be performed on other systems to confirm the applicability of this work on other systems as well. In addition, this analysis only considered fixed loads and faults in the transmission lines. Variations in the load profile and disturbances that do not constitute faults can also be studied to provide further insights into system operation.

Finally, to obtain more accurate results regarding the fault distance, additional more rigorous sampling can be done at shorter distances, which will significantly increase the available data points for training and validation, which

(a) Fault distance MAE vs. actual fault distance

(b) Fault distance MAE vs. fault type
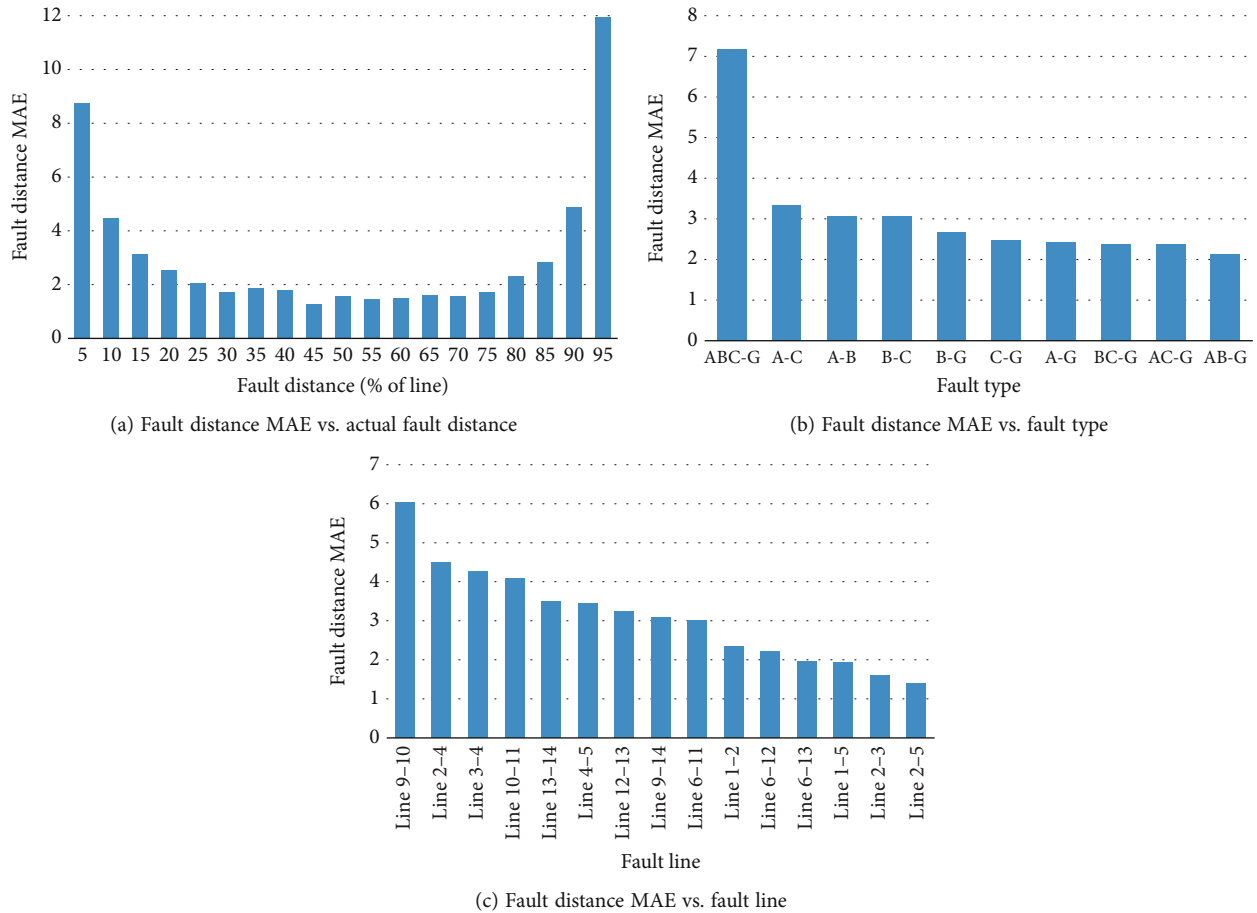

(c) Fault distance MAE vs. fault line

FIGURE 18: Analysis of fault distance MAE.

can lead to more accurate results. The simulation results presented in Figure 18 show variations in the MAE of the fault distance prediction, which can be further studied to obtain more insights about the effect of sampling on the correctness of the result, but for the purposes of this work, the results showed sufficient accuracy. Applying this work to other systems of a different size might require that some of the training and feature generation parameters be modified to better suit their properties.

In addition, while the effect of the sampling rate discussed in Section 3 showed sufficient accuracy of the sampling rate defined by the relevant IEEE standards, future improvements in the PMU technology would allow for the use of higher sampling rates, at which case it can be evaluated whether adding further redundancy in the sampling will have any additional benefits on the current prediction accuracy or not.

6. Conclusion

This paper presented a methodology for determining the optimal number and placement of PMUs to be used for classifying faults in transmission lines. Machine learning algorithms such as the support vector machine and random forest were used to classify the faults in terms of their type, the transmission line on which the fault occurred, and its distance on that line. It

was concluded that it is sufficient for the IEEE 14 bus system to install a PMU on only five buses to achieve high classification accuracy. Adding more PMUs will not add a significant improvement to the classification results.

Future directions for this work include analyzing and classifying other disturbances, which may require additional PMUs or different features. PMU signals can also be used to assess system stability, reliability, and power quality, all of which can be further analyzed and improved using ML techniques. Furthermore, the procedure used in this paper will be used on additional systems of various sizes to confirm the scalability of this method.

Data Availability

Data will be made available upon request.

Conflicts of Interest

The authors of this manuscript declare no conflicts of interest.

References

[1] S. M. Miraftabzadeh, M. Longo, F. Foiadelli, M. Pasetti, and R. Igual, "Advances in the application of machine learning

techniques for power system analytics: a survey," *Energies*, vol. 14, no. 16, p. 4776, 2021.

[2] S. M. Miraftabzadeh, F. Foiadelli, M. Longo, and M. Pasetti, "A survey of machine learning applications for power system analytics," in *2019 IEEE International Conference on Environment and Electrical Engineering and 2019 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*, pp. 1–5, Genova, Italy, June 2019.

[3] A. K. Ozcanli, F. Yaprakdal, and M. Baysal, "Deep learning methods and applications for electrical power systems: a comprehensive review," *International Journal of Energy Research*, vol. 44, no. 9, pp. 7136–7157, 2020.

[4] J. Xie, I. Alvarez-Fernandez, and W. Sun, "A review of machine learning applications in power system resilience," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, Montreal, QC, Canada, August 2020.

[5] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1694–1703, 2019.

[6] Y. Wang, M. Liu, and Z. Bao, "Deep learning neural network for power system fault diagnosis," in *2016 35th Chinese Control Conference (CCC)*, pp. 6678–6683, Chengdu, China, July 2016.

[7] N. Markovic, T. Stoetzel, V. Staudt, and D. Kolossa, "Hybrid fault detection in power systems," in *2019 IEEE International Electric Machines & Drives Conference (IEMDC)*, pp. 911–915, San Diego, CA, USA, May 2019.

[8] N. E. M. Barreto, R. Rodrigues, R. Schumacher, A. R. Aoki, and G. Lambert-Torres, "Artificial neural network approach for fault detection and identification in power systems with wide area measurement systems," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 6, pp. 1617–1626, 2021.

[9] S. A. Aleem, N. Shahid, and I. H. Naqvi, "Methodologies in power systems fault detection and diagnosis," *Energy Systems*, vol. 6, pp. 85–108, 2015.

[10] M. Wadi, "Fault detection in power grids based on improved supervised machine learning binary classification," *Journal of Electrical Engineering*, vol. 72, no. 5, pp. 315–322, 2021.

[11] M. S. Coutinho, L. R. G. S. Lourenço Novo, M. T. de Melo et al., "Machine learning-based system for fault detection on anchor rods of cable-stayed power transmission towers," *Electric Power Systems Research*, vol. 194, article 107106, 2021.

[12] Manojna, H. S. Shridhar, N. Nikhil, A. Kumar, and P. Amrit, "Fault detection and classification in power system using machine learning," in *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, pp. 1801–1806, Trichy, India, October 2021.

[13] S. Belagoune, N. Bali, A. Bakdi, B. Baadji, and K. Atif, "Deep learning through LSTM classification and regression for transmission line fault detection, diagnosis and location in large-scale multi-machine power systems," *Measurement*, vol. 177, article 109330, 2021.

[14] P. Tehrani and M. Levorato, "Frequency-based multi task learning with attention mechanism for fault detection in power systems," in *2020 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 1–6, Tempe, AZ, USA, November 2020.

[15] M.-T. A. M. R. Shadi, M.-T. Ameli, and S. Azad, "A real-time hierarchical framework for fault detection, classification, and location in power systems using PMUs data and deep learning," *International Journal of Electrical Power and Energy Systems*, vol. 134, article 107399, 2022.

[16] S. Zhang, Y. Wang, M. Liu, and Z. Bao, "Data-based line trip fault prediction in power systems using LSTM networks and SVM," *IEEE Access*, vol. 6, pp. 7675–7686, 2018.

[17] F. Zhang, Q. Liu, Y. Liu, N. Tong, S. Chen, and C. Zhang, "Novel fault location method for power systems based on attention mechanism and double structure GRU neural network," *IEEE Access*, vol. 8, pp. 75237–75248, 2020.

[18] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault location in power distribution systems via deep graph convolutional networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 119–131, 2020.

[19] H. Ren, Z. J. Hou, B. Vyakaranam, H. Wang, and P. Etingov, "Power system event classification and localization using a convolutional neural network," *Frontiers in Energy Research*, vol. 8, article 607826, 2020.

[20] B. Sreewirote and A. Ngaopitakkul, "Classification of fault type on loop-configuration transmission system using support vector machine," in *2017 6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)*, pp. 892–896, Hamamatsu, Japan, July 2017.

[21] M. S. Mandava, D. Jadhav, and R. R. Naik, "Fault classification using SVM," in *2015 IEEE International Circuits and Systems Symposium (ICSyS)*, pp. 17–21, Langkawi, Malaysia, September 2015.

[22] R. Pattanayak, S. Behera, and B. Parija, "Classification of faults in a hybrid power system using artificial neural network," in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, pp. 1–4, Bombay, India, March 2019.

[23] R. Resmi, V. Vanitha, E. Aravind, B. R. Sundaram, C. R. Aswin, and S. Harithaa, "Detection, classification and zone location of fault in transmission line using artificial neural network," in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, February 2019.

[24] A. Jain, A. S. Thoke, E. Koley, and R. N. Patel, "Fault classification and fault distance location of double circuit transmission lines for phase to phase faults using only one terminal data," in *2009 International Conference on Power Systems*, pp. 1–6, Kharagpur, India, December 2009.

[25] M. A. M. Nasrin, A. M. S. Omar, S. S. M. Ramli, A. R. Ahmad, N. F. Jamaludin, and M. K. Osman, "Deep learning approach for transmission line fault classification," in *2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pp. 164–169, Penang, Malaysia, August 2021.

[26] X. Jiang and Z. Ge, "Data augmentation classifier for imbalanced fault classification," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 3, pp. 1206–1217, 2021.

[27] N. H. Abbasy and H. M. Ismail, "A unified approach for the optimal PMU location for power system state estimation," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 806–813, 2009.

[28] A. A. Eladl, A. N. Sheta, M. A. Saeed, M. A. Abido, and M. A. Hassan, "Optimal allocation of phasor measurement units in distribution power systems," *Alexandria Engineering Journal*, vol. 61, no. 10, pp. 8039–8049, 2022.

[29] S. A. Taher, H. Mahmoodi, and H. Aghaamouei, "Optimal PMU location in power systems using mica," *Alexandria Engineering Journal*, vol. 55, no. 1, pp. 399–406, 2016.

[30] R. Sodhi, S. Srivastava, and S. N. Singh, "Optimal PMU placement method for complete topological and numerical observability of power system," *Electric Power Systems Research*, vol. 80, no. 9, pp. 1154–1159, 2010.

[31] V. Madani, M. Parashar, J. Giri et al., "PMU placement considerations—a roadmap for optimal PMU placement," in *2011 IEEE/PES Power Systems Conference and Exposition*, pp. 1–7, Phoenix, AZ, USA, March 2011.

[32] W. Li, D. Deka, M. Chertkov, and M. Wang, "Real-time faulted line localization and PMU placement in power systems through convolutional neural networks," *IEEE Transactions on Power Systems*, vol. 34, no. 6, pp. 4640–4651, 2019.

[33] P. Vidya and G. A. Shekhappa, "Optimal placement of PMU for enhancing power system monitoring," *I-Manager's Journal on Power Systems Engineering*, vol. 11, no. 2, pp. 23-24, 2023.

[34] Z. Jin, Y. Hou, Y. Yu, and L. Ding, "Optimal PMU placement in the presence of conventional measurements," *International Transactions on Electrical Energy Systems*, vol. 2022, article 8078010, 11 pages, 2022.

[35] A. Chowdhury, N. Sahu, and S. Shimi, "Optimal PMU placement in 14 bus system using integer linear programming," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, pp. 28–31, 2017.

[36] A. I. Al-Odienat, B. Malahmeh, and A. R. Tarawneh, "The optimal PMU placement in the power systems for the enhancement of state estimation," in *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)*, pp. 1–6, Istanbul, Turkey, June 2020.

[37] M. Baba, N. B. M. Nor, M. A. Sheikh, A. M. Baba, M. Rehman, and S. H. A. Shah, "A hybrid algorithm for optimal PMU placement," in *2020 8th International Conference on Intelligent and Advanced Systems (ICIAS)*, pp. 1–6, Kuching, Malaysia, July 2021.

[38] A. Almunif and L. Fan, "Optimal PMU placement for modeling power grid observability with mathematical programming methods," *International Transactions on Electrical Energy Systems*, vol. 30, no. 2, article e12182, 2020.

[39] "Power systems test case archive," 2023, January 2023, https://labs.ece.uw.edu/pstca/.

[40] "IEEE Standard for Synchrophasor Measurements for Power Systems," in *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1–61, 2011.

[41] A. Janssen, Q. Zhuang, S. McGuinness et al., "Operating experience with and future challenges for coordinating power plant and power system protection," *Cigre Science & Engineering*, vol. 7, p. 63, 2017.

[42] "Weka 3: Machine learning software in java," 2023, June 2023, https://www.cs.waikato.ac.nz/ml/weka/.

[43] M. A. Hall, *Correlation-based feature selection for machine learning, Master's thesis*, The University of Waikato, 1999.

[44] Google, "Welcome to colaboratory!," 2023, March 2022, https://colab.research.google.com/.