energies

MDPI

Article

Fault Detection and Classification in Transmission Lines Connected to Inverter-Based Generators Using Machine Learning

Khalfan Al Kharusi *, Abdelsalam El Haffar and Mostefa Mesbah

Electrical and Computer Engineering, Sultan Qaboos University, P.O. Box 33, Muscat 123, Oman; a.elhaffar@squ.edu.om (A.E.H.); m.mesbah@squ.edu.om (M.M.)

* Correspondence: s49626@student.squ.edu.om

Abstract: Integrating inverter-based generators in power systems introduces several challenges to conventional protection relays. The fault characteristics of these generators depend on the inverters' control strategy, which matters in the detection and classification of the fault. This paper presents a comprehensive machine-learning-based approach for detecting and classifying faults in transmission lines connected to inverter-based generators. A two-layer classification approach was considered: fault detection and fault type classification. The faults were comprised of different types at several line locations and variable fault impedance. The features from instantaneous three-phase current and voltages and calculated swing-center voltage (SCV) were extracted in time, frequency, and time–frequency domains. A photovoltaic (PV) and a Doubly-Fed Induction Generator (DFIG) wind farm plant were the considered renewable resources. The unbalanced data problem was investigated and mitigated using the synthetic minority class oversampling technique (SMOTE). The hyperparameters of the evaluated classifiers, namely decision trees (DT), Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), and Ensemble trees, were optimized using the Bayesian optimization algorithm. The extracted features were reduced using several methods. The classification performance was evaluated in terms of the accuracy, specificity, sensitivity, and precision metrics. The results show that the data balancing improved the specificity of DT, SVM, and k-NN classifiers (DT: from 99.86% for unbalanced data to 100% for balanced data; SVM: from 99.28% for unbalanced data to 99.93% for balanced data; k-NN: from 99.64% for unbalanced data to 99.74% for balanced data). The forward feature selection combined with the Bag ensemble classifier achieved 100% accuracy, sensitivity, specificity, and precision for fault detection (binary classification), while the Adaboost ensemble classifier had the highest accuracy (99.4%), compared to the other classifiers when using the complete set of features. The classification models with the highest performance were further tested using a new dataset test case. They showed high detection and classification capabilities. The proposed approach was compared with the previous methodologies from the literature.

Keywords: machine learning; fault detection; fault classification; inverter-based generators; power system protection; renewable energy; Bayesian optimization

Citation: Al Kharusi, K.; El Haffar, A.; Mesbah, M. Fault Detection and Classification in Transmission Lines Connected to Inverter-Based Generators Using Machine Learning. *Energies* **2022**, *15*, 5475. https://doi.org/10.3390/en15155475

Academic Editors: Nicu Bizon and Abu-Siada Ahmed

Received: 1 June 2022
Accepted: 26 July 2022
Published: 28 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The integration of inverter-based generators in modern power systems produces several challenges in power systems control, protection, operation, planning, and stability. Asymmetrical fault current has typically been represented as having positive, negative, and zero-sequence components in conventional rotating-machine power systems. This was based on the assumption that the positive and negative sequence networks are fully decoupled [1]. However, this is not the case for inverter-based generators (IBGs) that attempt to maintain a balanced current for unbalanced faults.

The fault detection and classification problem in power systems penetrated with IBGs has become a significant challenge for the following reasons:

- Only the positive current sequence is available for symmetrical and asymmetrical faults for fully converted renewable sources, like photovoltaic (PV) and type-4 wind turbines. The absence of a negative sequence current presents a challenge for the operation of protection devices that rely on negative sequence components. This challenge can be mitigated by specifying the requirement of negative sequence injection in the grid code using the decoupled sequence control mode of the inverter [2]. Furthermore, the difference between the phase angles of the negative-sequence voltage and current measured by a relay after an asymmetrical forward fault occurred was lesser for the system integrated with IBGs than the conventional power system with synchronous generators. This comparison was discussed in [3].
- The IBGs do not contribute to zero-sequence components because they are not grounded. In contrast, the coupling transformer grounding can obtain the zero-sequence component, a potent source supplying a high magnitude of zero-sequence current [4]. As a result, the zero-sequence component depends on the inverter, the IBG type, and the transformer connection [1].
- The stiffness of power systems with IBGs is reduced compared with conventional generation systems [5]. System stiffness (strength) can be evaluated by calculating the short circuit ratio (SCR). A power source is described as a weak system in the presence of IBGs (weak source means high SCR) [6]. IBGs have unique short circuit characteristics because of the integration of power electronics connected to the grid. When a short-circuit fault occurs, the inverter is switched to current-controlled mode (CCM), and the inverter behaves as a current source until the short-circuit fault is cleared by the protection devices [7]. Furthermore, the IBG output current increases as the voltage drops during faults to regulate it back to its P–Q setpoint. In this condition, the IBG becomes a current source [8].

Several protection strategies were proposed in the literature to enhance the ability of fault detection, classification, and localization for power systems connected with IBGs.

- *Adaptive protection schemes:* They are defined as the online protection schemes used to adapt relay settings and characteristics according to the system's current state [9]. Different adaptive schemes for microgrids were reviewed in [10]. The adaptive protection scheme depends on the communication infrastructure to exchange information in the form of measured network parameters such as voltage, current, and power. Therefore, the reliability of a viable adaptive protection scheme depends upon the redundancy of the communication system with the cybersecurity hazards [11,12]. Moreover, adaptive protection requires complex algorithms [13], which significantly increases the cost.
- *Modification of fault current level*: Fault contribution by IBGs could be modified by adding auxiliary devices on the IBG side to improve its performance during faults. Examples are crowbar rotor circuit (CRC), superconducting fault current limiter (SFCL), superconducting magnetic energy storage (SMES), and series dynamic braking resistor (SDBR). CRC was used to improve the stability of DFIG during faults and protect the rotor side converter [14]. SFCL aims to improve the low voltage fault ride-through (LV-FRT) capability [15]. SDBR was introduced to improve the LV-FRT capability of large wind turbines and the transient stability of DFIG during faults [16]. SMES stores the energy and handles its transfer caused by DFIG power fluctuation or grid fault to improve the LV-FRT [17]. As realized by the authors in [18], fault current-limiting devices introduced several challenges in the power system that require further analysis, such as interfering with communication lines, finding optimal design parameters, coordinated control design between these devices and other protective devices, feasibility analysis, field tests, and real-time grid operation.
- *Meta-heuristic techniques*: These are search algorithms capable of solving complex optimization problems. They include Genetic Algorithm, Annealing algorithm, Tabu Search, and Local Search algorithm. These techniques are high-level heuristics used to guide others for a better evolution in the search space [19]. Several researchers used the meta-heuristics for protection relay coordination to find the optimum relay

setting according to the system topology. Dynamic and flexible protection approach considering different grid operation modes of microgrids for earth and phase overcurrent coordination using charged system search (CSS) and Teaching-Learning-Based Optimization Algorithm (TLBO) was proposed in [20]. The studied microgrid was connected to a distributed generator without defining the type of generator technology. The inefficient numerical search is always a limitation of these techniques, especially for high-dimensional problems [21].

- *Machine learning techniques*: Many researchers proposed artificial intelligence techniques that utilize machine learning (ML) in power system protection for fault detection, classification, and localization. The implementation of machine learning for power system fault diagnosis was reviewed in [10,22,23]. The ultimate advantages of these techniques are the accuracy, self-adaptiveness, and robustness to parameter variations [24]. Existing ML techniques comprise the following stages: preprocessing, feature extraction, feature reduction, classification, and performance evaluation. For our focus, transmission line fault detection, classification, and localization using machine learning techniques are reviewed in this article.

Fault detection and classification for mutually coupled transmission lines using Discrete Wavelet Transformation (DWT) for three-phase current signals were proposed in [25]. ANN, k-NN, and DT classifiers were used to classify twenty-one classes for phase identification and four for ground faults identification. The accuracy was the classification metric used to evaluate the performance. The best performing classifier was the ANN, with 100% accuracy. The study did not consider the integration of IBGs and did not report the data balancing. With the same feature extraction technique (i.e., discrete wavelet transformation), the authors in [26] used three-phase currents and voltages to detect and classify the transmission line faults using k-NN and DT classifiers. The DT outperformed the k-NN with an accuracy of 100%. That study did not consider the IBGs integration, feature reduction, and data balancing. P. Ray et al. in [24] utilized the wavelet packet transformation to extract the features from three-phase voltages and currents to classify and localize the faults. The dataset consisted of eleven classes: one for non-fault events and ten for different fault types. The data samples were reported to be balanced and reduced using the forward feature selection technique. The accuracy and the absolute error were used to evaluate the SVM classifier. The results showed that the classification accuracy was 99.21%, and the fault localization absolute error was less than 0.21%. The incremental quantity of current signals was calculated as features to detect the faults during power swing in [27] using the Random Forest (RF) classification model. The reported accuracy was 99.8%. The authors in [28] compared different classifiers to detect symmetrical faults during power swing using the change in current magnitude, voltage magnitude, current angle, voltage angle, active power, reactive power, and apparent impedance. The mutual information feature selection algorithm was used to find the optimum subset of features. The boost ensemble outperformed the k-NN, DT, SVM, and Random Forest with an accuracy of 98.2%, and receiver operating characteristic (ROC) equals 1.0. In both studies, the IBG integration was not considered, and the data balancing was not used. Principal Component Analysis (PCA) was used in two studies for feature extraction and feature selection. PCA scores for three-phase line currents in [29] were used to detect and classify the faults. The Probabilistic Neural Network (PNN) was the best classifier and yielded 100% accuracy. The other study in [30] proposed the PCA indices to localize the faults by putting them as thresholds for different types of faults. The absolute deviation from the actual fault location was the performance metric of the proposed approach. The average absolute deviation was 0.1271%. None of these two PCA-based techniques considered the integration of IBGs. The percentage error (%) of fault location was reported as less than 1% when using Fast Fourier transform (FFT) and traveling wave frequencies for three-phase current signals with the Extreme learning machine (ELM) in [31]. The modal transformation was implemented in several research articles to extract the features from current and voltage for fault detection, classification, and localization. The mean vector of the voltage out of Clarke transformation

was proposed by [32] for fault classification. Fuzzy logic with Clarke transformation for ground fault detection and Generalized Neural Network for fault classification. In addition to Clarke transformation as a feature, the FFT for phase angle calculation is another feature. The authors in [33] proposed the entropy with fast discrete orthogonal S-transform (FDOST) for fault detection, classification, and localization on hybrid transmission lines (cables and overhead) with Support vector regression (SVR) for fault localization and SVM for fault detection and type classification. The SVM achieved a detection and classification accuracy of 98.2% and SVR with localization error between 0 and 0.47 km.

From the literature survey related to fault detection and classification in transmission lines, we can observe the following:

- Few feature selection techniques were investigated to find the optimum feature subset. In most cases, filter types (Information gain, Mutual information, etc.), wrapper type (forward feature selection), and feature transformation (PCA) were considered, but none of the researchers considered the embedded-type feature selection techniques.
- The issue of data imbalance was not highlighted, and the impact on the detection/classification performance was not investigated.
- Insufficient research addresses the problem of fault detection and classification in the presence of inverter-based renewables.
- The classification accuracy was dominantly used as the only metric to evaluate the classifier's performance, which could not be sufficient if the data was unbalanced.
- The classifiers' hyperparameters tuning with optimization algorithms were not considered.

This paper proposes a two-layer classification scheme using extracted features from different domains and several feature selection algorithms for transmission line fault detection and classification. The aim is to find the optimum combination of feature selection and classification model that results in the best classification performance. More specifically, the aims of this study are:

- to conduct a comprehensive study of ML-based transmission line fault classification involving many features extracted from different domains (time, frequency, and time–frequency), different feature selection/transformation algorithms, and many widely used ML-based classification models;
- to investigate the critical problem of data imbalance as faults are relatively rare events in power systems. The class unbalancing is addressed using the synthetic minority class oversampling technique (SMOTE) and by
- optimization of the classifiers' hyperparameters.

The remainder of this paper is organized as follows: Section 2 explains the system study, the simulation scenarios, data setup, and ML approach details. Results and discussion of the performance of classification models are discussed in Section 3. The conclusion is given in Section 4.

2. Methodology

This section describes the different steps used in fault detection and classification. These steps are preprocessing, feature extraction, feature reduction, decision making, and performance assessment.

2.1. System Study and Data Preparation

The 39 Bus New England System [34] shown in Figure 1 was used to generate the data in the present study. The system was modified by including a large-scale PV plant and a DFIG wind farm at bus 2. The protected line is line 01–02, and the signals are acquired from Bus 2. The dataset was constructed according to different generation types available at Bus 2. Five different combinations of generators were considered: G10 only, PV plant only, wind farm only, G10 with PV plant, and G10 with the wind farm.

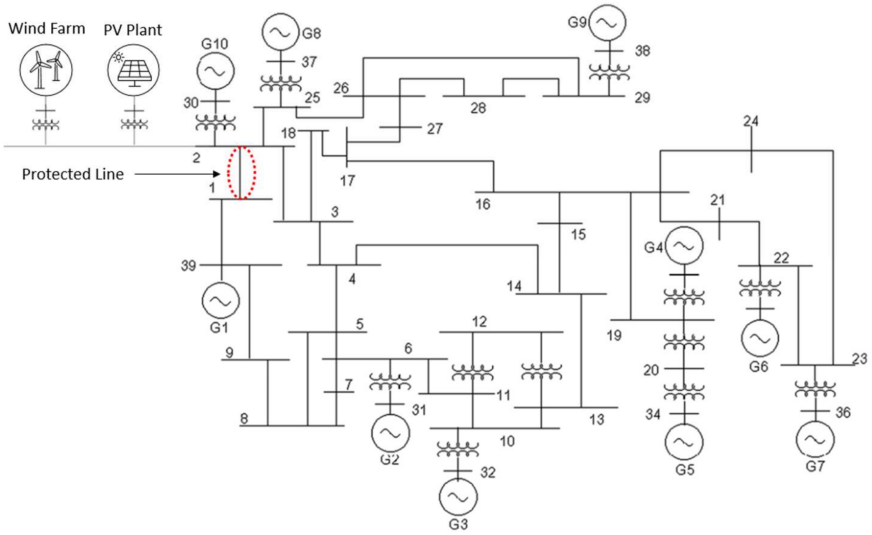

Figure 1. Modified 39- Bus New England Power System.

The models of the PV plant and wind farm were WECC large-scale PV plant, 300 MVA, 60 Hz [35], and WECC Type-3 Wind Turbine Generator (DFIG), 2.0 MVA, 60 Hz, and 150 units [36], respectively. The rated output of these plants was carefully selected to ensure the numerical stability of the system simulation. The simulation was generated using the Power Factory DigSilent software package (Power Factory 2019 SP6, DigSilent GmbH, Gomaringen, Germany). The different scenarios considered in lines 1–2 involved the following types of faults: phase-ground, two phases, two phases-to-ground, and three-phase faults at several locations (10, 50, and 90%) and different fault impedance (0 and 100 ohms).

The fault was incepted at 1.0 s and cleared after 100 ms. The power swing condition was simulated for post-fault events, as shown in Figure 2. At every generation connection status in Bus 2, sixty fault scenarios were created, as described in Table 1. The current signal of one simulated signal is similar to the one shown in Figure 2. The time frame of each simulated event was three seconds, which comprised normal, fault, and swing conditions.

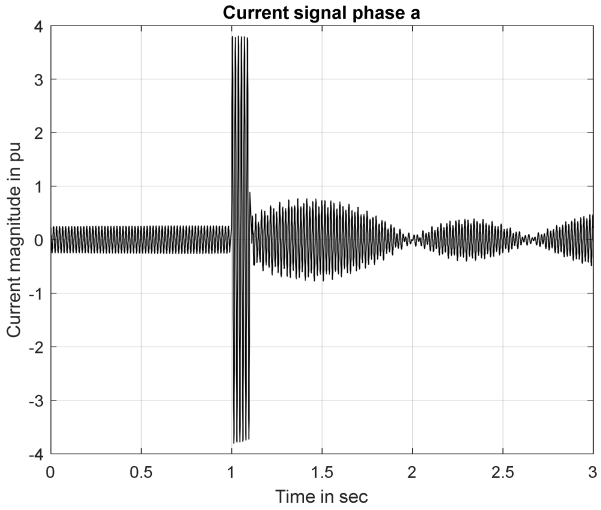

Figure 2. Phase A current signal of 3-ph fault at 50% with G10 only connected to Bus 2.

Table 1. Fault event scenarios description.

Fault Cases	Generator Type Connected to Bus 2	Fault Location (%)	Fault Impedance (ohms)	Fault Types
Cases from 1 to 10	G10, PV, WF *, G10 & PV, and G10 & WF	10	Zero	A-G, B-G, C-G, A-B, A-C, B-C, A-B-G, A-C-G, B-C-G and A-B-C
Cases from 11 to 20		50	Zero	
Cases from 21 to 30		90	Zero	
Cases from 31 to 40		10	100	
Cases from 41 to 50		50	100	
Cases from 51 to 60		90	100	

(* WF: Wind Farm).

The acquired signals at the measurement point in Bus 2 were the three-phase instantaneous currents (i_a, i_b, i_c), the three-phase instantaneous voltages (v_a, v_b, v_c), the phasor voltage, and the angle between voltage and current. The phasor voltage magnitude and angle were used to calculate the swing center voltage (SCV) as detailed in [37]. The sampling frequency considered in the simulation was 2 kHz.

Following the ML approach presented in Figure 3 and using the created dataset, the number of features extracted from the dataset was reduced using feature reduction algorithms reduce the computational burden of training the classification models. Four classification models were then used to classify the events: DT, K-NN, SVM, and Ensemble trees.

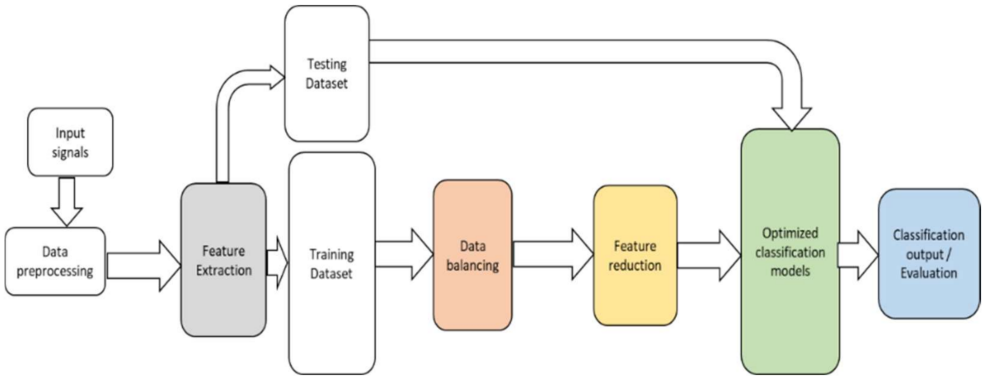

Figure 3. Machine learning approach for fault detection and classification.

The classifiers' hyperparameters were tuned using the Bayesian optimization algorithm. The performance of the classification models was evaluated using four classification metrics: accuracy, sensitivity, specificity, and precision. Further details of the proposed methods and algorithms are discussed in the following subsections.

The protection scheme suggested in this study is illustrated in Figure 4. There are two classification levels: fault detection (binary classification) and fault type classification (multi-class classification) models. The first classifier is intended to differentiate fault events from non-fault events. The non-fault events consisted of normal events and power swing events. If the first classifier detects a fault, the second classifier is used to classify it into one of the seven types of faults: A-G fault, B-G fault, C-G fault, A-B and A-B-G faults, A-C and A-C-G faults, C-B and C-B-G faults, and A-B-C faults. Identifying the fault type is vital for auto-reclose function activation or blocking. The input signals of the fault detection classifier are the three-phase instantaneous voltage and current signals and the SCV. On the other hand, the instantaneous current signals are the only input signals for the fault-type classifier. Many researchers proposed the current-only fault type classification method [38,39]. This scheme prioritizes fault detection, which is the ultimate action required to discriminate between faults and non-fault events. Then, the second layer is activated if the output of the first layer is "1".

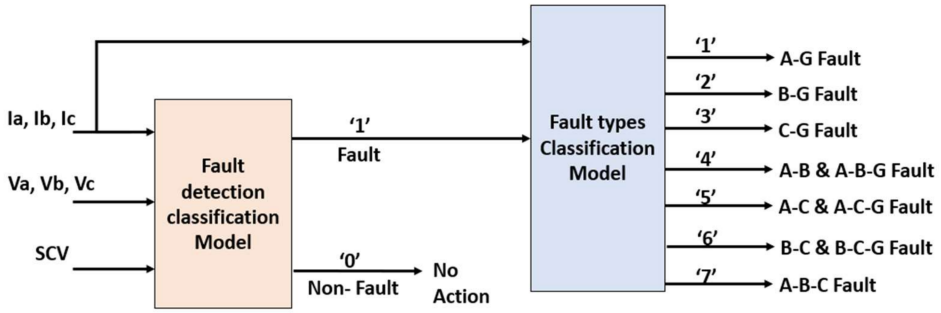

Figure 4. ML-based protection scheme.

2.2. Feature Extraction

All signals have been segmented in 8 ms (16 samples) epochs. Features were then extracted from the voltage (v_a, v_b, v_c), current (i_a, i_b, i_c), and SCV signals represented in the time domain, frequency domain, and time–frequency domain as shown in Table 2. These epochs extracted statistical features, including root-mean-square (RMS), maximum, minimum, mean, median, variance, standard deviation, kurtosis, and skewness. These time-domain features were extracted from the original signals and their first-order differences. In addition, the same statistical features were obtained from the measured signals' spectrograms (time-frequency representation). The statistical features were also extracted from the DWT's first and second detail coefficients. In addition, the estimated instantaneous frequency obtained from the Hilbert transformation algorithm was also extracted. The extracted frequency domain features spectral entropy and the mean and median frequencies. The total number of features was 49 for each signal epoch. As a result, the total number of extracted features was 343. A detailed description of these features is presented in Appendix A.

Table 2. Extracted features from selected signals in different domains.

Domain	Features	Number of Features for Each Signal
Time	Statistical features of original signals [40]	9
	Statistical features of first-order difference of the original signals [40]	9
Time-frequency	Statistical features of spectrogram [41]	9
	Statistical features of wavelet decomposition of first and second detail coefficients [42]	18
	Estimated instantaneous frequency [43]	1
Frequency	Spectral entropy [44]	1
	Mean and median frequency [45]	2

It is worth noting that the proposed fusion of the selected features was not considered in any previous study for transmission line fault detection and classification. All extracted features from current, voltage, and SCV signals were used for fault detection problems, whereas the extracted features from current signals are only used for the fault type classification stage. This was because the fault type classification requires the identification of faulty phases, which can be identified only by the current signals.

2.3. Data Balancing

Fault events in power systems form a minority class compared to normal conditions. Using unbalanced data tends to bias the classifier outputs toward the majority class. Two widely used approaches to balancing the datasets are under-sampling the majority class and over-sampling the minority one. Oversampling can be achieved by duplicating the samples in the minority class or synthetically adding new data samples. This approach is preferred when the majority class is not big enough or the minority class is too small. The

most widely used oversampling method, the Synthetic Minority Over-sampling Technique (SMOTE), is based on the k-nearest neighbor algorithm [7]. This method is considered in this research.

2.4. Feature Reduction

The number of features in the dataset is reduced to avoid the overfitting problem and the "curse of dimensionality". This can be done using either feature transformation or feature selection techniques. Feature transformation aims to transform the feature set into a lower dimension space without eliminating existing features. On the other hand, the feature selection methods rank the features by assigning them weights indicating their importance and selecting the ones with the highest weights. Feature selection techniques are of three types: filter, wrapper, and embedded. Filter methods are independent of any learning method and focus on the general characteristics of the data. Conversely, wrappers and embedded methods require a learning method to judge the importance of the features [46]. Figure 5 shows the different techniques used for feature reduction [40].

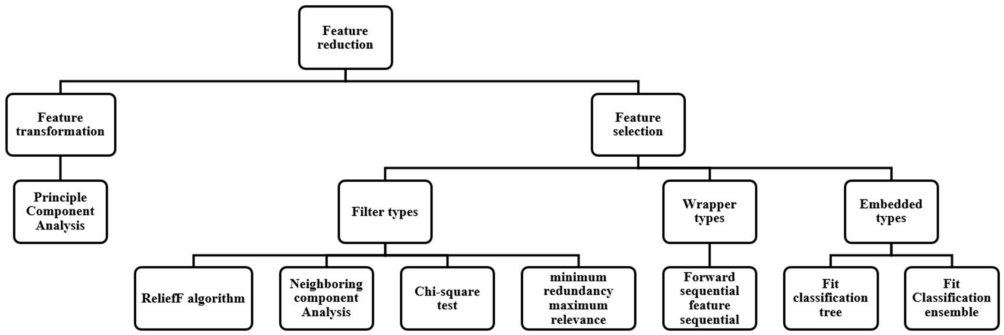

Figure 5. Proposed feature reduction algorithms.

2.5. Classification Models

In this paper, we selected five among the widely used classification models, namely decision trees [47], Support Vector Machines with Gaussian Kernel [48], k-nearest neighbors [49], and Ensemble trees [50]. The classifiers' hyperparameters can be tuned manually or automatically using different optimization algorithms like grid search, random search, and Bayesian optimization, among others. The objective of the optimization scheme is to minimize the classification error. The Bayesian optimization technique is proposed in this study. It is a sequential model-based optimization algorithm that uses the results from the previous iteration to decide on the following hyperparameter values. This process is performed until it converges to the optimum values or reaches a stopping criterion [35]. Bayesian optimization tends to converge faster to an optimal solution compared to grid/random search algorithms [51]. The optimizable classifiers with hyperparameter search options are presented in Table 3.

Table 3. Optimization search options of optimizable classification models [52].

Model	Hyper-Parameters	Search Options	Remarks
Optimizable Tree	Maximum number of splits	Search among integers log-scaled in the range [1, max(2, n − 1)]	Specify the maximum number of splits or branch points to control the depth of the tree.
	Split criterion	Gini's diversity index, Towing rule, and Maximum deviance reduction	This is to decide when to split the nodes.

Table 3. *Cont.*

Model	Hyper-Parameters	Search Options	Remarks
Optimizable GSVM	Box constraint level	Range [0.001,1000]	This is to keep the allowable values of the Lagrange multipliers in a box.
	Multiclass method Standardize data	Search between true and false	If predictors have widely different scales, standardizing can improve the fit.
Optimizable k-NN	Number of neighbors	Search among integers log-scaled in the range [1, max(2, n − 1)]	A fine kNN uses fewer neighbors, and a coarse kNN uses higher neighbors.
	Distance metric	Euclidean, City block, Chebyshev, Minkowski (cubic), Mahalanobis, Cosine, Correlation, Spearman, Hamming, Jaccard	These are metrics to determine the distance to points.
	Distance weight	Searches among Equal, Inverse, and Squared inverse.	Specify the distance weighting function.
	Standardized data	Search between true and false	Standardizing the data can improve the fit if predictors have widely different scales.
Optimizable Ensemble	Ensemble method	Search among AdaBoost, RUSBoost, LogitBoost, GentleBoost, and Bag	
	Maximum number of splits	Search among integers log-scaled in the range [1, max(2, n − 1)]	Specify the maximum number of splits or branch points to control the depth of the tree.
	Number of learners	Search among integers log-scaled in the range [10, 500]	Many learners can produce high accuracy but can be time-consuming to fit
	Learning rate	Search among real values log-scaled in the range [0.001, 1]	If the learning rate is set to less than 1, the Ensemble requires more learning iterations but often achieves better accuracy.

2.6. Evaluation Metrics

In many previous related works, authors commonly used the accuracy as the only metric to evaluate the performance of the proposed classifiers. However, the accuracy metric is not always good, especially for imbalanced data [53]. The classification performance metrics used in this paper are accuracy, sensitivity, specificity, and precision [54] defined in Table 4.

Table 4. Performance Metrics Definition.

Metric	Defined as	Confusion Matrix			
Accuracy	(TP + TN)/(TP + FN + FP + TN) *			Predicted Class	
Sensitivity	TP/(TP + FN)			Class 0	Class 1
Specificity	TN/(TN + FP)	Actual Class	Class 0	TP	FN
Precision	TP/(TP + FP)		Class 1	FP	TN

* TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative.

- **Accuracy** is the ratio of the number of correct predictions (fault and non-fault events) to the total number of input samples in the test dataset.
- **Sensitivity** is the percentage of true positives (non-fault events) that are correctly identified by the classifier.
- **Specificity** is the percentage of true negatives (fault events) that are correctly identified by the classifier.
- **Precision** indicates the percentage of instances the classifier detected as positives compared to the total positive instances.

3. Results and Discussion of Results

This section is divided into three main parts: the classification performance of the proposed two-layered classification scheme, the performance evaluation to detect and classify new fault scenarios, and a comparative analysis of previous studies in the literature. The classification performance is measured using the above-mentioned four performance metrics. More details about the setting of the proposed algorithms and techniques can be found in the Appendix B.

3.1. Performance of Fault Detection Model

The fault detection model was designed to classify fault from non-fault events. The performance evaluation covers the effect of data balancing on the classification models and the performance when using feature reduction techniques.

3.1.1. Performance of Balanced versus Unbalanced Datasets

First, the classification of the datasets discussed in Section 3 with all extracted features is considered. Figure 6 shows the classification performance of both balanced and unbalanced datasets using the complete set of features and the classifiers' hyperparameters are presented in Table 5 with the required training time. The following could be observed from these results:

- The classification performance of the four proposed classifiers was generally high. The Bag Ensemble and decision trees achieved better performance than k-NN and SVM.
- Balancing the dataset improved the specificity and sensitivity of the SVM and k-NN classifiers.
- The training time for the balanced dataset increased dramatically as the number of observations of the minority class (fault events) increased. In addition, the training time for the classifiers with more tuned parameters was higher, as in the case of tuning the k-NN and Ensemble.

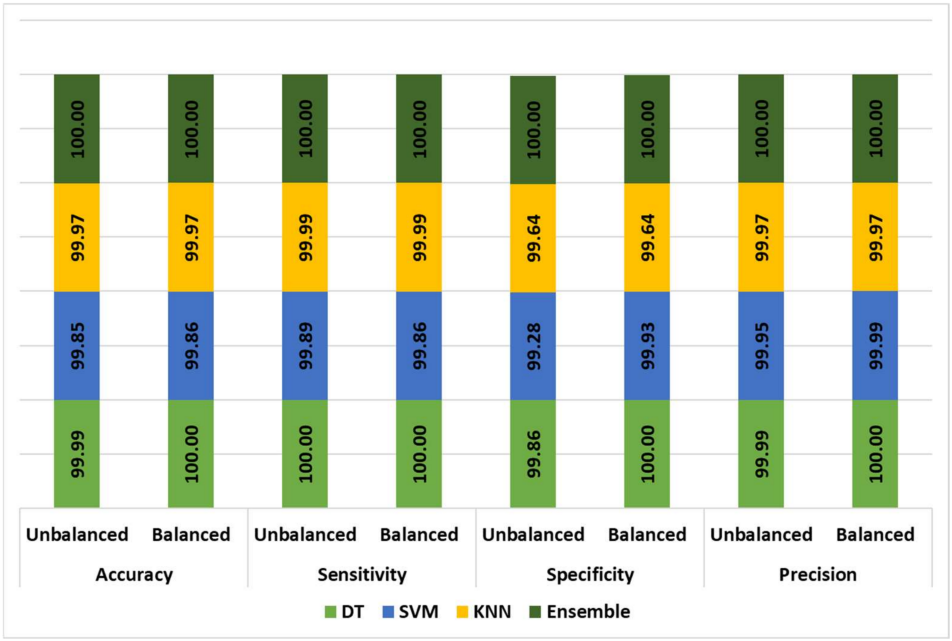

Figure 6. Classification performance of unbalanced/balanced datasets with a complete set of features.

Table 5. Classifiers' hyperparameters of unbalanced/balanced datasets.

Classifier	Unbalanced Dataset		Balanced Dataset	
	Hyperparameters	**Training Time (s)**	**Hyperparameters**	**Training Time (s)**
DT	Maximum number of splits: 428, Split criterion: Maximum deviance reduction	188	Maximum number of splits: 34,833, Split criterion: Maximum deviance reduction	227
SVM	Box constraint level: 47.0829, Standardize data: false	2058	Box constraint level: 995.6227, Standardize data: false	21,324
K-NN	Number of neighbors: 1, Distance weight: Inverse, Standardize data: true	5390	Number of neighbors: 1, Distance metric: Euclidian, Distance weight: Equal, Standardize data: true	15,187
Ensemble	Ensemble method: Bag, Maximum number of splits: 30,209, number of learners: 13, number of predictors to sample: 60	819	Ensemble method: Bag, Maximum number of splits: 16, number of learners: 228, number of predictors to sample: 145	17,680

3.1.2. Performance of Reduced Dataset

Table 6 presents the classification performance for each feature reduction technique. The best classifier selected for each method was the one with the best classification metrics. The selected features by the different feature reduction methods are shown in Figure 7. The selected features depend on each feature's weight score according to each method's criteria. The non-zero score selection is fundamentally used to choose the features. If the scores of all features are non-zeros (example of ReleifF), then the average score of all features was considered a threshold. Those which were greater than or equal to the average value were selected. The following notes could be highlighted from the results:

- With only 163 features selected using the sequential forward feature selection method, the classification performance was the highest using the Bag ensemble classifier. However, the training time was high.
- In the embedded-type feature selection methods (Fit trees and Fit Ensemble), single and two features were selected, respectively, with remarkable classification performance. The training time was low compared with others because of the small dimension of the training data.
- The mRMR feature selection algorithm with the k-NN classifier achieved the lowest performance.
- The chi-square test selected only 34 features, but the training time to tune the hyperparameters of the k-NN classifier was the highest. Its performance was quite good but less than NCA and ReleifF techniques.
- The Ensemble and DT classifiers were the best classifiers with most types of feature selection except for mRMR algorithm, where the k-NN was the best performer.
- The SVM with Gaussian kernel gave relatively poor results with all feature reduction algorithms.
- The PCA with GentleBoost Ensemble classifier performed well using 95% of the variance explained, but the training time was considerably high.

Table 6. Performance of reduced features for fault detection model.

Feature Reduction Method	Best Classifier	Number of Selected Features	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	Hyperparameters	Training Time (s)
NCA	Ensemble	8	99.98	99.98	100.00	100.00	Ensemble method: GentleBoost, Maximum number of splits: 283, number of learners: 274, Learning rate: 0.007819	804
Chi-square test	Ensemble	34	99.55	99.75	96.72	99.77	Ensemble method: AdaBoost, Maximum number of splits: 1087, number of learners: 136, Learning rate: 0.66932	14,327
ReliefF	DT	122	100.00	100.00	99.93	99.99	Maximum number of splits: 86,022, Split criterion: Gini's diversity index	104
mRMR	KNN	3	91.02	91.51	84.21	98.77	Number of neighbors: 28, Distance metric: Euclidean, Distance weight: Equal	552
Forward Sequential feature	Ensemble	163	100.00	100.00	100.00	100.00	Ensemble method: Bag, Maximum number of splits: 15,701, number of learners: 17, number of predictors to sample: 39	1223
Fit classification ensemble	Ensemble	1	99.63	99.61	99.86	99.99	Ensemble method: RUSBoost, Maximum number of splits: 1, number of learners: 10, Learning rate: 0.0073088	146
Fit classification trees	DT	2	99.99	100.00	99.78	99.98	Maximum number of splits: 49, Split criterion: Maximum deviance reduction	31
PCA	Ensemble	6 components	98.92	99.13	95.94	99.71	Ensemble method: GentleBoost, Maximum number of splits: 57, number of learners: 210	12,037

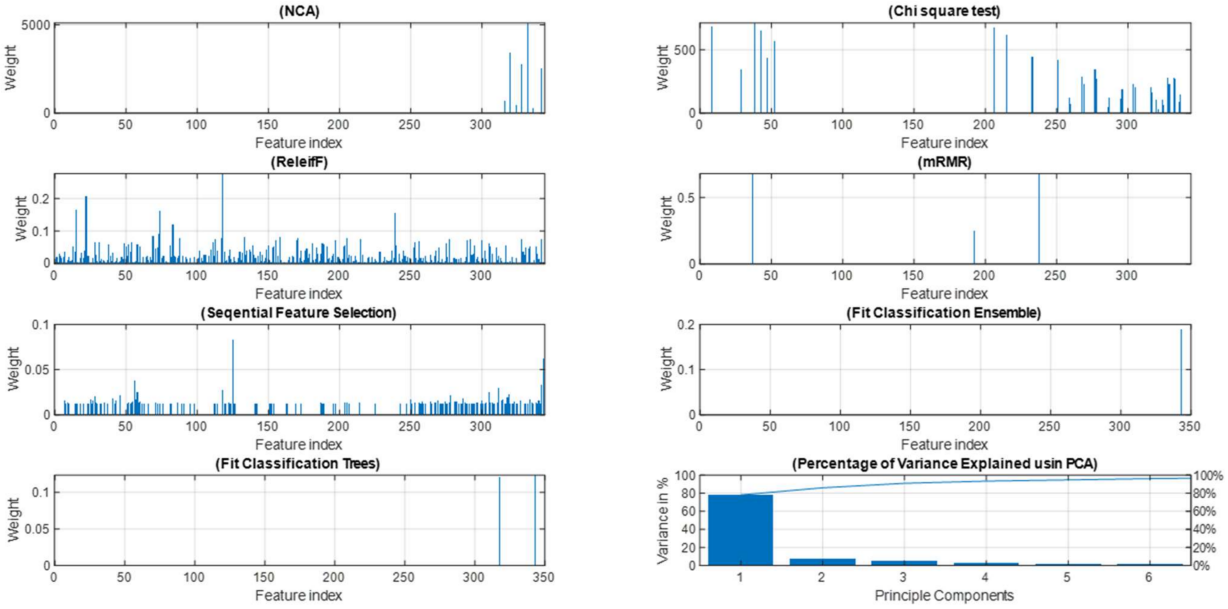

Figure 7. Feature reduction algorithms results for fault detection model (detailed description of the features is in Appendix A).

3.2. Performance of Fault Classification Model

The fault type classification model was designed to classify the type of fault after detecting it using the fault detection classification model, as illustrated in Figure 4. The optimum classification model was achieved by following the same ML approach stated in Figure 3 except for the data balancing because the groups were considered balanced during the fault simulation. The description of the dataset is presented in Appendix B.

3.2.1. Performance with a Complete Set of Features

The classification accuracy using the proposed classifiers with the complete set of extracted features of three-phase instantaneous current signals is presented in Figure 8. The maximum accuracy was 99.4% using the Ensemble classifier, whereas the lowest was 91.6% with k-NN.

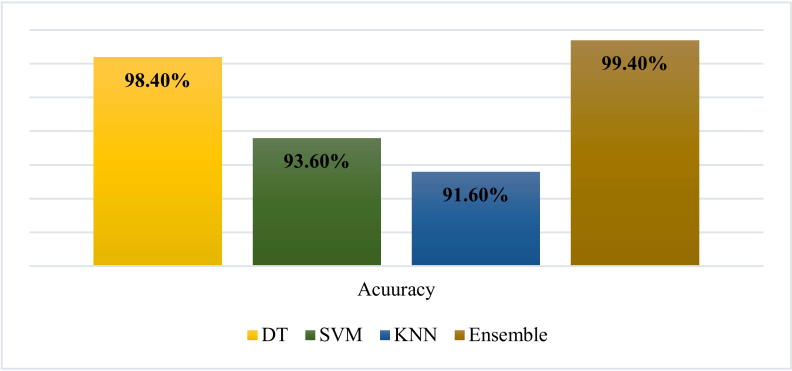

Figure 8. Classification accuracy of fault type classification model with the complete set of features.

The hyperparameters of the best classifier were as follows: Ensemble method: AdaBoost, the maximum number of splits: 5, the number of learners: 465, and the learning rate: 0.70288.

3.2.2. Performance of Reduced Dataset

The proposed feature-reduction algorithms presented in Figure 5 were also used with the proposed machine learning classifiers for the fault type classification model. The selected features using the proposed techniques are shown in Figure 9. It can be noticed that the number of features was reduced to 147, which represents the extracted features from three-phase current signals (3 signals × 49 features). The accuracy will be considered as the only classification performance metric as the fault type classes were considered balanced. Table 7 shows the detailed results of each feature-reduction method. The results showed that some selection techniques have classification accuracy close to the accuracy of a complete set of features. For example, the Fit ensemble method yielded an accuracy of 99.0% with 24 features. However, none of the feature-reduction algorithms could better improve the classification accuracy than using the complete set of features.

Table 7. Classification performance of reduced features for fault type classification.

Feature Reduction Method	Best Classifier	Number of Selected Features	Accuracy (%)	Hyperparameters	Training Time (s)
NCA	DT	103	98.5	Maximum number of splits: 2643, Split criterion: Maximum deviance reduction	23
Chi-square test	Ensemble	87	97.5	Ensemble method: AdaBoost, Maximum number of splits: 14, number of learners: 24, Learning rate: 0.0023388	145
ReliefF	Ensemble	41	98.2	Ensemble method: AdaBoost, Maximum number of splits: 7, number of learners: 449, Learning rate: 0.99393	324
mRMR	Ensemble	5	70.9	Ensemble method: Bag, Maximum number of splits: 618, number of learners: 245, number of predictors to sample: 2	236
Forward Sequential feature	Ensemble	57	98.8	Ensemble method: RUSBoost, Maximum number of splits: 112, number of learners: 11, Learning rate: 0.73117	133
Fit classification ensemble	Ensemble	24	99.0	Ensemble method: Bag, Maximum number of splits: 36, number of learners: 95, number of predictors to sample: 8	148
Fit classification trees	Ensemble	14	98.9	Ensemble method: Bag, Maximum number of splits: 220, number of learners: 12, number of predictors to sample: 5	68
PCA	Ensemble	3 components	74.6	Ensemble method: Bag, Maximum number of splits: 2619, number of learners: 473, number of predictors to sample: 3	202

Moreover, the lowest performance was with the selected features using mRMR, which resulted in around 80% accuracy. The highest training time was with the ReliefF selection technique and Adaboost ensemble classifier. In most cases, the Ensemble classifier with different ensemble methods was the best performer with the highest accuracy, except for NCA, where the DT was the best classifier.

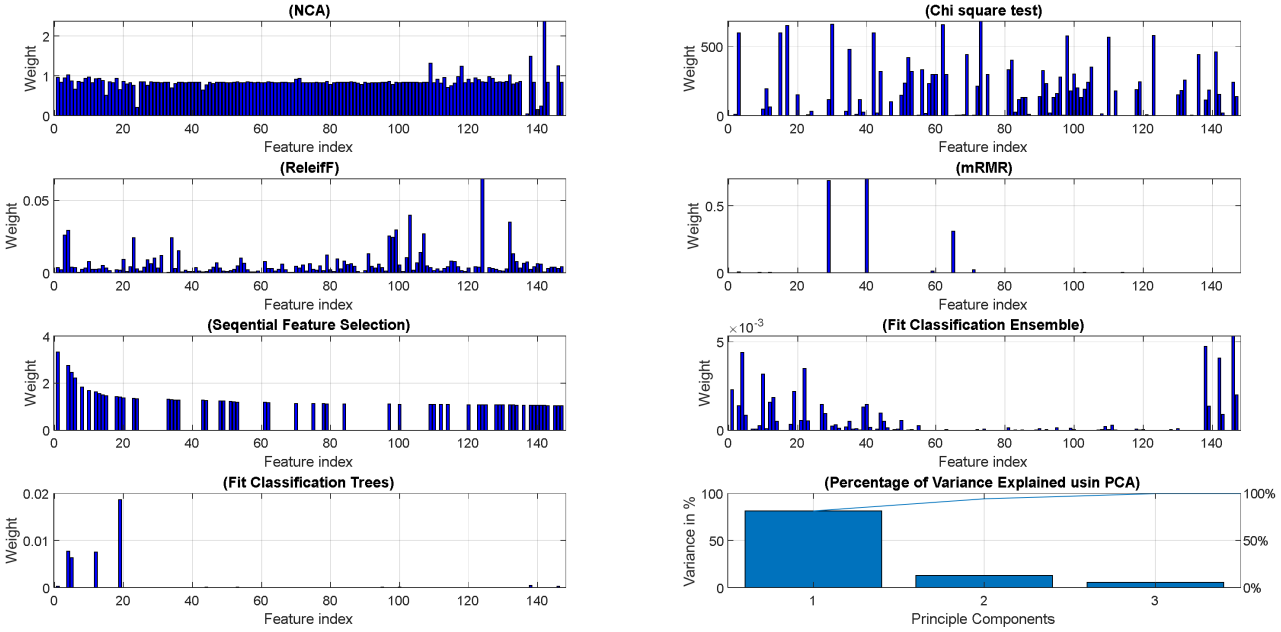

Figure 9. Feature reduction algorithms results for fault type classification model (detailed description of the features is in Appendix A, considering current signals only).

3.3. Performance Evaluation Using New Fault Scenarios

The performance of the classifiers was further evaluated using newer fault scenarios not included in the training and testing datasets. The classification models used for evaluation were *Bag ensemble* classifier with the complete set of features (343 features) for fault detection and the *Adaboost ensemble* classifier with the complete set of features (147 features) for fault classification.

The following lines describe the different fault scenarios and the results obtained using the proposed two-layer classifier (fault detection and fault type classifiers). Moreover, Figure 10 shows the output signals of detection and classification models.

Scenario 1: Three cascaded in-zone faults at 70% of the protected line (Line 1–2) from the measurement point were simulated. The first fault was A-G at 1.0 s, the second was A-B at 2.0 s, and the third was A-B-C fault at 3.0 s. The fault duration was 0.1 s. The generation connected to Bus 02 was G10 and PV plant. The faults were correctly detected and classified except for A-G fault, where the classifier was confused between class 1 (A-G) & class 2 (B-G).
Scenario 2: Out of zone fault in the line (1–39) at 50%. The fault was A-B-C fault incepted at 1.0 s for 100 ms. The generation connected to Bus 02 was G10 and Windfarm. The fault was detected. The fault type was initially detected as A-B-C fault at the inception of the fault, and then, during the fault, it was classified as A-C fault. Out-of-zone detection could be mitigated by introducing fault detection with ML for each line in the system.
Scenario 3: Out of zone fault in the line (2–25) at 50%. The fault was A-B fault incepted at 1.0 s for 100 ms. The generation connected to Bus 02 was G10 and Windfarm. The fault was detected and classified accurately, although the fault was not located at the protected line. Similar to Scenario 2, out-of-zone detection could be mitigated by introducing fault detection with ML for each line in the system.
Scenario 4: High impedance fault with fault resistance of 200 ohms was incepted at 70% of the line (01–02). The fault was A-B-C fault. The generation connected to Bus 02 was G10 and PV Plant. The fault was correctly detected.
Scenario 5: Fault during power swing was created at 2.0 s, and the power swing occurred due to fault clearance that happened at 1.0 s. Both faults were A-B-C faults created at 50% of the line (01–02). The generation connected to Bus 02 was G10 and PV Plant. The faults were correctly detected and classified.

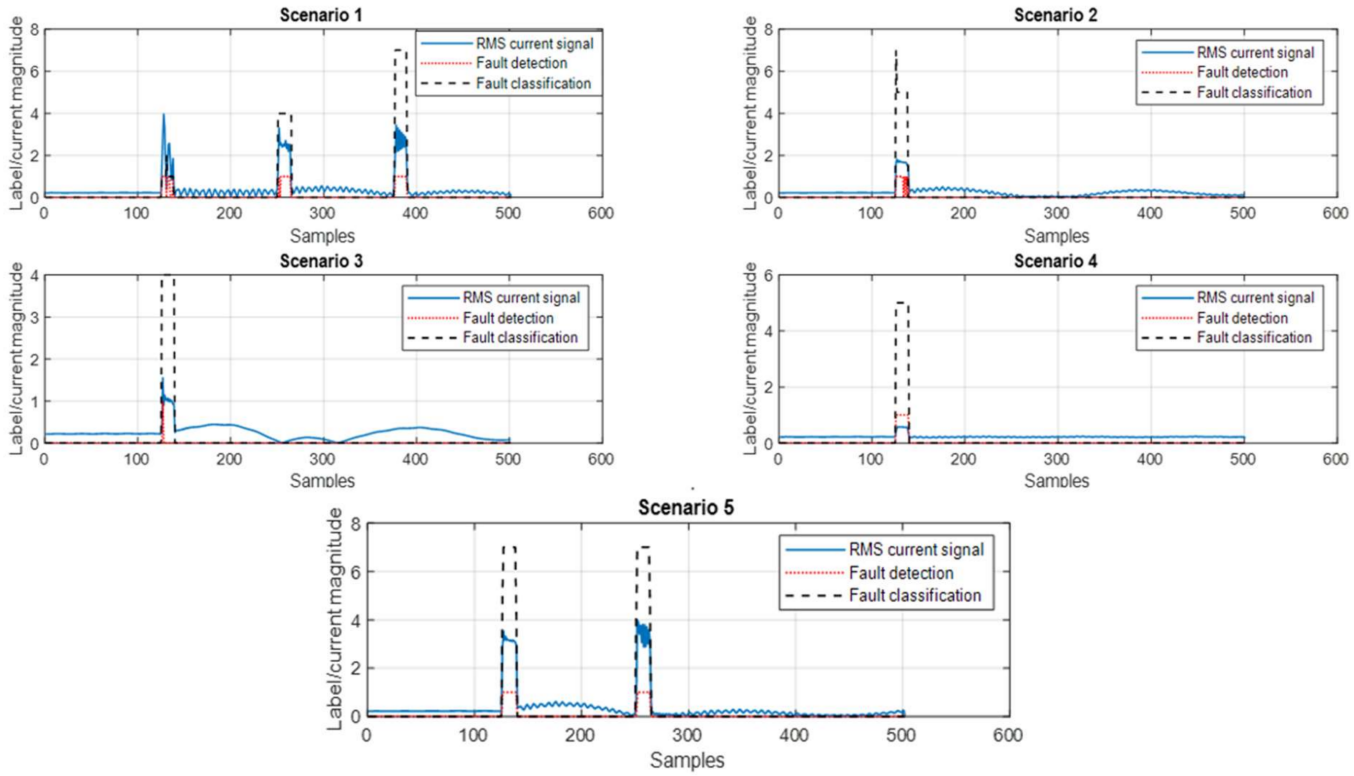

Figure 10. Fault detection and classification results related to the new fault scenarios.

3.4. Comparative Analysis of Different Methods in the Literature

This section compares the proposed two-layer classification model (fault detection then fault classification) with the previous classification methods in the literature. Table 8 shows the comparison of different fault detection and classification methods and the best model for detection and classification obtained from this study (Bag ensemble with forward-sequential feature selection for fault detection and Adaboost ensemble with the complete set of features of current signals for fault classification).

Compared with other methodologies from the literature, the implemented approach resulted in high detection and classification capability considering the integration of IBGs and reporting different classification metrics (accuracy, sensitivity, specificity, and precision), which are essential for the case of the unbalanced dataset. The 'accuracy' metric for fault classification was assumed sufficient as it was assured that the seven fault classes were created balanced.

Table 8. Comparative analysis of selected methodologies from the literature.

Reference	Objective	Methodology	IBG Consideration	Classifiers	Performance Metrics	Classification Results
[25]	Fault detection and classification for mutually coupled transmission lines	Discrete Wavelet Transformation was used to extract the features from three-phase currents. Twenty-one classes were considered for phase fault identification and four classes for ground fault identification. The data balancing was not considered, and no feature reduction technique was used.	No	ANN, k-NN, and DT.	Accuracy	Accuracy = 100% (ANN)
[26]	Fault detection and classification	Discrete Wavelet Transformation was used to extract the features from three-phase currents and voltages. Twelve classes were considered for normal events and different fault types. The data balancing was not reported, and no feature reduction technique was used.	No	k-NN and DT	Accuracy	Accuracy = 100% (DT)
[24]	Fault classification and localization	Wavelet packet transformation was used to extract the features from three-phase voltages and currents. The data was reported balanced, and the Forward feature selection algorithm was used to reduce the number of features. Ten classes were introduced for normal conditions and fault events with different types of faults.	No	SVM	Accuracy for classification, and absolute error for fault localization	Accuracy = 99.21% absolute error < 0.21%
[28]	Classification of symmetrical faults during power swing	Changes in current magnitude, voltage magnitude, current angle, voltage angle, active power, reactive power, apparent resistance, and reactance were considered the features of Voltage and current magnitude and angle, active and reactive power, and apparent impedance signals. Data balancing was not reported, and the Mutual information technique was used for feature selection. Two classes were considered for fault and swing events.	No	k-NN, DT, k-NN, Boost ensemble, SVM, and Random Forest	Accuracy, and receiver operating characteristic (ROC)	Accuracy = 98.2% (Boost ensemble) ROC = 1.0

Table 8. *Cont.*

Reference	Objective	Methodology	IBG Consideration	Classifiers	Performance Metrics	Classification Results
[29]	Fault detection and classification	Principle component scores of three-phase current signals were used as features. The dataset was created balanced. Eleven classes were used, one for non-fault events and ten for different fault types.	No	PNN	Accuracy	Accuracy = 100% (PNN)
[33]	detection, classification, and localization of faults on hybrid transmission lines (cables and overhead)	Entropy with fast discrete orthogonal S-transform (FDOST) was used to extract three-phase fault current signals. Eleven classes were introduced to represent a non-fault and different types of faults	No	Support vector regression (SVR) for fault localization SVM for fault detection and type classification	Accuracy	Accuracy = 98.2% Localization error = 0–0.47 km
Proposed approach	Fault detection and classification	As described in Section 2.	Yes (Large-scale PV and DFIG wind farm)	DT, k-NN, SVM, and ensembles	Accuracy, specificity, sensitivity, and precision	Fault detection: (Bag Ensemble with forward feature selection) Accuracy = 100%, specificity = 100%, sensitivity = 100%, and precision = 100% Fault classification: (Adaboost Ensemble) Accuracy = 99.4%

4. Conclusions

This study proposes a machine-learning approach to detect and classify faults in transmission lines connected to inverter-based generators (i.e., PV and DFIG wind farm plants). A two-layer classification scheme was implemented to detect a fault from non-fault events and then classify the type of detected faults. The features from measured three-phase voltages and currents were extracted in the time, frequency, and time-frequency domains. The main outcomes of this study are:

- The results showed that the data balancing using SMOTE improved the specificity and sensitivity metrics however the training time increased dramatically.
- Each proposed feature-reduction method developed a different selected subset of features and resulted in different classification performance.
- The Ensemble and DT classifiers performed better than others in most types of feature selection.
- The forward feature selection technique with the Bag ensemble classifier improved the classification metrics to 100% for fault detection using 163 features.
- The Adaboost ensemble classifier had the highest accuracy compared with other classifiers with 99.4% for fault type classification.
- The prediction capability for fault detection and classification was high using the complete set of features when tested with new test cases.
- Compared with other methodologies from the literature, the implemented approach resulted in high detection and classification capability considering the integration of IBGs.

The proposed approach uses classical machine learning models that learn from static, identically distributed, and well-labeled training data. That is not necessarily the case for the non-stationary behavior of power systems. Moreover, the assumption was that the simulated faults were stationary (time-invariant), which may not be the case in real-life systems exposed to environmental factors and aging. Tackling these issues requires using intelligent agents with the ability to continuously learn from real-time data. Incremental learning could be used to update the faults datasets online. As the new types of faults may not be identified in an online setting, unsupervised or semi-supervised learning techniques can be used. These directions are currently being investigated.

Author Contributions: Conceptualization, K.A.K.; methodology, K.A.K.; software, K.A.K.; validation, K.A.K., A.E.H. and M.M.; formal analysis, K.A.K., A.E.H. and M.M.; investigation, K.A.K.; resources, K.A.K.; data curation, K.A.K.; writing—original draft preparation, K.A.K.; writing—review and editing, A.E.H. and M.M.; supervision, A.E.H. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors are thankful to the Department of Electrical Engineering, Sultan Qaboos University, for providing facilities to conduct this research.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Features Description

($x_1 - x_9$): statistical features * of the squared signal of i_a, ($x_{10} - x_{18}$): statistical features of the squared signal of i_b, ($x_{19} - x_{27}$): statistical features of the squared signal of i_c, ($x_{28} - x_{36}$): statistical features of the squared signal of v_a, ($x_{37} - x_{45}$): statistical features of the squared signal of v_b, ($x_{46} - x_{54}$): statistical features of the squared signal of v_c,

($x_{55} - x_{63}$): statistical features of the squared signal of SCV, ($x_{64} - x_{72}$): statistical features of the first order difference signal of i_a, ($x_{73} - x_{81}$): statistical features of the first order difference signal of i_b, ($x_{82} - x_{90}$): statistical features of the first order difference signal of i_c, ($x_{91} - x_{99}$): statistical features of the first order difference signal of v_a, ($x_{100} - x_{108}$): statistical features of the first order difference signal of v_b, ($x_{109} - x_{117}$): statistical features of the first order difference signal of v_c, ($x_{118} - x_{126}$): statistical features of the first order difference signal of SCV, ($x_{127} - x_{135}$): statistical features of the first detail coefficients signal of i_a, ($x_{136} - x_{144}$): statistical features of the first detail coefficients signal of i_b, ($x_{145} - x_{153}$): statistical features of the first detail coefficients signal of i_c, ($x_{154} - x_{162}$): statistical features of the first detail coefficients signal of v_a, ($x_{163} - x_{171}$): statistical features of the first detail coefficients signal of v_b, ($x_{172} - x_{180}$): statistical features of the first detail coefficients signal of v_c, statistical features of the first detail coefficients signal of v_b, ($x_{181} - x_{189}$): statistical features of the first detail coefficients signal of SCV, ($x_{190} - x_{198}$): statistical features of the second detail coefficients signal of i_a, ($x_{199} - x_{207}$): statistical features of the second detail coefficients signal of i_b, ($x_{208} - x_{216}$): statistical features of the second detail coefficients signal of i_c, ($x_{217} - x_{225}$): statistical features of the second detail coefficients signal of v_a, ($x_{226} - x_{234}$): statistical features of the second detail coefficients signal of v_b, ($x_{235} - x_{243}$): statistical features of the second detail coefficients signal of v_c, statistical features of the second detail coefficients signal of v_b, ($x_{244} - x_{252}$): statistical features of the second detail coefficients signal of SCV, ($x_{253} - x_{261}$): statistical features of the spectrogram of i_a, ($x_{262} - x_{270}$): statistical features of the spectrogram of i_b, ($x_{271} - x_{279}$): statistical features of the spectrogram of i_c, ($x_{280} - x_{288}$): statistical features of the spectrogram of v_a, ($x_{289} - x_{297}$): statistical features of the spectrogram of v_b, ($x_{298} - x_{306}$): statistical features of the spectrogram of v_c, statistical features of the spectrogram of v_b, ($x_{307} - x_{315}$): statistical features of the spectrogram of SCV, ($x_{316} - x_{319}$): the mean, median, instantaneous frequency, and spectral entropy of i_a, ($x_{320} - x_{323}$): the mean, median, instantaneous frequency, and spectral entropy of i_b, ($x_{324} - x_{327}$): the mean, median, instantaneous frequency, and spectral entropy of i_c, ($x_{328} - x_{331}$): the mean, median, instantaneous frequency, and spectral entropy of v_a, ($x_{332} - x_{335}$): the mean, median, instantaneous frequency, and spectral entropy of v_b, ($x_{336} - x_{339}$): the mean, median, instantaneous frequency, and spectral entropy of v_c, ($x_{340} - x_{343}$): the mean, median, instantaneous frequency, and spectral entropy of SCV.

* Statistical features = [maximum, minimum, mean, median, standard deviation, variance, kurtosis, skewness, root mean square].

Appendix B

Experimental Settings

39-Bus Power System	As Detailed in [34].
PV inverter	10 kVA per inverter, local controller: constant Q, Short circuit model: Dynamic voltage support, Sub-transient short circuit: 1.21 kVA, R to X″ ratio: 0.1, K Factor: 2, Max. current: 1.1 pu, Td″ = 0.03 s, Td′ = 1.2 s
Wind Turbine	2 MVA, 1.0 power factor, local controller: constant Q, Short circuit model: Dynamic voltage support, Sub-transient short circuit: 2.39 MVA, R to X″ ratio: 0.1, K Factor: 2, Max. current: 1.1 pu, Td″ = 0.03 s, Td′ = 1.2 s
SPECTROGRAM	Window size: 16 samples, overlapping: 16 samples
SMOTE	Number of nearest neighbors: 5, Oversampling rate: 450%, Random seeds: 100
Optimizer settings	Optimizer: Bayesian Optimization, Acquisition function: Expected improvement per second plus, Iterations: 10

39-Bus Power System	**As Detailed in [34].**
Sequential feature selection algorithm	Criterion function: Residual Sum of Squares, Validation method: 5-fold cross-validation
ReliefF algorithm	Number of nearest neighbors: 5
PCA	Percentage of variance explained: 95%
Classifier-1 datasets (Fault Detection)	(1) **Unbalanced Dataset:** **Labels**: '**0**': non-fault events including normal system condition and power swing condition, and '1': fault events **Training Data**: Number of observations: 3240 (Fault) and 44729 (non-fault) **Testing Data**: Number of observations: 1318 (Fault) and 19239 (non-fault) Features: 343 features as defined in Table 2 (2) **Unbalanced Dataset:** **Labels**: '**0**': non-fault events including normal system condition and power swing condition, and '1': fault events **Training Data**: Number of observations: 28857 (Fault) and 31409 (non-Fault) **Testing Data: (kept same as for Unbalanced dataset)**
Classifier-2 datasets (Fault Classification)	**Labels**: '**1**': A_G fault, '**2**': B-G fault, '**3**': C-G fault, '**4**': A-B & A-B-G fault, '**5**': A-C & A-C-G fault, '**6**': B-C & B-C-G fault, and '**7**': A-B-C fault **Training Dataset**: Number of observations: **278** (A-Gfault), **277** (B-G fault), **263** (C-G fault), **521** (A-B & A-B-G fault), **540** (A-C & A-C-G fault), **524** (B-C & B-C-G fault), and **258** (B-C & B-C-G fault) **Testing Dataset**: Number of observations: **104** (A-G fault), **104** (B-G fault), **118** (C-G fault), **238** (A-B & A-B-G fault), **219** (A-C & A-C-G fault), **235** (B-C & B-C-G fault), and **122** (B-C & B-C-G fault) **Features**: 147 extracted features from only current signals (Ia, Ib, and Ic) as defined in Table 2.

References

1. IEEE Power & Energy Society. PES-TR81—Protection Challenges and Practices for Interconnecting Inverter Based Resources to Utility Transmission Systems. 2020. Available online: https://resourcecenter.ieee-pes.org/technical-publications/technical-reports/PES_TP_TR81_PSRC_WGC32_071520.html (accessed on 2 May 2020).
2. Liu, S.; Bi, T.; Liu, Y. Theoretical analysis on the short-circuit current of inverter-interfaced renewable energy generators with fault-ride-through capability. *Sustainability* **2017**, *10*, 44. [CrossRef]
3. Mahamedi, B.; Fletcher, J.E. Trends in the protection of inverter-based microgrids. *IET Gener. Transm. Distrib.* **2019**, *13*, 4511–4522. [CrossRef]
4. Haj-Ahmed, M.A.; Feilat, E.A.; Khasawneh, H.J.; Abdelhadi, A.F.; Awwad, A. Comprehensive Protection Schemes for Different Types of Wind Generators. *IEEE Trans. Ind. Appl.* **2018**, *54*, 2051–2058. [CrossRef]
5. North Americam Electric Reliability Corporation (NERC). Short-circuit modeling and system strength. In *White Papar*; North Americam Electric Reliability Corporation: Atlanta, GA, USA, 2018.
6. IEEE Power and Energy Society. *IEEE Std 2800*; IEEE Standard for Interconnection and Interoperability of Inverter-Based Resources (IBRs) Interconnecting with Associated Transmission Electric Power Systems. IEEE: Manhattan, NY, USA, 2022.
7. Liang, Z.; Lin, X.; Kang, Y.; Gao, B.; Lei, H. Short Circuit Current Characteristics Analysis and Improved Current Limiting Strategy for Three-phase Three-leg Inverter under Asymmetric Short Circuit Fault. *IEEE Trans. Power Electron.* **2018**, *33*, 7214–7228. [CrossRef]
8. IEEE/NERC. *Task force on Short-Circuit and System Performance Impact of Inverter Based Generation: Impact of Inverter Based Generation on Bulk Power System Dynamics and Short-Circuit Performance*; Technical Report; IEEE: Manhattan, NY, USA, 2018.
9. Usama, M.; Mokhlis, H.; Moghavvemi, M.; Mansor, N.N.; Alotaibi, M.A.; Muhammad, M.A.; Bajwa, A.A. A Comprehensive Review on Protection Strategies to Mitigate the Impact of Renewable Energy Sources on Interconnected Distribution Networks. *IEEE Access* **2021**, *9*, 35740–35765. [CrossRef]
10. Senarathna, S.; Hemapala, K.T.M.U. Review of adaptive protection methods for microgrids. *AIMS Energy* **2019**, *7*, 557–578. [CrossRef]

11. Kaur, G.; Prakash, A.; Rao, K.U. A critical review of Microgrid adaptive protection techniques with distributed generation. *Renew. Energy Focus* **2021**, *39*, 99–109. [CrossRef]
12. Adly, A.R.; Aleem, S.H.E.A.; Algabalawy, M.A.; Jurado, F.; Ali, Z.M. A novel protection scheme for multi-terminal transmission lines based on wavelet transform. *Electr. Power Syst. Res.* **2020**, *183*, 106286. [CrossRef]
13. Sakis, G.J.C.; Meliopoulos, A.P. Setting-less Protection. In Proceedings of the 2013 46th Hawaii International Conference on System Sciences, Wailea, HI, USA, 7–10 January 2013.
14. Okedu, K.E. Enhancing DFIG wind turbine during three-phase fault using parallel interleaved converters and dynamic resistor. *IET Renew. Power Gener.* **2016**, *10*, 1211–1219. [CrossRef]
15. Chen, L.; Li, G.; Chen, H.; Tao, Y.; Tian, X.; Liu, X.; Xu, Y.; Ren, L.; Tang, Y. Combined Use of a Resistive SFCL and DC-link Regulation of a SMES for FRT Enhancement of a DFIG Wind Turbine Under Different Faults. *IEEE Trans. Appl. Supercond.* **2019**, *29*, 1–8. [CrossRef]
16. Papaspilotopoulos, N.H.V.; Korres, G. An Adaptive Protection Infrastructure for Modern Distribution Grids with Distributed Generation. *Cigre Sci. Eng.* **2016**, *7*, 125–132.
17. Hossain, M.; Ali, M.H. Transient stability improvement of doubly fed induction generator based variable speed wind generator using DC resistive fault current limiter. *IET Renew. Power Gener.* **2016**, *10*, 150–157. [CrossRef]
18. Alam, S.; Abido, M.A.Y.; El-Amin, I. Fault Current Limiters in Power Systems: A Comprehensive Review. *Energies* **2018**, *11*, 1025. [CrossRef]
19. Wang, K.; Guo, M.; Dai, C.; Li, Z. Information-decision searching algorithm: Theory and applications for solving engineering optimization problems. *Inf. Sci.* **2022**, *607*, 1465–1531. [CrossRef]
20. El-Naily, N.; Saad, S.M.; Elhaffar, A.; Zarour, E.; Alasali, F. Innovative Adaptive Protection Approach to Maximize the Security and Performance of Phase/Earth Overcurrent Relay for Microgrid Considering Earth Fault Scenarios. *Electr. Power Syst. Res.* **2022**, *206*, 107844. [CrossRef]
21. Chopard, B.; Tomassini, M. Performance and limitations of metaheuristics. In *An Introduction to Metaheuristics for Optimization*; Natural Computing Series; Springer: Cham, Switzerland, 2018; pp. 191–203.
22. Rahman Fahim, S.; Sarker, S.K.; Muyeen, S.M.; Sheikh, M.; Islam, R.; Das, S.K. Microgrid Fault Detection and Classification: Machine Learning Based Approach, Comparison, and Reviews. *Energies* **2020**, *13*, 3460. [CrossRef]
23. Pérez-Ortiz, M.; Jiménez-Fernández, S.; Gutiérrez, P.A.; Alexandre, E.; Hervás-Martínez, C.; Salcedo-Sanz, S. A Review of Classification Problems and Algorithms in Renewable Energy Applications. *Energies* **2016**, *9*, 607. [CrossRef]
24. Ray, P.; Mishra, D.P. Support vector machine based fault classification and location of a long transmission line. *Eng. Sci. Technol. Int. J.* **2016**, *19*, 1368–1380. [CrossRef]
25. Swetapadma, A.; Yadav, A.; Abdelaziz, A.Y. Intelligent schemes for fault classification in mutually coupled series-compensated parallel transmission lines. *Neural Comput. Appl.* **2020**, *32*, 6939–6956. [CrossRef]
26. Wasnik, P.P.; Phadkule, N.J.; Thakur, K.D. Fault Detection and Classification in Transmission Line by using KNN and DT Technique. *Int. Res. J. Eng. Technol.* **2020**, *7*, 335–340.
27. Patil, D.; Naidu, O.D.; Yalla, P.; Hida, S. An Ensemble Machine Learning Based Fault Classification Method for Faults During Power Swing. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Asia, ISGT 2019, Chengdu, China, 21–24 May 2019; pp. 4225–4230. [CrossRef]
28. Lwin, M.; Min, K.W.; Padullaparti, H.V.; Santoso, S. Symmetrical fault detection during power swings: An interpretable supervised learning approach. In Proceedings of the IEEE Power and Energy Society General Meeting, Portland, OR, USA, 5–9 August 2018; pp. 1–5. [CrossRef]
29. Mukherjee, A.; Kundu, P.K.; Das, A. Application of Principal Component Analysis for Fault Classification in Transmission Line with Ratio-Based Method and Probabilistic Neural Network: A Comparative Analysis. *J. Inst. Eng. India Ser. B* **2020**, *101*, 321–333. [CrossRef]
30. Mukherjee, A.; Kundu, P.K.; Das, A. Transmission Line Fault Location Using PCA-Based Best-Fit Curve Analysis. *J. Inst. Eng. India Ser. B* **2021**, *102*, 339–350. [CrossRef]
31. Akmaz, D.; Mamiş, M.S.; Arkan, M.; Tağluk, M.E. Transmission line fault location using traveling wave frequencies and extreme learning machine. *Electr. Power Syst. Res.* **2018**, *155*, 106034. [CrossRef]
32. Sharma, S.K. GA-GNN (Genetic Algorithm-Generalized Neural Network)-Based Fault Classification System for Three-Phase Transmission System. *J. Inst. Eng. India Ser. B* **2019**, *100*, 435–445. [CrossRef]
33. Patel, B. A new FDOST entropy based intelligent digital relaying for detection, classification and localization of faults on the hybrid transmission line. *Electr. Power Syst. Res.* **2018**, *157*, 39–47. [CrossRef]
34. DIgSILENT Power Factory. *39 Bus New England System*; Power Factory: Gomaringen, Germany, 2015; pp. 1–16.
35. Lammert, G.; Ospina, L.D.P.; Pourbeik, P.; Fetzer, D.; Braun, M. Implementation and validation of WECC generic photovoltaic system models in DIgSILENT PowerFactory. In Proceedings of the IEEE Power and Energy Society General Meeting, Boston, MA, USA, 17–21 July 2016; pp. 3–7. [CrossRef]
36. Hiskens, I.A. Dynamics of Type-3 Wind Turbine Generator Models. *IEEE Trans. Power Syst.* **2012**, *27*, 465–474. [CrossRef]
37. Benmouyal, G.H.D.T. Zero-setting power-swing blocking protection. In Proceedings of the 3rd IEE International Conference on Reliability of Transmission and Distribution Networks (RTDN 2005), London, UK, 15–17 February 2005; pp. 249–254. [CrossRef]

38. Jarrahi, M.A.; Samet, H.; Ghanbari, T. Fast Current-Only Based Fault Detection Method in Transmission Line. *IEEE Syst. J.* **2019**, *13*, 1725–1736. [CrossRef]
39. Chiradeja, P.; Ngaopitakkul, A. Classification of Lightning and Faults in Transmission Line Systems Using Discrete Wavelet Transform. *Math. Probl. Eng.* **2018**, *2018*, 1847968. [CrossRef]
40. Nandi, H.A.K. *Condition Monitoring with Vibration Signals: Compressive Sampling and Learning Algorithms for Rotating Machines*; JohnWiley & Sons: Hoboken, NJ, USA, 2019; Volume 53.
41. Ramos-Aguilar, R.; Olvera-López, J.A.; Olmos-Pineda, I.; Sánchez-Urrieta, S. Feature extraction from EEG spectrograms for epileptic seizure detection. *Pattern Recognit. Lett.* **2020**, *133*, 202–209. [CrossRef]
42. Aliyu, I.; Lim, C.G. Selection of optimal wavelet features for epileptic EEG signal classification with LSTM. *Neural Comput. Appl.* **2021**, 1–21. [CrossRef]
43. Taheri, B.; Salehimehr, S.; Razavi, F.; Parpaei, M. Detection of power swing and fault occurring simultaneously with power swing using instantaneous frequency. *Energy Syst.* **2020**, *11*, 491–514. [CrossRef]
44. Kłosowski, G.; Rymarczyk, T.; Wójcik, D.; Skowron, S.; Cieplak, T.; Adamkiewicz, P. The Use of Time-Frequency Moments as Inputs of LSTM Network for ECG Signal Classification. *Electronics* **2020**, *9*, 1452. [CrossRef]
45. Phinyomark, A.; Thongpanja, S.; Hu, H.; Phukpattaranont, P.; Limsakul, C. The Usefulness of Mean and Median Frequencies in Electromyography Analysis. In *Computational Intelligence in Electromyography Analysis—A Perspective on Current Applications and Future Challenges*; IntechOpen: London, UK, 2012; pp. 195–220.
46. Remeseiro, B.; Bolon-Canedo, V. A review of feature selection methods in medical applications. *Comput. Biol. Med.* **2019**, *112*, 103375. [CrossRef] [PubMed]
47. Niyas, M.; Sunitha, K. Identification and classification of fault during power swing using decision tree approach. In Proceedings of the 2017 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems, SPICES 2017, Kollam, India, 8–10 August 2017; pp. 1–6. [CrossRef]
48. Yang, C.; Yang, J.; Ma, J. Adaptive Kernel Parameters. *Int. J. Comput. Intell. Syst.* **2020**, *13*, 212–222. [CrossRef]
49. Blagus, R.; Lusa, L. SMOTE for high-dimensional class-imbalanced data. *BMC Bioinform.* **2013**, *14*, 106. [CrossRef] [PubMed]
50. Yang, L.; Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **2020**, *415*, 295–316. [CrossRef]
51. Swamynathan, M. *Mastering Machine Learning with Python in Six Steps: A Practical Implementation Guide to Predictive Data Analytics Using Python*, 2nd ed.; Apress: New York, NY, USA, 2019; Volume 126.
52. Mathworks, C. Statistics and Machine Learning Toolbox TM User's Guide R 2016 b. 2016. Available online: https://www.mathworks.com/products/statistics.html (accessed on 25 May 2022).
53. Brownlee, J. Imbalanced Classification with Python: Better Metrics, Balance Skewed Classes, Cost-Sensitive Learning, V1.3; Machine Learning Mastery. 2020. Available online: https://www.scribd.com/document/517979097/Imbalanced-Classification-With-Python-by-Jason-Brownlee-Z-lib-org (accessed on 25 May 2022).
54. Gadekallu, T.R.; Khare, N.; Bhattacharya, S.; Singh, S.; Maddikunta, P.K.R.; Ra, I.-H.; Alazab, M. Early Detection of Diabetic Retinopathy Using PCA-Firefly Based Deep Learning Model. *Electronics* **2020**, *9*, 274. [CrossRef]