Machine Learning Model Applications for Fault Detection and Classification in Distributed Power Networks

Jose Eduardo Urrea Cabus*, [1], İsmail Hakkı Altaş[1]
[1]*Department of Electrical and Electronics Engineering, Karadeniz Technical University, 61080, Trabzon, Turkey*

***Abstract*—This paper compares various unsupervised feature extraction techniques and supervised machine learning models for fault detection and classification in a power distributed generation system. The modified IEEE 34 bus test feeder was implemented for the study case simulated through PowerFactory DigSILENT software. Data analysis results from three-phase voltages and currents collected were performed in Python. Simulation results confirm that by applying dimensionality reduction techniques such as feature extraction and wavelet family selection adequately, high identification and classification accuracy can be obtained, excluding the less essential characteristics and preventing the machine learning models from overfitting or underfitting the datasets.**

***Index Terms*— Data mining, fault diagnosis, feature extraction, machine learning.**

I. INTRODUCTION

One of the most demanding high-quality services, which is experiencing vast and rapid development nowadays, is the electricity supply [1]. Consequently, the electrical grid has been defined as the most extensive engineered system worldwide due to its indispensability in our daily lives and importance to the economies and progress of countries [2, 3]. Likewise, continuity and reliability have become essential requirements for customers that are particularly susceptible to power blackouts [4]. Therefore, fault condition detection is critical for reliable services [5]. Furthermore, detecting short circuits in distribution networks is much more difficult than in transmission networks because they are typically unbalanced and asymmetrical due to the increasing incorporation of renewable energy on the load side as distributed generation (DGs) [6]. This consequence has triggered the modernization and development of the smart grids, incorporating modern measurement and communication systems into the power systems' real-time monitoring [5, 7, 8]. Nowadays, the stakeholders' main goal is to improve the power grids by making them more intelligent, reliable, and sustainable. Fault detection systems yield a practical, fast, and reliable form of relaying operations. Additionally, they should perform satisfactorily under multiple operating conditions and diverse electrical grid parameters. When it comes to fault detections, they are supposed to be detected first, then correctly classified, and finally cleared in the shortest amount of time to maintain reliance and service continuity [4].

Methods for fault detection, classification, and location in power distribution systems in the presence of DGs and the intelligent agent's incorporation have been studied and published over the years [6, 10]. For instance, in [5] and [8], data mining has been incorporated into the protection functions to extend performance limits by producing enhanced conditions of states within the present protection technology limitations over the distribution network. On the other hand, in [11], they incorporate statistical machine learning methods for preventive maintenance, using historical electrical data collected that has been turned into models that aim to forecast the risk of failures for components and systems. Due to their versatility, high capability, and high accuracy performances, machine learning implementation for fault detection and classification in power systems has grown significantly in recent years [12]. In [1], [4], [6] and [18], authors have used Artificial Neural Networks (ANNs) for fault detection, classification, and location, collecting three-phase voltages and currents as input data for performing the output predictions. In [7], two techniques for fault identification and classification have been presented based on TA-QSSVM and A-QSSVM, respectively, a modification of the Support Vector Machine (SVM) algorithm; both methods are unsupervised and online with good performance during their accuracy calculation. Besides, [9] introduces a protection scheme using statistical models like energy, entropy, and standard deviation for microgrids using Wavelet Transforms (WTs) and Decision Trees (DTs) as a discriminating function. [12] has presented a semi-supervised machine learning approach based on co-training over a microgrid, where the harmony search algorithm is implemented to identify optimal wavelet transform families during the data pre-processing step. In [3, 14], the authors tested three machine learning algorithms and used the third level of decomposition for the wavelet transform in data pre-processing. The simulations showed the high capability of the Random Forest algorithm (RF) over the other algorithms. In [15], there are three families of discrete wavelet transforms (DWT) for feature extraction over the input data (e.g., motor current). After their extraction, classification results were performed by RF and XGBoost machine learning algorithms. [16, 17], they implemented WT and DWT for feature extraction,

Manuscript received July 07, 2021; accepted June 05, 2022.
*Corresponding author: joseeduardourrea@gmail.com

respectively. Zero-sequence current has been collected to perform fault detection in conjunction with Principal Component Analysis, SVM, and the Adaboost+CART algorithm as discriminant functions.

This research is based on the detection and classification of faults in a power-distributed network. Despite the previous contributions mentioned, a lack of interest was observed in selecting the WT or DWT families for data pre-processing. Furthermore, through this research, it has been concluded that the appropriate selection of WFT or DWT families and level of decomposition can improve the performance and accuracy of the machine learning algorithms. Additionally, unsupervised dimensionality reduction algorithms are applied to feature extraction techniques. Simulation results validate that using feature extraction techniques and wavelet packet transform selection effectively can achieve high identification accuracy by removing the less relevant features from consideration, preventing the machine learning algorithms from overfitting or underfitting the datasets.

The rest of the paper is organized as follows: Section II reviews the proposed methods. Simulation results are presented in Section III. Conclusions are presented in Section IV.

II. Proposed Method

The proposed method is based on three-phase voltages and currents collected from the IEEE 34 bus test feeder simulations: The zero-sequence voltage signal is implemented for wavelet family selection using the minimum entropy decomposition and the Support Vector Machine algorithm. After its identification, the model is set with the best wavelet mother and decomposition level results to perform the feature extractions. In the training and test subsets, the data is split and scaled. Six types of unsupervised dimensionality reduction are applied in order to reduce a high-dimension to a low-dimension, increasing the machine learning algorithm’s performance. The fault detection is performed by applying zero-sequence voltage components. For ground fault detection, the zero-sequence current is used for its identification. The three-phase voltage and current features are applied for the exact fault classification.

A. *Experimental Electrical Model*

A three-phase system with a balanced or unbalanced electrical grid has been described as the power system's backbone [3, 12]. It is also crucial to foresee the absence of results in terms of reliability and continuity in balanced or unbalanced three-phase systems. Similarly, distribution networks are seeing a surge in the penetration of distributed generation, which is being driven by financial considerations and cost savings. The advantages of adopting decentralized generation and grid topology designs as a long-term solution for the rising quantities of load demand and interconnections have been proven in previous studies [20].

B. *Case of Study*

The IEEE 34 bus test feeder, located in Arizona, USA, was selected as a test system in this paper. The system consisted of unbalanced loads distributed over the array of three-phases and one-phase (i.e., AN and BN) grid configurations; operational voltage levels are 24.9 kV and 4.16 kV, with a total load of 1769 kW and 1044 MVAr distributed over the grid; two capacitors installed at buses 844 and 848; and two regulators located in the line segments at 814-850 and 832-852, respectively. The test feeder is modelled and tested using the DIgSILENT PowerFactory software. Additionally, a three-phase meshed topology arrangement has been incorporated between nodes 816 and 832, with a length of 2.5km and 301 configurations [20]. Distributed generations have been modelled and placed at the weakest point over the feeder; two three-phase PV systems and one generator model available in the software's static generator library of the DIgSILENT software are used and placed in nodes 840, 848, and 890, respectively [20-22]. Fig. 1 shows the modelled IEEE 34 bus test feeder. Fault simulations have been carried out all over the electrical system and are stored in Excel files. The detection and classification stages have been performed in Jupyter Notebook and Python 3 with the implementation of Numpy, Skicit-learn, and Pandas libraries available in the software. Computer information Intel (R) Core (TM) i5-6400 CPU @2.70Hz, RAM 8GB, X64 bits Windows 10 Enterprise.

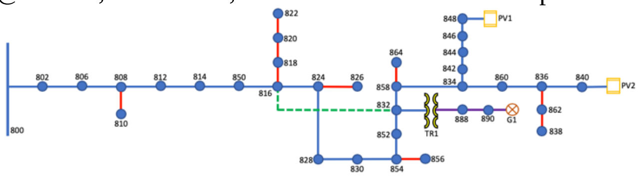


Fig. 1. Modified IEEE 34 bus test feeder

C. *Input Data*

Fig. 1 shows the modelled test feeder in DIgSILENT software, where different fault cases were simulated. Fault data was generated through the DPL code and different conditions such as random fault resistance, different fault locations, distributed generation level penetration, and various fault current inception angles. Table I shows a summary of the fault conditions in this paper.

TABLE I. FAULT DATA CONFIGURATIONS

Condition	Values
Fault type	LLL, LL, LLG, LG
Fault resistance in (Ω)	0, 20,50,80,100
Fault location in (%)	10, 25, 50,75,95
Fault angle inception in (θ)	0, 45,90,120
DGs’ level penetration in (%)	0, 25, 50, 75, 100

After simulating the model, three-phase voltages and currents are recorded. The collected data is then processed with the Clarke Transform (1) to obtain the voltage and current zero sequence component [3, 23]. It is essential to mention that the fault or non-fault presences during the energy harvesting process have been omitted. Through the signals’ analysis, it was concluded that different operating states exhibit different behaviors, which can be observed in the harmonic spectrum content or magnitudes of dominant frequencies. Therefore, a zero-sequence voltage component has been implemented to detect faults over the grid and select the best wavelet transform for data processing and feature extraction; additionally, the zero-sequence current component has been used for ground fault detection and classification.

$$\begin{bmatrix}\alpha\\ \beta\\ 0\end{bmatrix} = \frac{2}{3}\begin{bmatrix}1 & -\frac{1}{2} & -\frac{1}{2}\\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2}\\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2}\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix} \quad (1)$$

D. Wavelet Transform

A wavelet transform is a powerful tool consisting of low- and high-pass filters used in signal processing, voice recognition, and a wide variety of applications due to its good performance in frequency components (i.e. high and low-frequency) [13]. In this paper, the features contained in the voltage and current waveforms and zero-sequences are extracted with the discrete wavelet transform [3, 9, 12, 13, 15]. The discrete wavelet transform (DWT) coefficients are employed and collected using equation (2) derived from the discretization of CWT [15]. Where (j, k) are integer positives, ψ^* represents the mother wavelet's complex conjugate, and x(t) is the signal analyzed.

$$DWT_{j,k}(t) = \frac{1}{\sqrt{2^j}}\int_{-\infty}^{+\infty} x(t)\psi^*\left(\frac{t-2^jk}{2^j}\right)dt \quad (2)$$

The use of several wavelet families for signal analysis has been documented in previous research; however, no reason or technique for picking the mother wavelet has been presented. Aside from that, the mother wavelet selection and decomposition levels may have a significant impact on how regardless of its size the feature vectors are, which can result in significantly different outcomes from the study. As a consequence, in this paper, a combination of the grid searching method and the support vector machine is used to optimize the selection of wavelet family and grades of decomposition.

E. Feature Extraction

The data collected from the simulation is raw data presented in quite large dimensions and unscaled. Thus, to obtain high accuracy, robust results, and low computational complexity, the data must be as small as possible in dimensions. In this paper, the proposed method preprocesses the voltage and current signals through the discrete wavelet transform and extracts the most useful statistical features by applying feature extraction techniques that might contain essential data during transient events such as faults or perturbations [5, 9, 15, 24]. Then features are implemented to build the data-mining model using dimensionality reduction and machine learning algorithms for fault detection and classification. Twelve statistical features were used for feature extraction in this paper; the dataset has 7655 instances, which contain three-phase voltage and current with a 2-second duration. The details of the statistical features are shown in Table II.

TABLE II. STATISTICAL FEATURES

Feature	Functions	
Energy	$E = \sum_{i=1}^{N} x_i^2$	(3)
Mean Absolute	$X_{MAV} = \frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert$	(4)
R.M.S	$X_{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i)^2}$	(5)
Variance	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2$	(6)
Standard Deviation	$\sigma = \sqrt{\sigma^2}$	(7)
Kurtosis	$X_{Kurt} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^4}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^2}$	(8)
Skewness	$X_{Skew} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^3}{\left[\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2\right]^{3/2}}$	(9)
Shape Factor	$X_{SF} = \frac{X_{RMS}}{X_{MAV}}$	(10)
Impulse Factor	$X_{IF} = \frac{max_i\lvert x_i\rvert}{X_{MAV}}$	(11)
Crest Factor	$X_{crest} = \frac{max_i\lvert x_i\rvert}{X_{RMS}}$	(12)
Clearance Factor	$X_{clear} = \frac{max_i\lvert x_i\rvert}{\left[\frac{1}{N}\sum_{i=1}^{N}\sqrt{\lvert x_i\rvert}\right]^2}$	(13)
Shannon's Entropy	$SE = -\sum_{i=1}^{N} p_i \log p_i$	(14)

One of the essential procedures when dealing with data is that the data must be on the same scale because they might contain characteristics that vary significantly in magnitudes, units, and range, affecting the performance of the algorithms implemented for their analysis. Previous research has shown that rescaling the data increases the accuracy and performance of the algorithms. Therefore, the standardization function has been implemented, consisting of rescaling the data to have a mean of 0 and a standard deviation of 1. Equation (15) shows its representation.

$$x' = \frac{x_i - \bar{x}}{\sigma} \quad (15)$$

F. Dimensionality Reduction

Data compression is an essential subject in machine learning because it can help improve data storage, computational efficiency, and predictive performance. Unsupervised dimensionality reduction techniques have been implemented in this paper for feature extraction to identify the most relevant and not relevant patterns effectively, summarizing the original feature dataset from a high-dimensional space onto a low-dimensional feature subspace for their processing [10, 24-27]. A summary of the unsupervised dimensionality reduction techniques implemented in this paper is shown in Table III.

TABLE III. UNSUPERVISED DIMENSIONALITY REDUCTION TECHNIQUES

Symbol	Name
PCA	Principal Component Analysis
KPCA	Kernel Principal Component Analysis
LLE	Locally Linear Embedding
Isomap	Isometric Mapping
DL	Mini-Batch Dictionary Learning
ICA	Independent Component Analysis

G. *Machine Learning Algorithms for Decision Making*

Machine learning algorithms aim to infer comprehensible correlations or discover patterns between system variables in datasets that can be used later to forecast or comprehend system behaviors. Moreover, choosing the correct algorithm for those tasks often depends on the amount, quality, and correlation of their data features, which involves a process of trial and error [9, 24-27]. The proposed study uses five supervised machine learning algorithms (i.e. Logistic Regression (LR), Support Vector Machine (SVM), Decision Trees (DT), Random Forest (RF), and K-Nearest Neighbors (KNN)) in contrast with six unsupervised dimensionality reduction algorithms to perform the data analysis in DWT selection, ground-fault detection, detection, and classification faults over the grid.

The machine learning algorithms' robustness depends on their optimal hyperparameters' adjustment. Therefore, the proper values of the hyperparameters may boost the efficiency of the training model and test dataset accuracy [24-26]. In this paper, the Grid Search Algorithm has been proposed as an optimal model for hyperparameter selection. This algorithm is based on implementing a trial and error method to select the best parameters that give the best accuracy from a subset of hyperparameters. However, choosing the best hyperparameters sometimes leads the model to underfit or overfit the dataset, providing an unsatisfactory performance. Consequently, a K-Folds Cross-Validation model has been implemented to trade off the bias and variance to avoid underfitting or overfitting the training model [24-26].

H. *Model Evaluation*

The extracted features using pipeline models are fed into the five machine learning algorithms. Quantifying the quality of algorithms' predictions is assessed by applying the metrics and scoring listed in Table IV [26]. TN, TP, FP, and FN mean True Negative, True Positive, False Positive, and False Negative. For multiclass labels, a weighted average has been implemented; refer to [24-26]. Besides, Receiver Operating Characteristic (ROC) curves and the Area Under the Curve (AUC) percent have been implemented to visualize the algorithm's performance.

TABLE IV. METRICS AND SCORING LIST

Name	Functions	
Confusion Matrix	$CM = \begin{vmatrix} TN & FP \\ FN & TP \end{vmatrix}$	(16)
Accuracy	$Acc = \dfrac{TP + TN}{TP + TN + FP + FN}$	(17)
Precision	$PRE = \dfrac{TP}{TP + FP}$	(18)
Recall	$REC = \dfrac{TP}{TP + FN}$	(19)
F1	$F1 = 2 \times \dfrac{PRE \times REC}{PRE + REC}$	(20)
Log-loss	$L_{log}(y, p) = -(y \log(p) + (1 - y) \log(1 - p))$	(21)

III. Simulation Results

A. *Wavelet family selection*

After calculating the zero components from the 3-phase voltages and currents using equation (1), a subset of 600 instances was collected randomly to process the best wavelet family transform and decomposition levels using a zero-sequence voltage signal. In this paper, the Haar, db3, db4, db6, Sym4, and Coif2 wavelets were selected and compared in order to find the best wavelet family and decomposition level for signal analysis [3, 9, 12, 13, 15]. Table II: Twelve features were calculated from the subset. A proportion of 70:30 subset was used for training and testing models using the Grid Search algorithm and Support Vector Machine as a discriminant algorithm. The accuracy of detailed components at different decomposition levels is presented in Fig. 2. Accuracy results show that the db3 wavelet family at the seventh decomposition level provides the most distinctive accuracy value (i.e., 96.85%); thus, the wavelet decomposition step has been set to those values in order to determine the features of the three-phase voltages and currents for the next steps of the model.

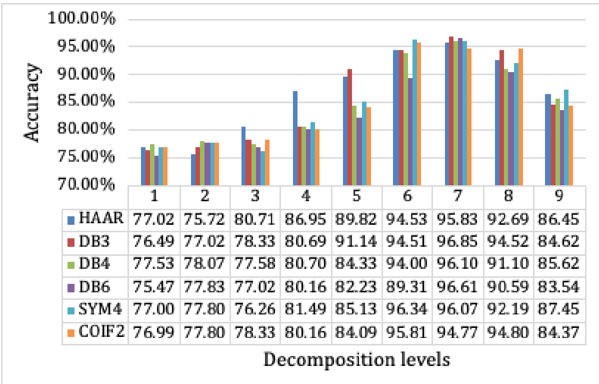
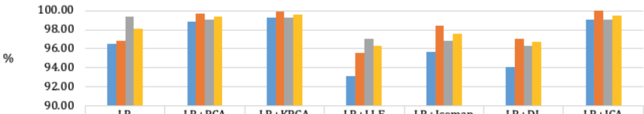


Fig. 2. Wavelet family results

B. *Fault Detection Results*

The proposed fault detection algorithm is based on the zero-sequence voltage signal. The training and test ratios were chosen as 65:35, consisting of an unbalanced dataset (i.e., no-faults and faults); twelve features were calculated and scaled before performing dimensionality reduction. The

Grid Search algorithm was used for the best hyperparameter selection. A 5-fold cross-validation approach was utilized in order to validate the trained model and trade-off the bias and variance. The most remarkable results are presented in Figs. 3–8; a summary of errors is presented in Table V.

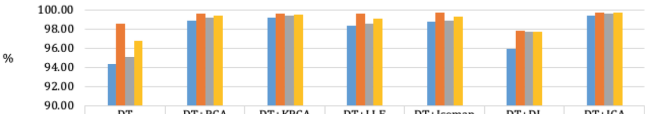


	LR	LR+PCA	LR+KPCA	LR+LLE	LR+Isomap	LR+DL	LR+ICA
Accuracy %	96.49	98.84	99.29	93.13	95.71	94.03	99.10
Precision %	96.84	99.71	99.92	95.52	98.46	97.07	100.00
Recall %	99.39	99.02	99.30	97.01	96.80	96.35	99.02
F1 %	98.10	99.36	99.61	96.26	97.62	96.71	99.51

Fig. 3. Fault detection results using Logistic Regression

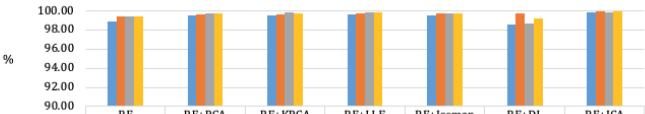
	SVM	SVM+PCA	SVM+KPCA	SVM+LLE	SVM+Isomap	SVM+DL	SVM+ICA
Accuracy %	99.03	99.55	99.29	99.55	99.25	94.66	99.85
Precision %	99.35	100.00	99.84	99.59	99.75	98.40	100.00
Recall %	99.59	99.51	99.39	99.92	99.43	95.70	99.84
F1 %	99.47	99.75	99.61	99.75	99.59	97.03	99.92

Fig. 4. Fault detection results using Support Vector Machine



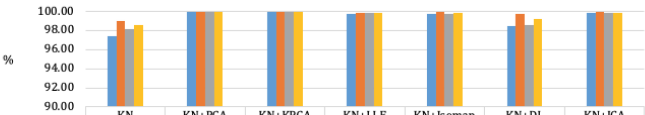
	DT	DT+PCA	DT+KPCA	DT+LLE	DT+Isomap	DT+DL	DT+ICA
Accuracy %	94.40	98.99	99.25	98.43	98.81	96.04	99.51
Precision %	98.64	99.63	99.71	99.63	99.75	97.87	99.75
Recall %	95.16	99.26	99.47	98.65	98.93	97.79	99.71
F1 %	96.87	99.45	99.59	99.13	99.34	97.83	99.73

Fig. 5. Fault detection results using Decision Tree



	RF	RF+PCA	RF+KPCA	RF+LLE	RF+Isomap	RF+DL	RF+ICA
Accuracy %	98.92	99.48	99.51	99.66	99.55	98.54	99.85
Precision %	99.39	99.63	99.59	99.75	99.71	99.75	100.00
Recall %	99.43	99.80	99.88	99.88	99.80	98.65	99.84
F1 %	99.41	99.71	99.73	99.82	99.75	99.20	99.92

Fig. 6. Fault detection results using Random Forest



	KN	KN+PCA	KN+KPCA	KN+LLE	KN+Isomap	KN+DL	KN+ICA
Accuracy %	97.43	99.93	99.93	99.74	99.78	98.51	99.81
Precision %	99.05	100.00	100.00	99.84	100.00	99.79	100.00
Recall %	98.11	99.92	99.92	99.88	99.75	98.57	99.80
F1 %	98.58	99.96	99.96	99.86	99.88	99.17	99.90

Fig. 7. Fault detection results using K-Nearest Neighbors

The best-adjusted model using dimensionality reduction techniques was analysed through the F1-score, where 99.61% of LR through K-PCA was reached. Furthermore, 99.92% of the SVM and RF and 99.73% of the DTs were reached using the ICA algorithm. Besides, 99.96% of the K-NN algorithm was reached using PCA and K-PCA algorithms.

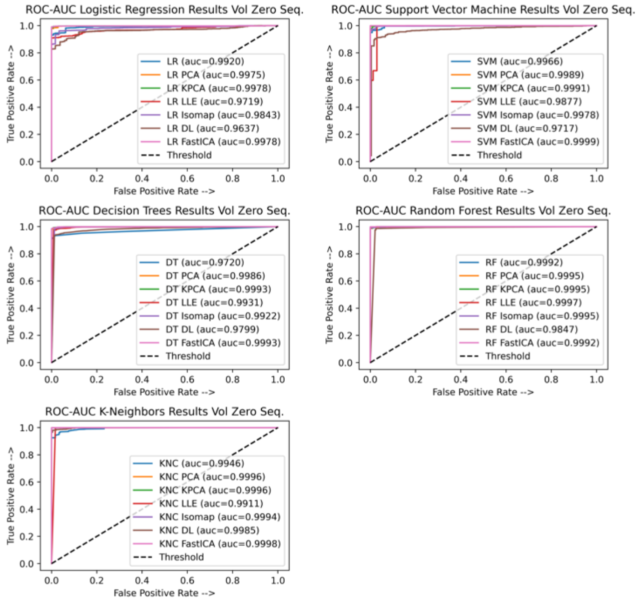
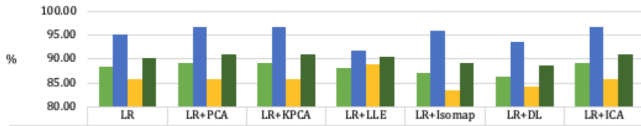


Fig. 8. Fault detection ROC-AUC curves results

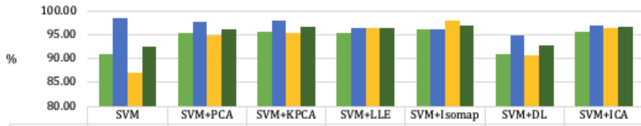
C. *Ground Fault Detection Results*

For future decision-making, it is essential to verify if the fault that occurred is or is not a ground fault. Hence, the zero-sequence current signal is implemented to identify whether a ground fault has happened in a feeder. Feature vectors are extracted based on Table II for their analysis. In addition to verifying its presence or not through machine learning algorithms, it is being investigated whether its detection can be improved by applying dimensionality reduction algorithms, as proposed in Table III. The results are presented in Fig. 9 to Fig. 14. In addition, a summary is presented in Table VI, where it presents the loss functions used for its identification and predictions.



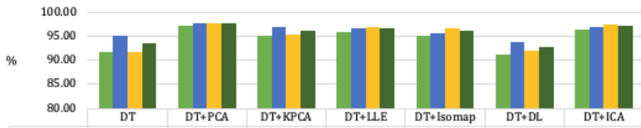
	LR	LR+PCA	LR+KPCA	LR+LLE	LR+Isomap	LR+DL	LR+ICA
Accuracy %	88.28	89.18	89.18	88.02	87.24	86.46	89.18
Precision %	95.03	96.80	96.80	91.77	95.86	93.63	96.80
Recall %	85.93	85.70	85.70	89.01	83.39	84.28	85.70
F1 %	90.25	90.92	90.92	90.37	89.19	88.71	90.92

Fig. 9. Ground fault detection using Logistic Regression



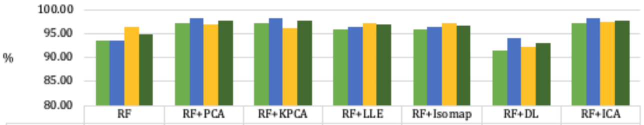
	SVM	SVM+PCA	SVM+KPCA	SVM+LLE	SVM+Isomap	SVM+DL	SVM+ICA
Accuracy %	90.97	95.34	95.75	95.45	96.12	91.01	95.75
Precision %	98.46	97.63	97.88	96.50	96.06	94.81	96.96
Recall %	87.06	94.92	95.33	96.28	97.87	90.72	96.28
F1 %	92.41	96.25	96.59	96.39	96.96	92.72	96.62

Fig. 10. Ground fault detection using Support Vector Machine



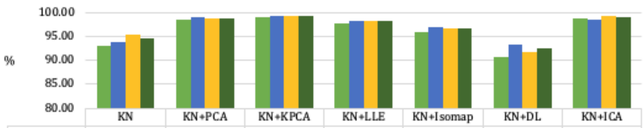
	DT	DT+PCA	DT+KPCA	DT+LLE	DT+Isomap	DT+DL	DT+ICA
Accuracy %	91.83	97.09	95.15	95.90	95.04	91.08	96.31
Precision %	95.05	97.81	96.88	96.69	95.56	93.74	96.83
Recall %	91.84	97.58	95.39	96.81	96.63	92.02	97.34
F1 %	93.42	97.69	96.13	96.75	96.09	92.87	97.08

Fig. 11. Ground fault detection using Decision Trees



	RF	RF+PCA	RF+KPCA	RF+LLE	RF+Isomap	RF+DL	RF+ICA
Accuracy %	93.51	97.05	97.13	95.93	95.90	91.38	97.24
Precision %	93.47	98.26	98.32	96.31	96.37	93.98	98.10
Recall %	96.45	97.04	96.10	97.28	97.16	92.26	97.52
F1 %	94.94	97.65	97.71	96.80	96.76	93.11	97.81

Fig. 12. Ground fault detection using Random Forest



	KN	KN+PCA	KN+KPCA	KN+LLE	KN+Isomap	KN+DL	KN+ICA
Accuracy %	93.13	98.54	99.03	97.80	95.82	90.67	98.69
Precision %	93.89	98.99	99.29	98.29	96.80	93.38	98.59
Recall %	95.33	98.70	99.17	98.23	96.57	91.73	99.35
F1 %	94.60	98.85	99.23	98.26	96.69	92.55	98.97

Fig. 13. Ground fault detection using K-Nearest Neighbors

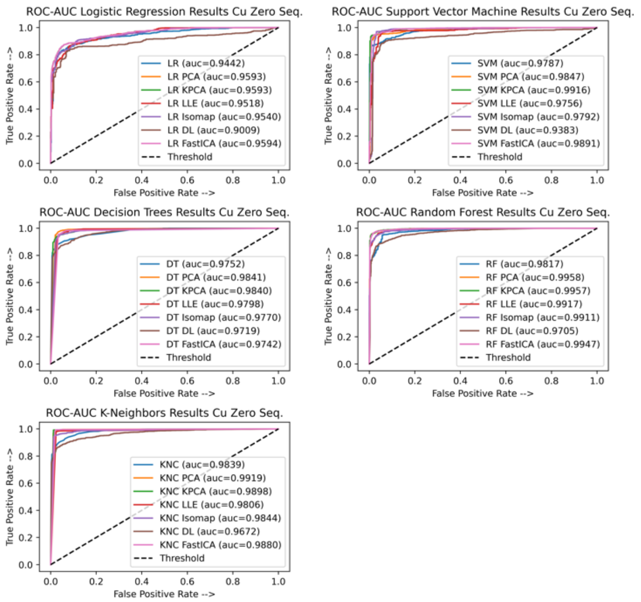


Fig. 14. Ground fault detection ROC and AUC curves results

Among the best results of F1-scores obtained with dimensionality reduction algorithms, the following can be highlighted: 96.62% obtained through the implementation of the ICA algorithm in SVM, 97.69% in DT implementing PCA, 97.81% in RF with ICA, and 99.23% in KNN with K-PCA. On the other hand, despite the improvements achieved in LR with dimensionality reduction algorithms, its poor performance in ground fault identification can be seen in the results, reaching 90.92%.

D. Fault Classification Results

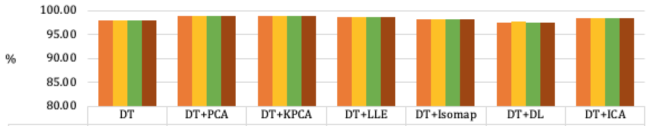
The first step for decision-making in situations of uncertainty is to know that something is happening, which leads to the identification of the situation discussed in the previous sections. After getting prior knowledge, it is necessary to identify the puzzle's pieces or situation experienced. Therefore, following the same model presented above, a fault classification model is proposed by implementing dimensionality reduction and machine learning algorithms. The results of the applications are presented below, from Fig. 15 to Fig. 19. A summary of cost functions and classification results is presented in Table VII.

	LR	LR+PCA	LR+KPCA	LR+LLE	LR+Isomap	LR+DL	LR+ICA
Accuracy %	95.93	99.81	98.06	96.94	96.64	96.46	99.63
Precision %	96.07	99.81	98.18	96.93	96.57	96.48	99.63
Recall %	95.93	99.81	98.06	96.94	96.64	96.46	99.63
F1 %	95.94	99.81	98.09	96.92	96.64	96.46	99.63

Fig. 15. Fault classification using Logistic Regression

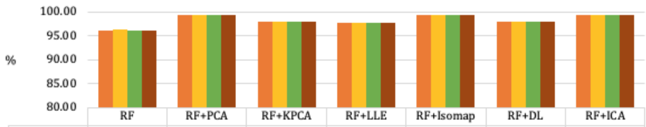
	SVM	SVM+PCA	SVM+KPCA	SVM+LLE	SVM+Isomap	SVM+DL	SVM+ICA
Accuracy %	97.16	98.10	98.32	96.38	97.46	99.59	99.85
Precision %	97.18	98.33	98.43	96.57	97.47	99.59	99.85
Recall %	97.16	98.10	98.32	96.38	97.46	99.59	99.85
F1 %	97.16	98.15	98.34	96.28	97.45	99.59	99.85

Fig. 16. Fault classification using Support Vector Machine



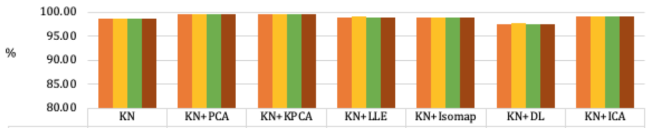
	DT	DT+PCA	DT+KPCA	DT+LLE	DT+Isomap	DT+DL	DT+ICA
Accuracy %	98.02	98.84	98.84	98.69	98.21	97.65	98.40
Precision %	98.04	98.85	98.85	98.72	98.23	97.66	98.40
Recall %	98.02	98.84	98.84	98.69	98.21	97.65	98.40
F1 %	98.03	98.84	98.84	98.70	98.21	97.65	98.39

Fig. 17. Fault classification using Decision Trees



	RF	RF+PCA	RF+KPCA	RF+LLE	RF+Isomap	RF+DL	RF+ICA
Accuracy %	96.16	99.44	97.95	97.84	99.40	97.91	99.51
Precision %	96.44	99.44	97.96	97.83	99.41	97.93	99.52
Recall %	96.16	99.44	97.95	97.84	99.40	97.91	99.51
F1 %	96.05	99.44	97.95	97.83	99.40	97.91	99.52

Fig. 18. Fault classification using Random Forest



	KN	KN+PCA	KN+KPCA	KN+LLE	KN+Isomap	KN+DL	KN+ICA
Accuracy %	98.73	99.66	99.66	99.03	98.88	97.65	99.25
Precision %	98.74	99.66	99.66	99.06	98.89	97.66	99.26
Recall %	98.73	99.66	99.66	99.03	98.88	97.65	99.25
F1 %	98.73	99.66	99.66	99.03	98.88	97.64	99.25

Fig. 19. Fault classification using K-Nearest Neighbors

From the results obtained, it can be highlighted that most of the proposed models achieved significant increases in the discriminant algorithms, which led to better performance in identifying faults on the electrical network. For instance, in the LR algorithm, 99.81% was reached by implementing the PCA algorithm; in SVM and RF, 99.85% and 99.53% were reached using the ICA algorithm, respectively. Besides, in DT and KNN, 98.84% and 99.66% were reached by applying PCA and K-PCA algorithms, respectively.

TABLE V. FAULT DETECTION SUMMARY RESULTS

	Log-Loss	Mean Abs.	TP	FP	TN	FN
LR	0.135	0.035	2424	79	162	15
LR+PCA	0.043	0.012	2415	7	234	24
LR+KPCA	0.030	0.007	2422	2	239	17
LR+LLE	0.122	0.069	2366	111	130	73
LR+Isomap	0.089	0.043	2361	37	204	78
LR+DL	0.132	0.060	2350	71	170	89
LR+ICA	0.045	0.009	2415	0	241	24
SVM	0.068	0.010	2429	16	225	10
SVM+PCA	0.019	0.004	2427	0	241	12
SVM+KPCA	0.026	0.007	2424	4	237	15
SVM+LLE	0.027	0.004	2437	10	231	2
SVM+Isomap	0.033	0.007	2425	6	235	14
SVM+DL	0.118	0.053	2334	38	203	105
SVM+ICA	0.008	0.001	2435	0	241	4
DT	0.105	0.056	2321	32	209	118
DT+PCA	0.058	0.010	2421	9	232	18
DT+KPCA	0.037	0.007	2426	7	234	13
DT+LLE	0.138	0.016	2406	9	232	33
DT+Isomap	0.116	0.012	2413	6	235	26
DT+DL	0.214	0.040	2385	52	189	54
DT+ICA	0.037	0.005	2432	6	235	7
RF	0.027	0.011	2425	15	226	14
RF+PCA	0.026	0.005	2434	9	232	5
RF+KPCA	0.033	0.005	2436	10	231	3
RF+LLE	0.020	0.003	2436	6	235	3
RF+Isomap	0.029	0.004	2434	7	234	5
RF+DL	0.246	0.015	2406	6	235	33
RF+ICA	0.018	0.001	2435	0	241	4
KN	0.055	0.026	2393	23	218	46
KN+PCA	0.026	0.007	2437	0	241	2
KN+KPCA	0.026	0.007	2437	0	241	2
KN+LLE	0.091	0.003	2436	4	237	3
KN+Isomap	0.041	0.002	2433	0	241	6
KN+DL	0.047	0.015	2404	5	236	35
KN+ICA	0.016	0.002	2434	0	241	5

TABLE VI. GROUND FAULT DETECTION SUMMARY RESULTS

	Log-Loss	Mean Abs.	TP	FP	TN	FN
LR	0.30	0.12	1454	76	912	238
LR+PCA	0.26	0.11	1450	48	940	248
LR+KPCA	0.26	0.11	1450	48	940	248
LR+LLE	0.26	0.12	1506	135	853	186
LR+Isomap	0.28	0.13	1411	61	927	281
LR+DL	0.34	0.14	1426	97	891	266
LR+ICA	0.26	0.11	1450	48	940	248
SVM	0.19	0.09	1473	23	965	219
SVM+PCA	0.16	0.05	1606	39	949	86
SVM+KPCA	0.12	0.04	1613	35	953	79
SVM+LLE	0.19	0.05	1629	59	929	63
SVM+Isomap	0.14	0.04	1656	68	920	36
SVM+DL	0.33	0.09	1535	84	904	157
SVM+ICA	0.15	0.04	1629	51	937	63
DT	0.26	0.08	1554	81	907	138
DT+PCA	0.48	0.03	1651	37	951	41
DT+KPCA	0.43	0.05	1614	52	936	78
DT+LLE	0.53	0.04	1638	56	932	54
DT+Isomap	0.67	0.05	1635	76	912	57
DT+DL	0.38	0.09	1557	104	884	135
DT+ICA	0.79	0.04	1647	54	934	45
RF	0.18	0.06	1632	114	874	60
RF+PCA	0.48	0.09	1642	29	959	50
RF+KPCA	0.11	0.03	1643	28	960	49
RF+LLE	0.18	0.04	1646	63	925	46
RF+Isomap	0.13	0.04	1644	62	926	48
RF+DL	0.24	0.09	1561	100	888	131
RF+ICA	0.13	0.03	1650	32	956	42
KN	0.22	0.07	1613	105	883	79
KN+PCA	0.25	0.01	1670	17	971	22
KN+KPCA	0.34	0.01	1678	12	976	14
KN+LLE	0.63	0.02	1662	29	959	30
KN+Isomap	0.45	0.04	1634	54	934	58
KN+DL	0.30	0.09	1552	110	878	140
KN+ICA	0.38	0.01	1681	24	964	11

TABLE VII. FAULT CLASSIFICATION SUMMARY RESULTS

	Log-Loss	Mean Abs.	Correct Classified	Mis-classified
LR	0.24	0.14	2571	109
LR+PCA	0.05	0.05	2675	5
LR+KPCA	0.13	0.18	2628	52
LR+LLE	0.12	0.31	2598	82
LR+Isomap	0.16	0.39	2590	90
LR+DL	0.15	0.40	2585	95
LR+ICA	0.07	0.08	2670	10
SVM	0.09	0.09	2604	76
SVM+PCA	0.09	0.09	2629	51
SVM+KPCA	0.75	0.06	2635	45
SVM+LLE	0.09	0.12	2583	97
SVM+Isomap	0.08	0.08	2612	68
SVM+DL	0.03	0.02	2669	11
SVM+ICA	0.02	0.01	2676	4
DT	0.32	0.07	2627	53
DT+PCA	0.40	0.05	2649	31
DT+KPCA	0.40	0.05	2649	31
DT+LLE	0.45	0.04	2645	35
DT+Isomap	0.62	0.07	2632	48
DT+DL	0.81	0.10	2617	63
DT+ICA	0.55	0.06	2637	43
RF	0.28	0.12	2577	103
RF+PCA	0.23	0.21	2665	15
RF+KPCA	0.20	0.07	2625	55
RF+LLE	0.12	0.07	2622	58
RF+Isomap	0.10	0.02	2664	16
RF+DL	0.21	0.07	2624	56
RF+ICA	0.14	0.02	2667	13
KN	0.09	0.04	2646	34
KN+PCA	0.12	0.02	2671	9
KN+KPCA	0.12	0.02	2671	9
KN+LLE	0.24	0.05	2654	26
KN+Isomap	0.15	0.04	2650	30
KN+DL	0.34	0.09	2617	63
KN+ICA	0.11	0.03	2660	20

IV. CONCLUSIONS

This paper studied the importance of wavelet family selection for feature extraction. Additionally, unsupervised dimensionality reduction models have been applied to the machine learning algorithms over a distributed network to improve predictions and classifications. The method developed utilizes three-phase voltages and currents as inputs to the model. Then, the zero-sequence components are calculated and implemented for wavelet selection, ground fault identification, and fault detection. The unsupervised dimensionality reduction algorithms extract the most relevant features for training the machine learning algorithms. The grid search algorithm and K-fold cross-validation were implemented to choose the appropriate hyperparameters and reduce overfitting or underfitting in the models. The simulation results prove that some of the proposed models have achieved satisfactory performance in fault detection and classification over the network studied.

ACKNOWLEDGMENT

Part of this work was presented orally at the IV. International Conference on Data Science and Applications 2021 with the title “Machine Learning Application for Fault Detection in Power Distributed Network”.

REFERENCES

[1] Karić A., Konjić T. Jahić A. Power System Fault Detection, Classification and Location using Artificial Neural Networks, In: Hadžikadić M., Avdaković S. (eds) Advanced Technologies, Systems, and Applications II. IAT 2017. Lecture Notes in Networks and Systems, vol 28. Springer, Cham. (2018) https://doi.org/10.1007/978-3-319-71321-2_8

[2] Y. Mo et al., Cyber–Physical Security of a Smart Grid Infrastructure, in Proceedings of the IEEE, vol. 100, no. 1, pp. 195-209, Jan. 2012, doi: 10.1109/JPROC.2011.2161428.

[3] H. Okumuş and F.M. Nuroglu, Wavelet Based Fault Detection and Classification Algorithm for a Real Distribution Feeder. EMITTER International Journal of Engineering Technology. 2019. 7. 10.24003/emitter.v7i1.382.

[4] Jamil, M., Sharma, S.K. and Singh, R. Fault detection and classification in electrical power transmission system using artificial neural network. SpringerPlus 4, 334 (2015). https://doi.org/10.1186/s40064-015-1080-x

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

[6] Dehghani, Farzad & Khodnia, Fereydoun & Dehghan, Esfandiar. Fault location of unbalanced power distribution feeder with distributed generation using neural networks. CIRED - Open Access Proceedings Journal. 2017. 1134-1137. 10.1049/oap-cired.2017.0007.

[7] N. Shahid, S. A. Aleem, I. H. Naqvi and N. Zaffar, Support Vector Machine based fault detection & classification in smart grids, 2012 IEEE Globecom Workshops, Anaheim, CA, USA, 2012, pp. 1526-1531, doi: 10.1109/GLOCOMW.2012.6477812.

[8] A. V. Masa, S. Werben and J. C. Maun, Incorporation of data-mining in protection technology for high impedance fault detection, 2012 IEEE Power and Energy Society General Meeting, San Diego, CA, USA, 2012, pp. 1-8, doi: 10.1109/PESGM.2012.6344893.

[9] S. Samantaray, D. Mishra and G. Joos, A Combined Wavelet and Data-Mining Based Intelligent Protection Scheme for Microgrid, 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 2018, pp. 1-1, doi: 10.1109/PESGM.2018.8586480.

[10] Chen, K., Huang, C. and He, J., Fault detection, Classification and location for transmission lines and distribution systems: a review on the methods. High Voltage,1:25,33,2016.https://doi.org/10.1049/hve.2016.0005.

[11] C. Rudin et al., Machine Learning for the New York City Power Grid, in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 2, pp. 328-345, Feb. 2012, doi: 10.1109/TPAMI.2011.108.

[12] T. S. Abdelgayed, W. G. Morsi and T. S. Sidhu, Fault Detection and Classification Based on Co-training of Semisupervised Machine Learning, in IEEE Transactions on Industrial Electronics, vol. 65, no. 2, pp. 1595-1605, Feb. 2018, doi: 10.1109/TIE.2017.2726961.

[13] C., Hwan Kim and R. Aggarwal, Wavelet transforms in power systems. II. Examples of application to actual power system transients, in Power Engineering Journal, vol. 15, no. 4, pp. 193-202, Aug. 2001, doi: 10.1049/pe:20010404.

[14] H. Okumuş And F. M. Nuroğlu, Power System Event Classification Based on Machine Learning, 3rd International Conference on Computer Science and Engineering (UBMK), Sarajevo, Bosnia And Herzegovina, pp.402-405, 2018.

[15] N. Toma, R. and Kim, J.-M. Bearing Fault Classification of Induction Motors Using Discrete Wavelet Transform and Ensemble Machine Learning Algorithms. Appl.Sci. 2020, 10,5251.https://doi.org/10.3390/app10155251Qwqwq

[16] R., Papia and M., Debani, Support vector machine-based fault classification and location of a long transmission line. Engineering Science and Technology, an International Journal. 2016, 19. 10.1016/j.jestch.2016.04.001.

[17] Guo, Mou-Fa, Z. Xiao-Dan, C. Duan-Yu and Y. Nien-Che. Deep-Learning-Based Earth Fault Detection Using Continuous Wavelet Transform and Convolutional Neural Network in Resonant Grounding Distribution Systems. IEEE Sensors Journal.2017, PP. 1-1. 10.1109/JSEN.2017.2776238.

[18] Ray, P., Panigrahi, B.K. and Senroy, N., Hybrid methodology for fault distance estimation in series compensated transmission line. IET Gener. Transm. Distrib.,7:431-439.2013,https://doi.org/10.1049 /iet-gtd.2012.0243

[19] J. Jiang et al., A Hybrid Framework for Fault Detection, Classification, and Location—Part I: Concept, Structure, and Methodology, in IEEE Transactions on Power Delivery, vol. 26, no. 3, pp. 1988-1998, July 2011, doi: 10.1109/TPWRD.2011. 2141157.Qwqw

[20] M. Davoudi, V. Cecchi and J. R. Agüero, Effects of stiffness factor on bus voltage variations in the presence of intermittent distributed generation, 2015 North American Power Symposium (NAPS), Charlotte, NC, USA, 2015, pp. 1-6, doi: 10.1109/NAPS.2015.7335187.

[21] R. C. Dugan and W. H. Kersting, Induction machine test case for the 34-bus test feeder -description, 2006 IEEE Power Engineering Society General Meeting, Montreal, QC, Canada, 2006, pp. 4 pp.-, doi: 10.1109/PES.2006.1709506.

[22] K. Balamurugan, D. Srinivasan, and T. Reindl, Impact of Distributed Generation on Power Distribution Systems, Energy Procedia, Volume 25, 2012, Pages 93-100,ISSN1876-6102, https://doi.org/10.1016/ j.egypro.2012.07.013.Sfds

[23] Sasa Mujovic, Snezana Vujosevic and Luka Vujosevic (2018) Zero-Sequence Voltage-based Method for Determination and Classification of Unloaded Overhead Line Operating Conditions at the Moment of Energization, Electric Power Components and Systems,46:2,162-176,DOI:10.1080/15325008.201 8.1433252

[24] Raschka S. and Mirjalili V., Python Machine Learning: Machine Learning and Deep Learning with Python, Scikit-learn, and TensorFlow, Packt Publishing,2017,ISBN1787125939/9781787125933

[25] Ankur A. Patel, Hands-On Unsupervised Learning Using Python, O'Reilly Media, Inc., 2019, ISBN9781492035640

[26] https://scikit-learn.org/stable/index.html (Accessing date: 9 Feb 2021)

[27] https://medium.com/ (Accessing date: 14 Feb 2021)

[28] https://www.mathworks.com/help/ (Accessing date: 20 Feb 2021)