Two-Tier Cascaded Classifiers to Improve Electrical Power Quality

Wisam Elmasry
Computer Engineering Department, Istanbul Kultur University, Istanbul, Turkey
w.elmasry@iku.edu.tr

Mohammed Wadi
Electrical & Electronics Engineering Department, Istanbul Sabahattin Zaim University, Istanbul, Turkey
mohammed.wadi@izu.edu.tr

Hossein Shahinzadeh
Department of Electrical Engineering, Amirkabir University of Technology, Tehran, Iran
h.s.shahinzadeh@ieee.org

2022 26th International Electrical Power Distribution Conference (EPDC) | 978-1-6654-5183-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/EPDC56235.2022.9817239

***Abstract*— Although the manifestation of faults in electrical distribution networks is likely scarce, they deemed to be one of the most severe threats which encounter the reliability and stability of power systems. Thereby, most of research efforts in the last decade have been directed to characterize the faults and propose detection systems. However, the state-of-art of these systems still undergo serious shortcomings such as automation, integration and validation. To surpass these shortcomings, a two-tier system is proposed for both fault detection and fault classification in electrical distribution networks. Prior to deal with faults, the captured voltage signals pass through a preparatory unit to process the raw voltage signals. Afterwards, the optimized classifiers detect and classify the faults in the first and second tiers, respectively. Furthermore, the performance of the classifiers in both tiers is validated using the VSB dataset (real-time data for fault detection). Finally, the proposed two-tier system proved its efficiency in detecting electrical faults and their type as well.**

Keywords*— *Electrical fault detection, Electrical fault classification, Feature and hyperparameters optimization, Electrical Power quality, Binary classifiers, Distribution networks.

I. INTRODUCTION

During the last decades, the attention has been turned to the rapid growth of electrical power grids in order to meet the increasing demands on electricity consumption. This raising trend poses serious threats and emerging challenges to all sections of modern power systems [1]. The most vulnerable section to such threats and challenges is the distribution network [2]. The distribution network in the electrical power system acts as a nerve in human body where signals are delivered to end-points. Therefore, preserving the stability and reliability of a distribution network is the most desired aim of any long-term planning process. This does not only improve the quality of electrical power but also it increases the contentment of users [1-2].

One of the most well-known conditions, which lead to unstable state in the distribution networks, is the unexpected occurrence of faults [3]. Basically, fault in the distribution networks is a case where the values of voltage or current exceed either the upper or lower level of acceptable normal values. Even though faults in the electrical distribution networks might occur unintentionally by human errors, environment bad conditions or equipment failure, they have destructive effects on both electrical devices and networks [4]. Therefore, prediction of these faults as soon as they appeared in the distribution networks could prevent such loses in running expenses and damages in infrastructures [3] [4].

Practically, the electrical power grids usually use a central system for monitoring and controlling the distribution networks. In other words, the main purpose of such systems in their entire life is to perform three different processes, namely, fault detection, fault classification, and fault localization in order to detect the occurrence of faults, detect the type of predicted fault, and detect the fault location, respectively [5]. The first two processes are the most critical and they are performed sequentially. On the other hand, the rapid evolution of computing systems helps the artificial intelligence to dominate in a wide range of practical applications [6]. Among of these applications is to automate the fault detection or fault classification problems [7]. Thereby, in the literature, there are dozens of studies which researched the faults in power grids [8-18]. In their works, they simulated electrical faults in power grids and utilized these simulated data to train and test one or more of machine learning methods. However, recently it has been reported that the majority of them have severe defect in three vital operations, namely, processing, validation, and automation [3-5] [7]. Processing in this context means that the raw data must be processed efficiently prior to be utilized in the training of classification methods [19]. On the other hand, the validation refers to which extent the used data and methods approach the real-world electrical power grids [17]. Moreover, the automation determines how the selecting of operating parameters and taking actions are taken place [18]. After deep understanding of fault occurrence in the electrical distribution networks and weakness points of previous studies, a two-tier cascaded classifiers are combined together in one system for both fault detection and classification. Specifically, a preparatory unit is added prior to the proposed system to analyze and process the raw data sufficiently. Furthermore, real-time data are utilized along with several widely-used machine learning classifiers to resolve the validation task efficiently. Moreover, many optimization methods are exploited to automate the selection of features and hyperparameters. All aforementioned merits of the proposed system are the major contributions of our study as illustrated in Figure 1.

The following sections of our paper are organized as follows. Section II shows the mechanism of the preparatory unit for data pre-processing. The full details of tier 1 of the proposed system are explained in Section III. Similarly, in Section IV, the function of tier 2 is described. Then, Section V discusses the obtained results using various analyses. Finally, the conclusions as well as the suggested future work are drawn in Section VI.

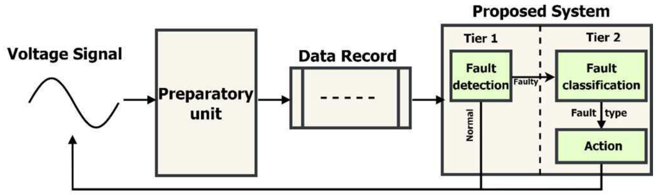

Fig. 1. The proposed two-tier system.

II. PREPARATORY UNIT

The proposed system in Figure 1 comprises of a preparatory unit followed by two successive tiers. Whereas the first tier performs electrical fault detection, the second tier completes the task by performing fault classification. It can be noticed that the preparatory unit is included as a preceding entry point to the proposed system. The aim of the preparatory unit is to process the raw data of electrical signals (either current or voltage signals) for being used with various machine learning models. This unit plays very important role in entire system since it converts the analog captured signal to a record of numerical features which helps the classifiers to grasp the characteristics of faults in distribution networks. In order to perform its duty to the fullest, the preparatory unit manages four consecutive phases, namely, data configuration, data filtering, feature extraction, and data normalization, as follows. Figure 2 depicts the structure of the proposed preparatory unit.

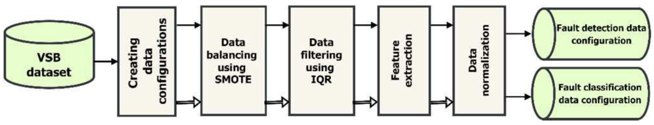

Fig. 2. Structure of the preparatory unit and its four phases.

A. Dataset and Configurations

With the aim of solving the scarcity of real-time data for training and testing classifiers in electrical fault detection or classification, the ENET research and development center [20] at Technical Ostrava University created the VSB dataset for a purpose of fault detection and classification. Unlike the simulated datasets, the VSB dataset contains real-time voltage measurements as integers which are captured by a sophisticated device plugged on the distribution lines of a real-world electrical power grid which has three phases running on 50 Hz. Originally, the VSB dataset has 8712 data points in which each of them consists of 800 thousands of voltage measurements for one complete cycle (20 ms) of a voltage signal. Bearing in mind these characteristics of the VSB dataset and it is also available publicly [21], it has been chosen to be the validation data in this study.

Since the proposed system performs fault detection and fault classification, two different data configurations must have been constructed to fit each of the two tasks. Regarding the fault detection task in the first tier, a data configuration of all samples in the VSB dataset is created where each sample accompanied with a class label has a value of 0 or 1 for the normal or faulty signals, respectively. However, the fault detection data configuration has a structural problem that only 6% of samples are faulty signals whereas the overwhelming rest samples are normal ones. Obviously, using such an unbalanced data configuration in fault detection will result in misleading results. Accordingly, a balancing process is applied to the fault detection data configuration by using the Synthetic Minority Oversampling Technique (SMOTE) [22]. The SMOTE oversampled the number of faulty samples until the final class distribution of the two classes is balanced.

On the other hand, the data configuration for fault classification in the second tier is created by considering only the faulty samples in the original VSB dataset. Subsequently, each three successive faulty samples are combined together in one data record to represent one of fault types. This can be explained by the fact that each sample in the VSB dataset includes voltage measurements for one phase of the power grid, in which three sequential samples are belong to a same voltage signal. As a result of that, each data record in the fault classification data configuration consists of 2.4 millions of voltage measurements accompanied with a class label of one of 7 fault types. Likewise, the SMOTE algorithm is utilized to balance the class distribution of all available seven classes.

B. Data Filtering

After creating data configurations, the values of voltage measurement in each sample need to be cleaned because when capturing voltage signals some noise measurements may be recorded [23]. Obviously, using data containing noise in training or testing machine learning classifiers will lead to poor results in fault detection and classification [24]. Hence, data filtering operation must be applied to data configurations to remove all noise values. In this paper, the Interquartile Range (IQR) technique is utilized to filter any possible noise in the data configurations. The reasons behind selecting this technique are its simplicity, generality, and robustness. Indeed, the IQR technique filters each sample in data configurations by executing five consecutive steps. Firstly, it arranges voltage measurements in an ascending order. Afterwards, it computes the IQR value by subtracting the first quartile (Q1) value of the voltage measurements from the third quartile (Q3) value of the voltage measurements. In the third step, the threshold value is computed by multiplying the value of IQR by 3 which is an adjustment factor. Thereafter, two boundary values, namely, lower bound (LB) and upper bound (UB) are calculated by subtracting the threshold value from Q1 for LB and adding the threshold value to Q3 for UB. Finally, the IQR technique removes any voltage measurement value that less than LB or greater than UB. The following equations demonstrate the IQR technique mathematically [17-18]:

$$IQR = Q3 - Q1 \tag{1}$$

$$Threshold = 3 * IQR \tag{2}$$

$$LB = Q1 - Threshold \tag{3}$$

$$UB = Q3 + Threshold \tag{4}$$

C. Feature Extraction

In this phase, the data configurations are clean from noise but they still have relatively a high number of features (thousands in fault classification data configuration and millions in fault classification data configuration). Such high-dimensional feature spaces are inadequate to be used with shallow machine learning classifiers in terms of performance. Thus, shrinking the feature space into a reasonable number of features is very vital process before proceeding in training and testing classifiers. This can be accomplished by extracting some useful and representative information from values of voltage measurement instead of using the values themselves. In this study, 19 of statistical information are extracted from each sample in the data configurations, namely, mean, standard deviation, maximum value, minimum value, (1%, 25%, 50%, 75%, and 99%) percentile values, (0%, 1%, 25%, 50%, 75%, 99%, and 100%) relative percentile values, lower band, upper band, and height [17] [18]. The following equations show how the values of relative percentile, lower band, upper band, and height are computed whereas the rest values are considered as a basic knowledge [17-18]:

$$Relative\_Percentile = Percentile - Mean \tag{5}$$

$$Lower\_Band = Mean - Standard\_deviation \tag{6}$$

$$Upper\_Band = Mean + Standard\_deviation \tag{7}$$

$$Height = Maximum - Minimum \tag{8}$$

Regarding the fault detection data configuration, it has 20 features for each sample, i.e., 19 extracted features and the class label. On the other hand, the fault classification data configuration has 58 features for each sample, i.e., 19 extracted features from each of the three phases of voltage signal and one class label.

D. Data Normalization

After finishing data balancing, data cleaning, and feature extraction, each sample in the data configurations must be normalized in a range of [0,1] in order to enable the classifiers to deal with them in a short domain. This can be accomplished using several data normalization methods, but among them the Minimum-Maximum (Min-Max) transformation is the mostly used in the literature. The basic idea of the Min-Max transformation is very obvious that each numerical feature (x) except the class label is subtracted from the minimum value of that feature and divided on the height value of that feature using Equation (9) [17-18]:

$$x = \frac{x - Min}{Max - Min} \tag{9}$$

III. TIER I

The first tier of the proposed system is dedicated to electrical fault detection. In other words, a record from the fault detection data configuration is inputted to the classifiers of the first tier in order to find out whether it represents a normal voltage signal or a faulty one. This task in the first tier can be accomplished by performing a binary classification, that is, the classifiers output ”0” if the sample is normal or ”1” otherwise. Figure 3 shows the methodology of the first tier. In addition to that, six well-known binary classification models will be exploited in parallel to detect faults, they are: Two Class Artificial Neural Network (TC-ANN) [6], Two Class-Naïve Bayes (TC-NB) [25], Two Class-Quantum Support Vector Machine (TC-QSVM) [26], Two Class-Boosted Decision Tree (TC-BDT) [27], Two Class-Decision Forest (TC-DF) [28], and Two Class-Decision Jungle (TC-DJ) [29].

Indeed, the are two common problems when training and testing machine learning classifiers, to be more specific, the existence of any redundant or irrelevant features in the feature space and the best value of hyperparameters for the used models that fit the underlying task. Solving these problems prior to training and testing is indispensable. Therefore, swarm intelligence metaheuristics are usually used in such a case. In this paper, the double Particle Swarm Optimization (PSO) algorithm [30] is utilized to optimize both features and hyperparameters in one process. The double PSO algorithm comprises of two levels: the upper level for feature selection using the Fitness Proportionate Selection Binary Particle Swarm Optimization and Entropy (FPSBPSO-E) method [30], and the lower level for hyperparameter selection using the PSO based method [31]. To begin with the FPSBPSO-E method, it uses the FPSBPSO [32] as the core of optimization algorithm and the single-objective filter-based entropy method [33] on the top of it in order to determine the optimal feature subset that minimizes the redundancy and maximizes the relevancy between the selected feature subset and the set of classes. On the other hand, the PSO-based method iterates until finding the optimal hyperparameters which maximizes the accuracy value of a particular machine learning model on the given training set.

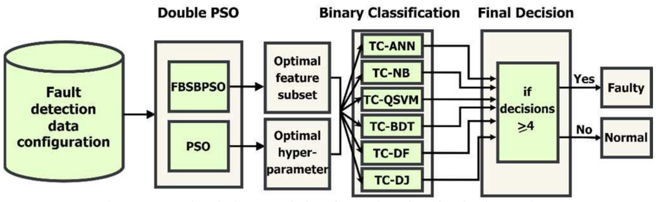

Fig. 3. Methodology of the first tier for fault detection.

To put all together, the fault detection data configuration is used along with the double PSO algorithm to find out the optimal feature subset and the optimal hyperparameter set for each of the used binary classification models. Thereby, every binary classification model will be tuned using its optimal hyperparameters and then will be trained and tested on the training and test sets of the optimal feature subset, respectively. The final decision of all binary classification models will be normal/faulty sample if and only if at least four of them decided it as a normal/faulty.

IV. TIER II

The second tier of the proposed system performs electrical fault classification only when a fault is detected in the first tier. In other words, as soon as the fault is detected in tier 1, the corresponding sample of that voltage signal in the fault classification data configuration will be forwarded to classifiers in tier 2 in order to specify the class of that fault, as illustrated in Figure 4. The fault classification process can be interpreted as a multiclass classification task.

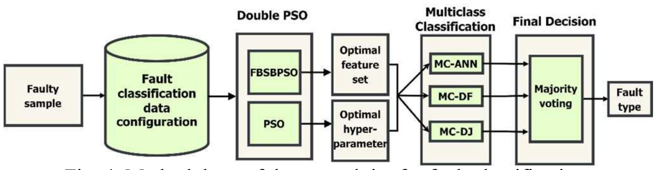

Fig. 4. Methodology of the second tier for fault classification.

In this study, the fault classification data configuration is created for a 3-phase electrical power grid and in this case seven possible classes of fault types are existed in the ”Class” label. Let A, B, C, and G denote the phase 1, phase 2, phase 3, and ground terminal, respectively. Then, the seven classes have the following labels: ”A,AG”, ”B,BG”, ”C,CG”, ”AB,ABG”, ”AC,ACG”, ”BC,BCG”, and ”ABC,ABCG”. It can be noticed that a fault in lines-only or in these lines with the ground is same, for instance, the fault in line ”A” is as same as the ”AG” fault. Furthermore, these seven fault types can be categorized in two main groups, namely, symmetrical fault (ABC, ABCG) and unsymmetrical faults (A,AG, B,BG, C,CG, AB,ABG, AC,ACG,and BC,BCG). The difference between symmetrical and unsymmetrical faults that the symmetrical faults are dangerous, rare, and stable whereas the unsymmetrical faults are safe, common, and unstable. Table I presents each class label and the corresponding states in the three phases.

TABLE I. THE SEVEN CLASSES OF THE FAULT TYPE.

Phase 1	Phase 2	Phase 3	Class
1	0	0	A,AG
0	1	0	B,BG
0	0	1	C,CG
1	1	0	AB,ABG
1	0	1	AC,ACG
0	1	1	BC,BCG
1	1	1	ABC,ABCG

Three multiclass classifiers are exploited in the second tier, namely, Multi Class-Artificial Neural Network (MCANN) [34], Multi Class-Decision Forest (MC-DF) [34] [35], and Multi Class-Decision Jungle (MC-DJ) [34-36]. Likewise, the fault classification data configuration is used along with the double PSO algorithm in order to find out the optimal feature subset as well as the optimal hyperparameters for the particular multiclass classifier. Afterwards, every multiclass classifier is constructed using its optimal hyperparameters and then it is trained and tested using the training and test sets of the optimal feature subset, respectively. The final decision will be considered as the majority voting for all multiclass classifiers if the majority for a particular class is guaranteed. Otherwise, when the majority is not existed, the decision of the MC-DJ classifier will be considered as the final decision.

V. RESULTS & DISCUSSION

To ensure high levels of computations and various resources, the Azure Machine Learning (AML) studio [37] is exploited as a cloud-based environment for designing and execution of our empirical experiments. In addition of that, all phases of the preparatory unit as well as the double PSO algorithm are implemented using the Python programming language and the NumPy library. It is worthy to mention that the creation of training and test sets of the two data configurations followed the method of a stratified 10-fold cross-validation. This means that each fold of the cross-validation split is balanced taking into consideration the class distribution.

In order to evaluate the performance of the used models, the outcome of the testing process is an array of four numbers, namely, True Negative (TN), True Positive (TP), False Positive (FP), and False Negative (FN). This array is deemed to be the key factor for computing various evaluation metrics. In this study, only the four standard evaluation metrics (Accuracy, Precision, Recall, and F1-Score) are calculated for the fault detection in Tier 1 using the following equations [31]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

$$Precision = \frac{TP}{TP+FP} \tag{11}$$

$$Recall = \frac{TP}{TP+FN} \tag{12}$$

$$F1\ Score = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{13}$$

On the other hand, TP, TN, FP, and FN values in the multiclass classification have a different meaning from those in the binary classification [34]. The distinction come from the fact that these values in the multiclass classification are the summation for each individual value of all available classes [34]. For example, the TP value in the multiclass classification in tier 2 is the summation of all TPs for the seven classes of fault type. The same concept is applicable to the other values. In this paper, only four basic evaluation metrics for multiclass classification (Macro-Averaged Accuracy, Macro-Averaged precision, Macro-Averaged Recall, and Macro-Averaged F1-Scaore) are calculated using the following equations [30]:

$$Accuracy_M = \frac{\sum_{i=1}^{7}\frac{tp_i+tn_i}{tp_i+tn_i+fp_i+fn_i}}{7} \tag{14}$$

$$Precision_M = \frac{\sum_{i=1}^{7}\frac{tp_i}{tp_i+fp_i}}{7} \tag{15}$$

$$Recall_M = \frac{\sum_{i=1}^{7} tp_i/tp_i+fn_i}{7} \tag{16}$$

$$F1\_Score_M = \frac{2\times Precision_M \times Recall_M}{Precision_M + Recall_M} \tag{17}$$

A. Fault Detection

Table II presents the results of fault detection in tier 1 of the proposed system in terms of percentage of binary classification evaluation metrics. Moreover, the comparison in Table II is between two cases: when using the proposed system and without it where the best results among all used models are highlighted as bold values. Obviously, the results confirm the significance of using the proposed system in the electrical fault detection where all the used metrics is enhanced by 27% to 29% if they are compared to the corresponding values without using the proposed system. Furthermore, the TC-DJ outperformed all the other binary classifiers in terms of all metrics and in all cases. This may imply that using the TC-DJ classifier is the best choice for fault detection process in tier 1.

Receiver Operating Characteristic (ROC) analysis is deemed to be an alternative method to asses the performance of the used models visually [38]. It draws a curve for each model by plotting the recall values as a function of the False Positive Rate (FPR) values [39], where the FPR metric can be calculated using the following equation:

$$FPR = \frac{FP}{FP+TN} \tag{18}$$

Figure 5 shows the ROC analysis of the used binary models when using the proposed system for fault detection in tier 1.

TABLE II. RESULTS OF FAULT DETECTION IN TIER 1.

Case	Metric	TC-ANN	TC-NB	TC-QSVM	TC-BDT	TC-DF	TC-DJ
Without proposed system	Accuracy	63.75	64.94	65.80	68.67	70.01	**70.71**
	Precision	69.32	69.72	70.30	71.59	72.53	**73.03**
	Recall	56.66	58.92	60.21	65.16	67.05	**68.03**
	F1-score	62.51	63.98	64.96	68.26	69.71	**70.46**
With proposed system	Accuracy	90.17	91.36	93.86	96.30	96.88	**97.34**
	Precision	95.78	96.17	97.97	98.16	98.45	**98.78**
	Recall	85.90	86.91	89.73	93.38	95.50	**96.30**
	F1-Score	90.61	91.51	93.93	95.79	96.66	**97.34**

In addition to that, the Area Under ROC (AUC) is considered as a performance metric that can interpret the ROC curve into a numeric value in [0,1] and it can be computed approximately using Equation (19). Table III presents the AUC values of all ROC curves in Figure 5. Notably, all used binary models are fit to the fault detection in tier 1 with the following descending order: TC-DJ, TC-DF, TC-BDT, TC-QSVM, TC-NB, and TCANN.

$$AUC = \frac{1}{2} \times (Recall + Specificity) \tag{19}$$

$$Specificity = \frac{TN}{TN + FP} \tag{20}$$

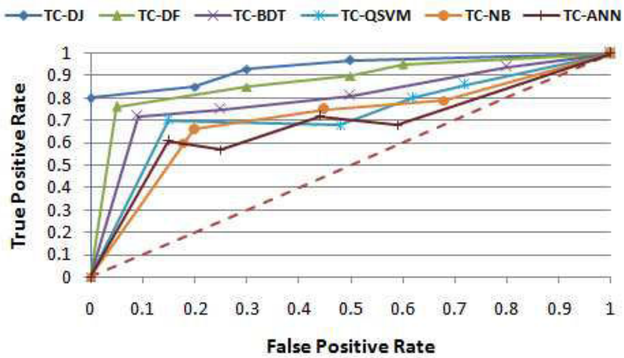

Fig. 5. ROC curve analysis for fault detection in tier 1.

TABLE III. AUC VALUES OF THE ROC CURVES ANALYSIS.

Binary model	AUC
TC-ANN	0.8657
TC-NB	0.8769
TC-QSVM	0.8951
TC-BDT	0.9260
TC-DF	0.9389
TC-DJ	0.9542

B. *Fault Classification*

Table IV shows the results of fault classification in tier 2 of the proposed system in terms of percentage of multiclass classification evaluation metrics. Furthermore, the comparison in Table IV is between two cases: when using the proposed system and without it where the best results among all used models are highlighted as bold values. Obviously, the results confirm the effectiveness of using the proposed system in the electrical fault classification where all the used metrics is increased by 16% to 21% if they are compared to the corresponding values without using the proposed system. Furthermore, the MC-DJ outperformed all the other multiclass classifiers in terms of all metrics and in all cases. This may imply that utilizing the MC-DJ classifier is the best choice for fault classification process in tier 2.

To show performance differences between the used multiclass classifiers, a comparison in the basis of Detection Rate (DR) percentage for each fault type is presented in Figure 6. From Figure 6, we can conclude that the MC-DJ classifier is superior to other classifiers in detecting fault types. However, all the used models are fit to the fault classification in tier 2 with the following descending order: MC-DJ, MC-DF, and MC-ANN.

TABLE IV. RESULTS OF FAULT CLASSIFICATION IN TIER 2.

Case	Metric	MC-ANN	MC-DF	MC-DJ
Without proposed system	$Accuracy_M$	70.65	73.36	**75.71**
	$Precision_M$	73.88	78.96	**79.68**
	$Recall_M$	69.11	74.44	**77.45**
	$F1\text{-}score_M$	72.32	76.87	**79.91**
With proposed system	$Accuracy_M$	91.35	95.28	**96.76**
	$Precision_M$	93.22	98.14	**98.46**
	$Recall_M$	89.69	93.04	**95.78**
	$F1\text{-}Score_M$	90.37	95.17	**96.65**

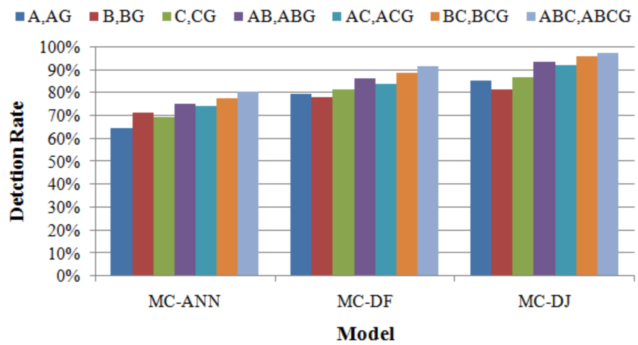

Fig. 6. Comparison between multiclass classifiers in detecting fault types.

VI. CONCLUSION

In this study, a deep investigation of electrical faults in the distribution networks is introduced. The integrity, preprocessing, validation, and automation are the common problems in the existing fault detection and classification systems. Accordingly, novel two-tier cascaded classifiers are proposed to overcome these problems and enhance the overall performance. The proposed system has a preparatory unit to process the raw data by performing four successive operations, namely, data configuration, data filtering, feature extraction, and data normalization. Afterwards, the processed data record is forwarded to first tier for fault detection and if the fault is detected it will be forwarded to the second tier for fault classification. Furthermore, a double PSO algorithm is utilized to optimize the feature set and hyperparameter values. Moreover, six binary classifiers and three multiclass classifiers are exploited to accomplish fault detection and fault classification, respectively. All these models are validated on the VSB dataset as realistic and modern data for fault detection and classification. Finally, the experimental results show a significant enhancement in both fault detection and fault classification when the used models benefited from the proposed system.

REFERENCES

[1] Wadi, M., & Elmasry, W. (2021). Statistical analysis of wind energy potential using different estimation methods for Weibull parameters: a case study. *Electrical Engineering*, *103*(6), 2573-2594.

[2] Wadi, M., & Elmasry, W. (2021, March). Modeling of wind energy potential in marmara region using different statistical distributions and genetic algorithms. In *2021 International Conference on Electric Power Engineering–Palestine (ICEPE-P)* (pp. 1-7). IEEE.

[3] Chen, K., Huang, C., & He, J. (2016). Fault detection, classification and location for transmission lines and distribution systems: a review on the methods. *High voltage*, *1*(1), 25-33.

[4] Raza, A., Benrabah, A., Alquthami, T., & Akmal, M. (2020). A review of fault diagnosing methods in power transmission systems. *Applied Sciences*, *10*(4), 1312.

[5] Prasad, A., Edward, J. B., & Ravi, K. (2018). A review on fault classification methodologies in power transmission systems: Part—I. *Journal of electrical systems and information technology*, *5*(1), 48-60.

[6] Elmasry, W., Akbulut, A., & Zaim, A. H. (2020). Comparative evaluation of different classification techniques for masquerade attack detection. *International Journal of Information and Computer Security*, *13*(2), 187-209.

[7] Prasad, A., Edward, J. B., & Ravi, K. (2018). A review on fault classification methodologies in power transmission systems: Part-II. *Journal of Electrical Systems and Information Technology*, *5*(1), 61-67.

[8] Jamil, M., Sharma, S. K., & Singh, R. (2015). Fault detection and classification in electrical power transmission system using artificial neural network. *SpringerPlus*, *4*(1), 1-13.

[9] Koley, E., Verma, K., & Ghosh, S. (2015). An improved fault detection classification and location scheme based on wavelet transform and artificial neural network for six phase transmission line using single end data only. *SpringerPlus*, *4*(1), 1-22.

[10] Uzubi, U., Ekwue, A., & Ejiogu, E. (2017, June). Artificial neural network technique for transmission line protection on Nigerian power system. In *2017 IEEE PES PowerAfrica* (pp. 52-58). IEEE.

[11] Eboule, P. S. P., Pretorius, J. H. C., Mbuli, N., & Leke, C. (2018, October). Fault detection and location in power transmission line using concurrent neuro fuzzy technique. In *2018 IEEE Electrical Power and Energy Conference (EPEC)* (pp. 1-6). IEEE.

[12] Gururajapathy, S. S., Mokhlis, H., & ILLIAS, H. A. B. (2018). Classification and regression analysis using support vector machine for classifying and locating faults in a distribution system. *Turkish Journal of Electrical Engineering & Computer Sciences*, *26*(6), 3044-3056.

[13] Dong, M., Sun, Z., & Wang, C. (2019, May). A pattern recognition method for partial discharge detection on insulated overhead conductors. In *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)* (pp. 1-4). IEEE.

[14] Qu, N., Li, Z., Zuo, J., & Chen, J. (2020). Fault detection on insulated overhead conductors based on DWT-LSTM and partial discharge. *IEEE Access*, *8*, 87060-87070.

[15] Wadi, M., & Elmasry, W. (2021, March). An anomaly-based technique for fault detection in power system networks. In *2021 International Conference on Electric Power Engineering–Palestine (ICEPE-P)* (pp. 1-6). IEEE.

[16] Wadi, M. (2021). Fault detection in power grids based on improved supervised machine learning binary classification. *Journal of Electrical Engineering*, *72*(5), 315-322.

[17] Elmasry, W., & Wadi, M. (2022). EDLA-EFDS: A Novel Ensemble Deep Learning Approach For Electrical Fault Detection Systems. *Electric Power Systems Research*, *207*, 107834.

[18] Elmasry, W., & Wadi, M. (2022). Enhanced Anomaly-Based Fault Detection System in Electrical Power Grids. *International Transactions on Electrical Energy Systems*, *2022*.

[19] Elmasry, W., Akbulut, A., & Zaim, A. H. (2021). A design of an integrated cloud-based intrusion detection system with third party cloud service. *Open Computer Science*, *11*(1), 365-379.

[20] ENET Centre "University of Technical Ostrava http://cenet.vsb.cz/en/, 2021.

[21] VSB Power Line Fault Detection , "Kaggle," https://www.kaggle.com/c/vsb-power-line-fault-detection/data, 2018.

[22] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.

[23] Naveed, K., Akhtar, M. T., Siddiqui, M. F., & ur Rehman, N. (2021). A statistical approach to signal denoising based on data-driven multiscale representation. *Digital Signal Processing*, *108*, 102896.

[24] García-Gil, D., Luengo, J., García, S., & Herrera, F. (2019). Enabling smart data: noise filtering in big data classification. *Information Sciences*, *479*, 135-152.

[25] Zeineddine, H., Braendle, U., & Farah, A. (2021). Enhancing prediction of student success: Automated machine learning approach. *Computers & Electrical Engineering*, *89*, 106903.

[26] Havlíček, V., Córcoles, A. D., Temme, K., Harrow, A. W., Kandala, A., Chow, J. M., & Gambetta, J. M. (2019). Supervised learning with quantum-enhanced feature spaces. *Nature*, *567*(7747), 209-212.

[27] Two-Class Boosted Decision Tree, "Microsoft docs," https://docs.microsoft.com/en-us/azure/machine-learning/studio module-reference/two-class-boosted-decision-tree, 2019.

[28] Two-Class Decision Forest, "Microsoft docs," https://docs.microsoft.com/en-us/azure/machine-learning/studiomodule-reference/two-class-decision-forest, 2019.

[29] Two-Class Decision Jungle, "Microsoft docs," https://docs.microsoft.com/en-us/azure/machine-learning/studiomodule-reference/two-class-decision-jungle, 2019.

[30] Elmasry, W., Akbulut, A., & Zaim, A. H. (2020). Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic. *Computer Networks*, *168*, 107042.

[31] Elmasry, W., Akbulut, A., & Zaim, A. H. (2018). Deep learning approaches for predictive masquerade detection. *Security and Communication Networks*, *2018*.

[32] Zhou, Z., Liu, X., Li, P., & Shang, L. (2014, December). Feature selection method with proportionate fitness based binary particle swarm optimization. In *Asia-Pacific Conference on Simulated Evolution and Learning* (pp. 582-592). Springer, Cham.

[33] Cervante, L., Xue, B., Zhang, M., & Shang, L. (2012, June). Binary particle swarm optimisation for feature selection: A filter based approach. In *2012 IEEE Congress on Evolutionary Computation* (pp. 1-8). IEEE.

[34] Elmasry, W., Akbulut, A., & Zaim, A. H. (2019). Empirical study on multiclass classification-based network intrusion detection. *Computational Intelligence*, *35*(4), 919-954.

[35] Rokach, L. (2016). Decision forest: Twenty years of research. *Information Fusion*, *27*, 111-125.

[36] Shotton, J., Sharp, T., Kohli, P., Nowozin, S., Winn, J., & Criminisi, A. (2013). Decision jungles: Compact and rich models for classification. *Advances in Neural Information Processing Systems*, *26*.

[37] Azure Machine Learning Studio, "Microsoft," https://studio.azureml.net/, 2022.

[38] Sahmoud, S., Elmasry, W., & Abudalfa, S. (2013). Enhancement the Security of AES Against Modern Attacks by Using Variable Key Block Cipher. *Int. Arab. J. e Technol.*, *3*(1), 17-26.

[39] Elmasry, W. (2018). New LSB-based colour image steganography method to enhance the efficiency in payload capacity, security and integrity check. *Sādhanā*, *43*(5), 1-14.