SAND2021-11515C

Testing Machine Learned Fault Detection and Classification on a DC Microgrid

Samuel T. Ojetola
Electric Power System Research
Sandia National Laboratories
Albuquerque, NM, USA
sojetol@sandia.gov

Matthew J. Reno
Electric Power System Research
Sandia National Laboratories
Albuquerque, NM, USA
mjreno@sandia.gov

Jack Flicker
Renewable, Distributed Systems Integration
Sandia National Laboratories
Albuquerque, NM, USA
jdflick@sandia.gov

Daniel Bauer
Block Energy Labs
Emera Technologies LLC
Burlington, ON, Canada
daniel.bauer@emeratechnologies.com

David Stoltzfuz
Block Energy Labs
Emera Technologies LLC
Burlington, ON, Canada
david.stoltzfus@emeratechnologies.com

***Abstract*—Interest in the application of DC Microgrids to distribution systems have been spurred by the continued rise of renewable energy resources and the dependence on electric loads. However, in comparison to AC systems, the lack of natural zero crossing in DC Microgrids makes the interruption of fault currents with fuses and circuit breaker more difficult. DC faults can cause severe damage to voltage-source converters within few milliseconds, hence, the need to quickly detect and isolate the fault. In this paper, the potential for Machine Learning (ML) multi-class classifiers to identify fault type and fault resistance in a DC Microgrid is explored. The ML algorithms are trained using simulated fault data recorded from a 750 V_{DC} Microgrid modeled in PSCAD/EMTDC. The performance of the trained algorithms are tested using real fault data gathered from an operational DC Microgrid located on the Kirtland Air Force Base. The result shows that ML algorithms can detect fault with 100% accuracy, determine the fault type with 100% accuracy, and estimate the fault resistance with 99% accuracy. By performing a self-learning monitoring and decision making analysis, protection relays equipped with ML algorithms can quickly detect and isolate faults to improve the protection operations on DC Microgrids.**

***Index Terms*—DC Microgrids, fault detection, machine learning, support vector machines, neural network, decision tree.**

I. Introduction

The increasing emergence of DC loads and generation sources together with advances in power electronic technologies in modern power networks has prompted investigation into the potential benefits and application of DC Microgrids [1]–[8]. Most modern electronic circuits require a DC power supply, and distributed energy resources such as solar panels and batteries generate DC power. The application of DC microgirds as power sources for telecommunication stations [1], data centers [2] and electric vehicle systems [7] have been explored. The increased efficiency, safety of power electronics component and simplicity in the operation have all contributed to the advances of DC Microgrids.

DC Microgrids are potentially sensitive to disturbances and faults due to low inertia and converter behaviors. Disturbances on DC systems are usually due to input power variations, fluctuations in loads, temporary faults and communication failures/delays [9]. Protecting the DC Microgrid from these disturbances is a major challenge mainly because it is more difficult to interrupt fault current with fuses and circuit breakers as there is no zero crossing of current in DC systems. Additionally, because of the presence of power electronics switches, DC fault currents are associated with a high magnitude and a peak that is sustained for multiple milliseconds. Such large currents can cause severe damage to the components of converters [10]. Hence, fast detection and isolation of faults in a DC Microgrid is critical. In literature, several protection schemes for DC systems have been studied. This includes overcurrent protection, high-speed differential schemes, rate of change of current, undervoltage protection [11]–[15]. In recent years, studies have examined how machine learning (ML) based protection schemes can improve fault detection in AC systems [16], [17], but these protection schemes have not been adequately examined on DC Microgrids.

In [18], hidden markov model were proposed to discriminate between nominal transient behavior and DC arc fault behavior across a variety of conditions. Support Vector Machines (SVM) was implemented to detect faults in DC systems in [19] and [20]. These studies have presented successful and high accuracy performance results, however, they have primarily used data from only simulated models of DC systems for fault classification.

This material is based upon work supported by the U.S. Office of Electricity.

Sandia National Laboratories is a multi-mission laboratory managed and operated by the National and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract NA0003525.

By utilizing real data gathered from an actual DC Microgrid, this paper explores the potential for ML algorithms embedded in a protective relay to identify and classify faults. Five different algorithms were used to classify fault data based on fault type and fault resistance. The ML algorithms are trained using simulated data from PSCAD and their performance is tested using the data gathered from an operational DC Microgrid.

The rest of this paper is organized as follows: Section II introduces the ML classification algorithms examined for DC fault classification in this paper. The details of the DC power system model used is described in Section III. Section IV explains the distribution of the training and testing data. Section V presents the result of the study and Section VI gives a conclusion.

II. Machine Learning Algorithm

ML is a potential tool for intelligent fault detection because of its effective pattern recognition capability as well as its adaptability in systems with versatile operating conditions. There are two types of learning, supervised and unsupervised learning. In supervised learning, the algorithm learns the mapping function of the input to the output data provided. Meanwhile, in unsupervised learning, there is only an input data and the algorithm learns the underlying structure or distribution of the data. The following supervised classification algorithms are considered in this paper;

A. Support Vector Classifier

The Support vector classifier (SVC) method classifies events with the help of a linear or non-linear function. It is based on an estimation of the most appropriate function for separating data. The SVC method aims at finding a special linear line separating between classes. There is a possibility to draw this line more than once during the classification. The SVC identifies the farthest line to both classes, and thus maximum error tolerance is determined. Upon identification of training data and the border line, test data is classified based on their places in reference to the border [21].

B. Bernoulli Naive Bayes

Naive Bayes (NB) methods are a set of supervised learning algorithms based on application Bayes' theorem with the "naive" assumption of conditional independence between every pair of features given the value of the class variable [22]. Bernoulli NB implements the naive Bayes training and classification algorithms for data that is distributed according to multivariate Bernoulli distributions.

C. Decision Trees

Decision Trees (DT) are a non-parametric supervised learning method used for classification and regression. It predicts the value of a target variable by learning simple decision rules inferred from the data features. A tree can be seen as a piecewise constant approximation [22].

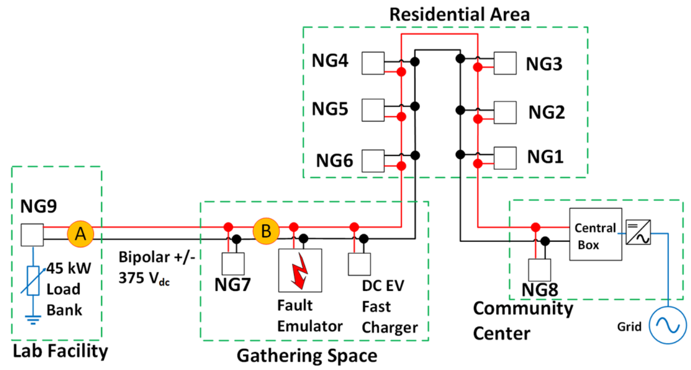


Fig. 1: ETL-KAFB DC Microgrid.

D. Nearest Centroid

The Nearest Centroid (NC) classifier is a simple classification algorithm that represents each class by the centroid of its members. It is based on the assumption that the distances between the samples belonging to the same class should be minimal. NC calculates the center (centroid) of each class and then assigns the unknown sample to the class which centroid is the closest [23].

E. Multi-layer Perception

Multi-layer Perception (MLP) is a supervised neural network model that trains using back propagation [22]. It consists of at least three layers: an input layer, a hidden layer, and an output layer. The output from the MLP structure depends on the weighted sum of its input pattern [24]. This concept of weighted sum can also be applied in every layer. MLP is sensitive to feature scaling.

III. Power system Model

This research effort uses the Emera Technologies Kirtland Airforce Base (ETL-KAFB) DC Microgrid located at the South-eastern part of Albuquerque, New Mexico. The ETL-KAFB DC Microgrid consists of a hierarchical, modular power electronics-based interface at each node that contains power conversion, control, protection and storage. The DC Microgrid serves nine nanogrids (NG) which includes a residential area of six duplex buildings, a community center, a gathering space and laboratory facilty as shown in Fig. 1. Each of the NGs include a load, a PV system, a Battery Energy Storage System (BESS) and a DC-to-DC converter to integrate the NG to the rest of the Microgrid. The size of the PV system in each NG varies from 5 kW to 13 kW while the size of the BESS are 9kWhr. The Microgrid is operated at a bipolar voltage of $\pm 375 V_{DC}$. A fault emulator installed at the gathering space of the Microgrid. The fault emulator can introduce fault of various impedance values between poles or a pole and ground on the Microgrid bus. The fault is actuated via a contactor. Fault data is collected at two locations on the Microgrid. The first batch of data is collected from a scope connected at node A, and the second batch from a scope connected to node B.

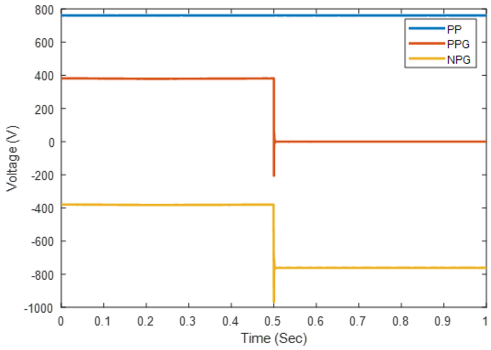


Fig. 2: Simulation data: 1.0 Ω positive pole to ground fault.

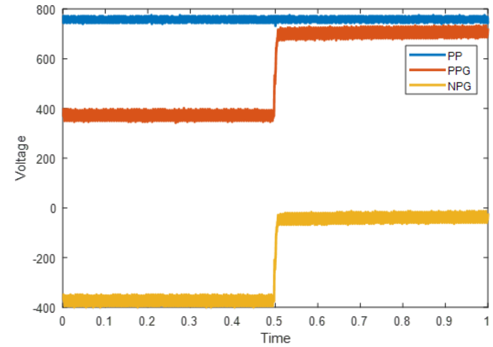


Fig. 3: Real data: 500 Ω negative pole to ground fault.

Additionally, the Emera DC Microgrid is modeled in PSCAD/EMTDC software package. Several faults conditions are simulated and the fault data from node A and node B are recorded. The simulation data gathered from PSCAD is used as training data, while the real data gathered from the field is used as the testing data for the ML classification algorithms.

A. Fault conditions

In this paper, positive pole to ground (PPG) and negative pole to ground (NPG) faults with three different fault resistances, (F_r) were considered:

- Fault 1: PPG fault with $F_r = 1.0$ ohms
- Fault 2: NPG fault with $F_r = 1.0$ ohms
- Fault 3: PPG fault with $F_r = 500$ ohms
- Fault 4: NPG fault with $F_r = 500$ ohms
- Fault 5: PPG fault with $F_r = 1000$ ohms
- Fault 6: NPG fault with $F_r = 1000$ ohms

B. Training Data

Using the PSCAD model of the DC Microgrid, six different fault conditions are simulated. The fault is applied at the location of the fault emulator shown in Fig. 1. Voltage measurement are recorded from node A and node B for each fault scenario. The total length of the simulation is 1 second, with a time step of 2.5μs. Fig. 2 shows a plot of the voltages measured at the node A after a PPG fault with a fault resistance (F_r) of 1.0 ohms is applied. The fault is applied at $t = 0.5$ sec. The measured data from the PSCAD simulation is used as the training data on the ML classification algorithms to identify and classify faults in the system.

C. Testing Data

To test the ML classification algorithms, real data gathered from the Microgrid is used as testing data. For the first batch of data gathered from node A, fault conditions 1, 2, 4, and 5 were considered. For the second batch of data gathered from node B, fault conditions 1, 2, 3, and 4 were considered. The total length of the data is 1 second with a time step of 1 μs. The plot of the voltages measured at the node A after a 1.0 Ω PPG fault is applied is shown in Fig. 3.

TABLE I: Data distribution by fault type

Data Type	Fault Type		
	No Fault	***PPG***	***NPG***
D1: Training data (Node A)	50%	25%	25%
D2: Training data (Node B)	50%	25%	25%
D3: Testing data (Node A)	20%	35%	45%
D4: Testing data (Node B)	50%	25%	25%

IV. Data Analysis

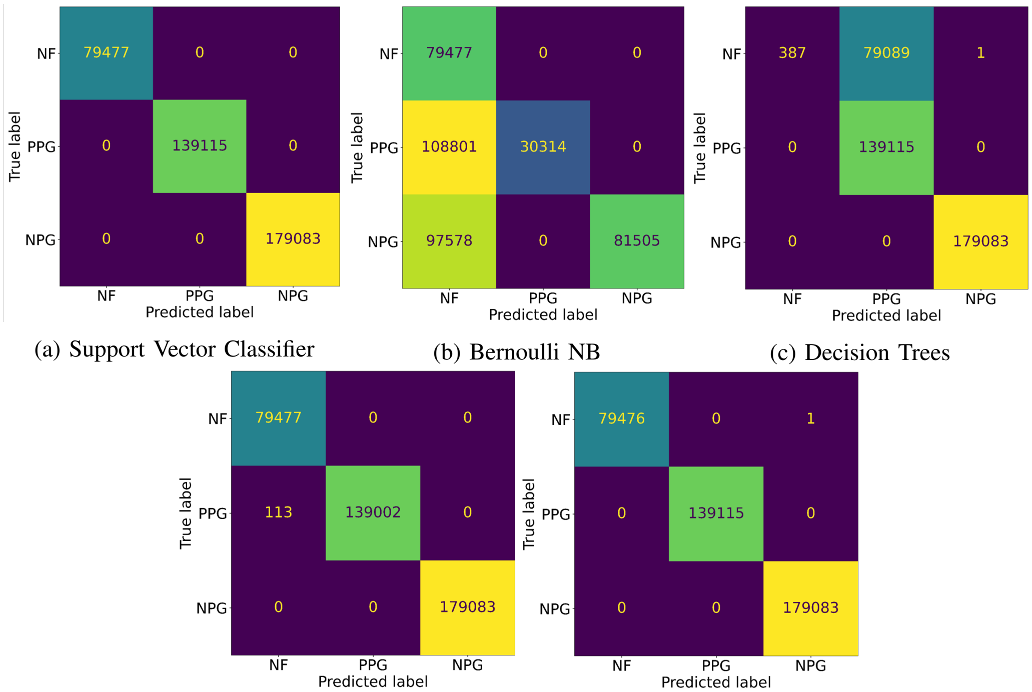
The simulation data generated 2.4 million data points for each measurement node which includes 400,000 points for each of the 6 fault scenario listed above. From the field, the real data gathered has a total number of 4 million data points for each measurement node. The distribution of the data by fault type is shown in Tab. I while the distribution of the data by their fault type and fault resistance i.e. fault 1 through 6 (F1 to F6) is shown in Tab. II.

A. Feature/Data Input

The measured voltage from the PSCAD simulation data, and the captured voltage measurement from the real data can be reduced to three components, the PPG, NPG, and pole to pole (PP) voltages. Each of the three components were provided as individual features to the ML classification algorithms to classify the data based on fault type and fault resistance. Each feature is then transformed such that the absolute values are mapped in the range of [0,1], i.e. the maximal absolute value is 1.0.

B. Evaluation Metrics

The performance of each of the ML classification algorithms is evaluated with a confusion matrix by comparing the predicted and actual outcomes. This format allows for the review of true positive, false positive, true negative and false negative. In addition, an accuracy score is also computed as the ratio of the number of correct prediction to the total number of prediction. The accuracy score ranges from 0 to 1. The higher the score, the more accurate the classification.



(a) Support Vector Classifier (b) Bernoulli NB (c) Decision Trees

(d) Nearest Centroid (e) Multi-layer Perception

Fig. 4: Confusion matrix showing the ML classification based on fault type

TABLE II: Data distribution by fault type and fault resistance

Data Type	Fault Type and Resistance					
	No Flt	***F1***	***F2***	***F3***	***F4***	***F5***	***F6***
D1	50%	8.3%	8.3%	8.3%	8.3%	8.3%	8.3%
D2	50%	8.3%	8.3%	8.3%	8.3%	8.3%	8.3%
D3	20%	22.5%	22.5%	0%	22.5%	12.5%	0%
D4	50%	12.5%	12.5%	12.5%	12.5%	0%	0%

V. Results

To explore the potential of ML algorithms in classifying DC faults based on fault type and fault resistance, the simulation data from PSCAD is used in training all five algorithms described in Section II. The performance of these algorithms in classifying the faults are tested with the real data from the actual DC power system.

A. Fault type classification

First, the ML models are used to classify the fault data into three categories: No fault (NF), PPG fault and NPG fault. Fig. 4 shows the performance of all five ML algorithms in classifying the data based on their fault type. Out of the five classifiers, the SVC, nearest centroid and the multi-layer perception accurately classified the data based on fault type.

B. Fault type and Fault resistance classification

The SVC, NC and the MLP algorithms accurately classified the data based on the fault type. These three classifiers are used to classify the data based on both the fault type and fault resistance. The data is classified into 5 labels: No fault, 1 ohm PPG fault, 1 ohm NPG fault, 500 PPG fault and 500 ohm NPG fault. Fig. 5, shows the performance of the classifiers in classfying the data based on the fault type and fault resistance.

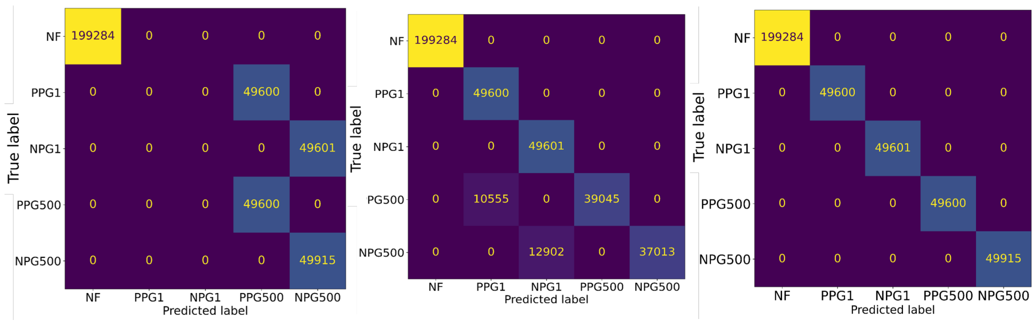
Table III shows the accuracy score for each ML classifier. In classifying by fault type, the SVC, NC and MLP classifiers had a score of at least 0.99. When classifying by fault resistance, only the MLP classifier performed accurately with a perfect score of 1.0.

VI. Conclusions

In this paper, the potential for machine learning classifiers embedded in a protective relay to identify and classify fault in a DC Microgrid is explored. A variety of algorithms such as SVC, NB, DT, NC, and MLP are applied on data gathered from an actual DC Microgrid. The fault data is classified based on fault type and fault resistance. The SVC, NC and MLP algorithms can detect fault and classify the fault type with 100% accuracy. Among all the algorithms presented in this paper, the MLP algorithm had the best performance in classifying the fault based on both the fault type and fault resistance

TABLE III: Accuracy Score of ML Classifiers

ML Classifier	Node A Accuracy score		Node B Accuracy score	
	Fault type	***Fault Res.***	***Fault Type***	***Fault Res.***
SVC	1.00	0.75	1.00	0.75
NB	0.48	0.48	0.75	0.74
DT	0.80	0.77	0.49	0.25
NC	0.99	0.94	1.00	0.94
MLP	0.99	0.99	1.00	1.00



(a) Support Vector Classifier (b) Nearest Centroid (c) Multi-layer Perception

Fig. 5: Confusion matrix showing the ML classification based on fault type and fault resistance

with an accuracy of 99%. Machine learning algorithms, when deployed inside a protective relay can accurately detect and classify faults in DC Microgrids.

Acknowledgment

The authors would like to thank Emera Technologies LLC for the use of one of their DC Microgrids for testing. The authors gratefully acknowledge the contributions of Dr. Ali Bidram of the University of New Mexico and Munim Bin Gani of Clemson University.

References

[1] T. M. Gruzs and J. Hall, "Ac, dc or hybrid power solutions for today's telecommunications facilities," in *INTELEC. Twenty-Second International Telecommunications Energy Conference (Cat. No.00CH37131)*, 2000, pp. 361–368.

[2] F. Bodi, ""dc-grade" reliability for ups in telecommunications data centers," in *INTELEC 07 - 29th International Telecommunications Energy Conference*, 2007, pp. 595–602.

[3] A. B. M. J. R. Rudy Montoya, Binod P. Poudel, "Dc microgrid fault detection using multiresolution analysis of traveling waves," *International Journal of Electrical Power & Energy Systems*, 2022.

[4] R. Darbali-Zamora, J. E. Quiroz, J. Hernández-Alvidrez, J. Johnson, and E. I. Ortiz-Rivera, "Viability assessment of a real-time simulation model for a residential dc microgrid network to compensate electricity disturbances in puerto rico," in *2018 IEEE ANDESCON*, 2018, pp. 1–6.

[5] S. Ojetola, J. Wold, D. Trudnowski, F. Wilches-Bernal, and R. Elliott, "A real power injection control strategy for improving transient stability," in *2020 IEEE Power and Energy Society General Meeting (PESGM)*, 2020, pp. 1–5.

[6] R. Darbali-Zamora, J. E. Quiroz, J. Hernández-Alvidrez, J. Johnson, and E. I. Ortiz-Rivera, "Implementation of a dynamic real time grid-connected dc microgrid simulation model for power management in small communities," in *2018 IEEE 7th World Conference on Photovoltaic Energy Conversion (WCPEC) (A Joint Conference of 45th IEEE PVSC, 28th PVSEC & 34th EU PVSEC)*, 2018, pp. 1179–1184.

[7] A. Emadi, Y. J. Lee, and K. Rajashekara, "Power electronics and motor drives in electric, hybrid electric, and plug-in hybrid electric vehicles," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 6, pp. 2237–2245, 2008.

[8] S. T. Ojetola, J. Wold, and D. J. Trudnowski, "A multi-loop transient stability control via power modulation from energy storage device," *IEEE Transactions on Power Systems*, pp. 1–1, 2021.

[9] S. Augustine, J. E. Quiroz, M. J. Reno, and S. Brahma, "Dc microgrid protection: Review and challenges." United States, Tech. Rep., 2018. [Online]. Available: https://www.osti.gov/servlets/purl/1465634

[10] K. A. Saleh, A. Hooshyar, and E. F. El-Saadany, "Ultra-high-speed traveling-wave-based protection scheme for medium-voltage dc microgrids," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1440–1451, 2019.

[11] M. E. Baran and N. R. Mahajan, "Overcurrent protection on voltage-source-converter-based multiterminal dc distribution systems," *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 406–412, 2007.

[12] K. A. Saleh, A. Hooshyar, and E. F. El-Saadany, "Hybrid passive-overcurrent relay for detection of faults in low-voltage dc grids," *IEEE Transactions on Smart Grid*, vol. 8, no. 3, pp. 1129–1138, 2017.

[13] S. D. A. Fletcher, P. J. Norman, K. Fong, S. J. Galloway, and G. M. Burt, "High-speed differential protection for smart dc distribution systems," *IEEE Transactions on Smart Grid*, vol. 5, no. 5, pp. 2610–2617, 2014.

[14] M. Monadi, C. Gavriluta, A. Luna, J. I. Candela, and P. Rodriguez, "Centralized protection strategy for medium voltage dc microgrids," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 430–440, 2017.

[15] A. A. S. Emhemed, K. Fong, S. Fletcher, and G. M. Burt, "Validation of fast and selective protection scheme for an lvdc distribution network," *IEEE Transactions on Power Delivery*, vol. 32, no. 3, pp. 1432–1440, 2017.

[16] N. S. Coleman, C. Schegan, and K. N. Miu, "A study of power distribution system fault classification with machine learning techniques," in *2015 North American Power Symposium (NAPS)*, 2015, pp. 1–6.

[17] C. B. Jones, A. Summers, and M. J. Reno, "Machine learning embedded in distribution network relays to classify and locate faults," in *2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2021, pp. 1–5.

[18] R. D. Telford, S. Galloway, B. Stephen, and I. Elders, "Diagnosis of series dc arc faults–a machine learning approach," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1598–1609, 2017.

[19] Z. Wang and R. S. Balog, "Arc fault and flash detection in photovoltaic systems using wavelet transform and support vector machines," in *2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC)*, 2016, pp. 3275–3280.

[20] S. Roy, S. Nayar, S. Kumar, A. Alam, and T. Ghose, "Bidirectional power flow in dc microgrid and its islanding detection using support vector machine," in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*, 2019, pp. 42–47.

[21] H. I. Bulbul and . Unsal, "Comparison of classification techniques used in machine learning as applied on vocational guidance data," in *2011 10th International Conference on Machine Learning and Applications and Workshops*, vol. 2, 2011, pp. 298–301.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[23] W. Chmielnicki, "Combining k-nearest neighbor and centroid neighbor classifier for fast and robust classification," in *Hybrid Artificial Intelligent Systems*, F. Martínez-Álvarez, A. Troncoso, H. Quintián, and E. Corchado, Eds. Cham: Springer International Publishing, 2016, pp. 536–548.

[24] S. Yibin and W. Peijin, "A multilayer perception with the adaptive learning factor," in *2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering. TENCOM '02. Proceedings.*, vol. 3, 2002, pp. 1546–1549 vol.3.