



## Electronic and Automation Engineering

Vol: 2(1), 2023

REST Publisher, ISBN: 978-81-956353-5-1

Website: <http://restpublisher.com/book-series/ea/>

DOI: <https://doi.org/10.46632/ea/2/1/8>



# Data Processing Method for Artificial Neural Network ANN Based Microgrid Protection Model

\*Sanghita Baidya, Champa Nandi

Tripura University, Suryamaninagar, Agartala, India

\*Corresponding author Email: [sanghita.ee@tripurauniv.in](mailto:sanghita.ee@tripurauniv.in)

**Abstract.** *Effective fault detection and isolation technologies are very necessary for uninterrupted power supply and for making a flexible protection scheme. Almost all protection schemes in the power system are based on data exchange among protection units through a strong communication structure. Thus, it is important to deal with a large amount of data. Artificial Intelligence (AI) is one of the key factors in this regard. AI has several sections and Artificial Neural Network (ANN) is one of them. It is suggested to implement the ANN-based models while working with big data. The existing protection models are facing difficulties while trying to deal with big data. Thus ANN-based approaches have come into the front line in advanced power system networks. The performance of the ANN model is depending on the training of the data set. Hence in this work, we are focusing on preparing the data to provide input in the ANN model. The principal component analysis (PCA) method is applied here for reduced the dimension of a large number of data sets. The new data set is used to run the k-means clustering algorithm. It is shown that the clustering is more accurate with the processed data set by PCA. Therefore, the prepared data set is used to run the ANN model that has a smaller size with higher information and minimum computational time. This study shows the data preparation part to train the ANN model.*

## 1. INTRODUCTION

Microgrids are now in trend for supplying reliable power in the energy system. But it has faced lots of issues with its technology. The DC current does not have any zero crossing in its waveform, which increase the risk of faults, especially arc-type faults. As the current required for arc fault generation is very low the commonly used circuit breakers are not suitable to protect the system. Some other issues in microgrid systems are the sudden increase of fault current, improper information, and guidelines, and lack of standards and practical implementations [1]. Besides, the use of distributed generators (DGs) causes the change of fault current amplitude and its direction during the faulty condition. Thus, a differential protection unit with high speed is obtained for managing these issues in the microgrid. However, this differential protection scheme is not suitable, if the amplitude of the short circuit current is exceedingly transient and the synchronization error is lasting for several microseconds [2]. The existing safety devices are not capable of effectively differentiating the various types of faults by discovering, categorizing, and localizing them that appear in the microgrids because of their dynamic nature, especially when the grid is operating in the islanded mode. There are numerous difficulties in research in the area of microgrid protection for the dynamic nature of the microgrid [4], [13], [15]. A cluster of decentralized/distributed energy systems that is a set of non-conventional energy resources (like wind, hydro and solar), storage the energy and supply it in the demand side (local customers) based on the need, is known as a microgrid [1]. The microgrid structure can operate in the mode of both grid-connected (while connected to the utility grid) as well as in islanded mode (when the connection with the utility grid is disturbed due to undesirable events such as faults, programming errors, etc.) [2], [3], [39]. Usually, the microgrid protection designs rely upon some vital factors such as the type of microgrid, topology of the microgrid, category of DG resources, communication type, details about the faults, relay model, data preprocessing and grounding method [2], [4], [5]. This makes a requirement for a reasonable and intelligent microgrid protection scheme that must be appropriate to integrate with the traditional protection units [10]. As the majority of the conventional protection architectures uses microprocessor-based computers or embedded method, sophisticated station protection architecture is required. In a protection scheme, the present measuring devices are used to give input data and to provide inclusive station protection. Thus, researchers began to research this zone utilizing numerous techniques, where artificial neural networks (ANNs) appeared as a promising method to detect different types of faults [6], [7], [8]. The ANN technology includes a data processing step that set the input data for the ANN [9], [11], [12]. Nonetheless, a very a smaller number of researchers have actualized the concept of overall station protection in

microgrids with functionalities that incorporate classification and detection of faults, faulty feeder isolation, and assessment of the defective area. Therefore, it is an important area for research that may bring an affordable, suitable and intelligent microgrid protection model. Here the use of artificial intelligence has come into the structure. The neural network doesn't need any complex programming, because it can be trained using training data and the obtained output is generalized and correlated data [14]. Thus, the use of ANN may give a better solution to the protection-related issues of microgrids. For implementing the ANN approach, it is important to prepare the input data because the ANN model depends on the training data. Hence in this paper, we focus on preparing the input data, so that the ANN model will get suitable information in the form of training data. The main aim of this work is to categorize the data set that we further used for training the ANN model. In this paper, we have used PCA and K-means clustering algorithm to process the data. So that the ANN model may get a highly informative data set as input to train the system. The paper is categorized into the following sections, such as theoretical background, overview of the proposed scheme, and data preparation with results. Finally, directions for future work and conclusions are added.

### **Artificial Intelligence (AI)**

To develop the stability, operation, and reliability of a microgrid an alternative robust solution scheme is discovered based on the concepts of artificial intelligence (AI) that includes machine learning to control the operation of big-scale data in power systems (conventional grids, microgrids, and smart grids) [16], [17], [18],[19]. The important properties of advanced protection systems should have a quick response, simplicity in nature, flexibility, and controllability in all complicated configurations. Recently ANN has had different unique applications in energy network that includes studies regarding transient stability, security assessment, voltage stability, primary warning/alert of the connecting power systems, etc. [20], [21], [22], [23], [24]. Deep Learning (DL) techniques may utilize in those situations that have colossal or intricate information handling challenges of ML or conventional methods for information preparation. DL techniques are unable to be reliant on human intercession. The subset of AI and DL comprises various layers of calculations; it gives an alternate understanding of the information it benefits from [25]. DL is primarily unique to ML since it presents information in the framework alternately. ANN layers are also used in ML calculations, DL networks, etc. Dissimilar to managed realizing which is the assignment of learning a capacity planning a contribution to a yield based on model info yield sets, unaided learning is set apart by least human oversight and could be depicted as such an AI looking for undetected examples in an informational index where no earlier names exist. The initial phase is information setup that is essential for data mining during data understanding, data learning, and large data [26],[27]. The information being talked about here comprise electrical boundaries and other different types of data which are changed into information that can be perceived through a mechanism. Before handling and examination comes **data**, learning which is the cycle through which basic information is sophisticated and changed over. It is a cycle wherein information is reformatted, amended, and joined to advanced information. Gathering, investigating, and utilizing the information closes in large information [25]. A deep Neural Network DNN is a kind of neural network that has various concealed layers of neurons between the inputs and outputs. It is broadly adjusted to display difficult non-linear frameworks in the research of the engineering field [28]. Likewise, the calculation of DNN just includes basic arithmetical conditions, delivering a quick computation speed. This trademark makes DNN fit for dealing with real-time issues [29], [30].

### **Machine Learning (ML)**

Two types of ML approaches are available, such as supervised machine learning and unsupervised machine learning. The input data used for supervised learning includes training data sets and testing data sets. And the output data should ideal for assembling a preparation model. The preparation model is worked by an information investigator or data researcher. Criticism is done regarding the exactness of the prepared model along with additional presentation measurements throughout the training period. Modification is done according to the requirement. When the preparation stage is finished, the model can anticipate results for new information [31], [32]. Classification of big data into different categories goes below the supervised learning process that assumes machine learning over models in Grouping. On the way of characterization, each case of the respective dataset is ordered into an objective worth. Characterization can either be parallel or multi-name. Here and there, one specific occasion can similarly have numerous classes well-known as a multi-class arrangement. Grouping calculations are significantly utilized to forecast the data for learning. On the other hand, the unsupervised learning technique is utilized to create deductions of the information that doesn't take any marked reactions. But the information isn't arranged, marked, or characterized into classes. Bunching investigation, quite possibly the most widely recognized unsupervised learning strategy is utilized to discover concealed examples in the information or to frame bunch dependent on the information. While ML models have been around for quite a long time, they have acquired energy with the ascent of AI. DL models are presently utilized in a large portion of intelligent AI applications [33], [34].

### **Supervised Learning:**

The ML is comprehensively ordered into two: unsupervised learning and supervised learning. Recorded or trained data is required in supervised learning that can contribute to the machine and work as a classifier model framed. A supervised learning algorithm requires an objective worth/ target. In the case of unsupervised learning, neither the train information nor the target is required [35]. Logistic regression is an algorithm based on grouping that used supervised learning at the

time of implementation. And it is utilized for prescient learning. This algorithm is utilized for representing information. The algorithm turns out superior for binary classification [36].

## 2. BLOCK DIAGRAM OF ARTIFICIAL INTELLIGENCE-BASED PROPOSED MODEL

AI systems help to take care of issues including the learning capacities of PCs, language handling, and recognizing different addresses in the climate in an exceptionally dependable and exact way. Smart cities idea bears the number of urban areas with knowledge and different ideas including brilliant vehicles, smart networks, and intelligent medical care utilizing the idea of deep learning. By and large, in smart cities where circulated things approach computational assets, transferring of data becomes inescapable due to high latency, consequently bringing about essential circumstances. Initially to achieve high dependability artificial intelligence-based K-means neural network (KNN) and convolution neural network (CNN) is required for processing and cleaning of the control node [28].

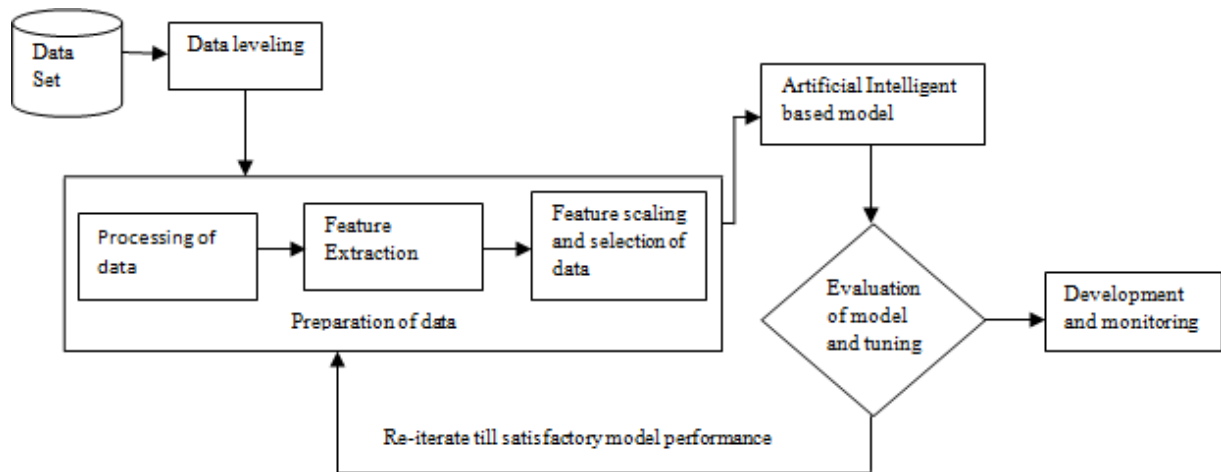


FIGURE 1. Block diagram of AI-based proposed scheme

Figure 1 illustrates the block diagram of the AI-based proposed scheme. Here the dataset consists of various data collected from the environment (such as temperature, humidity, and other parameters), the amount of solar irradiance whole over the day, and the measured data of voltage, current, etc, from the microgrid. The leveling of data is required to differentiate the characteristics of the data. Now the data should be prepared to put in the AI-based model as input. The model is now evaluated with the trained data. At last, the monitoring process is developed if the model can able to give a satisfactory performance; the else re-iteration process is activated until the desired output is obtained.

## 3. DATA PREPARATION

To work with big data sets machine learning-based supervised algorithm is considered in this work. A simulation model is developed to collect the data. To build AI based model we need to train the system with the collected information. Hence initially we process the data to make it ready for training. For that PCA is used. The obtained new data set from PCA is used as input in the k-means clustering algorithm. This algorithm is used to classify the data set into required groups. The data of each group has similar features to that group. Hence this classification of data set may apply to identify the faulted section. By using the clustering method four different groups are obtained here. After that, by using python programming we leveled the processed data into binary range (i.e., 0 or 1). Thus, we set the input of the supervised learning model with a target of 0 or 1.

### Principle Component Analysis (PCA)

PCA is a method based on a statistical analysis of a large number of data sets. The method is implemented to reduce the dimension of any large number data set with the least loss of information. It is a suitable method to decrease the size of the big number of variables to a lesser number of components, i.e., if the data set is having the size of  $(n \times m)$  number of large data, the PCA technique will reduce the size at  $(n \times k)$  number of small data. Where 'n' is the observation number, 'm' is the nos. of parameters and 'k' is the reduced components number of the large 'm' nos. parameters [36], [37]. The use of the PCA method helps to identify those components that are having close relation and interconnection with one another in a large number of data sets by summarizing them in different groups. The scores of the components are calculated

through PCA for each observation. This method offers independent and linear principal components (PCs) from the input variables. Instead of using the raw data (input variables) use of PCs can provide better accuracy. Thus, the implementation of PCA is very essential to process the data before using them in the ANNmodel [37].

#### ***K-means Clustering Algorithm***

Clustering is the task of dividing the population or data points into several groups such that data points in the same groups are more similar to other data points in the same group than those in other groups. In simple words, the aim is to segregate groups with similar traits and assign them into clusters. Clustering aims at finding smaller, more homogeneous groups from a large heterogeneous collection of items. The main purpose behind the study of classification is to develop a tool or an algorithm, which can be used to predict the class of an unknown object, which is not labeled. This tool or algorithm is called a classifier. Clustering deals with finding a structure in a collection of unlabelled data. Here K-Means has been adopted as the prototype of iterative model-based clustering because of its speed, simplicity, and capability to work within the format of a very large database. There is no limitation on the size of the data sets on the k-mean algorithm [40]. It also does not explicitly restrict the dimensions of the data and it is believed that it applies to a large variety of mixtures. The flow chart of k means algorithm is shown here in Fig.2:

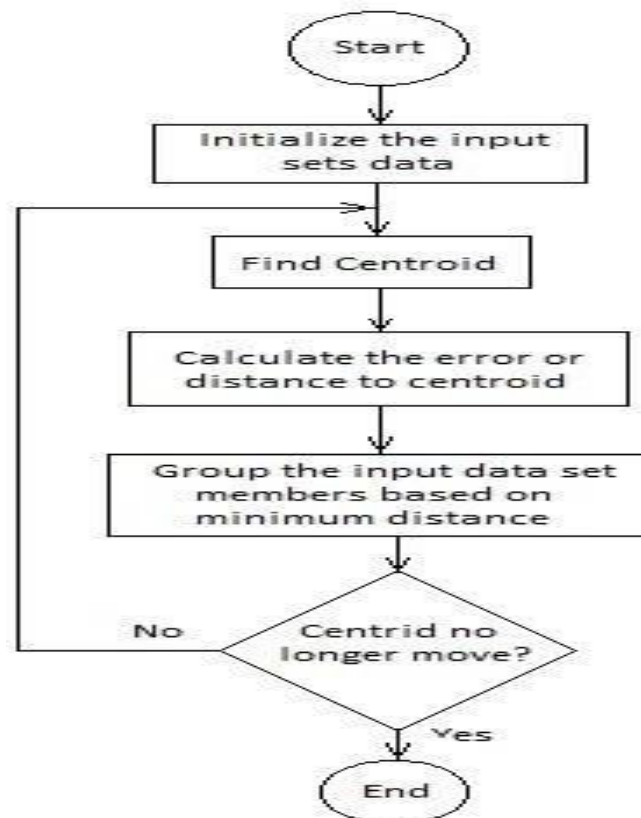
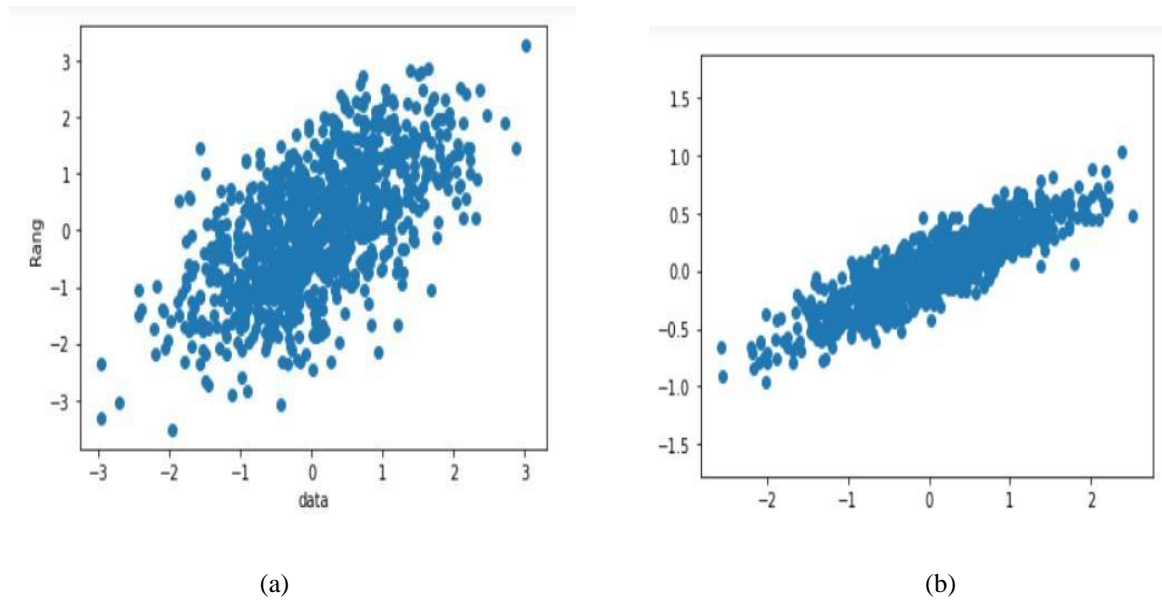


FIGURE 2. Flowchart of k-mean clustering algorithm

## 4. RESULT AND DISCUSSIONS

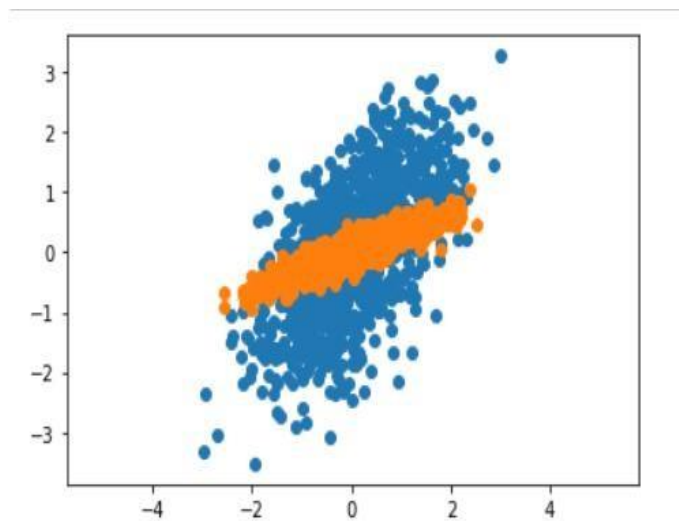
The first step is to reduce the dimension of the data with the maximum feature. So that the visualization of the data is proper and also the running time of the model will be less which further increases the efficiency of the system. Thus, we implement the PCA algorithm in the data set that reduces the dimension of the data. Figure 3 (a) shows the shape of the original data set.

The PCA algorithm is applied in the above-visualized data set to reduce the dimension of the data. In the original shape, the data has 6 features. After implementing the PCA algorithm 2 features are selected from the original set. In Figure.3(b) the reduced data set is shown. The data set in Fig.3(b) is restricted to (2, 2) rang because the shape is reduced to 2 features i.e., we have 2 columns here. From the above figure, it is clear that the implementation of PCA can eliminate the unnecessary data from our data set and make the data set clearer which also has maximum features.



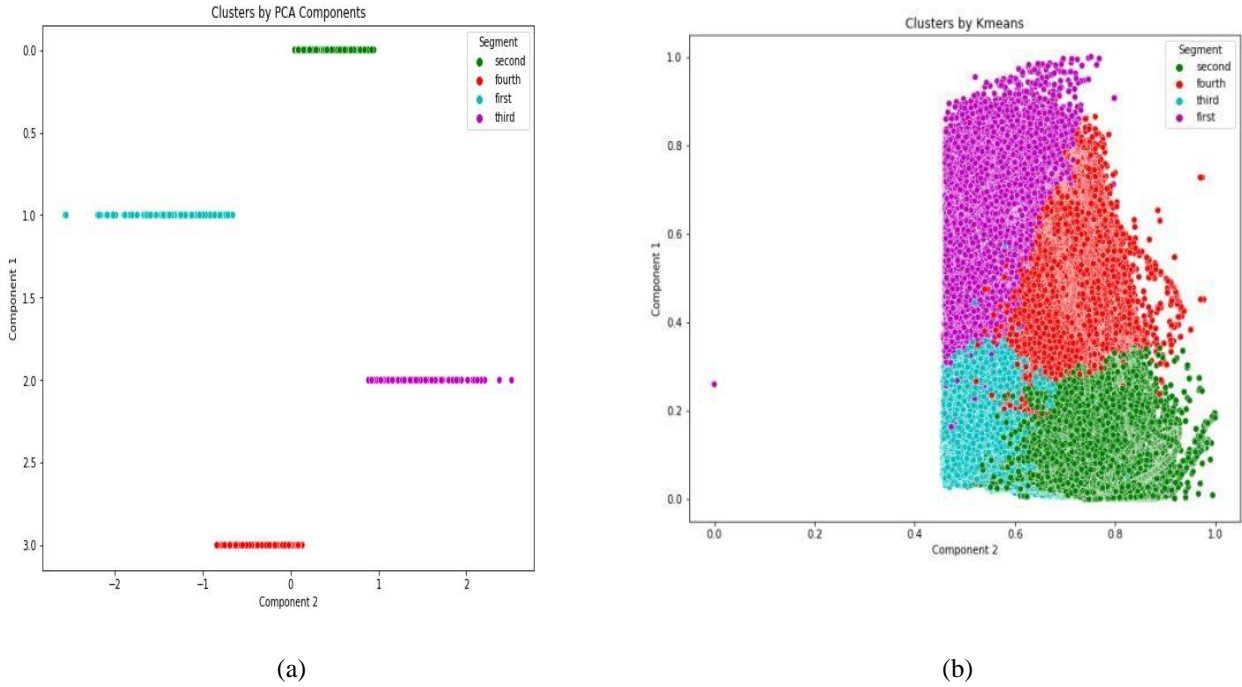
**FIGURE 3.** Visualization of data set. (a) Original data set, (b) Visualization of data set with PCA algorithm

In Figure 4 the original data set is compared with the processed data set. It is shown that the PCA approach can reduce the dimension of the data with a maximum no. of features. Those data which are less important or have less information are eliminated. Thus, it can further help to run the machine learning model faster.



**FIGURE 4.** Visualization of original data set with processed data set of data set

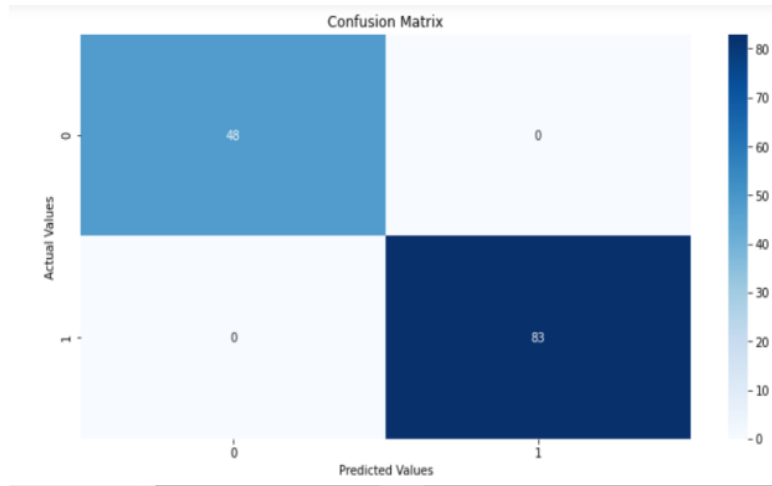
Now k-means clustering algorithm is applied to that data set. The algorithm successfully clustered the data set with 4 numbers of clusters. To obtain the number of clusters we used the Elbow method. In the next step, the number of centroids for each cluster is determined. After that labels of clusters are selected such that 0, 1, 2, 3. Thus the four clusters are clearly shown in Fig. 5.



**FIGURE 5.** Clustering of data set. (a) Clustering of data set with processed data from PCA algorithm, (b) Clustering of the original data set by K-means (without PCA algorithm)

From Figure 5(a), it is shown that four clusters are successfully generated by using the k-means clustering algorithm with the processed data collected from the PCA algorithm. The entire data set is separated into four numbers of clusters. This is the main contribution of the PCA algorithm in the case of dimension reduction of a largenumber of data sets with maximum features.

Figure 5(b), shows the clustering of the original data set by using a k-means clustering algorithm. The used data set is not processed by the PCA algorithm. Hence it is clearly shown from the figure that all four clusters are not completely separated from each other, instead, they are overlapping at some points. This creates confusion on the ML approach to taking decisions for further action.



**FIGURE 6.** Confusion Matrix of actual and predicted values

In Figure 6 the confusion matrix is shown that was obtained by the processed data set. It is shown from the figure that the false positive and false negative values are zero. Thus, the prediction accuracy is high. Thus, this data set is suitable to train the ANN based microgrid protection model. In this part of work, we only prepare the data for training our ANN based microgrid protection model. This data set consists of the value of current, voltage and power of different situations in our microgrid system. The next step of this work is to run the ANN based microgrid model with this pre-proceed data set in simulink platform.

## 5. CONCLUSION

As the traditional protection schemes are not so advanced to deal with a large number of data, the introduction of Artificial intelligence in this field is an important concept. The ANN is a part of AI that can offer better solutions regarding the issues of microgrid protection. The ANN network is a simple platform to work with big data. This scheme only uses two types of data, i.e., training data and testing data. The study includes the dimension reduction method of the large number of data sets using the principal component analysis technique. The reduced data set is used to make clusters using the k-means clustering algorithm. The evaluation of the work is investigated using python-based programming for the proposed system. Hence a confusion matrix is generated with the processed data set that shows zero false positives and zero false negative values. Thus, it consists of higher prediction accuracy. In this work, we try to prepare the input data for the ANN model with the help of PCA. The PCA method provides us with principal components that have maximum information with a reduced size. To analyze the PCA method we use a big data set collected from a simple simulation model of the microgrid. The implementation of the method provides us with principal components with the least loss of information which are used to make clusters of different levels. In this study, we only preprocess the data to train the ANN model. After that confusion matrix is evaluated. The extension of the work may include real-time data processing to train the neural network for protecting microgrids. **Acknowledgments.** The authors are thankful to Tripura University, Tripura, India for their support and required facilities for the study.

## REFERENCES

- [1]. D.K.J.S. Jayamaha, N.W.A. Lidula and A.D. Rajapakse, Protection and grounding methods in DC microgrids: Comprehensive review and analysis, *Renewable and Sustainable Energy Reviews*, Vol. 120, 1-21 (2020).
- [2]. J. Kennedy, P. Ciuffo, and A. Agalgaonkar, A review of protection systems for distribution networks embedded with renewable generation, *Renewable and Sustainable Energy Reviews*. Vol.58, 1308-1317 (2016).
- [3]. B. J. Brearley and R. R. Prabu, A review on issues and approaches for microgrid protection, *Renewable and Sustainable Energy Reviews*. Vol.67, 988-997 (2017).
- [4]. S. Baidya. and C. Nandi, *Green Energy Generation Using Renewable Energy Technologies*, edited by Prof. A. K. Bhoi et. al., (Springer, Singapore, 2020), pp 259–276.
- [5]. S. A. Hosseini, H. A. Abyaneh, S. H. Hesamedin Sadeghi, F. Razavi and A. Nasiri, An overview of microgrid protection methods and the factors involved, *Renewable and Sustainable Energy Reviews*, Vol. 64, 174-186 (2016).
- [6]. T. S. Ustun, C. Ozansoy and A. Zayegh, Recent developments in microgrids and example cases around the world - A review, *Renewable and Sustainable Energy Reviews*. Vol.15, 4030-4041 (2011).
- [7]. H. Lin, J. M. Guerrero, C. Jia, Z. Tan, J.C. Vasquez, and C. Liu, "Adaptive overcurrent protection for microgrids in extensive distribution systems", *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, ( IEEE, Lorence, Italy, 2016), pp.1-28..
- [8]. S. Zarrabian, R. Belkacemi and A. Babalola, "Intelligent mitigation of blackout in real-time microgrids: Neural network approach", in *2016 IEEE Power and Energy Conference at Illinois (PECI)*, ( IEEE, Urbana, IL, USA, 2016), pp.1-6.
- [9]. H. Lin, K. Sun, Z. Tan, C. Liu, J.M. Guerrero, and J.C. Vasquez, Adaptive protection combined with machine learning for microgrids, *IET Generation, Transmission & Distribution*, Vol.13, 770 – 779 (2019).
- [10]. S. Baidya, V. Potdar, P.P. Roy and C. Nandi, Reviewing the opportunities, challenges and future directions for the digitalization of energy, *Energy Research & Social Science*, vol.81, 1-24 (2021).
- [11]. Y. Y. Hong, Y.H. Wei, Y.R. Chang, Y.D. Lee and P.W.Liu, Fault detection and location by static switches in microgrids using wavelet transform and adaptive network-based fuzzy inference system. *Energies*, vol.7, 2658- 2675 (2014).
- [12]. S. Baidya and C. Nandi, "Arc Fault Analysis, Monitoring with Internet of Things (IoT) and Protection Scheme of a Solar PV Model", *Proceedings of the 2nd International Conference on Information Systems & Management Science (ISMS)*, (Tripura University, Agartala, Tripura, India, 2019), pp.1-6.
- [13]. S. Bhattacharjee and C. Nandi, "Design of an Industrial Internet of Things-Enabled Energy Management System of a Grid-Connected Solar-Wind Hybrid System-Based Batter Swapping Charging Station for Electric Vehicle", *Applications of Internet of Things, Lecture Notes in Networks and Systems*, edited by Dr. J. K. Mandal et. al., (Springer, Singapore, 2020), pp. 1-14.
- [14]. B. K. Panigrahi, P. K. Rout, P.K. Ray and A. Kiran, "Fault Detection and Classification Using Wavelet Transform and Neuro Fuzzy System", in *Proceedings of the 2018 International Conference on Current Trends towards Converging Technologies, ICCTCT*, (IEEE, Coimbatore, India, 2018), pp.1-5.
- [15]. M. Manohar, E. Koley and S. Ghosh, "A reliable fault detection and classification scheme based on wavelet transform and ensemble of SVM for microgrid protection", in *Proceedings of the 2017 3rd International Conference on Applied and Theoretical Computing and Communication Technology, iCATccT*, (IEEE, Tumkur, India, 2017), pp. 1-5.
- [16]. S. P. Puthenpurakel and P. R. Subadhra, "Identification and classification of microgrid disturbances in a hybrid distributed generation system using wavelet transform", in *2016 International Conference on Next Generation Intelligent Systems, ICNGIS*, (IEEE, Kottayam, India, 2016), pp, 1-5.
- [17]. J. J. Q. Yu, Y. Hou, A. Y. S. Lam and V. O. K. Li, Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks, *IEEE Transactions on Smart Grid*, vol.10, 1694 – 1703 (2019).

- [18]. T. Hubana “Artificial Intelligence based Station Protection Concept for Medium Voltage Microgrids”, 19th International Symposium INFOTEH-JAHORINA (INFOTEH), (IEEE, East Sarajevo, Bosnia and Herzegovina, 2020), pp.1-6.
- [19]. L. Galbusera, G. Theodoridis and G. Giannopoulos, Intelligent Energy Systems: Introducing Power-ICT Interdependency in Modeling and Control Design, *IEEE Trans. Industrial Electronics*, vol.62, 2468-2477 (2015).
- [20]. SK. Razibul Islam, K. M. Muttaqi and D. Sutanto, A Decentralized Multi agent-Based Voltage Control for Catastrophic Disturbances in a Power System, *IEEE Trans. Industry Applications*, Vol.51, 1201-1214 (2015).
- [21]. Y. Zhang, M. D. Ilic and O. K. Tonguz, Mitigating blackouts via smart relays: A machine learning approach, *Proceedings of the IEEE*, vol.99, 94-118 (2011).
- [22]. S. Gupta, F. Kazi, S. Wagh and R. Kambli, “Neural network based early warning system for an emerging blackout in smart grid power networks”, *Advances in Intelligent Systems and Computing*, edited by Rajkumar Buyya et. al., (Springer, Cham, 2015), pp. 173-183.
- [23]. S. Gupta, R. Kambli, S. Wagh and F. Kaziet, Support-Vector-Machine- Based Proactive Cascade Prediction in Smart Grid Using Probabilistic Framework, *IEEE Trans. Industrial Electronics*, vol.62, 2478-2486 (2015).
- [24]. R. Belkacemi, S. Zarrabian, A. Babalola and R. Craven., “Experimental Transient Stability Analysis of MicroGrid Systems: lessons learned”, in proceedings of IEEE Power and Energy Society (PES) General Meeting, (IEEE, Denver, CO, USA, 2015), pp.1-5.
- [25]. C.I. Chen, and Y.C. Chen, A neural network based data-driven nonlinear model on time-and frequency – domain voltage-current characterization for power quality study, *IEEE Trans. on Power Delivery*, vol.30, 1577-1584 (2015).
- [26]. B.Y. Vyas, B. Das and R. P. Maheshwari, Improved fault in series compensated transmission line: comparative evaluation of Chebyshev neural network training algorithms, *IEEE Transactions on Neural Networks and Learning Systems*, vol.27, 1631 – 1642 (2016).
- [27]. R. Sunitha, R. S. Kumar and A. T. Mathew, Online static security assessment module using artificial neural networks, *IEEE Trans. On Power Systems*, vol.28, 4328-4335 (2013).
- [28]. Y.Y Hong and M. T. A. M. Cabatac, Fault Detection, Classification, and Location by Static Switch in Microgrids Using Wavelet Transform and Taguchi-Based Artificial Neural Network, *IEEE Systems Journal*, vol.14, 2725 – 2735 (2020).
- [29]. S. Rahman Fahim, S. K. Sarker, S. M. Muyeen, Md. R. I. Sheikh and S. K. Das, Microgrid Fault Detection and Classification: Machine Learning Based Approach, Comparison, and Reviews, *Energies* vol.13, 1-21 (2020).
- [30]. J. J. q. Yu, Y. Hou, A. Y. S. Lam and V. O. K. Li, Intelligent Fault Detection Scheme for Microgrids with Wavelet-based Deep Neural Networks, *IEEE Transactions On Smart Grid*, vol.10, 1694 – 1703 (2019).
- [31]. I. Phafula, E. D. M. Koch and K. Nixon, “Preliminary Study of Fault Detection on an Islanded Microgrid Using Artificial Neural Networks”, *International SAUPEC/RobMech/PRASA Conference* , (Cape Town, South Africa, 2020), pp. 1-6.
- [32]. <https://www.igi-global.com/chapter/machine-learning-in-python/234122>.
- [33]. M. S. Kadhm, I. W. Ghindawi and D. E. Mhawi, An accurate diabetes prediction system based on K-means clustering and proposed classification approach, *International Journal of Applied Engineering Research*; vol.13, 4038–4041 (2018).
- [34]. P. Prabhu and N. Anbazhagan, Improving the performance of K-means clustering for high dimensional data set. *International Journal on Comput Science and Engineering*, vol.3, 2317-2322 (2011).
- [35]. R. Noori, A. Khakpour, B. Omidvar and A. Farokhnia, Comparison of ANN and principal component analysis- multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic, *Expert system with applications*, vol.8, 5856-5862 (2010).
- [36]. H. Abdi and L.J. Williams, Principal component analysis, *Wiley Interdisciplinary Reviews, WIREs Computational Statistics*, vol. 2, 433-459 (2010).
- [37]. A. Alver, and Z. Kazan, Prediction of full-scale filtration plant performance using artificial neural networks based on principal component analysis, *Separation and purification technology*, vol. 230, 1-12 (2020).
- [38]. S. Baidya, C. Nandi, A comprehensive review on DC Microgrid protection schemes, *Electrical Power Systems Research*, Vol- 210, 1-12 (2022).
- [39]. A. Saxena, M. Prasad, A. Gupta and N. Bharill, A review of clustering techniques and developments, *Journal of Neuro computing*, Volume 267, 664-681 (2017).