Indonesian Journal of Electrical Engineering and Computer Science
Vol. 32, No. 1, October 2023, pp. 353~362
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v32.i1.pp353-362

Fault classification and detection for photovoltaic plants using machine learning algorithms

Shaimaa Naser Kabour[1], Raghad Adnan Almalki[2], Lujain A. Alghamdi[2], Wujud Alharthi[2], Nisreen B. Alshagi[3]
[1]Computer and Communication Engineering, Lebanese International University, Zahle, Lebanon
[2]Computer Engineering, Taif University, Taif, Saudi Arabia
[3]Electrical and Computer Engineering, Libyan Academy, Tripoli, Libya

Article Info

Article history:

Received Jun 28, 2022
Revised Jun 12, 2023
Accepted Jun 17, 2023

Keywords:

Fault detection
Machine learning
Monitoring system
Photovoltaic plants
Renewable energy

ABSTRACT

Using photovoltaic (PV) energy has increased in recently, due to new laws that aim to reduce the global use of fossil fuels. The efficiency of a PV system relies on many types of malfunctions which may cause significant energy loss during the system's operation, besides the ecological factors. Consequently, a monitoring system (MS) capable of measuring both the environmental and electrical factors is described in order to gather real-time and historical data and estimate the plant efficiency metrics. Additionally, a recursive linear model for detecting problems in the system is presented, where the input is the irradiance and temperature of the PV module, whereas the output is the power, using the same MS. The achieved fault detection's accuracy for the 5-kW power plant reached 93.09 percent, based on 16 days and 143 hours of failures under various situations. After detecting a defect, a machine-learning-based algorithm categorizes each defect problem as short circuit, partial shadowing, deterioration, or open-circuit. The performance of the four most prevalent supervised machine learning (ML) approaches for this assignment (Naïve Bias, decision tree, LDA, and KNN) was evaluated according to their results.

Corresponding Author:

Shaimaa Naser Kabour
Computer and Communication Engineering, Lebanese International University
Zahle, Lebanon
Email: shaimaakabour@gmail.com

1. INTRODUCTION

During the last decade, the rise of renewable energy (RE) sources has accelerated, including solar energy, geothermal energy, biomass energy, hydro energy, and others. More specifically, an impressive total new capacity of 190 GW was added globally in 2018 to major hydro-power plants, which accounts for approximately half of the total power capacity installed during that year [1]. The number of solar photovoltaic (PV) plants had the most growth among the other commonly used renewable energy sources. In 2018, PV plants made up about 39 percent of the installed capacity [1].

The most convenient source of renewable energy is the sun, since it's large and its heat and light are distributed all over the globe, depending, of course, on the alignment between the earth and the sun. For this reason, the sunrays can be harvested and converted into electrical energy through what are called solar panels, or PV panels [2]. This energy collected and transformed by PV panels can then be used in many industrial applications, as well as residential and commercial applications, which explains the fact that solar energy is becoming widely popular [3]. However, there are numerous performance challenges that must be addressed for the purpose of enhancing the overall performance of PV. PV performance can be compromised for a variety of reasons, including excessive exposure to various weather conditions, such as soiling [4], wich leads to electrical and mechanical failures such as cracked cells and short-circuits [5]. Other factors that may have an impact on PV performance include PV material, battery type, panel orientation, and panel degradation. With such a high level of exposure, it becomes necessary to implement safety measures for maintaining the performance, reducing revenue losses and downtime, ensuring quick problem detection, problem classification, localization, and mitigation in PV systems [6]. Figure 1 depicts the challenges that influence the performance of the PV plate.

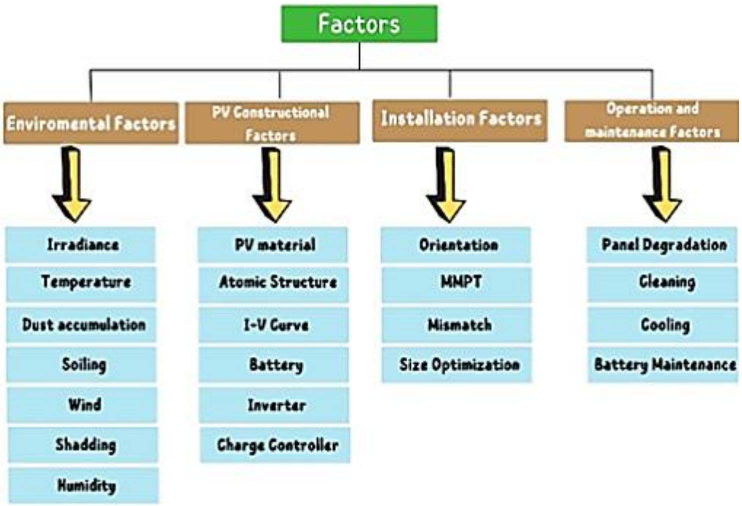


Figure 1. Factors influencing the effectivity of a PV panel

To protect the PV plate from the problems listed above, the PV plant should have a monitoring system (MS) that manages plant tasks, measures meteorological and electrical variables, finds errors and malfunctions, then reports benchmarking and performance to the grid operator either remotely or locally via a communication system [7], [8]. However, the MS alone is not sufficient to address the issues [9], since PV faults need the use of particular approaches to identify and categorize them using monitored data [8], [10]. In order to complete such duties, techniques are often separated into detecting and classifying the PV faults, with an emphasis on those that occur most, such as module mismatch, short circuit, and open-circuit [11]. There have been various ideas in the literature for defect detection. For example, in [12], fault detection according to satellite data is suggested. In another paper [13], the identification of PV panel faults through thermal imaging and the Canny edge detector is discussed. Various approaches based on PV system modeling have recently been suggested [14], [15], resulting in state-of-the-art outcomes in actual PV plants. On the other hand, the current models are primarily static, ignoring crucial dynamic modeling which leads to the difficulty in recognizing actions that occur in short time periods [16].

There are several ways to fault classification, including optical methods [17], mathematical techniques, and thermographic picture analysis techniques [13] by employing theoretical and virtual PV plant models [18]. Recently, various machine learning-based strategies are suggested to improve the classification accuracy in a collection of scenarios, namely shadowing and PV module deterioration [8], [19], [20]. Despite this, the majority of the approaches are based on simulated data, where the researchers do not provide a comprehensive examination of online fault classification methods.

Moreover, the limitations of detection and classification methods are compounded by them not having the results of the solution included in a devoted system or hardware, combined with a monitoring model, and whenever that happens, there is a limited power output for the PV plant otherwise the detection could only be done by disrupting the system's normal function. However, this study aims to overcome the mentioned limitation, where our contributions can be listed as: i) integrating a fault classification and detection technique to an embedded PV plant monitoring model, which allows identifying and classifying non-intrusively online the different PV defects, as well as offering an MS integrated to the plant; ii) providing a comprehensive comparison of dynamic models and ML algorithms for detecting and classifying actual fault situations in a 5 kW PV plant; and iii) deducing the best system for classifying and detecting faults in PV models online. Such a suggestion has not yet been offered in the associated literature, to our knowledge.

2. THE COMPREHENSIVE THEORETICAL BASIS

This section is divided into four primary subsections, each of which serves as the foundation for the proposed work study. First of all, defects in PV systems are examined, with a focus on the most important faults and descriptions of each problem. Then, a discussion of monitoring systems is done, followed by fault detection techniques, along with a fault classification discussion aligned with the key constraints identified in current work. Moreover, in the next section, the original components of this study, such as linked monitoring systems, defect detection, and identification algorithms, will be described in depth.

2.1. Faults in PV systems

Examining some of the predominant faults in solar systems is imperative for the sake of evaluating fault incidences and effects. In this context, the reference [21] presents a comprehensive examination of such faults, classifying them into two categories: direct current (DC) faults and alternating current (AC) faults where the fault of this kind is caused by issues with the system's inverter or the power grid. On the other hand, faults in the bypass diode, issues with the maximum power point tracking (MPPT) method, arc faults, ground faults, open circuit, cell or module mismatch (permanent or temporary), and short-circuit are among the most common DC defects. An open-circuit defect happens when there is a disconnect at some point in the system which has a big influence on power generation [22] where it may damage anything from one string of panels to reaching the whole system. Moreover, mismatching cells or modules happens if certain cells or panels in a photovoltaic model have electrical characteristics which are drastically altered from the rest, causing the system to malfunction [23].

On the other hand, whenever a low impedance route arises in the system, that results in a short-circuit problem occuring. This may happen at numerous locations in PV systems, such as between two terminals of the same module. The occurrence of short-circuit between two places on the same string will be considered in this study, especially between the negative pole of a panel and the positive pole of the next panel.

2.2. Monitoring systems

Some of the measures are discussed in [20], [24] which are utilized in detecting and classifying faults in the grid-connected PV systems such as wind speed, output voltage, total irradiance, wind direction, current of every PV array, ambient temperature, energy, and output power of every PV array [25]. These variables provide the foundation for evaluating the performance of the plant in real-time and improving system dependability, as described in [24]. Several systems were presented with the goal of creating a sensor network that is low-cost to monitor huge PV plants while avoiding a heavy reliance on separate sensors. For example, a collection of temperature sensors that is low cost, as well as current and voltage sensors were used to identify key problems in a single PV panel, including dirtying, permanent and temporary shadowing, and abnormal aging. Short-circuits and other sorts of defects, on the other hand, are not addressed in said study [5].

2.3. Fault detection

Generally, detecting faults in PV systems depends on modeling the system to make the results of modeling comparable to the data that is actually acquired [15]. The modeling stage is usually separated into two categories: dynamic and static. Static systems don't recognize time to be an independent variable, thus they are often called non-memory systems. Conversely, dynamic systems take into consideration the model's temporal fluctuations. In the modeling method of [6], [26], [27], a static system relying on a single-diode model is used to identify flaws and anticipate energy production. It is possible to ignore individual features and the dynamics of distinct PV models by reducing the PV cell to a static and generic one, which affects how certain phenomena are modeled and accordingly it affects the detection of defects happening in short periods of time.

The Hammerstein-wiener model, on the other hand, is used in [28] to simulate a system with nonlinearities. The input and output signals were irradiance and DC power, respectively. The selected sample duration was 15 minutes, which made it difficult to identify short-term occurrences like partial shadowing. An ARMAX model is suggested in [29] for predicting the produced power of a PV system one day ahead.

2.4. Fault classification

Artificial intelligence models, particularly machine learning classifiers, are among the techniques used to conduct fault classification, where such models gained growing attention and acceptance in recent literature [8], [30] and are also the major approach in this study. For example [10] shows how artificial neural networks (ANN) may be used to categorize the performance of a solar system into four stages where this approach was examined and trained in a simulation and achieved approximately 88.89% accuracy. A two-stage system is mentioned in [31], with the first step being for defect detection and the second level being for categorization. A multilayer perceptron ANN is employed for fault classification, with a 90.3 percent total accuracy. Moreover, this model trains the network with solely simulated data and is examined with an actual plant depending on the model's V×I graph. On the other hand, a technique is described for classifying PV faults in these instances:

short circuit, open-circuit, deterioration, and shadowing, utilizing a kernel extreme learning machine and information from V×I graphs [11]. Three case studies were conducted to assess the system's performance. Due to the restricted quantity of data obtained in the actual plant, the last strategy utilizes simulated data for to train the model and real data to test it. Using the true data, the accuracy ranged from 97.9% to 99.0%, whereas for mixed data, the ultimate accuracy reached 98.9%. However, the PV model should be disconnected due to the fact that the method depends on V×I curve, in order to perform the proposed fault classification procedure. The PV system must first be disconnected in order to carry out the suggested fault classification technique since it depends on an external device that has to be connected to the plant in order to produce V-I graphs despite having a pretty good performance (>95.0%). Table 1 displays the outcomes of the various methods related with the current study that were used to assess the proposed study's issues.

Table 1. Similar systems comparison

Section	Approach	Accuracy	Reference
Faults in PV systems	Study on short circuit, module mismatch, and open-circuit faults.	94%	[21]
Fault detection	In order to detect defects and predict energy output, a static model based on a single diode model is taken into account.	98.5%	[6]
	The system's nonlinearities are emulated using the hammerstein-wiener model.	84.67%	[32]
	ARMAX model for a day ahead prediction of the generated power.	95%	[33]
Fault classification	Using ANN for classifying the operation of a PV model.	88.89%	[11]
	A multilayer perceptron ANN is applied for fault classification.	90.3%	[34]
	using a kernel extreme learning machine and information from V×I graphs to classify PV faults.	98.9%	[10]

3. METHOD

This section describes in detail the 5 sequential stages which the proposed algorithm passes through, where each stage considers the inputs from the previous stage. Figure 2 depicts the five sequential stages that the proposed system consisted of. The first stage is data gathering, in which the data from the PV plant is simulated. Next comes the second stage, which is the preprocessing stage, where the data is prepared and cleaned and defects and irregularities are inserted into the PV plant to cause maloperation and therefore separate the data set into five different classes that are shown in Table 2. On the other hand, even though solar panels contain cells that convert solar energy into electrical energy, sensitive sensors have been added to those solar panels to measure the electric field, power, and temperature. Furthermore, to classify the voltaic cell as a normal operation cell, open circuit cell, short circuit, degradation, or shadowing. The third stage in the proposed algorithm is "data classification," in which we have used four machine learning algorithms, which are the naive bayes algorithm, the decision tree algorithm, LDA, and K nearest neighbor to classify the data into different categories. The fourth stage is the evaluation process where F1 score, precision score, accuracy score, and recall scores are used to test the classification accuracy using the four different machine learning algorithms. Finally, the last stage is the benchmarking stage that validates and checks the numbers calculated in stage 4.

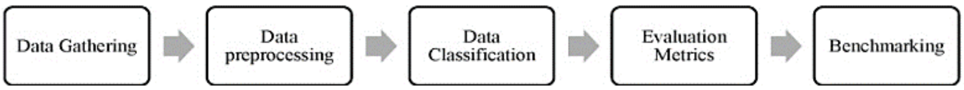


Figure 2. The five sequential stages that the proposed system consisted of

Table 2. The five different classes of the dataset

Value	Description
0	Normal operation (no faults)
1	Short-circuit (short circuit between 2 modules of a string)
2	Degradation (there is a resistance between 2 modules of a string)
3	Open circuit (one string disconnected from the power inverter)
4	Shadowing (shadow in one or more modules)

3.1. Dataset

The suggested systems for detecting and classifying faults use system identification and ML approaches, that need a huge dataset of historical operational data for training them, especially when using the machine learning techniques. The variety of the training dataset, containing operating data under all the considered faults over the whole range of environmental circumstances, is critical to the correctness of these algorithms in various scenarios. The proposed paper developed a simulator for PV plant which is capable of quickly producing the requisite dataset, since waiting for all of these environmental combinations to naturally occur in order to produce the required faults is unfeasible for most PV systems. The created model, on the other hand, must precisely capture the behavior of an actual system. The process for introducing operational failures in an actual installation is first presented. The suggested electrical simulator that replicates the actual installation behavior is provided in the sequence.

3.2. Fault classification

The faults can occur naturally such as shadowing which is produced by nearby buildings, or they can be artificial like open circuit, degradation, and short-circuit. Degradation was induced by introducing a resistive load bank between two modules. Opening one of the string's main circuit breaker induced the open circuit, while cable connected between the positive connection of a module and the negative connection of the adjacent module was used to induce a short circuit.

If a fault is discovered at a certain time (f(k)=1), the user is alerted to it through a fault classification block which identifies the most probable reason for the abnormal operation. We tested the accuracy of the four most prevalent supervised ML approaches for this assignment which are (decision tree, linear discriminant analysis for machine learning, k-nearest neighbors (KNN), and Naive Bayes algorithm). The variables used as input for these methods are those which represent the behavior of the DC side of the PV plant, which is wherever the faults happen, and creating a feature vector as in (1).

$$\mathrm{FV}(\mathrm{k}) = [\mathrm{g}(\mathrm{k}) \quad \mathrm{t}(\mathrm{k}) \quad \mathrm{V}_{\mathrm{dc},1}(\mathrm{k}) \quad \mathrm{V}_{\mathrm{dc},2}(\mathrm{k}) \quad \mathrm{i}_{\mathrm{dc},1}(\mathrm{k}) \quad \mathrm{Vi}_{\mathrm{dc},2}(\mathrm{k})\ ] \tag{1}$$

4. RESULTS AND DISCUSSION

All the findings are presented and discussed in this section. First, individual defect detection findings are shown, with recursive techniques compared and the best applicable model developed in this study indicated. The simulated data are emphasized such that the simulated data's validation of for classification is shown in the sequence. Following that, the outcomes of individual fault categorization for several ML models are provided.

The evaluation of classification algorithms began with deciding the most suitable algorithm for this step. The acquired results, accuracy, loss, F1 score, precision, and recall are shown in Table 3. The categorization of the actual dataset was used to determine test accuracy. The discriminant algorithm and the nave bayes method have the least accuracy for our use case, as well as being the ones that suffer the biggest performance loss when actual data is used. Remarkabley, the classification tree and KNN classifiers outperformed the validation data using actual data. The fact that there are extreme temperature values in the training set that are not recorded in actual data may justify these results. Those circumstances caused dissatisfying classification outcomes in the training set, but fortunately, they are not present in the test set as the actual data requires a temperature between five and 32 degrees Celsius.

Some criteria are usually used to evaluate the model's performance whatever it was. Often, F1 score, recall, precision, as well as accuracy are used for evaluating the model's performance. The majority class is often named the negative outcome (e.g., such as "*no change*" or "*negative test result*"), and the minority class is often named the positive outcome (e.g., "change" or "positive test result"). Precision measures the fraction of the data points that our model says are relevant actually and are in fact relevant. The following formula can be used to calculate precision (2).

$$precision = \frac{\text{true positive}}{\text{true positive+false positive}} \tag{2}$$

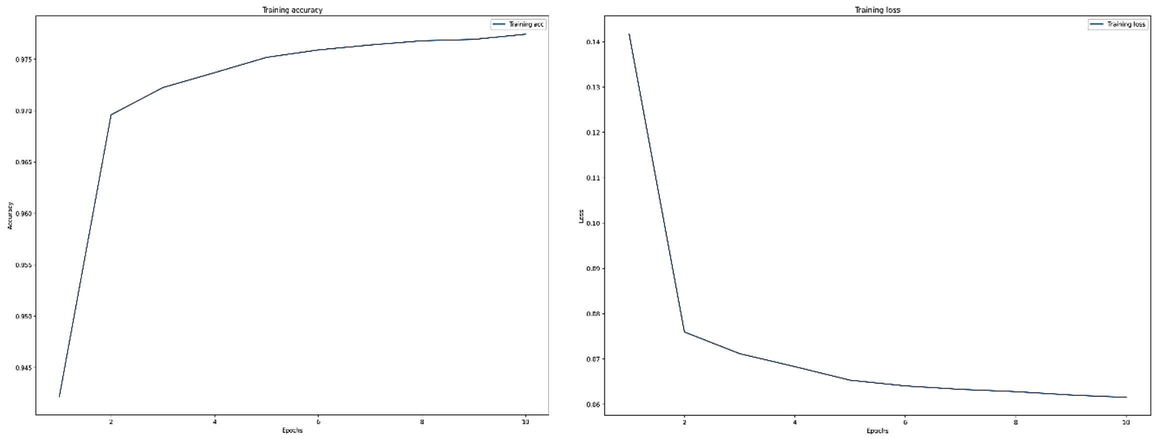
Recall is the measurement of the total relevant results correctly classified by our model in percentage. The following formula can calculate recall (3) [32].

$$recall = \frac{true\ positive}{true\ positive + false\ negative} = \frac{true\ positive}{total\ actual\ positive} \tag{3}$$

F1 score is explained as being the weighted average of the precision and recall. The best value for F1 score is 1 while the worst is 0 [32]. The following formula calculates F1 score (4) [32].

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \tag{4}$$

The overall results of our proposed model are gathered in Table 3, such that the accuracy is 99.0% and the F1 score is 0.98. Figure 3 shows the training and validation (Figure 3(a) for accuracy and Figure 3(b) for loss). Figure 4 shows comparison of the four algorithms based on recall, precision, F1 score, loss, and accuracy.



(a) (b)

Figure 3. Training and validation (a) accuracy and (b) loss

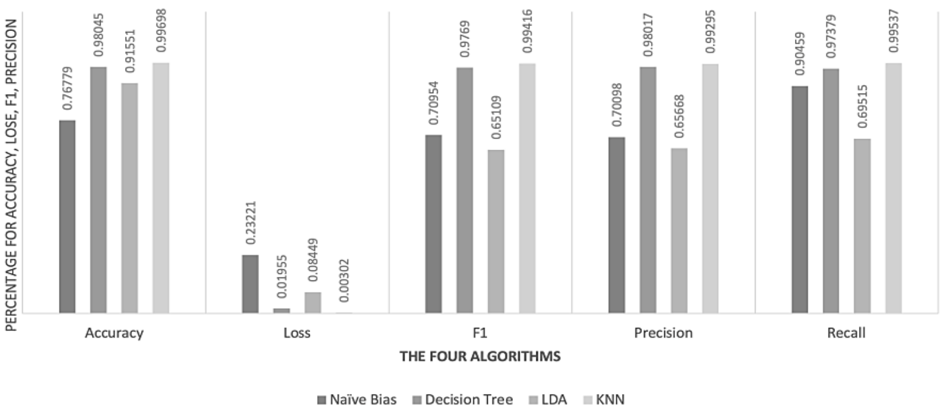


Figure 4. Comparison of the four algorithms based on recall, precision, F1 score, loss, and accuracy

Table 3. Results for machine learning algorithms in terms of accuracy, loss, F1 score, precision, and recall

Algorithm	Accuracy	Loss	F1	Precision	Recall
Naïve bias	0.76779	0.23221	0.70954	0.70098	0.90459
Decision tree	0.98045	0.01955	0.97690	0.98017	0.97379
LDA	0.91551	0.08449	0.65109	0.65668	0.69515
KNN	0.99698	0.00302	0.99416	0.99295	0.99537

Detection accuracy (DA) refers to the ratio between the number of fault nodes that have been correctly identified and the total number of actual fault nodes, whereas loss represent the ratio between the number of fault nodes that have been not identified or missed and the total number of actual fault nodes. Figure 5 shows the data retrieved form the solar panels. However, this data might indicate the presence or absence of errors. The values of errors and their classification are shown above. The retrieved data include the normal operation, in addition to: shadowing, open circuit, degradation and short circuit. From Figure 5, the bar plot resembles the output of the model and the frequency of each possible output. Thus, it is clear that the most frequent output is normal operation, whereas shadowing occurs significantly less times that normal shadowing, and rarely any degradation is shown alongside short and open circuits. In Figure 6, we illustrate a box plot showing the outliers in the data. The data in our study are continuous, thus it becomes clear when outliers are existing. This figure shows a lack of outliers which positively influences the obtained results and it signifies the good quality of our model's result.

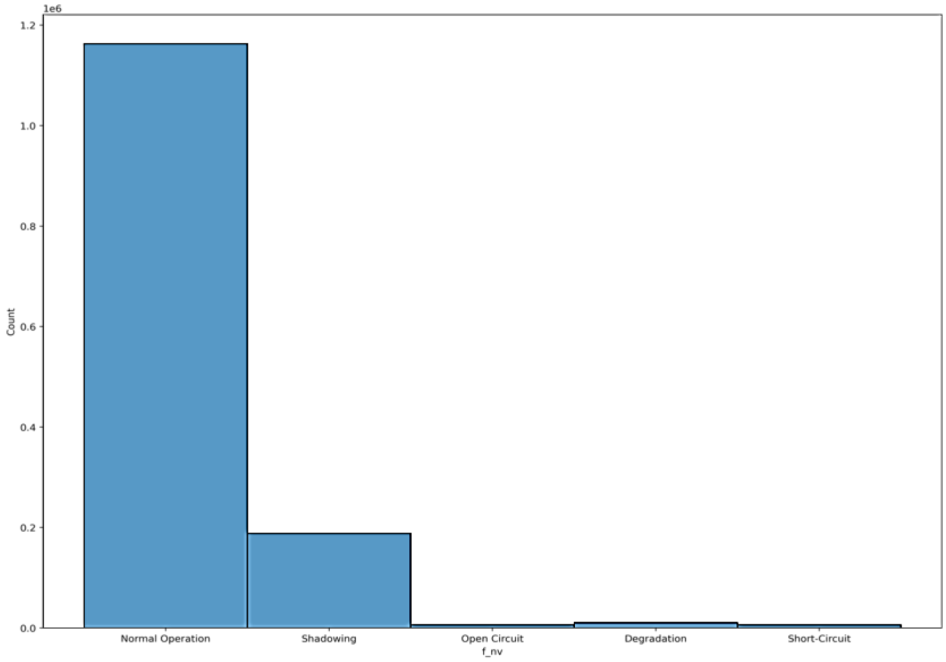


Figure 5. Histogram of the received data reveling normal operations, shadowing, open circuit, degradation, and short circuit frequencies

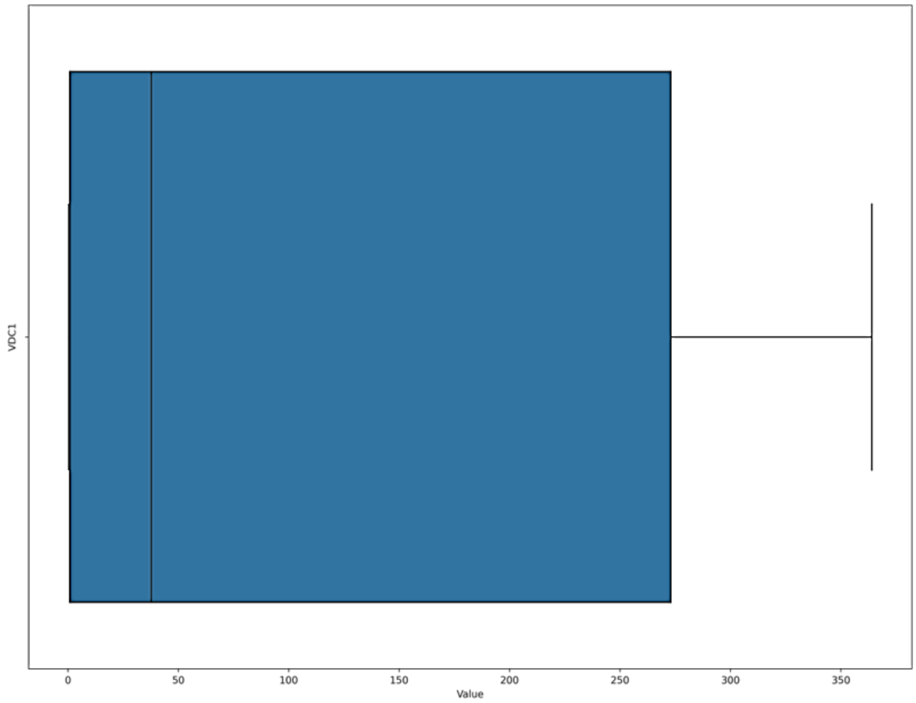


Figure 6. The outliers of data

Figure 7 shows the relation between error occurrences as an output with respect to the input parameters. The figure contains four sub-figures, each for a specific input parameter: vdc1 (Figure 7(a)), vdc2 (Figure 7(b)), idc1 (Figure 7(c)), and idc2 (Figure 7(d)). For instance, in Figure 7(a) it is evident that normal operations are happening between 0 and 250 volts, whereas shadowing is mainly taking place between 250 and 300 volts. Similarly, in Figure 7(b), normal operations, and shadowing occur at the same levels as vdc1. Yet, the difference between vdc1 and vdc2 is evident when it comes to short circuit, such that the short circuit occurs beyond 250 volts in vdc1 but at 200 volts in vdc2. Figure 7(c) and Figure 7(d) show the relationship between the occurrence of errors and the current parameters. Similar to the voltage parameters, normal operations occur at lower current levels, while shadowing mainly occurs at higher current levels. Moreover, the short circuit occurs at high current levels in idc1 and idc2.

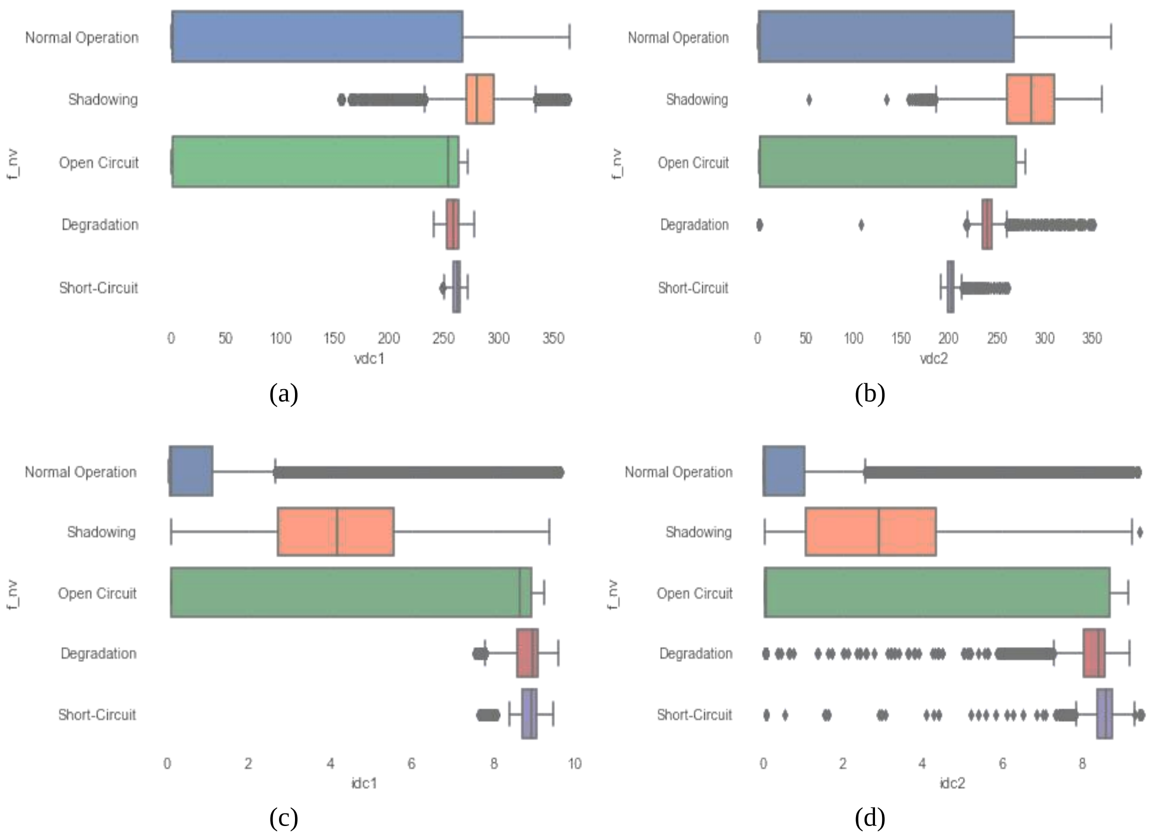


Figure 7. The relationship between error occurrences (a) vdc1, (b) vdc2, (c) idc1, and (d) idc2

5. CONCLUSION

Machine learning has shown a promising result in classifying the nature of the PV panel along with its defects. However, the classification algorithm should be selected based on the nature of the data present. Moreover, based on the different criterion measured in section four such as F1 score, precision, and recall measures, we can infer that the absence of extreme temperature values in the training set causes a skyrocketing performance of the KNN and decision trees compared to the LDA and Naïve bayes algorithm. As a result, choosing the optimal algorithm for performing the classification is based on the data characteristics such as its variance along with-it containing data from all the different circumstances and conditions.

REFERENCES

[1] A. Jäger-Waldau, "PV Status Report 2019," *Publications Office of the European Union: Luxembourg*, pp. 7-94, 2019, doi:10.2760/326629.
[2] M. Saad, H. M. Sultan, A. Abdeltwab, and A. A. Z. Diab, "Feasibility study of a small-scale grid-connected PV power plants in egypt; case study: new valley governorate," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 624–631, 2021, doi: 10.14569/IJACSA.2021.0120673.
[3] S. M. Babbar, C. Y. Lau, and K. F. Thang, "Long term solar power generation prediction using adaboost as a hybrid of linear and non-linear machine learning model," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, pp. 536–545, 2021, doi: 10.14569/IJACSA.2021.0121161.
[4] M. Abdolzadeh and M. Ameri, "Improving the effectiveness of a photovoltaic water pumping system by spraying water over the front of photovoltaic cells," *Renewable Energy*, vol. 34, no. 1, pp. 91–96, Jan. 2009, doi: 10.1016/j.renene.2008.03.024.
[5] M. Dhimish, V. Holmes, and M. Dales, "Parallel fault detection algorithm for grid-connected photovoltaic plants," *Renewable Energy*, vol. 113, pp. 94–111, Dec. 2017, doi: 10.1016/j.renene.2017.05.084.
[6] A. Chouder, S. Silvestre, B. Taghezouit, and E. Karatepe, "Monitoring, modelling and simulation of PV systems using LabVIEW," *Solar Energy*, vol. 91, pp. 337–349, May 2013, doi: 10.1016/j.solener.2012.09.016.

[7] B. Ando, S. Baglio, A. Pistorio, G. M. Tina, and C. Ventura, "Sentinella: smart monitoring of photovoltaic systems at panel level," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 8, pp. 2188–2199, Aug. 2015, doi: 10.1109/TIM.2014.2386931.
[8] V. S. B. Kurukuru, A. Haque, M. A. Khan, and A. K. Tripathy, "Fault classification for photovoltaic modules using thermography and machine learning techniques," in *2019 International Conference on Computer and Information Sciences, ICCIS 2019*, Apr. 2019, pp. 1–6, doi: 10.1109/ICCISci.2019.8716442.
[9] C. H. D. Costa *et al.*, "A monitoring and management system for medium-scale photovoltaic plants," in *2019 IEEE PES Conference on Innovative Smart Grid Technologies, ISGT Latin America 2019*, Sep. 2019, pp. 1–6, doi: 10.1109/ISGT-LA.2019.8895463.
[10] Z. Li, Y. Wang, D. Zhou, and C. Wu, "An intelligent method for fault diagnosis in photovoltaic array," in *Communications in Computer and Information Science*, vol. 327 CCIS, no. PART 2, 2012, pp. 10–16, doi: 10.1007/978-3-642-34396-4_2.
[11] Z. Chen, L. Wu, S. Cheng, P. Lin, Y. Wu, and W. Lin, "Intelligent fault diagnosis of photovoltaic arrays based on optimized kernel extreme learning machine and I-V characteristics," *Applied Energy*, vol. 204, pp. 912–931, Oct. 2017, doi: 10.1016/j.apenergy.2017.05.034.
[12] A. Drews *et al.*, "Monitoring and remote failure detection of grid-connected PV systems based on satellite observations," *Solar Energy*, vol. 81, no. 4, pp. 548–564, Apr. 2007, doi: 10.1016/j.solener.2006.06.019.
[13] J. A. Tsanakas, D. Chrysostomou, P. N. Botsaris, and A. Gasteratos, "Fault diagnosis of photovoltaic modules through image processing and Canny edge detection on field thermographic measurements," *International Journal of Sustainable Energy*, vol. 34, no. 6, pp. 351–372, Jul. 2015, doi: 10.1080/14786451.2013.826223.
[14] E. Garoudja, F. Harrou, Y. Sun, K. Kara, A. Chouder, and S. Silvestre, "A statistical-based approach for fault detection and diagnosis in a photovoltaic system," in *2017 6th International Conference on Systems and Control, ICSC 2017*, May 2017, pp. 75–80, doi: 10.1109/ICoSC.2017.7958710.
[15] R. S. Kulkarni and D. B. Talange, "Modeling of solar photovoltaic module using system identification," in *2017 7th International Conference on Power Systems, ICPS 2017*, Dec. 2018, pp. 782–784, doi: 10.1109/ICPES.2017.8387395.
[16] M. P. Rodrigues *et al.*, "A MISO nonlinear model of photovoltaic panel based on system identification," in *2019 IEEE PES Conference on Innovative Smart Grid Technologies*, 2019, pp. 1–6, doi: 10.1109/ISGT-LA.2019.8895325.
[17] C. E. Packard, J. H. Wohlgemuth, and S. R. Kurtz, "Development of a visual inspection data collection tool for evaluation of fielded pv module condition," Golden, CO (United States), Aug. 2012. doi: 10.2172/1050110.
[18] S. R. Madeti and S. N. Singh, "Monitoring system for photovoltaic plants: a review," *Renewable and Sustainable Energy Reviews*, vol. 67, pp. 1180–1207, Jan. 2017, doi: 10.1016/j.rser.2016.09.088.
[19] Y. Chouay and M. Ouassaid, "An intelligent method for fault diagnosis in photovoltaic systems," in *Proceedings of 2017 International Conference on Electrical and Information Technologies, ICEIT 2017*, Nov. 2018, vol. 2018-January, pp. 1–5, doi: 10.1109/EITech.2017.8255225.
[20] C. H. Da Costa *et al.*, "A comparison of machine learning-based methods for fault classification in photovoltaic systems," in *2019 IEEE PES Conference on Innovative Smart Grid Technologies, ISGT Latin America 2019*, Sep. 2019, pp. 1–6, doi: 10.1109/ISGT-LA.2019.8895279.
[21] S. R. Madeti and S. N. Singh, "A comprehensive study on different types of faults and detection techniques for solar photovoltaic system," *Solar Energy*, vol. 158, pp. 161–185, Dec. 2017, doi: 10.1016/j.solener.2017.08.069.
[22] A. Mellit, G. M. Tina, and S. A. Kalogirou, "Fault detection and diagnosis methods for photovoltaic systems: a review," *Renewable and Sustainable Energy Reviews*, vol. 91, pp. 1–17, Aug. 2018, doi: 10.1016/j.rser.2018.03.062.
[23] A. Triki-Lahiani, A. Bennani-Ben Abdelghani, and I. Slama-Belkhodja, "Fault detection and monitoring systems for photovoltaic installations: A review," *Renewable and Sustainable Energy Reviews*, vol. 82, pp. 2680–2692, Feb. 2018, doi: 10.1016/j.rser.2017.09.101.
[24] "Std. 61724, Photovoltaic system performance monitoring-guidelines for measurement, data exchange and analysis," in International Electro-technical Commission IEC 61724, 1998.
[25] K. J. Ali, A. H. Mohammad, and G. T. Hasan, "An empirical correlation of ambient temperature impact on PV module considering natural convection," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 19, no. 2, pp. 627–634, Aug. 2020, doi: 10.11591/ijeecs.v19.i2.pp627-634.
[26] B. Cojocariu, C. Petrescu, and D. Stefanoiu, "Photovoltaic generators-modeling and control," in *Proceedings - 2015 20th International Conference on Control Systems and Computer Science, CSCS 2015*, May 2015, pp. 32–37, doi: 10.1109/CSCS.2015.96.
[27] W. D. Soto, S. A. Klein, and W. A. Beckman, "Improvement and validation of a model for photovoltaic array performance," *Solar Energy*, vol. 80, no. 1, pp. 78–88, Jan. 2006, doi: 10.1016/j.solener.2005.06.010.
[28] A. Alqahtani, S. Marafi, B. Musallam, and N. E. D. A. El Khalek, "Photovoltaic power forecasting model based on nonlinear system identification," *Canadian Journal of Electrical and Computer Engineering*, vol. 39, no. 3, pp. 243–250, 2016, doi: 10.1109/CJECE.2016.2584081.
[29] Y. Li, Y. Su, and L. Shu, "An ARMAX model for forecasting the power output of a grid connected photovoltaic system," *Renewable Energy*, vol. 66, pp. 78–89, Jun. 2014, doi: 10.1016/j.renene.2013.11.067.
[30] Y. Hammoudi, I. Idrissi, M. Boukabous, Y. Zerguit, and H. Bouali, "Review on maintenance of photovoltaic systems based on deep learning and internet of things," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 26, no. 2, pp. 1060–1072, May 2022, doi: 10.11591/ijeecs.v26.i2.pp1060-1072.
[31] W. Chine, A. Mellit, V. Lughi, A. Malek, G. Sulligoi, and A. M. Pavan, "A novel fault diagnosis technique for photovoltaic systems based on artificial neural networks," *Renewable Energy*, vol. 90, pp. 501–512, May 2016, doi: 10.1016/j.renene.2016.01.036.
[32] A. Wills, T. B. Schön, L. Ljung, and B. Ninness, "Identification of hammerstein-wiener models," *Automatica*, vol. 49, no. 1, pp. 70–81, Jan. 2013, doi: 10.1016/j.automatica.2012.09.018.
[33] R. S. Tsay, *Analysis of Financial Time Series*, John Wiley & Sons, 2005.
[34] S. Abirami and P. Chitra, "Energy-efficient edge based real-time healthcare support system," in *Advances in Computers*, 2020, vol. 117, no. 1, pp. 339–368, doi: 10.1016/bs.adcom.2019.09.007.

BIOGRAPHIES OF AUTHORS

Shaimaa Naser Kabour is certified instructor at cisco networking academy. She is skilled in artificial intelligence, python, robotics, IOT, networking, and scratch. Strong education professional with a master's degree focused in communication and computer engineering from lebanese International University, and teaching diploma degree. Her research interests include software engineering, operating systems, machine learning, photovoltaic system, industrial applications, industrial informatics, and intelligent systems. She can be contacted at email: shaimaakabour@gmail.com.

Raghad Adnan Almalki received her B.Sc. degree in computer engineering from Taif University, Taif, Saudi Arabia. She is currently working as a system analyst. She is skilled in HTML, CSS, JavaScript, Python, PHP and Sql. Her research interests include machine learning, industrial applications, industrial informatics, and intelligent systems. She can be contacted at email: raghadnanmalki@gmail.com.

Lujain A. Alghamdi received the B.Sc. degree in computer engineering from Taif University, Taif, KSA. Her skills include MySQL, C++, Matlab, adaptability and networking. Her research interests include software engineering, operating systems, computer architecture, industrial applications, machine learning, and databases. She can be contacted at email: lojain_54@outlook.com.

Wujud Alharthi received her B.Sc. degree in computer engineering from Taif University, Taif, Saudi Arabia. Her skills include MySQL, C++, Java and Matlab. Her research interests include cybersecurity, artificial intelligence and cloud computing, soft computing, machine learning, industrial informatics, and intelligent systems. She can be contacted at email: alharthiwa@gmail.com.

Nisreen B. Alshagi certified CCNAv7: introduction to networks, IoT fundamentals: connecting things, internet of things workshop (IoT) and cybersecurity essentials from cisco networking academy 2021-2022. She is skilled in python, android, java, C#, vb.net. She is currently a master student at the thesis stage. She can be contacted at email: nisren.alshagi@academy.edu.ly.