IoT-Enabled Cloud-Based Industrial Monitoring and Management Framework

Dasaraju Chandra Mohan*[1], Yogesh Rajkumar R[2],
[1]Department of Computer Science and Engineering Bharath Institute of higher education and research Chennai, India,
[2]Department of Information Technology Bharath Institute of higher education and research Chennai, India
chandraraju9427@gmail.com

Abstract. The given paper introduces Intelligent Monitoring System (IMS) dedicated to Photovoltaic (PV) plants with the usage of the low-cost hardware and lightweight software that would make its deployment face-free in a variety of PV installations. The system has a platform that is based on the Internet of Things (IoT) platform, which allows the platform to communicate seamlessly, interoperate, and handle data in real-time. An embedded personal cloud server is added to perform effective computation and safe storage of PV system data, and a web-based monitoring interface is given to allow the visualization of several users. IMS enables the use of advanced deep ensemble-based learning in detecting faults and predicting power. The model forecasts PV output at different environmental conditions due to a long short-term memory (LSTM) ensemble model to enable optimal energy production and early detection of malfunctions. Fault diagnosis is done based on features found in Current voltage (I V) characteristics and with the help of an ensemble of Naive Bayes, K Nearest Neighbors and Support Vector Machine models with addition of a feature selection algorithm. An actual PV installation was used to test the system and proved the IMS is scalable, interoperable, and useful in the overall monitoring of PV plants, both data acquisition, performance measurement, fault identification, and predictive analytics.

Keywords: Intelligent Monitoring System (IMS), IoT Platform, Cloud Computing, LSTM Ensemble, Deep Learning, Fault Detection, Ensemble Learning.

1 Introduction

Photovoltaic (PV) technology and its installations have seen a booming growth over the recent years and it is expected to continue growing with cumulative installations of over 758 GW in 2020 and exceeding 1 TW in the near future. With the accelerated pace of PV deployment, the matters of system reliability, the stability of the operations, and long-term performance have become even more significant. The PV modules are key to the efficiency, cost-effectiveness, and lifespan of the solar power plants, but the

S. P. Vijayaragavan et al. (eds.), *Proceedings of the Global Conference on Sustainable Energy Systems, Smart Electronics and Intelligent Computing (GCSESEIC 2025)*, Advances in Engineering Research 297,
https://doi.org/10.2991/978-94-6239-654-8_18

environmental factors, as well as the degradation of the components and the issues within the working process, tend to influence them. These problems can greatly lower the yield of energy when they are not early detected and corrected. The conventional PV monitoring methods also encounter a number of issues such as expensive nature of commercial data acquisition systems, use of special computing hardware and reliance on heavy proprietary software platforms. Also, the power of PV output is subject to changes in weather conditions and system failures, thus real-time monitoring and predictive analytics are necessary to ensure successful operation and maintenance. Lack of proper supervision may lead to reduced life span of the system, more downtimes and losses that are significant economically [1].

In the past 20 years, there are many proposed monitoring systems that include data loggers that are part of LabVIEW systems, wireless acquisition units, and also those that are based on power-line communication. Even though these solutions enhanced data collection and visualization, a lot of them are costly to implement, complex in hardware, and have less flexibility, which does not favor small-scale or remote PV installations. The recent developments in the low-cost embedded system like Raspberry Pi and the advent of Internet of Things (IoT) technologies has brought forth new opportunities of real-time, scalable, and affordable PV monitoring. The use of machine learning methods to predict output power and fault diagnosis has also been investigated, but the current methods tend to be susceptible to bad data, are computationally costly, or limited to a particular environment [2].

To overcome these shortcomings, this paper presents an Intelligent Monitoring System (IMS) that incorporates the IoT technology, cloud computing, and lightweight web-based interfaces to offer an overall, but cost-effective, monitoring system of PV plants. The suggested system comprises a low-cost data acquisition hardware platform, a cloud-based data processing and storage service, and a simple web monitor that can be viewed by more than one user. To analyze its performance, a group long short-term memory (LSTM) neural network is employed in predicting one-day-ahead power at the last stage to provide a better estimate of the energy yield and detect faults at an early stage. Moreover, a machine learning ensemble model that incorporates Naive Bayes (NB), K-Nearest Neighbors (KNN) and Support Vector Machine (SVM) is used in the multi-type fault detection and classification such as the open-circuit fault, string degradation fault and array degradation fault. The main task of the IMS is to reduce hardware expenses, at the same time increasing the precision of monitoring and the flexibility of the system. It is intended to be scalable, interoperable and applicable to both stand-alone and remote PV installations. Comprehensively, the suggested monitoring framework provides a comprehensive solution that facilitates the data collection, performance analysis, predictive analytics, and auto fault diagnosis as an efficient and viable way of making PV plants more reliable and efficient [3].

2 Literature review

Conventional PV surveillance and primitive systems. The initial PV monitoring was based on commercial data loggers and PC based software suites. These systems were dependable in measurement and visualization but were more often than not costly, hardware intensive and reliant on proprietary platforms. Numerous of them utilized multi-channel data reclamation units and desktop software (e.g., LabVIEW) to obtain climatic and electrical measurements and create records. Although these solutions proved to be useful in research and large utility-scale installations, small, decentralized, and remote installations did not fit very well, because of their cost, complexity, and limited scalability [4].

System constraints and technology communication. To transmit sensor data to remote servers, researchers investigated various communication media wired solutions, power-line communication (PLC), and short-range wireless (ZigBee) to transmit sensor data. PLC had the benefit of leveraging existing cabling but was limited by costs and at the panel level, ZigBee and other wireless solutions had minimized wiring but were usually dependent on a PC or heavy back-end to handle the received streams. Typical shortcomings in these studies are the high cost of sensors and loggers (e.g. precision pyranometers), reliance on centralized PCs and inappropriateness in remote or low budget PV plants [5].

IoT and cheap hardware platforms. Due to the availability of cheap microcontrollers and single-board computers (such as ESP8266/ESP32 and Raspberry Pi), small, energy-saving data-acquisition nodes were now able to be connected directly to the Internet. The IoT platforms are built based on these cheap nodes, cloud storage and lightweight backends to provide increased scalability and remote access. Experiments of this type have shown that large cost savings can be achieved and still achieve reasonable measurement fidelity, but hardware integration and stable connectivity in distant locations are still viable issues [6].

Light weight backends and cloud computing. Backends on the cloud have the benefit of being scaled, allowing parallel processing as well as remote deployment of models. As opposed to using proprietary heavyweight cloud services, some recent publications indicate that more basic, open-source backends (Python/Flask, REST APIs, lightweight databases or CSV logs) are effective in analytics driven by ML in PV monitoring. Cloud computing eliminates local compute capacity, streamlines updates to models and consolidates historical data to evaluate long term performance.

Connectivity dependence and the outage requires a strong synchronization strategy are some of the trade-offs [7].

Strategies of cloud data logging and resilience. Good field loggers have to ensure that the data is intact even when they lose connectivity. Store-and-forward Systems with local buffering logic and retry/synchronization logic (store-and-forward) guarantee that no data is lost and that the data sets will be consistent in their future analysis. Lightweight and open logging designs that log to a simple database or CSV on the server-side have been embraced due to ease of implementation and readability. In general, the literature justifies the shift of monolithic PC-based loggers to distributed IoT-enabled data loggers having local persistence [8].

Time-series forecasting: ANN to LSTM ensembles. Planning and fault detection are based on accurate short-term power forecasting. There have been successful and inconclusive results to the application of classical ANNs, SVMs, and KNN to predicting PV power: ANNs are able to model nonlinear relationships but are sensitive to noise, SVMs are able to generalize very well on small datasets but do not scale well to large feature spaces, and KNN is simple and unable to behave in large dimensionality. Recurrent architectures, in particular LSTM networks are geared towards resolving the long-term dependency issue in sequential data and have become highly popular in PV time-series forecasting. It is also found that ensemble strategies involving many predictors minimize the variance and enhance robustness especially in changeable weather conditions [9].

Machine learning fault detector and diagnosis. Fault diagnostic methods based on machine learning compute discriminative signals based on measured signals and use supervised classifiers to indicate faulty conditions. Research activity in the area is between model-based comparisons (measured vs. modeled outputs) and data-driven classifiers like fuzzy systems, MLP and SVM. Solutions based on models are highly intuitive, however being sensitive to modeling error and threshold sensitivity, purely data-driven techniques must have representative fault datasets and take care with feature engineering. Recent studies have emphasized the advantage of ensemble classifier which includes the complementary strength of multiple learners into better classification accuracy along with robustness of the classification [10].

Loopholes and impetus to the current work. Whereas the literature of successful components of IoT loggers, cloud backends, LSTM forecasting, and ML-based fault diagnosis are abundant, little of the literature has brought these components together in an affordable, end-to-end framework that meets the needs of small and remote PV installations. The main gaps are that most previous systems were high-component cost, relied on PC-based processing, did not give much consideration to robust data buffering and synchronization and did not explore the ensemble LSTM design as well as feature-selected EN classifiers to practical field applications. Such gaps are the reasons why this current paper is devoted to developing a lightweight, low-cost IMS combining ESP8266-based logging, cloud-hosted ML processing, physics-driven I-V feature extraction, SFFS selection, and ensemble LSTM/EN models to provide accurate forecasting and fault classification in the various PV settings [11].

3 Proposed methodology

This study is aimed at developing an inexpensive, lightweight yet high precision monitoring system that will be able to accurately monitor the performance of photovoltaic (PV) plants. In this regard, an Intelligent Monitoring System (IMS) is suggested, which incorporates the IoT-based sensing, cloud-based data processing, and machine learning models to predict power tomorrow and diagnose faults autonomously. Millions of meteorological and electrical parameters are collected in the system with the help of sensors on PV modules, arrays, and the environment. These sensors obtain real-time irradiance, temperature, current, and voltage values that are necessary to estimate the PV performance and detect anomalies [12].

At the center of the data acquisition unit is a cloud-based data logger that is made of ESP8266 and microcontroller boards. The ESP8266, a cheap, IoT-based Wi-Fi device, is used to first process electrical measurements to provide capabilities to be used as a real-time monitor. ESP8266 sends the sensor data which has been processed to a cloud server via a normal internet connection after preliminary formatting. The module will have a configuration to have several access points, which will provide continuity in the transmission of data. The proposed IMS uses simple communication, unlike the traditional systems, and utilizes the common protocols used in the IoT to facilitate fast, seamless, and low-cost data transfer [13].

Backup mechanism is used at the level of PV plant in order to ensure reliability of the system. A local storage of sensor data is in place in the event of a temporary communication loss or hardware disruption, which is later updated with the server when the system is reinstated. The privacy of the end-users regarding data is also observed since sensor readings are stored in the environment of the servers that are controlled by the plants and are not disseminated to external servers. Continuous data streams pass through the server, which carries out preprocessing and implements the suggested ensemble deep learning model to predict power output of PV in a day ahead. The features derived on the basis of I V characteristics and environmental measurements are used to perform fault detection. These characteristics are categorized based on a collection of machine learning algorithms to obtain possible patterns of malfunctions in PV modules and arrays

3.1 IoT platform

The Internet of Things (IoT) is particularly important in helping to turn PV systems into real time monitoring by developing a network of connected sensors, microcontrollers and cloud services that can communicate with one another independently. The IOT platform in the proposed IMS will ensure the collection of environmental and electrical parameters and generate them to the cloud server where they can be stored and analyzed. An IoT system starts with the microcontroller unit (MCU), which the program is loaded on to manage sensor data acquisition routines and communication routines. This program has the instructions required to read sensor values and transmit them to the cloud by a normal HTTP POST request [14].

The ESP8266 module once the IoT firmware has been installed on the microcontroller creates an internet connection with a configured access point. Once connected, the module starts sending data packets or voltage, current, temperature, and irradiance measurements to the cloud server in regular intervals. In case of a network failure, the ESP8266 tries to regain connection automatically, and the process of transmission is not so sensitive that it cannot continue without human intervention. In case of the complete transfer failure, the system will create an error message in the system and save the data locally until communication is restored. This means that no sensor data is left out and the cloud server will continue to get synchronized and full datasets to be further processed.

3.2 Cloud server

With the offered IMS, the cloud server will serve as the main computing and storage system, which will make it possible to process PV system data delivered to the IoT platform efficiently. The system is built on a simple and inexpensive Python, Flask, and simple hosting, cloud infrastructure instead of using more sophisticated commercial cloud infrastructures or SaaS-level services. This would be affordable and yet has enough computational power to run machine learning jobs like power prediction, fault diagnosis, etc. The cloud server receives sensor data that is received, preprocesses it where necessary, and implements the ensemble learning models that have been created to work with the IMS.

Scalability and the availability of data are also the main purposes of cloud computing. The system will be able to consume a large amount of historical and real-time sensor data without being constrained by the local hardware because the backend is hosted on a cloud server. Some of the activities that the server handles include the execution of prediction algorithms, updating of the model outputs as well as maintaining a synchronized log of all measurements that are captured. Even though the platform does not implement more sophisticated models of cloud services, such as SaaS, PaaS, or IaaS, the platform enjoys the essential features of the cloud setting, like remote access, automated data processing, and fault tolerance against local device failures. This diminutive cloud-based architecture is guaranteed to have high-performance reliability at low implementation and operational cost [15].

3.3 Cloud logger

Within the suggested IMS, the cloud data logger is the communication interface between the sensor network at the field level and the computation unit on the cloud. The data logger sends real-time environmental and electrical measurements of the microcontroller to a remote cloud data use an active Wi-Fi connection. In contrast to a traditional commercial logging system based on Google Scripts or a proprietary system, the proposed architecture would be based on a completely customized and low-cost system based on open-source software and cheaper microcontroller hardware. This makes sure that the system is flexible, lightweight, and it cannot easily be topped in various configurations of PV plants. The data logger is initiated to a process where the firmware that is stored in the microcontroller is executed and is coded in C/C++ and is set to read the sensor values within a set time interval [16].

After the start, ESP8266 module also connects to the Wi-Fi network at hand. The logger transfers the data collected to the cloud server through RESTful HTTP request once a successful connection is achieved. In case of the loss of a connection, the gadget reconnects automatically until the connection has been regained to avoid any loss of data. Once the data has been received, the cloud server authenticates the request and archive the data in a structured format in the form of CSV or database. This archived data is subsequently employed in machine learning activities such as prediction of power and malfunctioning. The proposed data logger will not require any complex authentication processes, third-party scripts, or specific cloud services because of a lightweight cloud-based backend that is implemented using Python and Flask. This ensures that the system is very inexpensive, maintainable, and reliable and yet still provides the PV monitoring application with the accurate and continuous data recording [17].

3.4 Using ML techniques

Forecasting Architecture of Power

Recurrent neural networks (RNNs) are also common in modeling sequential and time-dependent data which is why they can be used to predict power generation in PV systems. Nevertheless, the traditional RNNs are commonly subject to severe problems like the vanishing and exploding gradients that inhibit the possibility of learning long-term dependencies using historical data. In order to overcome these shortcomings, the Long Short-Term Memory (LSTM) networks have been proposed as an improved version of the RNNs that can store information during longer durations in time. This paper constructs an ensemble-based LSTM forecasting model to attain correct and strong forecasting of PV output power. Ensemble learning methods involve the use of several models in order to enhance generalization, reduce the errors of prediction, and decrease the vulnerability of the system to the noise of environmental conditions. The suggested ensemble model combines the strengths of various learners of LSTM to increase the robustness of the forecast to different weather conditions and irradiance.

LSTM Structure

The LSTM networks are an improvement of the traditional RNNs in that they have gate-based mechanisms, which control information flow. The model makes decisions on which information should be stored, updated, or discarded in the process of training with these gate input, forget, and output. This allows LSTMs to address gradient concerns and can be effectively used in the time-series prediction of problems like PV power prediction. The nonlinear activation functions and gating operations make up the internal structure of an LSTM unit where the information is stored in a memory cell that retains long term information. All these components contribute to the learning of the network of the complicated dependence between the past environmental variables and the future power outputs.

LSTM Ensemble Architecture

Ensemble learning methods are applied extensively to enhance the accuracy and the robustness of machine learning models in predictions. To improve the accuracy of forecasting the power of PVs in diverse settings and working conditions, an ensemble-based LSTM model is created in the current research. The proposed architecture is based on the amount of multiple LSTM models that are trained independently on the same input data rather than basing on a single neural network. This enables the system to learn various patterns in the time-series data, which is not very sensitive to noise and is better in generalization.

The ensemble architecture has three layers, which are input layer, training layer and forecasting layer. Each LSTM model is fed with historical electrical parameters and environmental measurements in the input layer. At the training stage, each LSTM network is trained individually, which allows each model to learn various elements of the underlying data distribution. The results of all the trained LSTM models are fused in the last forecasting layer using weighted averaging strategy. Each of the models is given a weight coefficient which indicates its accuracy in prediction such that more dependable models will have a higher weight in the final output. The assignment of weights is carried out on the basis of normalized error of each individu al .LSTM model over the training time. The weights are allocated to the models with less forecasting errors and those with greater errors respectively. This study takes into consideration two ensemble setups, one with mean absolute error (MAE) and the other with root mean square error (RMSE) to compute weight. It was experimentally shown that the combination of five LSTM models was the most suitable (M = 5) balance of computation and forecasting. The overall ensemble forecast at day d +1 is given by the weighted average of the outputs of the individual LSTM learners.

4 Fault detection and classification

In this research, the ensemble learning (EN) framework is utilized in enhancing the reliability and the accuracy of fault detection of PV arrays. The suggested model incorporates several monitored machine learning models to constitute a sound classification plan to locate various types of faults in various operating circumstances. As shown in Figure 10a, the process of building the EN-based diagnostics model includes a series of major steps, the first of which is the extraction of the corresponding fault- related features of the electrical and environmental data gathered on the PV modules. A feature selection method is used to select the most informative parameters in order to simplify the computations involved in as well as increase the discriminative strength of the classifier. This makes the model work effectively without undermining the diagnostic performance of the model. The dataset is then separated into training and validation sets, where the training set is applied to get the fault patterns learnt and the validation to adjust and test the accuracy of the classifier. Once the training stage is complete, the model is tested on an unknown test set of new and previously unknown input conditions. This measure is necessary to test the generalization property of the ensemble model and to make it reliable in the real-life PV systems. The ensemble method enables the prediction strengths of multiple base learners to combine and achieve a higher fault classification accuracy as opposed to single machine learning models.

4.1 Feature extraction

The proposed fault diagnosis model involves the extraction of diagnostic features whereby the feature extraction process is conducted in two steps namely: (1) fault-related attribute identification and (2) diagnostic features extraction. It all starts with the analysis of current voltage (I V) characteristics of the PV array since the curves will give valuable knowledge regarding the electrical behavior of the system in healthy and faulty conditions. To achieve stable I- V characteristics, output current of PV array and voltage are monitored continuously. Based on every IV curve, five critical operating points are selected, namely short-circuit current, open-circuit voltage, maximum power point (MPP), half of short-circuit current and half of open-circuit voltage. Based on these reference points, 16 diagnostic features normalized to Standard Test Conditions (STC) are obtained allowing effective and repeatable fault detection. After the extraction of features, classification occurs by a two-layer ensemble learning model. The initial layer is binary classification to determine the normal and faulty operating conditions. When a fault has been identified, the second layer performs multi-classification to detect the type of fault. The model is trained to classify three categories of PV faults of great significance, namely, open-circuit faults, string degradation faults, and array degradation faults.

4.2 Feature selection

The feature selection technique is important in enhancing the accuracy and efficiency of machine learning models since it is possible to determine the most informative set of features and remove those whose contribution to classification performance is insignificant. Such reduction not only improves the predictive power of the model, but also reduces the computational cost as well as training time. This study uses a wrapper-based method as one of the feature selection strategies because it allows one to evaluate features directly with respect to its effect on the performance of the classifier. In particular, the best feature subset is identified by the help of the Sequential Floating Forward Selection (SFFS) algorithm. SFFS would prove to be especially beneficial since it involves both forward movement and conditional backward elimination which would enable the algorithm to revisit options and optimize on the features already selected. It starts with an empty set, and then the most significant feature in terms of improving the classification accuracy is added. After every iteration, the algorithm tests whether any of the features earlier chosen during the procedure became unnecessary and useless, and discards them where needed. The process of adding and conditionally removing features by this iterative process goes on until the optimal trade of features and their accuracy is found. The SFFS algorithm can dynamically modify the subset of features used to identify the most relevant features to high fault detection and classification in PV systems.

4.3 Proposed algorithm of ensemble learning

The paper presents a study that proposes a powerful ensemble learning framework to enhance the accuracy of fault classification in photovoltaic (PV) systems. Rather than using a single classifier and having to perform worse the proposed method combines three complementary algorithms Support Vector Machine (SVM), Naive Bayes (NB), and K-Nearest Neighbors (KNN). The ensemble works based on a mechanism of probabilistic decision making, wherein every classifier is able to forecast the likelihood of a sample fitting in every fault class. The averages are then taken over all the classifiers. The label of the class that carries the greatest mean probability is declared as the final prediction. This method can capitalize on the individual strengths of individual classifiers, decrease the individual weaknesses and greatly improve the overall reliability and accuracy of PV fault diagnosis.

5 Results

5.1 Fault detection

The Fault Detection (FD) layer is a binary classifier that is developed to differentiate between the normal and faulty states of the PV system. In the case of training and evaluation, there are 400 normal events and 1400 faulty events. Of these, 1300 samples will be taken at random as targets to be trained and 400 as targets to be validated. In order to test the ability of generalization, a second sample of unseen data with 60 normal and 156 faulty events is tested again. The performance of FD layer is contrasted with majority voting-based ensemble model and performance of each individual classifier (SVM, Naive Bayes and KNN). The suggested FD layer will achieve better accuracy in all comparisons. In order to further assure the reliability of the results, 10-fold cross-validation process is used, which ensures that the feature selection procedure, and the classification model are stable. Fig.1 shows the EN classifier graph. The findings indicate that there is an improvement in accuracy with a smaller but more informative set of features and generally FD layer has a stronger and more accurate fault detection as compared to other methods that were tested.Fig.2 shows the proposed method result graph.

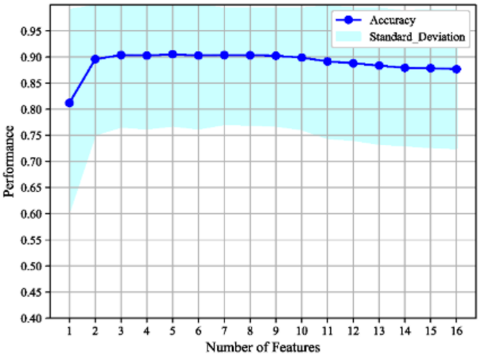

Fig. 1. EN classifier graph

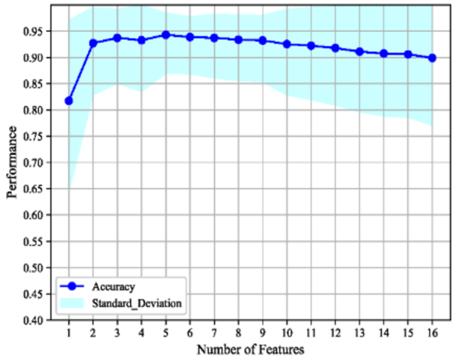

Fig. 2. proposed method result graph

5.2 Fault classification

In this section, the authors introduce a multi-class fault classification layer that can distinguish three large categories of PV system faults, which include open-circuit faults, array degradation faults, and string degradation faults. A total of 1600 labeled occurrences of faults are present in the entire dataset, which is split into three classes. In developing the model, 1280 samples are trained and 300 samples are validated and the classes are balanced. An extra hidden dataset of 160 fault events is employed to test generalization to real world in order to have samples of all three categories of faults. Two experimental conditions are carried out in order to evaluate the usefulness of various learning algorithms, both the proposed ensemble model and the specific classifiers, the SVM, Naive Bayes, and KNN. In case of the former, the proposed ensemble algorithm is much more accurate (89.11) than the majority voting ensemble algorithm, although the performance of the SVM and Naive Bayes is poorer. The suggested approach is even more effective in the second situation where the accuracy rate is 94.74% with low variation showing its greater adaptability to a wide range of operating systems and conditions shown in Fig.3 and Fig.4. Table 1 shows the performance comparison of proposed model.

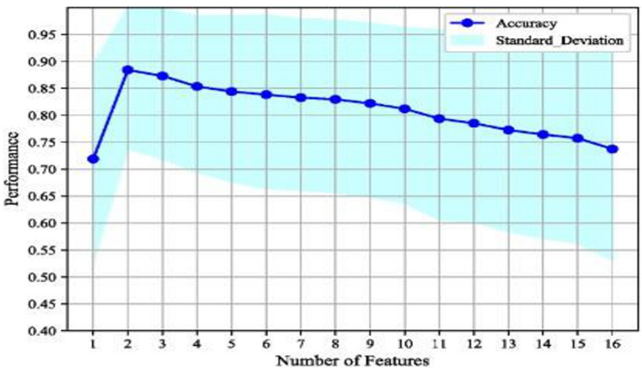

Fig. 3. EN classifier graph

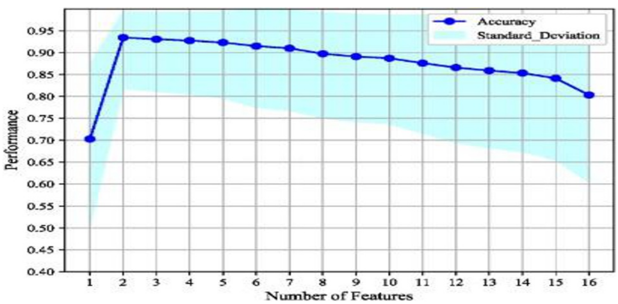

Fig. 4. proposed method result graph

Table 1 .Model performance Comparison

Model	Accuracy (%)	Precision	Recall	F1-Score
SVM	92.5	0.91	0.93	0.92
NB	88.2	0.86	0.87	0.86
KNN	90.4	0.89	0.90	0.89
Ensemble Model	**96.8**	**0.95**	**0.97**	**0.96**

6 Conclusion

Information safety is one of the main issues in cloud computing This paper shows how an IoT-based Intelligent Monitoring System (IMS) can offer smart, independent, and economical monitoring of photovoltaic (PV) plants. The system combines the lightweight software and low-cost hardware and takes deep ensemble learning models to detect fault and predict power. The Fault diagnosis of IMS is conducted in two distinct steps, and they are as follows: Firstly, key features are extracted, based on Current-voltage (I V) Characteristics in a series of normal and faulty operating conditions, and secondly, the use of an ensemble model, consisting of Naive Bayes (NB), Support Vector Machine (SVM), and K-Nearest Neighbors(KNN), is used to classify these operating conditions. In order to improve diagnostic performance, feature selection algorithm is used to determine the most significant features. The IMS is interoperable, scalable and applicable in remote and rural and urban deployments. It enables end-to-end PV monitoring, i.e. the acquisition and storage of data, pre/ post-processing, fault diagnosis, performance analysis, and the prediction of output power through cloud-based processing and IoT-enabled communication to allow access to the user. The experimental conclusions show that the LSTM ensemble model has a better accuracy in power forecasting and fault classification, which detects faults with rate of accuracy of 96% and classifies faults with an accuracy of 96% and also has a very low forecasting error. Future development will be aimed at expanding the system to identify other electrical faults in the system like diode bypass, line-ground fault and shading. In order to further improve the accuracy, one will consider the optimization algorithms such as genetic algorithms or particle swarm optimization to choose the most appropriate learners in the ensemble model, and will also test other loss functions to obtain better power prediction

References

1. López-Vargas, A., Fuentes, M., Vivar, M.: IoT application for real-time monitoring of solar home systems based on Arduino with 3G connectivity. In: IEEE Sensors Journal, vol. 19, pp. 679–691 (2018)
2. Das, U.K., Tey, K.S., Seyedmahmoudian, M., Mekhilef, S., Idris, M.Y.I., van Deventer, W., Horan, B., Stojcevski, A.: Forecasting of photovoltaic power generation and model optimization: A review. In: Renewable and Sustainable Energy Reviews, vol. 81, pp. 912–928 (2018)
3. Samara, S., Natsheh, E.: Intelligent real-time photovoltaic panel monitoring system using artificial neural networks. In: IEEE Access, vol. 7, pp. 50287–50299 (2019)
4. Wang, F., Zhen, Z., Wang, B., Mi, Z.: Comparative study on KNN and SVM based weather classification models for day-ahead short-term solar PV power forecasting. In: Applied Sciences, vol. 8, p. 28 (2017)
5. Raza, M.Q., Mithulananthan, N., Li, J., Lee, K.Y., Gooi, H.B., Nadarajah, M.: Ensemble framework for day-ahead forecast of PV output power in smart grids. In: IEEE Transactions on Industrial Informatics, vol. 15, pp. 4624–4634 (2018)
6. Berghout, T., Benbouzid, M., Bentrcia, T., Ma, X., Djurović, S., Mouss, L.-H.: Machine learning-based condition monitoring for PV systems: State of the art and future prospects. In: Energies, vol. 14, p. 6316 (2021)

7. Basnet, B., Chun, H., Bang, J.: Intelligent fault detection model for photovoltaic systems. In: Journal of Sensors, vol. 2020, p. 6960328 (2020)
8. Dhimish, M., Holmes, V., Mehrdadi, B., Dales, M., Mather, P.: Photovoltaic fault detection algorithm based on theoretical curves modelling and fuzzy classification system. In: Energy, vol. 140, pp. 276–290 (2017)
9. Garoudja, E., Chouder, A., Kara, K., Silvestre, S.: Enhanced machine learning-based approach for failures detection and diagnosis of PV systems. In: Energy Conversion and Management, vol. 151, pp. 496–513 (2017)
10. Appiah, A.Y., Zhang, X., Ayawli, B.B.K., Kyeremeh, F.: Review and performance evaluation of photovoltaic array fault detection and diagnosis techniques. In: International Journal of Photoenergy, vol. 2019, p. 6953530 (2019)
11. Balaji, A., Sathyasri, B., S, V.V.R., Indumathy, D., Krishnan, R., Vanaja, S.: Intruder Alert System in Smart Home based on IoT Technique. (2022).
12. Harrou, F., Sun, Y., Taghezouit, B., Saidi, A., Hamlati, M.-E.: Reliable fault detection and diagnosis of photovoltaic systems based on statistical monitoring approaches. In: Renewable Energy, vol. 116, pp. 22–37 (2018)
13. Eskandari, A., Milimonfared, J., Aghaei, M.: Line-line fault detection and classification for photovoltaic systems using ensemble learning model based on IV characteristics. In: Solar Energy, vol. 211, pp. 354–365 (2020)
14. Sinthia, P., M, Malathi., T, Sripriya., Krishnan, R., G, Gurumoorthy., Jalaldeen, K.: Monitoring vital parameters of comatose patients using smart sensors integrated with cloud storage. (2024). https://doi.org/10.1109/i-smac61858.2024.10714845.
15. Zhao, Y., Ball, R., Mosesian, J., de Palma, J.F., Lehman, B.: Graph-based semi-supervised learning for fault detection and classification in solar photovoltaic arrays. In: IEEE Transactions on Power Electronics, vol. 30, pp. 2848–2858 (2014)
16. Vanitha, V., Joe, S.B., Krishnan, R., Fletcher, A.S.A., Anju, M., Akila, V.: Cognitive Threats Detection Model using Nature Inspired Chimpanzee Optimization for IoT Networks (CCM-COM). In: Atlantis highlights in engineering/Atlantis Highlights in Engineering. pp. 629–637 (2025). https://doi.org/10.2991/978-94-6463-754-0_55.
17. Yi, Z., Etemadi, A.H.: Line-to-line fault detection for photovoltaic arrays using multiresolution signal decomposition and two-stage support vector machine. In: IEEE Transactions on Industrial Electronics, vol. 64, pp. 8546–8556 (2017).