Solar Panel Fault Detection Using Data Mining Process

Velvizhi.G[1], Sandhiya.A[2], Swetha.T[3], Vinodhini.B[4]
[1, 2, 3, 4] Dept of Computer Science and Engineering and Technology
[1, 2, 3, 4] RAAK College of Engineering and Technology, Pondicherry, Pin-605010,India

Abstract- *For future energy demands, it is critical to progress toward a dependable, cost-effective and sustainable renewable energy source. Solar power generation is expanding globally as a result of growing energy demands and depleting fossil fuel reserves, which are presently the primary sources of power generation in recent days. Data mining is a process of discovering interesting patterns and knowledge from large amount of dataset. The data sources can include data bases, data warehouse, the web and other information repositories are data that steamed into the system dynamically. Data mining process will extract relevant information from raw data such as transactions, flat files, live streaming of data. Data is increasing day by day, when a new data is found, it can change the result. By using data mining process in solar panel repository we can mining the history of the energy produced in the solar panel.so, we can easily identify the faults present in the solar panel. In photovoltaic arrays, we can use data mining process to analyse the solar cells to detect faults and fault locatization using data source present in solar panel data repository.*

Keywords- Data Mining, Photovoltaic array, Fault detection, Fault localization and Fault classification.

I. INTRODUCTION

Data mining process will extract relevant information from raw data source. It also is a buzzword and is frequently applied to any form of large -scale data or information processing (collection, extraction, warehousing, analysis, and statistics) as well as any application of computer decision support systems, including artificial intelligence and business intelligence. In recent years, Python has become more and more used for the development of data centric applications to the support of a large scientific computing community and to the increasing number of libraries available for data analysis.

In particular, we will see how it can be used:
In particular, we will see how it can be used:

- Import and visualize data.
- Classify and cluster data.
- Discover relationships between the data using regression and correlation measures.
- Reduce the dimensionality of the data in order to using regression and correlation measures.

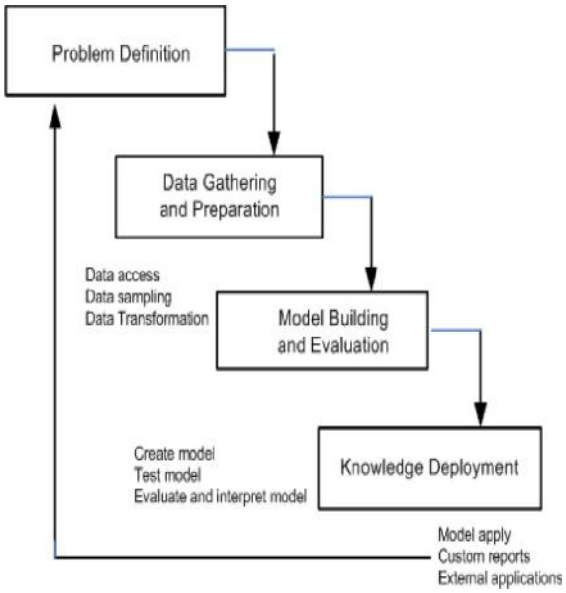

Fig.1 : The Data Mining Process

Data mining

Data mining is the process of analyzing a large batch of information to discern trends and patterns. Data mining can be used by corporations for everything from learning about what customers are interested in or want to buy to fraud detection and spam filtering.Data mining programs break down patterns and connections in data based on what information users request orprovide.In many companies uses the data mining techniques to commodify their users in order to generate high profit.Solar panels are devices which can be used to absorbs the sun's rays and convert them into electrical energy through the photovoltaic effect.

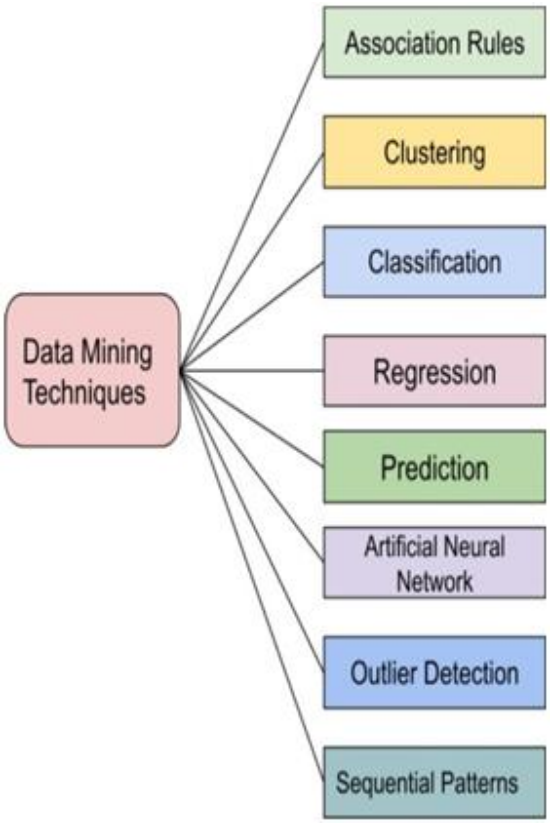

Fig.2 : Techniques of Data Mining

II. PROPOSED SYSTEM

A.Solar Energy

Solar energy is most efficient renewable energy resource .when on day 1367(W/m2), the sun provides energy to the satellite [SMAD].This process uses semiconductors which when excited by the photons, release a free electron that is then lost to flow as current. Manufacturing processes and material choices have improved over the years so that higher grade cheaper terrestrial solar cells made from silicon can reach efficiencies of around 18% and more expensive space grade solar cells made from triple junction gallium arsenide can reach about 32% efficiency in production Models [SMAD].

Some cells still in the R&D phases have been known to go up to 40% efficiency.This section is meant as a step-by-step walkthrough of the design process which took place during the Fall 2011 semester to construct the solar panels. Each decision will be discussed to the best of the author's abilities so that the reasons and facts necessary for each decision are documented and commented on.

B.Design Process in Solar Panel

The project began with a set of constraints. Since the satellite will not have a G&C (Guidance and Control) System, this frequency has to be high enough so that given random rotation, the satellite will still flash sufficiently toward earth to be seen when overhead .These cells generally range from 12%-18% efficient and come in two crystal types, mono crystalline and polycrystalline.

Mono crystalline cells are usually more efficient due to the presence of just a single crystal but are a bit more expensive while poly crystal lines are less expensive but generally less efficient. The efficiency for these solar cells is about 25%-30% for production models, but some R&D cells have been known to reach 40%.

III. FUNCTIONS IN A PROPOSED SYSTEM

Classification & Prediction

There are two methods of data processing that can be used to derive models that represent significant groups or to forecast future patterns in knowledge. The modes are as follows:

- Classification
- Predicting

Categorical type labels are predicted by the classification models; prediction models forecast continuous-valued functions. For example, we can construct a classification model to classify bank loan requests as either secure or dangerous or a forecast model to estimate future customers' expenses in USD on electronic equipment given their incomes and profession.

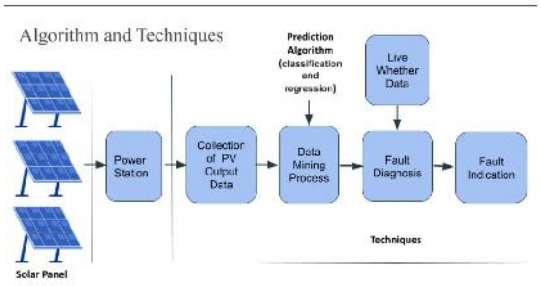

Fig.3: Algorithms and Techniques by using proposed system

Classification

Examples of situations where the role of data processing is classified are below

In order to know if the client is risky or which one is secure, a bank loan officer needs to evaluate the details.

A brand manager at a business has to assess a buyer who would purchase a new device with using a specific profile.

A model or classifier is built to predict the categorical labels in each of the above mentioned cases. For loan applicant data, these labels are dangerous or stable and for marketing data, yes or no is used

Prediction

The following are examples of situations where prediction is the goal of data processing.

Suppose the marketing manager has to estimate how much a single client can pay at his business after the transaction. We are bothered to forecast a numerical value in this case. Therefore an example of numeric prediction is the data processing activity. In this scenario, a system or a predictor that forecasts the continuous-valued-function or ordered value will be created .The study of regression is a mathematical technique most commonly used in numeric prediction.

Working of the Classification

Let us appreciate the functioning of classification with the assistance of the bank loan application that we have mentioned above. There are two stages in the Data Classification system.

- The Classifier or Model construction.
- Using Classification Classifier.

a).The Classifier or model classifier.

- The classification algorithms construct the classifier in this stage.

  The classifier is constructed from a training set composed of tuples of the databases and their corresponding class names.
- Each category that makes up the training set is referred to as a category or class. You may also refer to these tuples as samples, objects, or data points.

b).Using Classification Classifier

The classifier is used for classification in this stage. Test data is used here to approximate the consistency of the rules for classification. If the precision is considered appropriate, classification rules can be extended to new data tuples.

Fault Classification

- The developed fault classification algorithm is very efficient in terms of training, validation, testing, and fault detection time.
- Faults that may be hidden in PV arrays without being noticed can also be detected using the developed algorithm. This makes the system more understandable and reliable.
- The case study verifies the simulation results and proves the developed algorithm. Advantages of Proposed System
- A subfield of control which concerns itself with monitoring a Solar panel system.
  - Identifying when a fault has occurred.
  - Identifying when pollution formed Pin pointing the type of fault and its location.

Advantages should correlated with our project

- While high pollution formed on solar panel our project is easily detect the type of fault and alert through the CCS (Central Control Station)

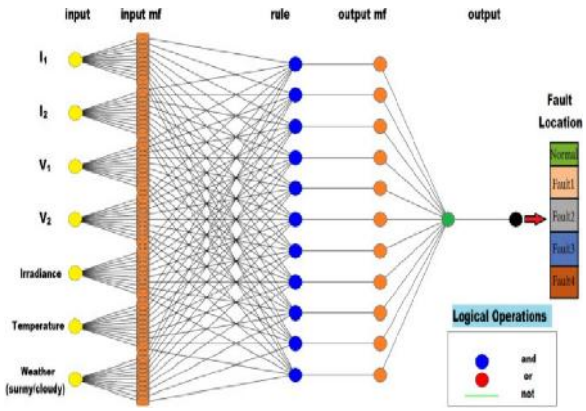

Fig.4 : Fault Classification

- The local climate, seasons and weather conditions expected output value is unpredictable.

In this case our tool is embedded with live weather support to identify the fault by machine learning algorithm.

- Advantages should correlated with our Proposed System:
- A subfield of control which concerns itself with monitoring a Solar panel system.

- Identifying when a fault has occurred.
- Identifying when pollution formed Pin pointing the type of fault and its location.
- While high pollution formed on solar panel our project is easily detect the type of fault and alert through the CCS (Central Control Station)
- The local climate, seasons and weather conditions expected output value is unpredictable. In this case our tool is embedded with live weather support to identify the fault by machine learning algorithm.

IV. CONCLUSION

The aim of the project was to create a system that can detect faults in PV systems with the Use of data mining processes and meteorological data. This was achieved by creating an expected output of a PV system which is compared to the actual output.

The extent of what the system can detect turned out to depend on three parameters: The size of the power decreases, the threshold and the time horizon. The results found in this report shows that given a certain decrease, the threshold and time horizon can be appropriately tweaked in order to yield the wanted results of zero false

Positives, average days the same as the time horizon and good detection percentage (true Positives compared to false negatives).In total, it is possible to detect faults in PV systems by utilizing machine learning and Meteorological data.

V. FUTUREWORK

The data mining process will extract the relevant data from the data sources. It will be used to classify and display the analysis of the data set. In photovoltaic arrays, all the generated power data is stored in the data repository (solar power station). In our project, we can use the software to identify the fault in the solar panel using the data mining process. The software can work on any type of system. It can be used to identify the fault in the solar panel based on the previous data generated report, and it can also classify the type of the fault, like shorting of module wires and diodes, layer module pollution formed, and maintenance required. It also provides live streaming of the weather to identify and determine the faults in PV cells. By using this software, we can reduce manpower and costs. This can be more useful in solar panel power generation to identify the fault easily. In the future, it can be operated by mobile phones, tablets, and other devices to monitor solar panel power generation at any time and anywhere. It can analyze the data and show the power generation results.

REFERENCE

[1] Masoud Alajmi, Osama Aljasem, Niaz Ali, Adnan Alqurashi , Ikhlas Abdel-Qader, “Fault Detection and Localization in Solar Photovoltaic Arrays Framework: Hybrid Methods of Data-Analysis and a Network of Voltage-Current Sensors ”IEEE 2018

[2] Xingke Guo, Zhixiong Na, Dayan Ma, Yudong Lu and Xin Luo,”Fault diagnosis of photovoltaic system based on machine learning model fusion”IEEE 2019

[3] Mohamed M. Bard, Mostafa S. Hamad2, (SENIOR,IEEE), Ayman S. Abdel-Khalik,(SENIOR, IEEE), Ragi A. Hamdy, (SENIOR, IEEE), Shehab Ahmed, (SENIOR, IEEE), and Eman Hamdan,”Fault Identification of Photovoltaic Array Based on Machine Learning Classifiers”IEEE 2021

[4] Shazia Baloch and Mannan Saeed Muhammad ,’’An Intelligent Data Mining-Based Fault Detection and Classification Strategy for Microgrid”IEEE 2021

[5] Vishal R. Shinde, Mr.Vipinkumar Gautam, Mr.Chinmay Gawande, Miss.Pooja Mahale , “Solar Data Fault Detection using Machine Learning” IJREAM 2019

[6] M. R. Alam, M. T. A. Begum, and K. M. Muttaqi, ‘‘Assessing the performance of ROCOF relay for anti-islanding protection of distributed generation under subcritical regions of power imbalance,’’ IEEE Trans. Ind. Appl., vol. 55, no. 5, pp. 5395–5405, Sep. 2019.

[7] A. Mohamed, B. Younes, T. Lamhamdi, H. El Moussaoui, and H. El Markhi, ‘‘Fault location and isolation technique in smart distribution systems with distributed generation,’’ in Proc. 1st Int. Conf. Innov. Res. Appl. Sci., Eng. Technol. (IRASET), Apr. 2020

[8] J. J. Q. Yu, Y. Hou, A. Y. S. Lam, and V. O. K. Li, ‘‘Intelligent fault detection scheme for microgrids with wavelet-base deep neural networks,’’ IEEE Trans. Smart Grid, vol. 10, no. 2, pp. 1694–1703, Mar. 2019

[9] S. Jamali, S. Ranjbar, and A. Bahmanyar, ‘‘Identification of faulted line sections in microgrids using data mining method based on feature discretisation,’’ Int. Trans. Electr. Energy Syst., vol. 30, no. 6, Jun. 2020, Art. no. e12353.

[10] H. M. Sharaf, H. H. Zeineldin, and E. El-Saadany, ‘‘Protection coordination for microgrids with grid-connected and islanded capabilities using communication assisted dual setting directional overcurrent relays,’’ IEEE Trans. Smart Grid, vol. 9, no. 1, pp. 143–151, Jan. 2018.

[11] Radu Platon, Jacques Martel, Norris Woodruff, and Tak Y. Chau - “Online Fault Detection in PV Systems” IEEE 2017.

[12] Dipak Saha, Soham Adhya, “An IoT Based Smart Solar Photovoltaic Remote Monitoring and Control unit” CIEC 2016.
[13] Mohamed Fezari, Ali Al-Dahoud, “Remote Monitoring System for Solar Power Panels using Intelligent Sensors Network” IEEE Nov. 2016.
[14] Khurum Nazir Junejo, Faizan Jawaid, “Predicting Daily Mean Solar Power Using Machine Learning Regression Techniques” IEEE Aug. 2016.
[15] Andreas S. Spanias, “Solar Energy Management as an Internet of Things (IoT) Application” IEEE 2016.