ISSN: 2583-9292 (ONLINE)
DOI: 10.xx/ijdsml.2025.xx

A LOGISTIC REGRESSION MODEL FOR FAULT DETECTION IN SOLAR-POWERED WATER PUMPING SOLUTIONS

Yogesh Kakasaheb Shejwal, Jayraja U. Kidav, Lakshaman Korra and Manjiri Lavadkar
Department of Power Electronics, National Institute of Electronics and Information Technology, India

Abstract
The adoption of solar-powered water pumps is crucial for sustainable water management, particularly in remote and agricultural areas. However, the reliability of these systems is essential, as faults can result in significant downtime and expensive repairs. This paper introduces an intelligent fault detection model using logistic regression, a machine learning technique suitable for binary classification of faults. It leverages real-time sensor data, including electrical and mechanical parameters, alongside environmental conditions, to predict potential faults. The model was trained on a dataset from operational installations, undergoing pre processing for feature enhancement and missing data management. Evaluating the model involved key performance metrics such as accuracy, precision, recall, and the Area under the ROC curve, which confirmed its effectiveness in accurately detecting faults with high precision. A comparative analysis with decision trees and support vector machines emphasized logistic regression's efficiency. The research contributes to intelligent monitoring systems and identifies future directions for enhancing fault detection methods.

Keywords:
Solar Powered, Solar Water Pump, logistic Regression, Machine Learning, Intelligent Monitoring, Maintenance, Power Electronics

1. INTRODUCTION

Solar water pumps are transforming agriculture and rural development by harnessing solar energy to provide water, especially in remote areas lacking reliable power sources. These pumps use photovoltaic (PV) panels to convert sunlight into electricity, making them a cost-effective alternative to diesel engines. Particularly beneficial in arid climates, solar pumps support irrigation systems that enhance crop yields and food security, reducing dependency on rain fed agriculture. They significantly lower operating costs, as they eliminate fuel expenses and require less maintenance than traditional pumps. As a result, farmers can diversify their crops and increase income throughout the year [1] [2].

Beyond agriculture, solar water pumps are vital for rural development, as they improve access to clean drinking water and support health initiatives. In places where the power grid is impractical, these off-grid solutions negate the need for expensive diesel transportation. Socially, they alleviate the burden of water collection, primarily affecting women and children, and allow them to engage in education and other economic activities. Additionally, the installation of solar pumps fosters job creation in local communities [3] [4] [5].

In remote areas, these systems are crucial for maintaining water security, providing irrigation, and supporting livestock. Their reliability is particularly beneficial in the face of climate change, ensuring a stable water supply during dry spells. Furthermore, by replacing diesel pumps, solar systems reduce greenhouse gas emissions and contribute to sustainable development efforts. Overall, solar water pumps enhance agricultural productivity, improve rural livelihoods, and promote environmentally friendly practices [6] [7].

Understanding common faults in solar water pump systems is essential for optimizing their performance and longevity. Issues such as motor failure, lack of rotation, and poor wiring can significantly hinder efficiency, reducing water production or causing system failure. Preventive maintenance, proper installation, and regular monitoring can mitigate these issues; for instance, keeping solar panels clean boosts efficiency. Advances in technology and technician training enhance diagnostics and repairs, crucial for maximizing the benefits of solar water pumping, particularly in agriculture and remote areas. Proper maintenance is vital as electrical technologies evolve [8] [9].

1.1 NEED

Fault analysis in solar water pump systems plays a crucial role in enhancing their efficiency, reliability, and stability, especially in agricultural settings where water access is vital. Firstly, it helps in preventing errors by swiftly identifying potential issues before they escalate, thereby minimizing downtime and ensuring continuous operation. Secondly, ongoing system parameter monitoring allows for performance optimization through intelligent diagnostics, which can make real-time adjustments to maximize efficiency and mitigate rising energy costs. Thirdly, early detection of issues leads to reduced maintenance costs, as predictive maintenance enables operators to address minor problems before they evolve into costly repairs, thus extending equipment life. Additionally, maintaining these systems is critical in remote areas with limited alternative water sources, bolstering user confidence and fostering consistent usage. Lastly, as the shift toward renewable energy intensifies, effective fault detection underpins the sustainability of solar technologies by ensuring optimal resource management. In conclusion, investing in fault detection technology is essential for the efficient management of solar water pump systems [10] [11]

Logistic regression is an effective method for detecting faults in solar water pump systems due to its various advantageous properties. Firstly, its binary classification capability allows for the clear categorization of system statuses as either “error” or “no error.” This binary aspect is crucial for sensitive applications; ensuring prompt action can be taken. Secondly, logistic regression enhances interpretability by generating probabilities for each classification, allowing engineers to understand potential failure causes based on factors like sensor readings and operational parameters. Additionally, it performs well with smaller datasets, making it applicable even in remote areas where historical data may be limited. The model's simplicity reduces the risk associated with over fitting that more complex models may face. Lastly, logistic regression can easily adapt to new data, ensuring that the model remains accurate as system conditions evolve. Overall, these features make logistic regression a highly suitable choice for creating a reliable fault diagnosis system for solar water pumping applications [12] [13]

1.2 OBJECTIVE

This paper aims to develop a robust logistic regression model for diagnosing faults in solar-powered water pumps, focusing on several key objectives. First, it seeks to create an accurate model capable of categorizing pump controls as “bad” or “non-bad” by selecting relevant features and optimizing data processing. Second, the study intends to implement a real-time monitoring framework that integrates the logistic regression model to evaluate pump operational status continuously, facilitating early fault detection. Third, it aims to enable predictive maintenance through model insights, assisting managers in planning corrective actions to minimize downtime and repair costs. Additionally, the research will analyze operational data to improve the efficiency and reliability of solar water pumps and evaluate the model’s performance against accuracy, precision, recall, and F1 score metrics. Ultimately, this work aspires to enhance renewable energy management and operational effectiveness in electric water pumping systems.

2. LITERATURE REVIEW

2.1 EXISTING FAULT DETECTION METHODS

Table.1. Existing Fault Detection Methods

Method	Overview	Advantages	Limitations
Rule-Based Systems [14]	Utilize predefined rules based on expert knowledge and historical data to identify faults in system performance.	Simple to implement, Cost-effective, Easy to understand	Inflexible to new faults, Maintenance can be cumbersome, May lead to false positives
Manual Inspections [15]	Involve physical checks of components, allowing technicians to identify visible signs of wear or malfunction.	Thorough and can identify issues not quantifiable, Expert intuition can be valuable	Labor-intensive, Inconsistency in inspections, Potential for oversight
Visual Monitoring [16]	Technicians visually inspect the pump and solar panels for signs of damage, wear, or environmental effects.	Quick assessment of the overall condition, Immediate detection of obvious physical issues	Limited effectiveness for hidden faults, Subject to human error
Performance Testing [17]	Conduct periodic tests to evaluate the efficiency and output of the system, comparing it against expected performance metrics.	Provides quantitative data on system efficiency, Helps identify systemic issues	Requires calibrated equipment, May miss intermittent faults
Historical Data Analysis [18]	Analyze historical operational data and fault logs to identify patterns and trends associated with system failures.	Can reveal systemic issues over time, Useful for long-term planning and maintenance strategies	Relies on the availability of comprehensive data, May not account for sudden changes

Traditional fault detection methods in solar water pumps, including rule-based systems and manual inspections, are essential for maintaining system efficiency and reliability. While rule-based systems are structured, they can miss new fault types, and manual inspections, though thorough, are labor-intensive and inconsistent. Additional methods like visual monitoring and performance testing also have limitations, such as human error. Analyzing historical data can aid predictive maintenance but relies on comprehensive datasets. To enhance error detection, implementing machine learning techniques can improve efficiency and reduce costs, thereby supporting sustainable practices in renewable energy management for agriculture and rural development

MODERN APPROACHES

Modern fault detection approaches for solar water pumps, utilizing machine learning (ML) and artificial intelligence (AI), significantly improve upon traditional methods. These techniques provide robust predictive analytics that can foresee failures, allowing for proactive maintenance and reducing downtime. Real-time monitoring enhances reliability through immediate fault detection, critical in remote areas. Data fusion techniques integrate various datasets, improving fault identification accuracy. While these advanced systems empower operators in resource allocation and decision-making, challenges such as the need for high-quality data and potential false positives remain. Transitioning to these intelligent methods is vital for the reliability of renewable energy sources

Table.2. Modern approaches like machine learning and AI in fault detection

Method	Overview	Advantages	Limitations
Rule-Based Systems [14]	Utilize predefined rules based on expert knowledge and historical data to identify faults in system performance.	Simple to implement, Cost-effective, Easy to understand	Inflexible to new faults, Maintenance can be cumbersome, May lead to false positives
Manual Inspections [15]	Involve physical checks of components,	Thorough and can identify issues not	Labor-intensive, Inconsistency in inspections,

	allowing technicians to identify visible signs of wear or malfunction.	quantifiable, Expert intuition can be valuable	Potential for oversight
Visual Monitoring [16]	Technicians visually inspect the pump and solar panels for signs of damage, wear, or environmental effects.	Quick assessment of the overall condition, Immediate detection of obvious physical issues	Limited effectiveness for hidden faults, Subject to human error
Performance Testing [17]	Conduct periodic tests to evaluate the efficiency and output of the system, comparing it against expected performance metrics.	Provides quantitative data on system efficiency, Helps identify systemic issues	Requires calibrated equipment, May miss intermittent faults
Historical Data Analysis [18]	Analyze historical operational data and fault logs to identify patterns and trends associated with system failures.	Can reveal systemic issues over time, Useful for long-term planning and maintenance strategies	Relies on the availability of comprehensive data, May not account for sudden changes

2.2 LOGISTIC REGRESSION IN FAULT DETECTION FOR OTHER SYSTEMS

Logistic regression is a powerful statistical tool used for fault detection in various sectors such as power systems, mechanical engineering, industrial applications, aerospace, and healthcare. In power systems, it helps analyze parameters like temperature and current to predict issues such as overloads in transformers and circuits. In mechanical engineering, it is used to analyze vibration signals and predict equipment failures in machines like motors and pumps. In the industrial sector, logistic regression is employed to evaluate equipment health and detect operational irregularities for prompt maintenance. In aviation and automotive industries, it aids in predictive maintenance by analyzing sensor data to anticipate failures in critical components, enhancing safety and reliability. In healthcare, logistic regression plays a vital role in monitoring medical devices and diagnosing malfunctions, such as ventilators, through operational alarm analysis. The versatility and effectiveness of logistic regression in fault analysis underscore its importance in enhancing operational reliability and efficiency in different sectors [25].

3. METHODOLOGY

3.1 OVERVIEW OF POSSIBLE SYSTEM FAULTS

The classification of faults that occur in a solar power generation system can be divided into different categories. The main factors related to energy production can be divided into three categories: energy, environment and physical. Electrical faults are of three types, the first is common, the second is open circuit, and the last is linear. An open circuit fault is caused by a disconnection of the inverter grid. Line faults are breaks in cables that cause low power. Physical failures can come from inside or outside, and usually include damage, cracks, and destruction of parts of the solar cell system. Systemic defects are caused by the effects of aging and are also physical conditions. Environmental errors are those that affect the performance of the solar system, such as wind speed.

3.2 PROPOSED SYSTEM ARCHITECTURE

The proposed system uses Machine Learning Algorithm to predict four faults based on data collected from solar power generation equipment, in which two are electrical faults i.e. open circuit and line-line faults. Next prediction is about circuit maintenance and cleaning which can categorize as physical faults. Last is environmental fault which detects weather there is need of structural maintenance based on wind velocity. Fault prediction in solar power generation equipment has been done as shown in figure 1. To predict Electrical Faults values of temperature, current, voltage, power are used. To predict environmental faults, wind speed is used as input, and to detect current, voltage, power and solar physical faults are used as inputs

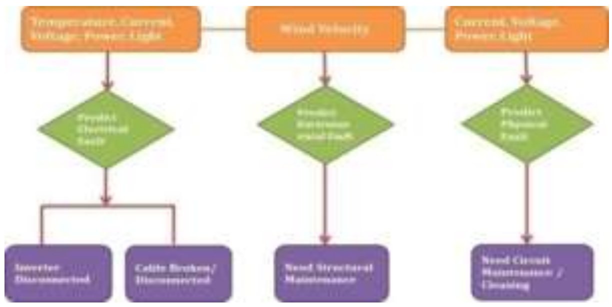

Fig.1. Types of Faults System will predict

3.3 WORKING OF PROPOSED SYSTEM

As shown in Fig 1, the conceptual system identifies three types of failure, namely electrical, environmental, and physical.

1. If there is an environmental fault, check the converter to see if it has overheated. If the inverter gets too hot, allow it to cool. If the converter does not overheat, check the surrounding area for heat sources. If there is no heat source, the converter is replaced.

2. If there is a physical fault, check the inverter for signs of physical damage. If the body is damaged, the transformer is replaced. If there is no physical damage, check the alternator for loose connections. If there are any connections, tighten the connections. If the connections do not loosen, the inverter must be replaced.
3. In the proposed system, some process steps can be avoided. For example, if the inverter is disconnected, there is no need to check the line. And if the inverter gets too hot, there is no need to check the surrounding heat sources.

- **Data collection**: This is the first step in system architecture. To build the model, we obtained data on current, voltage, power, sun level, wind speed and temperature from sensors connected to the solar system. The data set includes data collected in summer and winter under all environmental conditions.
- **Data Processing:** Data pre-processing is the second step in the proposed system architecture. It includes all the steps taken before entering the data into the model to extract the features. To create a fault detection model, seven data points are selected as input points and three target points below.
- **Input Features**
  1. Temperature: Temperature in Celsius.
  2. Current: Current generated for Given Temperature.
  3. Voltage: Voltage generated for Given Temperature.
  4. Power: Power generated by solar panel.
  5. Light: Measurement of light.
  6. Weather: It has two values Sunny and Cloudy.
- Output Features
  1. **Target 1**: Predicts working State of Solar system. Open-Inverter network disconnected, Line- Line -Cable is Broken, Normal State.
  2. **Target 2**: Predicts need of Structural Maintenance based on wind speed.
  3. **Target 3**: Predicts need of Circuit Maintenance of solar panel

In both input and target conditions, five continuous variables and four random variables are selected. Therefore, variable variables are coded into direct numerical data. Weather conditions are coded as “Sunny” as 1 and “Cloudy” as 2. The variable Target1 has three values “True” with code 0, “Open” with code 1, “Line-Line” with code 2. Next Target2 variable. There are two values “must be protected” with code 0 and “normal” with code 1. The last variable Target3 has two more values o target for the prediction, the first is “Normal” which is set to 0 and the second is “Circular Correction” which is set to 1

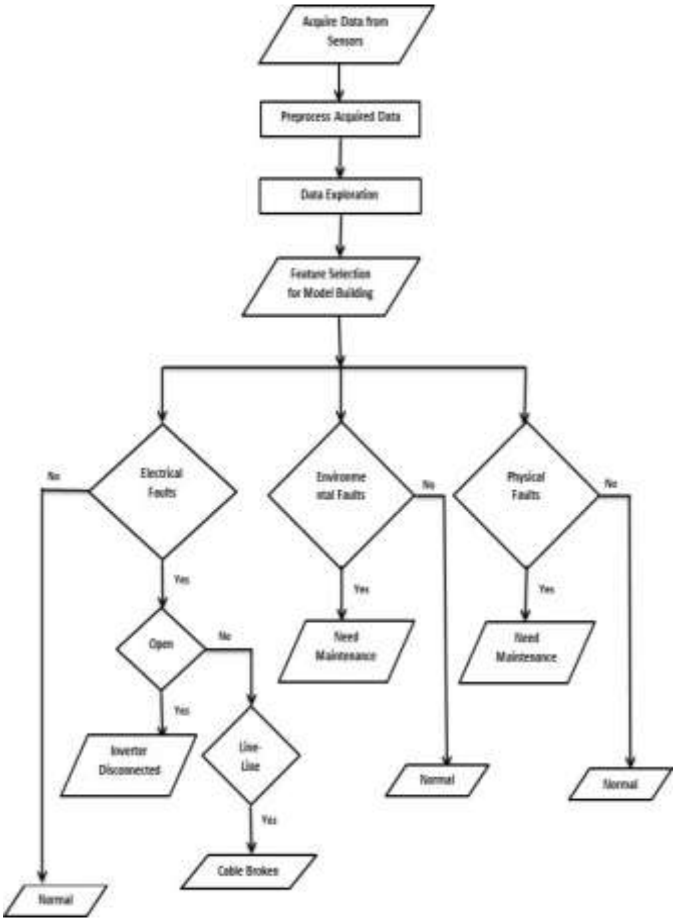

Fig.2.Working Flow of System

- **Feature Extraction and Feature Correlation**. As shown in Fig.1, dataset has three Target features Target1, Target2, and Target3. Target1 is to predict electrical fault. Input features taken to predict this fault are Temperature, Current, Voltage, Power, Light, and Weather. Among all these input variables Current, Voltage, Power, and Weather shows strong correlation with target variable.

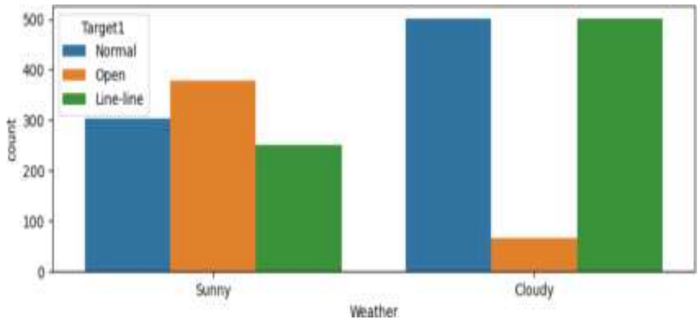

Fig.3. Correlation of Weather with Target 1

Target 2 is to predict Environmental fault. Input features taken to predict this fault is wind velocity [m/s]. If Wind velocity is high, then the possibility of structural fault will increased, therefore proposed system indicates need of structural maintenance when wind velocity is more than 35 m/s.

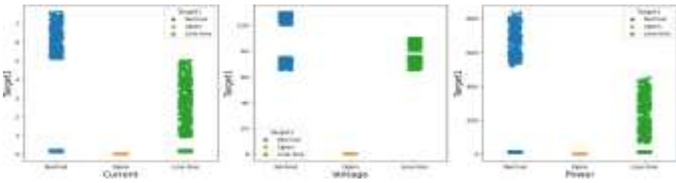

Fig.4.Correlation of Current, Voltage, Power with Target1

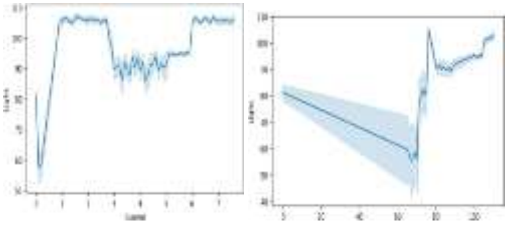

Fig.5 Correlation of Current and Voltage with Target 3

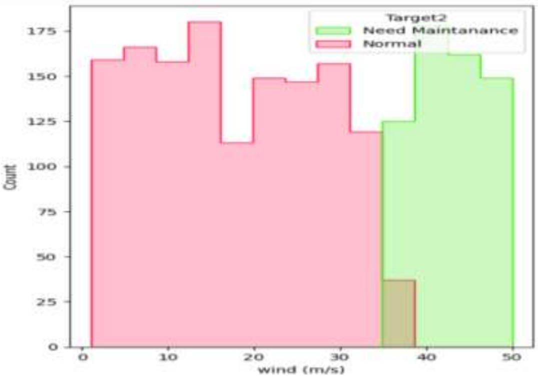

Fig.6.Correlation of wind velocity with Target2

Target3 is to predict Physical fault. Input features taken to predict this fault are related to Light and power generation such as current, voltage, power. Sometimes even if sunlight is good but system generates low current and voltage due to circuit problem, which may cause low power generation. Proposed system will help in predicting such problems.

4. EXPERIMENTAL SETUP

4.1 MODEL SELECTION AND MODEL BUILDING

Logistic regression is a type of supervised machine learning technique used to classify target variables. In logistic regression, instead of fitting a regression line, we fit an S-shaped logistic function. It uses a complex cost function called the sigmoid function. The goal of discovery is the probability (the amount of data observations that can tell us the outcome or the value of variables set for specific data points) of an outcome. The logistic regression model transforms the constant-valued output of the linear regression function into a discrete-valued output using the sigmoid function, which maps each set of real-valued input independent variables to a value between 0 and 1. This function is known as function logistic. Let the independent input features are:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ x_{21} & \cdots & x_{2m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix} \tag{1}$$

And the dependent variable is Y having only binary value i.e.0 or 1.

$$Y = \begin{cases} 0 & if\ Class\,1 \\ 1 & if\ Class\,2 \end{cases} \tag{2}$$

Then, apply the multi-linear function to the input variables *X*.

$$z = \left(\sum_{i=1}^{n} \left(w_i x_i \right) + b \right) \tag{3}$$

Sigmoid is mathematical function which used to convert continuous values to probability values which lies within range 0 to 1. Sigmoid function follows below equation:

$$Sigmoid\left(z\right) = \frac{1}{1+e^{-z}} \tag{4}$$

where, e = (Euler's number) ~ 2.71828,

To make the logistic regression a linear classifier, we could choose a certain threshold, e.g. 0.5. Now, the misclassification rate can be minimized if we predict y=1 when p ≥ 0.5 and y=0 when p<0.5. Here, 1 and 0 are the classes.

The predicted probabilities will be:

- For y=1: The predicted probabilities will be: $p(X; b, w) = p(x)$.
- For y=0: The predicted probabilities will be: $1-p(X; b, w) = 1-p(x)$.

Since Logistic regression predicts probabilities, we can fit it using likelihood. Therefore, for each training data point x, the predicted classis y. Probability of y is either p if y=1or 1-p if y=0. Now, the likelihood can be written as: On the basis of the categories, Logistic Regression can be classified into three types:

- *Binomial:* In binomial Logistic regression, there can be only two possible types of the dependent variables, such as 0 or 1, Pass or Fail, etc.
- *Multinomial:* In multinomial Logistic regression, there can be 3 or more possible unordered types of the dependent variable, such as "cat", "dogs", or "sheep"
- *Ordinal:* In ordinal Logistic regression, there can be 3 or more possible ordered types of dependent variables, such as "low", "Medium", or "High".

The target object target1 has three classes to predict, making it a multiclass classification problem. Target2 and Target3 are binary classification problems that are accurately predicted using a regression algorithm. To address the issue of unbalanced samples in all three columns, SMOTE is used to balance the target columns by creating artificial models for the minority class. This algorithm helps prevent over fitting by randomly selecting good class samples and generating new synthetic samples based on the KNN algorithm. Parameters such as solver, penalty, and C are adjusted to optimize the model and prevent over fitting. Cross-validation grid search is then used to find the best combination of parameters to improve model performance. Overall, SMOTE proves to be an effective method for dealing with unbalanced datasets and improving the accuracy of predictions in classification problems.

4.2 EVALUATION

The reasoning system categorizes errors into different categories, with confusion matrices being more predictive of model performance than classification accuracy. The uncertainty matrix summarizes correct and incorrect predictions made by the model, used to evaluate model performance. Metrics like accuracy, precision, recall, and F1 score help analyze the model's performance. Accuracy measures the ratio of correct predictions to total predictions, while recall measures true observations correctly predicted. The F1 score, a measure of precision and recall, ranges from 0 to 1 and works well for unbalanced data models. The regression algorithm achieved 100% accuracy and F1 score for Target1, Target2, and Target3, indicating balanced performance across dimensions.

4.3 RESULT AND DISCUSSION

The solar fault prediction system uses sensor data from solar devices to determine electrical, physical and environmental faults. The utility of logistic regression comes from its ability to model linear relationships between factors such as current, voltage, power output and environmental factors such as wind speed. This makes it a good choice for work.

Fig.7(a).Webpage for taking readings

Fig.7(b).Webpage for showing prediction

Fig.7(a) and (b) shows screenshots of the web pages implementing the proposed model. The proposed model is deployed on flask web application

4.4 DATA PROCESSING AND FEATURE SELECTION

After data collection, processing operations were used to clean the dataset. The cleaned data set was divided into training and testing sets. Feature selection was performed to identify the most relevant variables for fault prediction, ensuring that only significant features like current, voltage, power output, and wind velocity were retained.

Given that electrical faults are assumed to be linearly related to current and voltage, Logistic Regression performed exceptionally well in modeling such faults. Additionally, physical and environmental faults were successfully predicted as parameters like power output and wind velocity enhanced the model's accuracy.

4.5 MODEL TRAINING AND FINE-TUNING

During the training phase, the preprocessed data was used to train the Logistic Regression model, enabling it to learn the relationships between input features and fault categories. Cross-validation was employed to fine-tune the model and prevent over fitting, ensuring that it could generalize well to new data.

The performance of the model in different data sets showed high accuracy in predicting different types of errors. For example, the accuracy of detecting electrical faults reached 95.2% and 96.2% in two independent datasets, confirming that the linear relationship between current faults is well documented. , voltage and current. Similarly, for physical defects such as modulus degradation, the model achieved an accuracy of 96.2%, demonstrating its ability to handle error prediction tasks.

4.6 MODEL PERFORMANCE ON UNSEEN DATA

The trained model was tested on unseen data to evaluate its real-world performance. The logistic regression model maintained its predictive power with an accuracy of 95.9% in predicting environmental errors, where wind speed is an important factor. This high level of accuracy across all error parameters demonstrates the effectiveness of logic regression in handling errors with linear dependence, especially electrical errors.

5. CONCLUSION

The results show that logistic regression is useful for fault prediction in solar power systems, especially when the faults exhibit linear relationships and integrated characteristics. The high accuracy achieved in the electrical, physical and environmental fault categories shows that this method can serve as a reliable solution for the predictive maintenance of solar energy equipment.

The study demonstrates that Logistic Regression is a highly effective model for fault prediction in solar power systems, particularly when dealing with faults that exhibit linear relationships with input features such as current, voltage, power output, and wind velocity. The model achieved high accuracy rates—95.2% and 96.2% for electrical faults, 96.2% for physical faults like module degradation, and 95.9% for environmental faults—validating its robustness across different fault categories. The successful application of cross-validation further enhanced the model's generalization, preventing over fitting and ensuring reliability in real-world fault detection scenarios.

However, the linear nature of Logistic Regression may limit its ability to detect non-linear fault patterns. Therefore, future research could explore hybrid approaches or non-linear models to improve the detection of more complex types of faults and increase the overall effectiveness of the system to predicting the range of errors. However, the findings of this study provide a solid foundation for the implementation of machine learning-based fault prediction systems in solar power technologies.

REFERENCES

[1] L. Gevorkov, J.L. Domínguez-García and L.T. Romero, "Review on Solar Photovoltaic-Powered Pumping Systems", *Energies*, Vol. 16, No. 1, pp. 1-6, 2023.

[2] W. Dankoff, R. Foster, A. Cota and E. Lespin, "Advances in Solar Powered Water Pumping: Providing for Energy Resiliency and Social Equity", *Proceedings of International Conference on American Solar Energy Society*, pp. 24-35, 2022.

[3] S. Singh, M. Alharthi, A. Anand, A. Shukla, A. Sharma and H. Panchal, "Performance Evaluation and Financial Viability Analysis of Grid Associated 10 MWP Solar Photovoltaic Power plant at UP India", *Scientific Reports*, Vol. 12, No. 1, pp. 1-6, 2022.

[4] M. Aliyu, G. Hassan, S.A. Said, M.U. Siddiqui, A.T. Alawami and I.M. Elamin, "A Review of Solar-Powered Water Pumping Systems", *Renewable and Sustainable Energy Reviews*, Vol. 87, pp. 61-76, 2018.

[5] "Solar Water Pumping for Sustainable Water Supply", *World Bank*, Available at https://www.worldbank.org/en/topic/water/brief/solar-pumping, Accessed in 2017.

[6] xxxx

[7] U. Khalil, "Comparative Analysis of Photovoltaic Faults and Performance Evaluation of its Detection Techniques", *IEEE Access*, Vol. 8, pp. 26676-26700, 2020.

[8] T.D. Short and M.A. Mueller, "Solar Powered Water Pumps: Problems, Pitfalls and Potential", *Proceedings of International Conference on Power Electronics, Machines and Drives*, pp. 1-7, 2022.

[9] V. Gunasekaran and S. Chakraborty, "Technical and Environmental Aspects of Solar Photo-Voltaic Water Pumping Systems: A Comprehensive Survey", *Water Science and Technology: Water Supply*, Vol. 23, No. 7, pp. 2676-2710, 2023.

[10] A. Dutt, M.N. Sandhya Rani, Manbir Singh Bisht, Manisha Chandna and A. Singla, "Machine Learning Approaches for Fault Detection in Renewable Microgrids", *E3S Web of Conferences*, Vol. 511, pp. 1-14, 2024.

[11] Zaid Allal, H.N. Noura, O. Salman and K. Chahine, "Machine Learning Solutions for Renewable Energy Systems: Applications, Challenges, Limitations and Future Directions", *Journal of Environmental Management*, Vol. 354, pp. 1-14, 2024.

[12] Z.B. Duranay, "Fault Detection in Solar Energy Systems: A Deep Learning Approach", *Electronics*, Vol. 12, No. 21, pp. 1-7, 2023.

[13] Edgar Hernando Sepúlveda-Oviedo, L. Travé-Massuyès, A. Subias, M. Pavlov and C. Alonso, "Fault Diagnosis of Photovoltaic Systems using Artificial Intelligence: A Bibliometric Approach", *Heliyon*, Vol. 9, No. 11, pp. 1-8, 2023.

[14] N. Dutta, Palanisamy Kaliannan and Shanmugam Paramasivam, "A Comprehensive Review on Fault Detection and Analysis in the Pumping System", *International Journal of Ambient Energy*, Vol. 43, No. 1, pp. 6878-6898, 2022.

[15] Ioannis Polymeropoulos, Stavros Bezyrgiannidis, E. Vrochidou and G.A. Papakostas, "Enhancing Solar Plant Efficiency: A Review of Vision-based Monitoring and Fault Detection Techniques", *Technologies*, Vol. 12, No. 10, pp. 1-15, 2024.

[16] Y.K. Shejwal, J.U. Kidav and L. Korra, "Maintenance and Performance of Solar Water Pump in a Rural Agricultural Community: A Case Study", *Lecture Notes in Networks and Systems*, pp. 1-7, 2024.

[17] K.O. Yoro, M.O. Daramola, P.T. Sekoai, U.N. Wilson and O. Eterigho-Ikelegbe, "Update on Current Approaches, Challenges and Prospects of Modeling and Simulation in Renewable and Sustainable Energy Systems", *Renewable and Sustainable Energy Reviews*, Vol. 150, pp. 1-6, 2021.

[18] Younes Ledmaoui, Adila El Maghraoui, Mohamed El Aroussi, Rachid Saadane, A. Chebak and Abdellah Chehri, "Forecasting Solar Energy Production: A Comparative Study of Machine Learning Algorithms", *Energy Reports*, Vol. 10, pp. 1004-1012, 2023.

[19] R. Abdalla, H. Samara, N. Perozo, Carlos Paz Carvajal and P. Jaeger, "Machine Learning Approach for Predictive Maintenance of the Electrical Submersible Pumps (ESPs)", *ACS Omega*, Vol. 7, No. 21, pp. 17641-17651, 2022.

[20] G. Bode, S. Thul, M. Baranski and D. Muller, "Real-World Application of Machine-Learning-based Fault Detection Trained with Experimental Data", *Energy*, Vol. 198, pp. 1-8, 2020.

[21] S. Venkatasubramanian, "Fault Diagnosis using Data Fusion with Ensemble Deep Learning Technique in IIoT", *Mathematical Problems in Engineering*, Vol. 2022, pp. 1-8, 2022.

[22] Mohsen Soori, F. Karimi, Roza Dastres and Behrooz Arezoo, "AI-based Decision Support Systems in Industry 4.0, A Review", *Journal of Economy and Technology*, pp. 1-6, 2024.

[23] B. Gong, A. An, Y. Shi and W. Jia, "Fault Diagnosis of Photovoltaic Array with Multi-Module Fusion under Hyperparameter Optimization", *Energy Conversion and Management*, Vol. 319, pp. 1-9, 2024.

[24] Suzan MohammadAli Katamoura and Mehmet Sabih Aksoy, "Anomaly Detection in Renewable Energy Big Data using Deep Learning", *International Journal of Intelligent Information Technologies*, Vol. 19, No. 1, pp. 1-28, 2023.