IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 11, April 2025 Impact Factor: 7.67

Fault Prediction and Awareness for Power Distribution in Grid-Connected Renewable Energy Systems Using Hybrid Machine Learning

Mayuri Padgelwar and Prajakta Kasulkar

Department of Electrical and Power System

Ballarpur Institute of Technology, Chandrapur, India

Abstract: *This study addresses defects in electrical power systems, focusing on short circuits that can disrupt normal operation. The method emphasis is on hybrid microgrid systems connected to the grid, crucial for efficient power management. The novel method proposes an adaptive electricity management technique for fault scenarios in grid-tied conversions, minimizing sensor requirements. This paper presents a hybrid machine learning framework for fault prediction and situational awareness in power distribution systems integrated with Renewable Energy Systems (RES). This research proposes a hybrid machine learning approach that combines the strengths of multiple algorithms to predict faults and enhance awareness in grid-connected RES environments. By leveraging historical and real-time data, the model aims to identify fault patterns, classify fault types, and trigger timely alerts*

Keywords: Hybrid Machine Learning, Matlab Blocks ,Distribution Line , Renewable Sourses , Fault Prediction

I. INTRODUCTION

Day by day the conventional energy sourses are lack due to the environmental crises and they are not fulfil the demand of human beings we used the demand for human being the renewable sourses are used .The Renewable sourses are clean and pure they not harmful the environment as well as human boby .The global shift toward renewable energy has increased the complexity of power distribution networks. Fault prediction and situational awareness have become critical in ensuring grid reliability and stability. Traditional fault detection methods struggle with the variability and nonlinearity introduced by RES. This study introduces a hybrid ML-based approach to improve predictive accuracy and operational awareness .The increasing demand for renewable energy has also led to an increase in the use of solar and wind energy. Solar and wind energy production is a complex process whose performance depends on many factors such as precipitation, solar radiation, temperature, humidity, wind, and lightning. Accurately measuring solar and wind energy is crucial for energy companies to balance supply and demand, reduce costs, and increase energy efficiency. Machine learning-based approaches have shown great results in directly estimating solar and wind energy production. However, achieving a high level of detail similar to the lower 99th percentile requires sample selection, training, evaluation, and guidance.

II. MATERIALANDMETHOD

System Architecture

The solar and wind power prediction system will consist of several components working together to collect, process, and analyze data to make accurate predictions. The following is a high-level overview of the system architecture:

Data Collection: The first part of the system will be responsible for collecting data from various sources, such as weather forecasts, satellite imagery, and historical solar production data. This data will be used to train and validate forecast models.

Pre-processing: Once the data is collected, it needs to be cleaned, curated, and converted into a format suitable for analysis. These preliminary steps will include removing missing data, normalizing values, and converting the data into a standard format.

Feature Engineering: The next step is to extract relevant features from the preliminary data. This will involve identifying patterns in the data, removing outliers, and identifying relationships between multiple variables

Machine Learning Model: At the heart of the solar forecasting system is a machine learning model that can learn to predict solar production based on previously processed and feature-engineered data. The model will be trained using historical data and will be continuously updated as new data becomes available.

Model Evaluation and Selection: After the model is trained, it should be evaluated to check its accuracy and performance. Various measurements such as standard error, root mean square error and correlation coefficient will be used in the evaluation of the model. According to the evaluation results, it will be decided to use the most effective model.

System Implementation

Using solar and wind prediction will involve many components working together to collect, process, and analyze data to make an accurate prediction., Scikit-Learn Python machine learning library. A simple linear regression model was used as a meta-study and examined on four-fold cross-validated predictions of the base model with the original input specifications. The stacked regressor uses the cross_val_predict function, which returns for each example in the training data the prediction that the example received when it was in the validation set. The predictions between the different models are used as input for the meta-learner. This method will reduce the risk of overfitting. Linear regression: Linear regression is one of the simplest and most popular machine learning algorithms. It is a method used for predictive modeling purposes. Linear regression makes predictions about continuous, real or numerical variables. The linear regression algorithm describes the relationship between a variable (y) and one or more independent variables (y), hence the name linear regression. Since linear regression shows a linear relationship, it can be seen that the value of the variable varies according to the value of the individual variable. The linear regression model provides a straight line to represent the relationship between variables.

Decision tree: In a decision tree, the algorithm starts from the root of the tree to predict the class of a given dataset. This algorithm compares the value of the root attribute with the value of the data (real data) attribute and jumps to the next node along the branch based on the comparison result. For the next node, the algorithm compares the attribute value again with other child nodes and moves on. It continued this process until it reached a leaf in the tree.

Random Forest: Random forest is a classifier that consists of multiple decision trees on subsets of a given dataset and takes averages to increase the prediction accuracy of the dataset. The goal is to take the predictions from each tree and predict the final outcome based on the majority vote of the predictions.

Machine Learning in Inverter Side Machine learning is used on the inverter side to improve performance and efficiency. Machine learning algorithms that analyze hybrid system data can optimize the operation of the inverter, converting direct current to alternating current with high efficiency. These algorithms can adapt to environmental changes and predict when the power output will change, allowing the inverter to adjust its operation accordingly. If there is a fault or inefficiency in the system.

This allows early detection of problems, reduces the risk of operational failures, and increases the overall reliability of the power generation system.perspectives (57%), followed by studies on smart grid factors (18%), experimental studies (15%), and examinations of challenges and benefits (9%) [15]. The survey emphasizes the growing significance of predictive maintenance, especially in the context of the Industry 4.0 revolution, and underscores the potential of Artificial Intelligence (AI) techniques for enhancing the reliability of smart grid systems.

Overview of Block Diagram

The system'soverall block diagram for Hybrid energy production with energy storage is shown in Figure1 In the proposed method, Hybrid Machine Learning is integrated to enhance fault prediction and awareness. The adaptable variable identifier interacts with MATLAB simulations to dynamically adapt and optimize the hybrid renewable energy system's response to grid faults . By combining data-driven in sights from machine learning with the adaptability of the variable identifier, the system achieves robust fault analysis, effectively mitigating disruptions and improving overall reliability indiverse operational modes. This innovative approach demonstrates superior performance compared to traditional and single-machine learning methods, ensuring efficient and adaptive renewable energy integration.

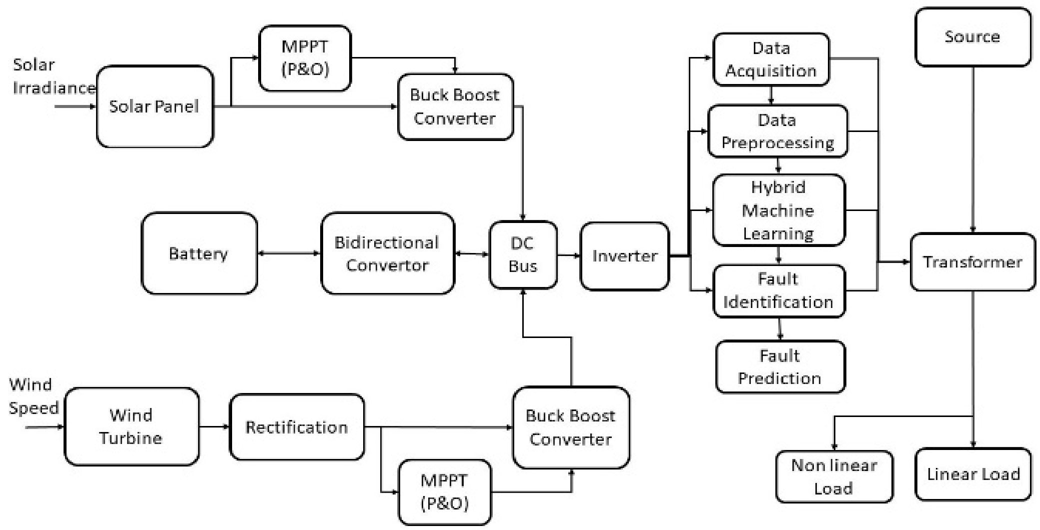

FIGURE 1. Hybrid machine learning-based fault diagnosis model.

SVM Approach

From the selected optimum relay time, the relay time is estimated using SVM approach. The model aims to draw decision boundaries between data points from different classes and separate them with maximum margin. The reason for choosing the SVM approach is, a) High-dimensional spaces had more benefits than SVM. b) Also, it handles both classification and regression on linear and non-linear data and finds the complex relationship between the inputs. This is because many relay operating time is presented in the adaptive protection process of the microgrid. c) The SVM will perform when the number of features for each data point is greater than the number of training data samples. d) Since the support vector classifier places data points above and below the classification hyperplane, there is no probabilistic justification for the classification. e) SVM performs excellently when there is a significant gap between classes and when there are more dimensions than samples. f) Diverse kernel functions can be offered for the decision function, making it versatile. The structure for SVM is shown in Figure 5, The data given for classifying is represented as a unique point in a space where each point is represented by some FIGURE 5. Structure for SVM. vector u. ui ∈ Q n (6) Here, Q n specifies a vector space with n dimension. Another vector value is described in equation (7), v ∈ Q l (7) where, l specifies the vector space with l dimension. For constructing an optimal separating hyperplane: min 1 2 a T . a + D Xδi (8) Subject to, vi a T . ui + η ≥ 1 − δi (9) δi ≥ 0, i = 1, ..., l (10) where, δi is an additional variable that characterizes the magnitude of the error ui belonging to the training set, D is the parameter of the algorithm, and l indicates the number of classes. Then, the training of the data class is defined as in equation (11), min 1 2 γ T . Zγ − e T γ (11) Subject to, v T γ = 0 (12) 0 ≤ δi ≤ D, i = 1, ... , l (13) where, γ defines a dual variable, e T is the vector of all ones, the parameter D > 0 is the upper bound and Z is a l positive semi-definite matrix. Thus, the matrix is derived as follows, Zij = vivju T i uj (14) The decision function is derived by the relation of the weight function, which is derived in equation (15), and the decision function is given in equation (16), a = Xvγ u (15) sgn Xvγ u + η (16)

Begin Initialize vector u, additional variable δi , and weight a For each u do Construct hyperplane if vi a T . ui + η ≥ 1 − δi { min 1 2 a T . a + D Xδi } else { Take iteration } end if Construct decision function sgn () if loss = satisfied { Return } else { Iteration it = it + 1 } end if End for End where, u is a vector and it is categorized as η = a T uk − vk . In this, γk = 0. Finding the exact location and grid segment where the fault occurred is a crucial next step after finding a failure in the AC microgrid. For appropriate fault isolation and to reduce downtime, fault location is essential. The SVM algorithm is particularly effective for this task due to the following reasons: The objective of the SVM algorithm is to determine the best decision boundary that maximizes the group margin. Classifying faults into different regions and sections of the microgrid (such as a specific distribution line and bus) is frequently the challenge when it comes to fault location.

SVM works well for this reason because it can clearly define the limits of choice between different fault locations based on sensor data (such as voltage and current measurements collected by different grid nodes). Even with complex and nonlinear fault location data, the kernel method in SVM allows these high-dimensional features to be transformed into a higher-dimensional space that facilitates fault classification. By building several binary classifiers, each specific to identifying the fault's location (e.g., which line and bus the fault is on), the SVM algorithm can effortlessly handle such multiclass problems. The pseudocode for the SVM approach is given above. It gathers all input and target sample data, as well as sampling information for various generating variables (also known as input sample data) and faulty location lines on the variables affecting fault currents. After that, the further process is carried out, and the performance is analyzed in the further section. These power electronics switches will have a short-time withstand capability in a millisecond range. The AC Microgrid requires instant fault isolation within milliseconds. The protection device needs maintenance due to its sensitivity and faster response. The operational and program execution times of the power system and AC Microgrid's system protective devices can be operated in milliseconds or microseconds. The faculty clearing time is needed in milliseconds for both grid-connected and islanded modes on the AC Microgrid. These factors have led to the proposed research method, utilizing the proposed SVM algorithm approach to focus on the system's capacity to operate at a suitable time. Highlighted the importance of AC microgrids to modern distributed power systems, particularly for facilitating the integration of renewable energy sources, enhancing reliability, and reducing energy expenses. Emphasize how Support Vector Machines are currently utilized in microgrids with operations involving fault location, stability analysis, load forecasts, and power quality enhancement.

The results show that the proposed approach beats conventional SVM algorithm implementations' speed, accuracy, and robustness for particular AC microgrid demands. The benefits of the widespread utilization of the SVM algorithm of AC microgrid operation and current methods struggle with real-time implementation, flexibility, and managing dynamic operational conditions. The proposed SVM technique involves a dynamic kernel function designed specifically to deal with stochastic and nonlinear characteristics of AC microgrids. With error rates of 0.0% and 0.183%, this method improves fault location and identification accuracy while minimizing execution time to 2.02 milliseconds compared to conventional models.

We have demonstrated its scalability, resilience, and adaptability by testing real-time information, resulting in a useful tool for modern AC microgrid management systems. In addition to addressing current issues, the proposed SVM algorithm research project provides the method for more sustainable and effective energy systems. In situations where there are more dimensions (features) than samples, SVM performs exceptionally well. Finding the hyperplane that optimizes the margin between classes is the main goal of the procedure, which guarantees a strong classification border. When dealing with non-linear data, SVM employs the ''kernel trick'' to translate it to higher dimensions where it can be separated linearly. This adaptability enables the proposed SVM techniques to work with different kinds of data. Using the regularization parameters voltage, current, and impedances, one may manage the trade-off between optimizing the margin and lowering error on training data. The proposed SVM method is useful for application, data classification, and fault location because it avoids the curse of dimensionality and performs well with sparse data. The SVM techniques based on convex optimization concepts ensure .

that, unlike techniques that could converge to local minima, a globally optimal solution will be found (for specific kernels and situations). SVM performs effectively with linearly separable data when using the linear kernel and with advanced kernels for non-linear data. This application helps to avoid overfitting, especially with noisy datasets. A lot of machine learning models, particularly neural networks, are opaque. It can be challenging to comprehend their decision-making process, which can be problematic in crucial fields such as law and healthcare. While complex models might be effective in training data, they might not generalize to new, untested data. Applications of machine learning in data mining and surveillance may violate people's privacy. ML model integration into current systems can be challenging and necessitate extensive re-engineering. Regular retraining and updates are necessary because models may deteriorate over time due to concept drift and changes in data distributions. A key element of various machine learning techniques, including support vector machines and kernel-based learning methods, are kernel functions. Selecting the right kernel function and optimizing its parameters is crucial for accurate and reliable fault location results. A common option for fault location is the radial basis function (RBF) kernel due to its effective modeling of non-linear relationships. This technique employs distance to compute feature space similarity and the accuracy of the kernel's fault location estimation during validation. The RBF conditions and polynomial kernels are more suitable for non-linear complex situations, while a linear kernel is proven effective for simpler cases. The accuracy of fault location is greatly impacted by the kernel choice and its parameters. Because of their flexibility in non-linear situations, RBF kernels are frequently used; however, domain-specific data and performance measures should always be the basis for selection. Viability concerning resource utilization and execution, particularly in real-time systems.

III. CONCLUSION

In summary, this paper presents a machine learning-based approach to predict solar and wind energy production with high accuracy using the 99% AUC metric. The planning process includes data collection, pre-screening, selection, sample selection, training, evaluation, and implementation. Collect and preprocess high-quality data from multiple sources, including weather data, solar radiation data, and historical solar and wind data, which are designed to eliminate outliers, solve vanishing problems, and optimize data. Select key features such as temperature, humidity, wind speed, and solar radiation for model training. Support vector machines (SVM), random forests, and gradient boosting are used as machine learning algorithms to produce accurate predictions. Models are trained on large datasets of historical solar radiation data and other important data. Use AUC and other metrics like precision, recall, and F1 score to evaluate model performance. Machine learning models are used in production to make real-time predictions for solar and wind energy

IV. ACKNOWLEDGMENT

Authors are thankful to P.G. Master of Technology In-charge and Department of Electrical and Power System , Ballarpur Institute of Technology College for providing necessary Department Facilities for conducting this study. The authors express sincere thanks to Professor Prajakta Kasulkar and all collaborative staff, and all friends and family members for constant moral support

REFERENCES

[1].N. K. Roy, A. Kumar, and B. Singh, "Solar power forecasting using hybrid LSTM-SVR model with feature engineering techniques," Energy, vol. 216, p. 119117, 2021.

[2]. S. Ali, S. T. Iqbal, and A. Javaid, "A novel approach for solar power forecasting using deep learning and ensemble methods," Journal of Cleaner Production, vol. 316, p. 126315, 2021.

[3]. Aayushi Gupta1, Jyotirmay Patel2, Mansi Gupta1, Harshit Gupta1 (2017), "Issues and Effectiveness of Blockchain Technology on Digital Voting", International Journal of Engineering and Manufacturing Science. ISSN 2249-3115 Vol. 7, No. 1. [4]. H M Diagne, M David, P Lauret, J Boland and N. Schmutz., ―Review of solar irradiance forecasting methods and a proposition for smallscale insular grids ‖ , Renew. Sustain. Energy Rev., vol. 27, pp. 65 – 76, Nov, 2013.

[5]. Chaum, D., Essex, A., Carback, R., Clark, J., Popoveniuc, S., Sherman,, A. and Vora, P. (2008) Scantegrity: "End-to-end voter-veri_able optical- scan voting", IEEE Security Privacy,

[6]. Hao, F., Ryan, P. Y. A., and Zielinski, P. (2010) "Anonymous voting by two-round public discussion", IET Information Security, vol. 4, no. 2, pp. 62-67, June 2010.