JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH

E-ISSN: 2980-9231
https://jce.bushehr.iau.ir

https://doi.org/10.30495/jce.2023.1996606.1224

Vol. 13/ No. 52/Summer 2024

Research Article

Optimizing Solar Radiation Prediction Based on The Internet of Things Platform in Photovoltaic Power Plant

Neda Ashrafi Khozani, Instructor[1] | Maryam Mahmoudi, Instructor[2*] | Shabnam Nasr Esfahani, Assistant Professor[3]

[1]Department of Computer Engineering, Meymeh Branch, Islamic Azad University, Meymeh, Iran, anahid.shaabani@gmail.com

[2]Department of Computer Engineering, Meymeh Branch, Islamic Azad University, Meymeh, Iran, Mahmoudi.m174@iau.ac.ir

[3]Department of Electrical Engineering, Meymeh Branch, Islamic Azad University, Meymeh, Iran, nasr.sh.2010@gmail.com

Correspondence
Maryam Mahmoudi, Instructor of Computer Engineering, Meymeh Branch, Islamic Azad University, Meymeh, Iran, Mahmoudi.m174@iau.ac.ir

Received: 16 September 2023
Revised: 15 December 2023
Accepted: 22 December 2023

Abstract

The solar radiation value parameter is one of the most important parameters in determining the output power value of photovoltaic panels. Accurate prediction of this parameter is crucial for dispatching and load management planning. Managers and designers encounter economic and managerial challenges due to the uncertainty and difficulty in predicting solar radiation levels. This research introduces a highly accurate prediction method utilizing tree-based methods, enhanced by meta-heuristic algorithms to boost performance. The proposed method emphasizes preventing overfitting and ensuring high reliability for use in Internet of Things systems. Meta-heuristic algorithms are utilized for optimizing tree-based methods, as well as for feature and instance selection. Employing meta-heuristic methods as the main innovation in this research not only optimizes machine learning model settings but also mitigates the impact of noise, outliers, and ineffective inputs, thereby enhancing the final output quality. Utilizing an innovative fitness function in model optimization enhances prediction accuracy and adaptability to real photovoltaic power plant environments. The final outcome is a strong model that has a score of 0.95 with the R-square criterion and is optimal model.

Keywords: Internet of Things, Decision Tree, Machine Learning, Bat Algorithm, Photovoltaic Power Plants.

Highlights

- Accurate prediction of the amount of solar radiation as an important parameter in determining the amount of output power of photovoltaic panels.
- Optimizing tree-based models with meta-heuristic algorithms for prediction of solar radiation parameters.
- Balance the accuracy, complexity, and suitability of the final model for IoT device implementation.
- In the end, a strong model that has a score of 0.95 with the R-square criterion was obtained in this research.
- The final model can be implemented in the environment of power plants based on the Internet of Things.

Citation: N. Ashrafi Khozani, M. Mahmoudi, and Sh. Nasr Esfahani, “Optimizing Solar Radiation Prediction Based on The Internet of Things Platform in Photovoltaic Power Plant,” Journal of Southern Communication Engineering, vol. 13, no. 52, pp. 33–44, 2024, doi: 10.30495/jce.2023.1996606.1224, [in Persian].

1. Introduction

Solar radiation is a fundamental determinant of the power output in photovoltaic (PV) systems, making accurate solar radiation predictions is a critical component of energy management and planning in PV power plants. The highly variable and unpredictable nature of solar radiation introduces significant challenges in maintaining consistent energy output, which is essential for efficient power system dispatch and load balancing. As the demand for renewable energy sources continues to grow, so does the need for advanced predictive models that can effectively manage the inherent uncertainties of solar radiation. The unpredictable nature of solar radiation can lead to significant inefficiencies in power generation, making it difficult to meet energy demand reliably. Traditional prediction models often fall short in accuracy and adaptability, especially when applied to the dynamic and data-rich environments typical of IoT systems in modern PV plants.

While there are some attempts to predict this parameter accurately using methods such as Artificial Neural Networks [1], LSTM [2], or Support Vector Machines [3], there is still a research gap to make predictive models for this parameter keeping a balance between accuracy and generality and useability in real-time systems.

This study aims to address the challenges of solar irradiance prediction by exploring advanced machine learning techniques. Specifically, it focuses on developing a robust model that can accurately forecast solar radiation using meteorological data. The research leverages data preprocessing techniques such as outlier detection and normalization to ensure the quality and consistency of the input data. Additionally, the study compares various machine learning algorithms to identify the most effective approach for predicting solar irradiance.

2. Innovation and contributions

In this paper, the key innovations is the use of meta-heuristic algorithms (Bat Algorithm) to optimize tree-based predictive models as well as feature selection and instance selection. The research contributes to the field by developing models suitable for integrating machine learning models with IoT systems for real-time monitoring and error prevention in photovoltaic power plants. This integration is designed to handle the computational constraints of IoT devices while maintaining high prediction accuracy. Another innovative aspect of this research lies in its application of machine learning models that are not only optimized for accuracy but also tailored to minimize overfitting and improve generalizability in practical settings. This is achieved through the use of novel fitness functions and metaheuristic algorithms, which enhance the robustness and simplicity of the models for real-world deployment.

By enhancing the prediction models for solar radiation, the research supports better energy management and sustainability practices. The improved accuracy in predicting solar energy output allows for more efficient use of resources and helps in stabilizing voltage and frequency in power networks. Additionally, this innovation reduces economic and managerial challenges associated with the variability of solar power, facilitating more widespread adoption of clean energy technologies. The study also emphasizes the importance of noise reduction and the handling of outliers in data, which are critical for improving the quality of predictive models in energy applications.

Among the innovations of this paper, the following can be stated:

- Utilizing Bat Algorithm to optimize tree-based predictive models as well as feature selection and instance selection.
- Using an innovative fitness function in the optimization process to keep both the accuracy and generality of the final model.

3. Materials and Methods

The dataset used in this study comprises environmental data collected from a photovoltaic power plant over a significant period. The data include various parameters essential for accurate solar radiation prediction: Temperature, humidity, and wind Speed.

The proposed method machine learning approach for predicting solar radiation, and utilizing two main methods: Random Forests and AdaBoost. Both methods are based on decision trees but differ in their training and combination processes. Random Forests involve multiple decision trees making predictions independently, with their results combined to form the final model. In contrast, AdaBoost trains decision trees sequentially, with each new tree focusing on data points that previous trees predicted poorly.

A critical aspect of this approach is hyperparameter tuning, as the models require careful selection of parameters like the number of trees, tree depth, and more. These parameters significantly impact the model's performance. Additionally, the selection of features (inputs) and training samples is crucial to the model's success. The study introduces an innovative approach by using the Bat Algorithm, a nature-inspired metaheuristic, to optimize hyperparameters, feature selection, and sample selection. The Bat Algorithm is chosen for its rapid convergence, offering an advantage similar to that of Genetic Algorithms.

The dataset is processed to remove noise and outliers, ensuring that the input data for the machine learning models is of high quality. The cleaning process involves normalizing the data and applying statistical methods to detect and remove anomalies. The cleaned and processed dataset is then used to train and test various machine learning models, allowing for the optimization of prediction algorithms for solar radiation in photovoltaic power plants. The data is then split into training and test sets, with a 90:10 ratio, where only 10% of the data is used for testing.

The model selection process involves choosing an algorithm (either Random Forests or AdaBoost), tuning hyperparameters, and selecting features and samples. This approach optimizes a fitness function based on the model's accuracy, measured by the R^2 score, and its complexity, indicated by the number of trees and tree depth. The goal is to achieve a balance between high predictive accuracy and minimal complexity, making the final model efficient for implementation in IoT-based control systems.

The final model not only aims for high accuracy but also focuses on minimizing complexity to ensure efficient resource use when deployed in IoT systems. The model can generate a set of fixed rules for solar radiation prediction, which can be applied offline, providing robustness against communication failures in IoT environments. This approach ensures that the prediction process remains reliable, even in the face of connectivity issues, by replacing the need for repeated model development with a set of pre-defined predictive rules.

4. Results and Discussion

The study implemented a data preprocessing step using Cook's analysis, which identified and removed 2.7% of the data as outliers before model training. Following this, the Bat Algorithm was employed to optimize the configuration for predicting solar radiation intensity, resulting in a Random Forest model with specific hyperparameters.

The final model achieved an R^2 score of 0.956 on the training data and 0.948 on the test data, indicating high predictive accuracy. Although some candidate models showed higher accuracy during optimization, they were rejected due to significant overfitting or excessive complexity (e.g., a high number of decision trees). The chosen Random Forest model had a maximum depth of 25 and 89 trees, leading to a manageable number of rules that can be used offline in IoT devices.

The model's generalization ability was further validated using the k-fold cross-validation method (k=10), which yielded a low standard deviation of 0.00148 in the R^2 score, indicating high robustness.

When compared with other methods on similar datasets, the proposed method generally outperformed them in terms of accuracy, except in one case where it had a slightly lower R^2 score. However, this was justified by the proposed model's lower complexity and smaller standard deviation, highlighting its efficiency and simplicity.

5. Conclusion

This study successfully demonstrates the use of machine learning models, optimized through heuristic algorithms, for predicting solar radiation in photovoltaic power plants within an IoT framework. By addressing the variability in solar radiation, the research enhances the efficiency and reliability of photovoltaic systems. The optimized models, which incorporate advanced techniques such as feature selection and hyperparameter tuning, provide high prediction accuracy while remaining computationally efficient for real-time applications.

The integration of these models into IoT environments enables continuous and accurate forecasting, aiding in effective energy management and improving the stability of power networks. This work not only contributes to better predictive capabilities in renewable energy systems but also promotes more sustainable energy practices. Future research can build on these findings by exploring additional environmental factors and advanced modeling techniques to further refine prediction accuracy and system performance.

6. Acknowledgement

The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors have no relevant financial or non-financial interests to disclose.

7. References

[1] A. Angstrom, "Solar and terrestrial radiation. Report to the international commission for solar research on actinometric investigations of solar and atmospheric radiation," *Quarterly Journal of the Royal Meteorological Society*, vol. 50, no. 210, p. 121-126, 1924, doi: 10.1002/qj.49705021008.

[2]. X. Qing and Y. Niu, "Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM," *Energy*, vol. 148, pp. 461-468, 2018, doi: 10.1016/j.energy.2018.01.177.

[3] N. Sharma, P. Sharma, D. Irwin and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, Brussels, Belgium, 2011, pp. 528-533, doi: 10.1109/SmartGridComm.2011.6102379.

[4] A. Mellit, M. Menghanem and M. Bendekhis, "Artificial neural network model for prediction solar radiation data: application for sizing stand-alone photovoltaic power system," *IEEE Power Engineering Society General Meeting*, 2005, San Francisco, CA, USA, 2005, pp. 40-44 Vol. 1, doi: 10.1109/PES.2005.1489526.

[5] J. Lago *et al.*, "Short-term forecasting of solar irradiance without local telemetry: A generalized model using satellite data," *Solar Energy*, vol. 173, pp. 566-577, 2018, doi: 10.1016/j.solener.2018.07.050.

[6] N. Dong *et al.*, "A novel convolutional neural network framework based solar irradiance prediction method," *International Journal of Electrical Power & Energy Systems*, vol. 114, p. 105411, doi: 10.1016/j.ijepes.2019.105411.

Appendix

Table 1. Parameters of the Final Model

Hyperparameter	Value
Minimum samples per split	32
Maximum number of features	4
Maximum tree depth	25
Number of base decision trees	89

Table 2. Comparison with existing methods. (It should be noted that, due to the differences in datasets and evaluation metrics reported in existing research, these methods have been re-implemented based on the parameters provided in the respective articles, and the results have been reported on similar datasets in this table.)

Method	Average R^2 Score in 10 Repeats	Standard Deviation of R^2 in 10 Repeats
Proposed Method	0.956	0.00148
Artificial Neural Networks [4]	0.948	0.00942
Artificial Neural Networks and Wavelet Analysis [1]	0.893	0.01240
LSTM [2]	0.941	0.01107
Convolutional Neural Networks [5]	0.971	0.09849
Neural Networks and PSO/GA Hybrid Algorithm [6]	0.939	0.00144
Support Vector Machines [3]	0.921	0.00166

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Publisher's Note: All content expressed in this article is solely that of the authors, and does not necessarily reflect the views of their affiliated organizations or the publisher, editors, and reviewers. Any content or product that may be reviewed and evaluated in this article is not guaranteed or endorsed by the publisher.

Author Contributions: All authors reviewed the manuscript.