HUIT Journal of Science ***25*** *(1) (2025) 156-164*

DOI: https://doi.org/10.62985/j.huit_ojs.vol25.no1.164

FORECASTING POWER GENERATION OF SOLAR PANELS USING MACHINE LEARNING

Huynh Van Van*, Nguyen Ngoc Thanh Van

Faculty of Electrical and Electronics Engineering,

Ton Duc Thang University, Ho Chi Minh City, Viet Nam

*Email: *huynhvanvan@tdtu.edu.vn*

Received: 4 December 2024; Accepted: 20 January 2025

ABSTRACT

Solar energy, a renewable resource, is becoming an increasingly important part of the global energy source. This energy, which is known for its clean and abundant amount, has become the subject of numerous research studies. Advances in science and technology have made the use of solar energy more common and widely used. However, the operation of the solar system sometimes encounters difficulties due to constant weather fluctuations, which significantly affect the output capacity of the photovoltaic solar panels. This can lead to an electricity surplus when the load consumes less power or a lack of power supply when there is insufficient energy to produce it. To solve this problem, this research focuses on using Machine Learning methods to predict the output capacity of solar panels based on weather data. There are two main data sources used in the research: the data provided and the actual data collected through the measurement process achieved. The input data for the prediction model includes ambient temperature, wind speed, humidity, and total solar radiation from the provided dataset, as well as ambient temperature and solar illuminance from the collected dataset. The output of the prediction model is the power generation from solar panels. The results of the study will help the power system operate more efficiently, ensuring the balance between supply and demand, thus utilizing the potential of solar energy effectively.

Keywords: Machine learning, solar energy, forecast, renewable resource.

1. INTRODUCTION

Vietnam actively participates in the global trend of countries concentrating on renewable energy solutions. Given Vietnam's year-round high amount of solar radiation, solar energy has a lot of promise. The expansion of solar power plants is supported by this abundant sunlight, which lowers greenhouse gas emissions and dependency on fossil fuels. Optimizing the use of solar energy requires accurate solar radiation predictions. To forecast solar power and enhance the management of renewable energy, a variety of models, including regression and Machine Learning techniques, have been used [1, 2]. Further encouraging the incorporation of renewable energy sources are techniques such as Multigene Genetic Programming (MGGP) and Multilayer Perceptron (MLP), which have proven successful in predicting solar irradiance across various sites [3].

Investing in solar energy also brings economic benefits, such as job creation and improved quality of life, especially in rural and remote areas with limited access to reliable energy. As Vietnam continues to expand its solar power projects, the development of advanced forecasting models becomes increasingly important for managing variability in solar power output. Studies have highlighted the potential of automatic hourly solar forecasting using tree-based machine learning models and other algorithms to handle diverse climatic conditions [4]. Furthermore, deep learning methods like Extreme Gradient Boosting (XGBoost) [5] and optimized neural networks have shown promise in addressing challenges related to power output fluctuations [6, 7].

Solar power output depends significantly on solar radiation, but it can fluctuate due to factors like weather conditions. This variability can cause imbalances in the electricity grid, making short-term forecasting essential for maintaining stability. Research has explored the impact of environmental factors such as temperature on PV module surfaces during electricity generation, providing insights into managing these fluctuations [8]. Moreover, comparative studies on data mining methods and preprocessing techniques have identified the most effective approaches for predicting solar radiation and temperature, reducing forecasting errors [9, 10].

The goal of this research is to create a forecasting model that uses weather data to predict solar power generation in order to overcome these difficulties. Studies on solar power forecasting using deep learning have shown that MLP Regression approaches are especially well-suited for this purpose because of their sophisticated data processing capabilities [11]. The accuracy of these models is further improved by incorporating cutting-edge optimization techniques, including the Chimp Optimization Algorithm [12]. Additionally, hybrid architectures that combine the advantages of recurrent and convolutional networks, such as deep convolutional long short-term networks, have been proposed for probabilistic forecasting [13].

Another key component of this research is the integration of Internet of Things (IoT)-based solutions for real-time data collection and performance evaluation. IoT devices, such as those built on ESP32, have been successfully implemented in photovoltaic systems for remote monitoring and data management [14, 15]. These systems enable the efficient collection of weather and operational data, providing a robust foundation for model validation and refinement.

A comprehensive review of Machine Learning techniques for solar power forecasting highlights the importance of selecting the right model and tuning hyperparameters for accuracy [16, 17]. Advanced methods have been applied in Alice Springs, demonstrating their effectiveness in managing power fluctuations and optimizing energy distribution [18]. Studies also emphasize the role of Machine Learning in mitigating environmental risks and improving solar system performance [19]. Beyond solar energy, these techniques show potential for broader applications in renewable energy systems like wind power, supporting global sustainability goals [20].

In general, the research aims to develop a forecasting model using weather data to predict solar power generation. Initially, the study will use data from collaborative projects with the university in Taiwan to construct and validate various forecasting methods and processes. Once a model is developed, data from an IoT-based model will be collected to evaluate its performance in handling diverse data sources. The research will utilize MLP Regression techniques to enhance the model's accuracy in predicting solar power output, leveraging MLP's advanced data processing capabilities. Python programming language, with its robust libraries and tools for data analysis and Machine Learning, will be used to develop and implement the forecasting model.

2. METHODOLOGY

2.1. Machine Learning

Artificial Intelligence, especially Machine Learning, is gaining prominence and widespread use. Machine Learning enables computers to learn and adapt by analyzing data, rather than following pre-programmed instructions. It helps uncover patterns in large datasets and is used in areas like image recognition, spam filtering, complex modeling, and prediction. Machine Learning includes three main types: supervised learning, unsupervised learning, and reinforcement learning, each with its strengths and weaknesses. The method chosen depends on data characteristics and goals. Applying Machine Learning involves multiple stages of data processing and classification to develop an accurate predictive model.

Preprocessing: This stage is a critical initial step in Machine Learning, focused on improving data quality. Missing values are addressed using methods like mean imputation, while outliers, often caused by illogical input-output connections, are removed to prevent noise and enhance forecasting performance. Effective data preparation improves model performance, accuracy, and analysis efficiency.

Clustering: This technique is used to group data into meaningful clusters. Data within the same cluster will have more similar characteristics compared to others in different clusters. This stage helps uncover hidden structures within the data being analyzed.

Split data: A complete dataset is often divided into smaller parts to serve different purposes, aiming to build a model with good generalization and high performance. In Machine Learning, data is typically split into three common sets: train data, validation data, and test data.

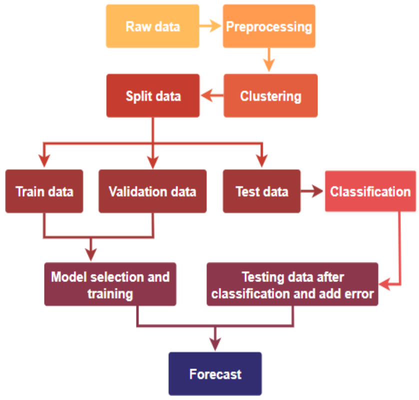

Figure 1. Flowchart of machine learning stages

Classification: In this stage, the model classifies test data when the data has been previously clustered. It learns from the data and then predicts new clusters for data points with high accuracy.

Model selection and training: This stage involves choosing and training the best model by evaluating various algorithms, tuning hyperparameters, and using validation and test data to ensure good performance. The selected model is then deployed and monitored for effectiveness in practice.

2.2. MLP – Multi-Layer Perceptron Regression

MLP Regression is a Machine Learning method featuring multiple layers of neurons, inspired by the human brain's neural network. Neurons in MLP are organized in layers, with each neuron connected to those in subsequent layers via adjustable weights. These weights are optimized during training to improve the model's ability to understand and represent data. MLP aims to uncover and model nonlinear relationships between input and output data, enabling accurate predictions or classifications.

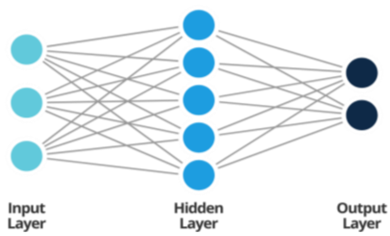

Figure 2. MLP neural network diagram

The structure of layers is typically composed of three main types of layers, each serving a specific function, working together to enable the model to learn from data and make accurate predictions:

Input Layer: This layer is responsible for taking in the raw data that will be processed by the network.

Hidden Layer: This layer consists of one or more layers of neurons that perform the essential task of processing and extracting features from the input data. The hidden layers enable the network to learn complex patterns and representations by transforming the input data through nonlinear activations.

Output Layer: This layer produces the final prediction or output based on the processed information from the hidden layers. It translates the network's learned features into the desired result, such as a classification label or a forecasting value.

In MLP, the process starts with the input layer receiving raw data. Each neuron in this layer computes a weighted sum of the inputs and applies an activation function to produce output values. These outputs are then passed to the neurons in the first hidden layer. This process continues through all hidden layers until the data reaches the output layer.

The optimal number of hidden layers in a neural network model is determined through iterative testing. This involves comparing different configurations to assess the model's performance using error metrics.

2.3. IoT Model

To assess the model's adaptability to different data sources, an IoT application model was developed for data collection. In this model, the ESP32 microcontroller plays the most crucial role. ESP32 is an electronic component that facilitates the collection and transmission of data, specifically related to weather conditions and solar panel output in this study. With its built-in Bluetooth and Wi-Fi connectivity, the ESP32 easily exchanges data with a server or cloud platform. As a result, the model can be controlled and monitored remotely, facilitating data collection.

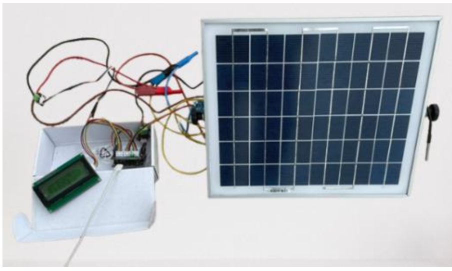

Figure 3. IoT Model

The components constituting the model include the following parts:

Solar Panel: Used to collect energy from sunlight and convert it into electrical energy. In the model, the solar panel acts as a small power source.

Load: A 12V small light bulb connected to the solar panel, receiving energy and functioning as an electrical load in the system.

Current-Voltage Sensor: Directly connected to the solar panel and the load, this sensor measures and records data on the current and voltage output from the panel, providing necessary information for calculating the panel's power output.

Temperature Sensor: Placed near the solar panel to collect data on the surrounding environmental temperature, which affects the power output value.

Light Sensor: Mounted parallel to the surface of the solar panel, this sensor measures the illuminance of sunlight, helping to assess the amount of light the panel receives.

LED Screen: Directly displays measured parameters such as illuminance, ambient temperature, current, and voltage, providing quick and clear information.

ESP32: The central electronic component that connects the LED screen and sensors, responsible for transmitting collected data to the server for analysis and storage.

Power Supply for ESP32: Powered by a rechargeable battery, ensuring the ESP32 operates continuously and stably within the model.

Connecting Wires: Used to connect components in the system, ensuring the transmission of data and electrical power among the various parts.

To facilitate data collection, the model will be placed on the rooftop of a building at the university. The data will be recorded continuously and sent to the server, with an interval of 1 minute between recordings.

3. RESULTS AND DISCUSSION

3.1. Data Description

The data source provided is available from collaborative scientific projects in Taiwan. However, it should be clarified that this dataset is solely used to validate the solar panel power forecasting process in this research and is not intended for forecasting in Vietnam. Once the research achieves its objectives, the method will be applied to weather data collected in Vietnam to forecast capacity that aligns with local conditions. The dataset includes information such as the following:

- The time interval between the two data points: 1 hour.

- The dataset consists of 5 columns (4 inputs, 1 output) with 12827 rows.
- Input data: Ambient temperature, wind speed, humidity, and total solar radiation.
- Output data: Power generation from solar panels.

The dataset will be used to evaluate the forecasting ability of the model, including testing the implementation processes, data filtering processes, and model-building selection methods. If the forecast results meet the error requirement (around 5% or less), this criterion can conclude that the forecasting model operates stably and is reliable in its forecasting ability.

After completing the forecasting model, the collected dataset through the IoT model will be used to verify the implementation steps and the chosen methods when applied to the new dataset. This process helps assess the model's capabilities in practical conditions and its ability to handle a variety of different types of data. This dataset includes the following information:

- The time interval between the two data points: 1 minute.
- The dataset consists of 3 columns (2 inputs, 1 output) with 14903 rows.
- Input data: Ambient temperature and solar illuminance.
- Output data: Power generation from solar panels.

3.2. Error Metrics

In this research, two error metrics are chosen as criteria for selecting the appropriate number of hidden layers when building the forecasting model. Key parameters for evaluation include:

- MSE (Mean Squared Error) is widely used in regression tasks to quantify the average squared differences between predicted and actual values. It is an important criterion for assessing and selecting effective predictive models. A lower MSE indicates higher accuracy and a better alignment with the actual data.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

- MRE (Mean Relative Error) is the average of the relative errors between the model's predicted values and the actual values, expressed as a percentage of the original values.

$$MRE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{2}$$

Where y_i is the actual value for the i-th observation; $\hat{y}_i$ is the predicted value for the i-th observation; n is the number of observations.

3.3. Result

The dataset will be processed following the steps outlined in the flowchart for building the Machine Learning forecasting model. The initial raw dataset, after being preprocessed, results in a complete dataset used for training the forecasting model. In the data clustering stage, it is necessary to determine the optimal number of clusters to ensure effective grouping and enhance the accuracy and interpretability of the results. The Elbow Method is utilized for this purpose, relying on the Within-Cluster Sum of Squares (WCSS) to determine the number of clusters. K, representing the number of clusters, the formula for calculating WCSS is defined as follows:

$$WCSS = \sum_{i=1}^{K}\sum_{x\in C_i}\left\|x - \mu_i\right\|^2 \tag{3}$$

Where K is the number of clusters; C_i are data points belonging to cluster i; μ_i is the centroid of cluster i; x is the data point within a cluster.

This formula shows how clustered the data points are, with a smaller WCSS value indicating greater proximity among the data points. The following graph illustrates the results of the Elbow Method for determining the optimal number of clusters.

The graph shows that WCSS decreases sharply until $K = 5$, after which the reduction becomes more gradual. Therefore, 5 is selected as the optimal number of clusters for effective data grouping.

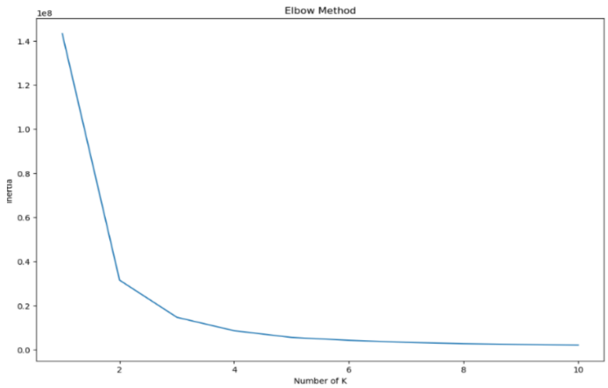

Figure 4. Elbow Method Graphic

Each cluster will be developed with a separate model instead of creating a common model for the entire dataset. The reason for this is that data points belonging to the same cluster often share similar characteristics; therefore, each cluster's model will be analyzed in detail and aligned better with their distinct attributes. This will help improve the accuracy of predictions.

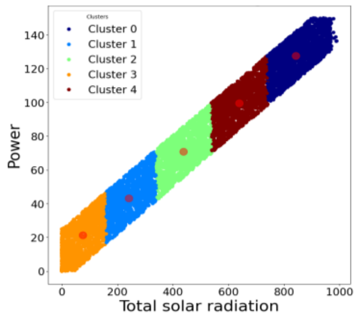

Figure 5. Clustering graphic

Each cluster will be developed with a separate model instead of creating a common model for the entire dataset. The reason for this is that data points belonging to the same cluster often share similar characteristics; therefore, each cluster's model will be analyzed in detail and aligned better with their distinct attributes. This will help improve the accuracy of predictions.

Set up neural network models with 1, 2, 3, 4, and 5 hidden layers containing 20 neurons for each data cluster. The number of neurons was determined through experimental surveys to ensure a balance between model complexity and computational efficiency. After training, the models will be compared with each other to evaluate their suitability, generalization ability, and predictive effectiveness. The results of the modeling process will be summarized in Table 1.

Table 1. Values for error metrics in clusters

Layer / Cluster	1		2		3		4		5	
	MSE	MRE	MSE	MRE	MSE	MRE	MSE	MRE	MSE	MRE
0	83.365	3.868%	84.029	3.881%	84.037	3.884%	84.728	3.899%	82.931	3.864%
1	71.402	3.626%	71.172	3.622%	71.239	3.624%	76.112	3.704%	70.923	3.618%
2	81.786	3.894%	81.350	3.876%	81.679	3.888%	82.321	3.904%	82.134	3.904%
3	82.630	4.004%	85.567	4.041%	84.302	4.024%	86.049	4.049%	83.343	4.013%
4	78.517	3.817%	76.512	3.792%	77.916	3.805%	80.784	3.858%	79.287	3.824%

Based on the data from Table 1, using the criterion of selecting the lowest error, we will sequentially choose the appropriate number of hidden layers for each data cluster. Specifically:

- Cluster 0: 5 hidden layers (MSE=82.931, MRE=3.864%)

- Cluster 1: 5 hidden layers (MSE=70.923, MRE=3.618%)
- Cluster 2: 2 hidden layers (MSE=81.350, MRE=3.876%)
- Cluster 3: 1 hidden layer (MSE=82.630, MRE=4.004%)
- Cluster 4: 2 hidden layers (MSE=76.512, MRE=3.792%)

Having developed models with a suitable number of hidden layers for each data cluster, the number of hidden layers varies depending on the complexity of the relationships between data points within each cluster. Clusters with a large amount of data often require a more complex model structure to effectively minimize forecasting errors.

To train the forecasting models effectively, the MLP Regressor was configured as follows:

- Hidden layer: 5 hidden layers for Cluster 0 and Cluster 1; 2 hidden layers for Cluster 2 and Cluster 4; 1 hidden layer for Cluster 3, each layer contains 20 neurons.
- Activation function: Rectified Linear Unit (ReLU)
- Loss function: MSE
- Optimizer: Adam optimizer
- Maximum iterations: 1500

We will execute the forecasting program and assess the results by comparing the predicted values with the actual power values, resulting in the following graph:

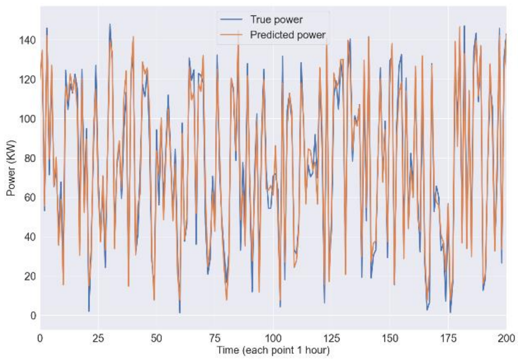

Figure 6. Comparison of predicted and actual values (Provided data)

Based on the graph, it can be observed that the power forecasting model has demonstrated a good ability to predict the actual electricity consumption trends. The forecasted data curve closely follows the actual data curve, indicating that the model has captured the main fluctuations in power and has a relatively good forecasting capability.

From the results obtained, we will apply all the processes to build a complete forecasting model on a new dataset - the collected dataset through the IoT model. Following the same steps, after executing the forecasting process, we obtain the following graph:

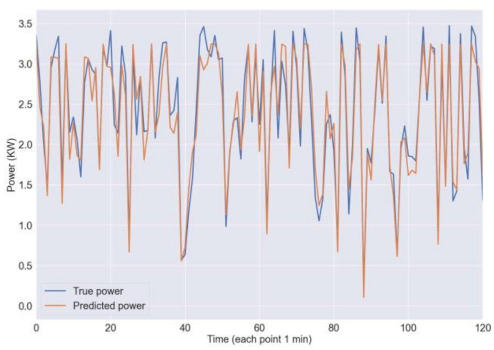

Figure 7. Comparison of predicted and actual values (Provided data)

Through the chart, we can see that the forecasted values closely follow the actual values. This indicates that the model has successfully captured the trend in actual power fluctuations. This result also demonstrates that the forecasting model can handle various types of datasets with different variables, proving its applicability in practice.

4. CONCLUSION

In conclusion, the research on applying IoT and Machine Learning in solar panel power output forecasting has been completed with positive results. Through thorough data analysis, processing, and calculation, the study emphasizes the importance of power forecasting and the challenges in building an accurate predictive model. The steps undertaken have helped develop an appropriate model, yielding forecasts with minimal errors and high accuracy. The research has achieved its objectives, contributing to the optimization of smart solar energy technology and supporting power output forecasting in practice. Furthermore, this research can be applied to predict actual power generation using comprehensive training datasets collected from renewable energy power plants in Vietnam. These findings provide a strong foundation for future advancements in renewable energy forecasting and its practical applications in the energy sector.

REFERENCES

1. Elsayed A., Elshinawy M. Y., Ibrahim N., Woodall R., Badawy A. H. A., & Ranade S. - Regression models for solar radiation prediction, 2023 Congress in Computer Science, Computer Engineering & Applied Computing (CSCE), IEEE (2023) 274-281. https://doi.org/10.1109/CSCE60160.2023.00048
2. Voyant C., Notton G., Kalogirou S., Nivet M. L., Paoli C., Motte F., & Fouilloy A. - Machine learning methods for solar radiation forecasting: A review, Renewable Energy **105** (2017) 569-582. https://doi.org/10.1016/j.renene.2016.12.095
3. Mendonça de Paiva G., Pires Pimentel S., Pinheiro Alvarenga B., Gonçalves Marra E., Mussetta M., & Leva S. - Multiple site intraday solar irradiance forecasting by machine learning algorithms: MGGP and MLP neural networks, Energies **13** (11) (2020) 3005. https://doi.org/10.3390/en13113005
4. Yagli G. M., Yang D., & Srinivasan D. - Automatic hourly solar forecasting using machine learning models, Renewable and Sustainable Energy Reviews **105** (2019) 487-498. https://doi.org/10.1016/j.rser.2019.02.006
5. Abumohsen M., Owda A. Y., Owda M., Abumihsan A., & Stergioulas L. - Forecasting solar power generation using extreme gradient boosting: A machine learning approach, 2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA), IEEE (2024) 505-510. https://doi.org/10.1109/ICSINTESA62455.2024.10748030
6. Inanlougani A., Reddy T. A., & Katiamula S. - Evaluation of time-series, regression and neural network models for solar forecasting: Part I: One-hour horizon, arXiv preprint arXiv:1708.08376 (2017). https://doi.org/10.48550/arXiv.1708.08376
7. Soleymani S., Mohammadzadeh S. - Comparative analysis of machine learning algorithms for solar irradiance forecasting in smart grids, arXiv preprint arXiv:2310.13791 (2023). https://doi.org/10.48550/arXiv.2310.13791
8. Karim S. R., Sarker D., & Kabir M. M. - Analyzing the impact of temperature on PV module surface during electricity generation using machine learning models, Cleaner Energy Systems **9** (2024) 100135. https://doi.org/10.1016/j.cles.2024.100135
9. Alay F. D., İlhan N., & Güllüoğlu M. T. - A comparative study of data mining methods for solar radiation and temperature forecasting models, Journal of Universal Computer Science **30** (6) (2024) 847. https://doi.org/10.3897/jucs.109080
10. Paoli C., Voyant C., Muselli M., & Nivet M. L. - Solar radiation forecasting using ad-hoc time series preprocessing and neural networks, Emerging Intelligent Computing Technology and Applications: 5th International Conference on Intelligent Computing (ICIC 2009), Ulsan, South Korea, Springer Berlin Heidelberg (2009) 898-907. https://doi.org/10.48550/arXiv.0906.0311
11. Elsaraiti M., & Merabet A. - Solar power forecasting using deep learning techniques, IEEE Access **10** (2022) 31692-31698. https://doi.org/10.1109/ACCESS.2022.3160484

12. Al-Dahidi S., Alrbai M., Alahmer H., Rinchi B., & Alahmer A. - Enhancing solar photovoltaic energy production prediction using diverse machine learning models tuned with the chimp optimization algorithm, Scientific Reports **14** (1) (2024) 18583. https://doi.org/10.1038/s41598-024-69544-8
13. Bai M., Zhao X., Long Z., Liu J., & Yu D. - Short-term probabilistic photovoltaic power forecast based on deep convolutional long short-term memory network and kernel density estimation, arXiv preprint arXiv:2107.01343 (2021). https://doi.org/10.48550/arXiv.2107.01343
14. Hercog D., Lerher T., Truntič M. & Težak O. - Design and implementation of ESP32-based IoT devices, Sensors **23** (15) (2023) 6739. https://doi.org/10.3390/s23156739
15. Allafi I., & Iqbal T. - Design and implementation of a low cost web server using ESP32 for real-time photovoltaic system monitoring, 2017 IEEE Electrical Power and Energy Conference (EPEC), IEEE (2017) 1-5. https://doi.org/10.1109/EPEC.2017.8286184
16. Gaboitaolelwe J., Zungeru A. M., Yahya A., Lebekwe C. K., Vinod D. N., & Salau A. O. - Machine learning based solar photovoltaic power forecasting: A review and comparison, IEEE Access **11** (2023) 40820-40845. https://doi.org/10.1109/ACCESS.2023.3270041
17. Markovics D., & Mayer M. J. - Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction, Renewable and Sustainable Energy Reviews **161** (2022) 112364. https://doi.org/10.1016/j.rser.2022.112364
18. Mahmud K., Azam S., Karim A., Zobaed S., Shanmugam B., & Mathur D. - Machine learning based PV power generation forecasting in Alice Springs, IEEE Access **9** (2021) 46117-46128. https://doi.org/10.1109/ACCESS.2021.3066494
19. Anuradha K., Erlapally D., Karuna G., Srilakshmi V., & Adilakshmi K. - Analysis of solar power generation forecasting using machine learning techniques, E3S Web of Conferences **309** (2021) 01163. https://doi.org/10.1051/e3sconf/202130901163
20. Alvarez L.F.J., González S.R., López A.D., Delgado D.A.H., Espinosa R. & Gutiérrez S. - Renewable energy prediction through machine learning algorithms, 2020 IEEE ANDESCON, IEEE (2020) 1-6. https://doi.org/10.1109/ANDESCON50619.2020.9272029

TÓM TẮT

DỰ BÁO CÔNG SUẤT PHÁT CỦA TẤM PIN NĂNG LƯỢNG MẶT TRỜI SỬ DỤNG PHƯƠNG PHÁP HỌC MÁY

Huỳnh Văn Vạn*, Nguyễn Ngọc Thanh Vân

Khoa Điện - Điện tử, Trường Đại học Tôn Đức Thắng, TP. Hồ Chí Minh, Việt Nam

*Email: *huynhvanvan@tdtu.edu.vn*

Năng lượng mặt trời, một loại nguồn năng lượng tái tạo, đang ngày càng trở thành một phần quan trọng của nguồn năng lượng toàn cầu. Năng lượng này, được biết đến với đặc tính sạch và nguồn cung dồi dào, đã trở thành đối tượng của nhiều nghiên cứu khoa học. Những tiến bộ trong khoa học và công nghệ đã thúc đẩy việc sử dụng năng lượng mặt trời trở nên phổ biến và được sử dụng rộng rãi hơn. Tuy nhiên, việc vận hành hệ thống năng lượng mặt trời đôi khi gặp nhiều khó khăn do sự biến động thời tiết liên tục, ảnh hưởng đáng kể đến công suất đầu ra của các tấm pin năng lượng mặt trời. Vấn đề này có thể dẫn đến thừa điện khi tải tiêu thụ ít năng lượng hơn hoặc thiếu nguồn điện cung cấp cho tải khi không đủ năng lượng để sản xuất. Để giải quyết vấn đề này, nghiên cứu tập trung vào việc sử dụng các phương pháp Machine Learning để dự đoán công suất đầu ra của các tấm pin mặt trời dựa trên dữ liệu thời tiết. Có hai nguồn dữ liệu chính được sử dụng trong nghiên cứu: dữ liệu được cung cấp và dữ liệu thực tế thu thập thông qua quá trình đo lường. Dữ liệu đầu vào cho mô hình dự đoán bao gồm nhiệt độ môi trường, tốc độ gió, độ ẩm không khí và tổng lượng bức xạ mặt trời từ tập dữ liệu được cung cấp, cũng như nhiệt độ môi trường và độ rọi ánh sáng mặt trời từ tập dữ liệu thu thập. Đầu ra của mô hình dự đoán là công suất từ các tấm pin mặt trời. Kết quả của nghiên cứu sẽ giúp hệ thống điện hoạt động hiệu quả hơn, đảm bảo sự cân bằng giữa cung và cầu, từ đó khai thác hiệu quả tiềm năng của năng lượng mặt trời.

Từ khóa: Học máy, năng lượng mặt trời, dự báo, nguồn năng lượng tái tạo.