

Predicting the Potential of Renewable Solar Energy Based on Weather Data in Indonesia Using the Random Forest Method

Noviyanti. P^{1✉}, Maya Sari², Kusnanto³

^{1,2} Information Technology, Institut Shanti Bhuana, Indonesia

³ Elementary School Teacher Education, Institut Shanti Bhuana, Indonesia

✉ **Corresponding Author:** Noviyanti P (e-mail: noviyanti@shantibhuana.ac.id)

Article Information	ABSTRACT
Article History Received: July 28, 2025 Revised : August 09, 2025 Published: October 02, 2025	<p>Renewable energy plays a crucial role in reducing greenhouse gas (GHG) emissions. Excessive use of fossil fuels, such as coal, can produce GHG emissions that trigger extreme weather and global warming. Therefore, efforts to increase renewable energy utilization are necessary, in line with the Government Work Plan (RKP) target, which targets renewable energy contributions to reach 23% by 2025. This study aims to predict the potential for solar renewable energy in an area based on radiation, temperature, and rainfall variables. The method used is a supervised learning-based Random Forest. Weather data was obtained through the Open Meteo API, then processed by assigning weights to variables to produce output labels, which were then used in the classification process and model performance evaluation. The results showed that the Random Forest model produced an accuracy of 99.82%, with predictions of low/no potential energy being completely correct, medium energy potential experiencing only one error, and high energy potential also experiencing only one error. Based on these findings, the Random Forest method has proven effective in predicting solar power potential with high accuracy and is able to identify variables with the highest to lowest levels of importance.</p>
Keywords: <i>Renewable energy; Weather Data; Random Forest; Solar Energy; Indonesia.</i>	

INTRODUCTION

Renewable energy is energy provided by nature and can be used continuously and sustainably. Renewable energy is currently a government focus to address the global economy and mitigate greenhouse gas (GHG) emissions. According to the government's work plan (RKP), renewable energy potential is targeted to continue increasing, reaching approximately 23% by 2025 (Directorate General of Electricity, 2021).

Indonesia is one of the countries that has abundant natural resources (SDA) (Putra et al., 2023). However, the potential of existing natural resources has not been fully utilized. One potential that can be utilized through natural resources is renewable energy. Limited data can be a barrier to the underutilization of natural resources for renewable energy potential, so most still produce energy from fossil fuels. Fossil fuels, such as coal, oil, and natural gas, can produce energy but have a negative impact on the environment (Febriani Irma, 2024). Excessive use of fossil fuels can increase greenhouse gas (GHG) emissions, resulting in extreme climate change (Tjiwidjaja & Salima, 2023). In addition, based on research conducted by (Abdelsamiea & Abd El-Aal, 2025) conducted a study using the Random Forest method to estimate industrial energy efficiency. The study explained that energy efficiency in industry is inefficient and ineffective, resulting in increased CO₂ emissions. The study also explained that this condition can cause environmental

damage and increase emissions even further. This is a particular concern for the government and society to gradually shift from fossil fuels to renewable energy by utilizing inexhaustible and sustainable natural resources, such as solar energy, hydropower, geothermal energy, wind energy, and bioenergy.

Renewable energy is the primary solution to addressing the challenges of climate change and dependence on fossil fuels. Energy sources such as solar, wind, and hydropower have significant potential for development, but their utilization still faces various obstacles, such as uncertain weather conditions, optimal power plant locations, and energy production efficiency (Taneza & Firdaus, 2025).

In recent years, developments in Artificial Intelligence technology using Machine Learning methods have opened up new opportunities in renewable energy management and optimization. Machine learning is a subset of artificial intelligence developed based on data and experience, so it doesn't require explicit programming (P et al., 2025). In addition, machine learning can also be used to process big data, analyze, and make predictions (Horhoruw et al., 2024). The research will be conducted based on weather data from the Open Meteo API. The Open Meteo API provides free data access and can be obtained in real-time. The weather data can be accessed by referring to locations in Indonesia based on Province, Regency/City, Latitude, and Longitude. The weather data obtained consists of environmental variables such as sunlight intensity, wind speed, and rainfall levels, to produce a model that can be used to predict renewable energy potential more accurately. This prediction helps better planning in energy resource management and increases the efficiency of power plants. The initial data obtained in this study is raw data and is not equipped with output labels, so this study utilizes a weighted score based on determining and supporting variables used to predict the potential of solar renewable energy (Sahin et al., 2023) (Nguyen et al., 2025).

According to Katadata Green data, renewable energy usage increased from 2015 to 2021, reaching 4.9% from 2015 to 12.16% in 2021 (Annur, 2022). The government also has a target to increase the use of renewable energy by 2025 to reach up to 23% (Annur, 2022). One contribution to achieving this target is to support it with data and a recommendation model for renewable energy management based on the available potential portfolio. This research aims to help increase the use of renewable energy which is currently being developed so that Indonesia can gradually shift from fossil fuels such as coal, oil, and natural gas which produce GHG emissions when burned to produce energy to renewable energy generated from nature, such as solar energy and wind energy. To achieve this goal, an in-depth analysis of the required reference data is required. This research uses the Random Forest method as one of the learning methods.

The Random Forest method is a classification method that can improve the accuracy and performance of the model in carrying out classification (Mahmuda, 2024) (Suci Amaliah et al., 2022). This method also has reliability, such as being able to produce relatively low errors and being able to provide performance on model performance (Pratiwi & Nugroho, 2024) (Efendi & Zyen, 2024). The research will use the Random Forest method. Random Forest is used for classification and performance analysis in predicting potential based on location. The results of this study are expected to contribute to more efficient decision-making in renewable energy planning and implementation, as well as support efforts to transition to sustainable energy.

Previous research was conducted by (Yağmur et al., 2023) discusses predicting renewable energy production using machine learning, with a case study in Turkey. The study aims to develop a predictive model for renewable energy production based on data from the Turkish government. The data used is based on socioeconomic, environmental, and energy data from 1990 to 2020. The study used machine learning to generate the predictive model. The results showed a success rate of 92%. The difference between previous research and the current study lies in data analysis. The previous study used Turkish government data, while the researcher's study uses Indonesian government data, specifically direct satellite data that can obtain real-time weather data.

Previous research was conducted by (Malakouti, 2023) discusses the use of machine learning algorithms to predict renewable energy from turbine power as a substitute for fossil fuel

energy. The study recorded 850,660 data points obtained in real time from power plants from January 1, 2020, to December 31, 2020. Based on the research, machine learning can be used to generate accurate predictions of turbine power production. This previous research can serve as a reference and support for future research using machine learning algorithms.

Previous research was conducted by (Febtiawan et al., 2024) (Ardianto et al., 2022) (Villegas-Mier et al., 2022) related to solar energy production forecasting using the Random Forest method. The study obtained an accuracy of 96%. These results demonstrate the reliability of the Random Forest method as a machine learning method for forecasting solar energy production. This research supports future research utilizing the reliability of the Random Forest method.

Previous research was conducted by (Liu et al., 2025) discusses a comprehensive review of wind energy prediction based on machine learning. The study explains the advantages of machine learning over other traditional prediction methods. The study explains that machine learning algorithms refer to several factors, including data preprocessing, feature selection strategies, and model optimization techniques. These factors are covered in machine learning, one of which is the Random Forest method.

Previous research conducted by (Mallala et al., 2025) this paper discusses renewable energy forecasting based on public data obtained through the Kaggle platform. The study aims to predict future sustainable renewable energy trends. The study used the Random Forest method and achieved a high accuracy of 99%. The difference between previous research and the current study lies in the data used. This study utilizes satellite data, which is relatively rare and can be obtained in real time in Indonesia, allowing for trends in optimizing renewable energy in potential areas.

RESEARCH METHODS

This research falls into the quantitative category and focuses on renewable energy optimization using big data and machine learning approaches. The data generated from the big data approach will be processed using machine learning to produce an optimization model. The research flow can be seen in Figure 1.

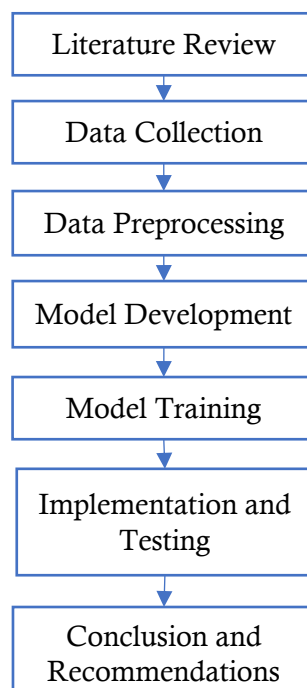


Figure 1. Research Flow

1. Literature Review
The initial process that has been carried out in this research is conducting a literature study by looking for references related to the hottest topics and identifying problems regarding the research topics raised by the researcher.
2. Data Collection
The collected data consists of Indonesian weather data filtered by district. The data used was obtained from Open-Meteo, which provides an open-source weather data API.
3. Data Preprocessing
Perform data preprocessing, such as normalization and removal of irrelevant data, as well as selecting data to be processed based on numeric data.
4. Model Development
Building a model using machine learning techniques. The model will be built using analysis, clustering, classification, and prediction techniques. This analysis will then produce a predictive model. The stage in achieving the output to produce a score-based predictive model uses a weighted assessment. The weights for the solar energy potential assessment are 70% solar radiation, 15% temperature, and 15% rainfall, with the opposite direction, because high rainfall reduces the effectiveness of the photovoltaic system. Furthermore, the data that has been categorized based on the assessment score is used for classification, prediction, and model performance.
5. Model Training
Train the model using the prepared training data and train the model using techniques that can optimize the model through a Machine Learning approach based on the Random Forest method.
6. Implementation and Testing
Once the model has been successfully trained and evaluated, it can be implemented and analyzed to identify renewable energy potential, thus obtaining a final conclusion that can be used to predict and optimize renewable energy production in several regions.
7. Conclusion and Recommendations
Make conclusions from the research results and provide suggestions for further development in the future.

The flowchart for solving problems in optimizing renewable energy potential using weighting and the Random Forest method is shown in Figure 2 below.

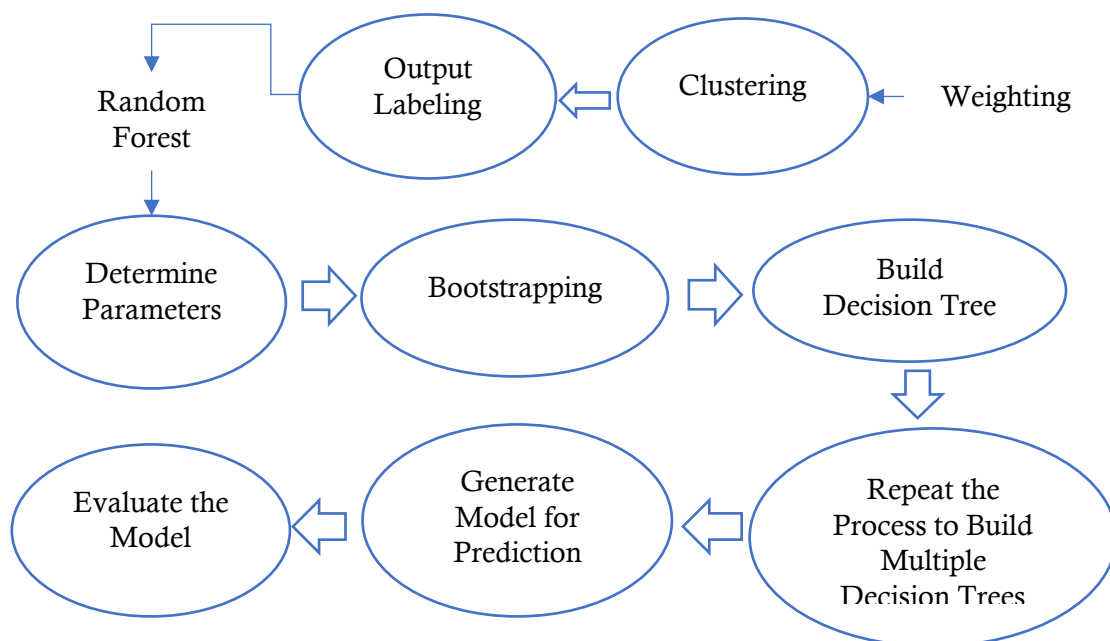


Figure 2. Flowchart of the Random Forest Method

Based on Figure 2, the study uses 2 processes, namely weighting and the Random Forest method with the following explanation (Sahin et al., 2023):

1. Weight assessment based on 70% Solar Radiation, 15% Temperature, and 15% Rainfall in the opposite direction.
2. Clustering into 3 categories, namely Low Potential/No Potential, Medium Potential, and High Potential.
3. Data obtained with labeled output, this data is used for processing using the Random Forest method.
4. Determine Parameters
At this stage, the researcher determines the basic parameters for the Random Forest model, such as:
 - Number of decision trees (n_estimators)
 - Maximum depth of each tree (max_depth)
 - Number of randomly selected features at each branch (max_features)
5. Build Decision Tree
The model builds a decision tree based on a random subset of data resulting from bootstrapping.
 - Each node in the tree will decide based on features such as: Temperature, Radiation, and Rainfall.
6. Repeat the Process to Build Multiple Decision Trees
The tree creation process is repeated a specified number of times (e.g. 100 trees).
Each tree:
 - Trained on different subsets of data
 - Using a randomly selected subset of features
 - The result is an ensemble of varying decision trees.
7. Generate Model for Prediction
After all the trees are formed, the model is used to predict the energy potential with the input given according to the variables used when processing the data.
8. Evaluate the Model
Evaluation is done with test data using metrics such as:
 - Check Accuracy
 - Classification Report
 - Confusion Metrics

The random forest method was chosen because of its reliability in determining features of importance in data, allowing it to be adapted to the research conducted by referring to the weighting of the main determining and supporting variables. This study used three variables solar radiation as the determining variable, and temperature and rainfall as supporting variables (Sahin et al., 2023).

RESULTS AND DISCUSSION

The data in this study consists of the variables Location, Temperature, Wind Speed, Rainfall, and Solar Radiation. These variables will be used to predict renewable energy potential, such as solar energy. The data used was obtained through Open-Meteo. Open-Meteo is an open-source API service that provides weather data throughout Indonesia. Data obtained through Open-Meteo does not include energy potential, so output determination or data labeling is required. The Random Forest method is a supervised learning method, so labeling renewable energy potential is necessary.

The output labeling is based on a weighted assessment with solar radiation at 70%, temperature at 15%, and rainfall at 15% in the opposite direction. The weighted scoring results have three categories: low/no energy potential, medium energy potential, and high energy potential. The data used in this study is 202,761 data obtained from 2022-2024. The 10 head data in this study are as follows:

	waktu	suhu	curah_hujan	kecepatan_angin	radiasi_gelombang	kabupaten_kota	provinsi	lintang	bujur	radiasi
0	2022-01-01	26.0	21.4	10.7	15.65	Malang (Kab)	Jawa Timur	-7.9008	112.5979	4.347570
1	2022-01-02	25.5	13.3	13.8	17.36	Malang (Kab)	Jawa Timur	-7.9008	112.5979	4.822608
2	2022-01-03	25.9	0.5	18.4	20.42	Malang (Kab)	Jawa Timur	-7.9008	112.5979	5.672676
3	2022-01-04	26.1	5.1	16.4	19.72	Malang (Kab)	Jawa Timur	-7.9008	112.5979	5.478216
4	2022-01-05	25.6	4.7	14.2	19.50	Malang (Kab)	Jawa Timur	-7.9008	112.5979	5.417100
5	2022-01-06	26.9	1.5	14.0	23.77	Malang (Kab)	Jawa Timur	-7.9008	112.5979	6.603306
6	2022-01-07	26.5	1.7	17.3	20.02	Malang (Kab)	Jawa Timur	-7.9008	112.5979	5.561556
7	2022-01-08	26.7	2.9	12.1	23.11	Malang (Kab)	Jawa Timur	-7.9008	112.5979	6.419958
8	2022-01-09	25.9	44.4	13.2	17.29	Malang (Kab)	Jawa Timur	-7.9008	112.5979	4.803162
9	2022-01-10	26.6	42.6	12.0	17.65	Malang (Kab)	Jawa Timur	-7.9008	112.5979	4.903170

Figure 3. Raw Data Head

Figure 3 is a raw data head displayed as many as 10 of 202,761 data. Data is obtained per day based on location. The locations used are 34 locations based on Provinces with a number of Regencies spread throughout Indonesia. Based on the data shown in Figure 3 that there is no output variable in determining renewable energy potential so that in this study using weighting for assessing solar energy potential, Solar radiation is 70%, Temperature is 15%, and Rainfall is 15% in the opposite direction, because high rainfall reduces the effectiveness of the photovoltaic system.

Weighting

The clustering process uses a weighted assessment based on the variables of temperature, rainfall, and solar radiation. The categories used are low/no energy potential, medium energy potential, and high energy potential.

Before weighting 202,761 data sets were collected, consisting of raw daily data for the past three years, from 2022 to 2024. This data was aggregated from daily to monthly data. This aggregation resulted in a total of 5,545 data sets. This data was used for the output labeling process.

	provinsi	kabupaten_kota	tahun_bulan	suhu	curah_hujan	radiasi
0	Aceh	Aceh Besar	2022-01	29.530645	253.0	5.213634
1	Aceh	Aceh Besar	2022-02	29.650000	450.6	4.891909
2	Aceh	Aceh Besar	2022-03	29.820968	422.7	4.947484
3	Aceh	Aceh Besar	2022-04	30.353333	329.4	5.424878
4	Aceh	Aceh Besar	2022-05	31.054839	249.1	5.168290
5	Aceh	Aceh Besar	2022-06	30.866667	250.8	5.115131
6	Aceh	Aceh Besar	2022-07	30.845161	241.9	4.957072
7	Aceh	Aceh Besar	2022-08	31.141935	246.1	5.054840
8	Aceh	Aceh Besar	2022-09	31.436667	242.5	5.312786
9	Aceh	Aceh Besar	2022-10	29.195161	719.6	4.380637

Figure 4. Head 10 Data with Labeled Columns

Figure 4 shows 10 data sets that were successfully aggregated from daily to monthly data based on location. The next step is to label the output based on the weighting assessment.

```

scaler = MinMaxScaler()
df_norm = pd.DataFrame(

scaler.fit_transform(df_bulanan[['temperature', 'rainfall', 'radiation'
]]),
columns=['temperature', 'rainfall', 'radiation']
)

```

```

df_norm['rainfall'] = 1 - df_norm['rainfall'] # reverse rainfall

# Solar score with weight
df_monthly['surya_score'] = (
    0.7 * df_norm['radiation'] +
    0.15 * df_norm['temperature'] +
    0.15 * df_norm['rainfall']
)

def potential_label(row):
    if row['radiation'] >= 6 and row['temperature'] <= 35
    and row['rainfall'] <= 300:
        return 'High Potential'
    elif row['radiation'] >= 5 and row['temperature'] <= 37
    and row['rainfall'] <= 400:
        return 'Medium Potential'
    else:
        return 'Low Potential/No Potential'

df['potential_label'] = df.apply(potential_label, axis=1)

```

Based on the labeling process using the code above, the labeling data is obtained with 3 categories: low energy potential/no potential, medium energy potential, and high energy potential. As explained in the previous Introduction chapter, the Random Forest method is a supervised method, so labeling is important to do at the beginning as part of data preprocessing. The labeled data will then be used to generate classification and model performance in predicting potential using the Random Forest method.

Random Forest Method

The Random Forest method is used to predict renewable energy potential based on labeled data. The data used is aggregated into monthly data sets, totaling 5,545 data points.

	provinsi	kabupaten_kota	tahun_bulan	suhu	curah_hujan	radiasi	kecepatan_angin	skor_surya	label_potensi
0	Aceh	Aceh Besar	2022-01	29.530645	253.0	5.213634	12.374194	0.563242	Potensi Sedang
1	Aceh	Aceh Besar	2022-02	29.650000	450.6	4.891909	11.953571	0.496464	Potensi Rendah/Tidak Potensial
2	Aceh	Aceh Besar	2022-03	29.820968	422.7	4.947484	9.845161	0.509220	Potensi Rendah/Tidak Potensial
3	Aceh	Aceh Besar	2022-04	30.353333	329.4	5.424878	9.996667	0.598758	Potensi Sedang
4	Aceh	Aceh Besar	2022-05	31.054839	249.1	5.168290	12.982258	0.569867	Potensi Sedang
5	Aceh	Aceh Besar	2022-06	30.866667	250.8	5.115131	13.070000	0.559433	Potensi Sedang
6	Aceh	Aceh Besar	2022-07	30.845161	241.9	4.957072	14.677419	0.534342	Potensi Rendah/Tidak Potensial
7	Aceh	Aceh Besar	2022-08	31.141935	246.1	5.054840	14.888710	0.552507	Potensi Sedang
8	Aceh	Aceh Besar	2022-09	31.436667	242.5	5.312786	15.470000	0.597224	Potensi Sedang
9	Aceh	Aceh Besar	2022-10	29.195161	719.6	4.380637	12.127419	0.388149	Potensi Rendah/Tidak Potensial

Figure 5. Head Data with Output Labeling

Figure 5 shows the labeled head data. This data will then be used to predict renewable energy potential. Based on the Random Forest algorithm, the first step is to determine the base model for the Random Forest model.

The Random Forest method first performs a classification process. Classification is achieved by training the model, evaluating it, and then predicting energy potential. Therefore, the initial step is to determine the variables used in the classification and the target variable.

```
features = ['temperature', 'radiation', 'rainfall']
target = 'potential_label'

X = df[features]
y = df[target]

X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)
```

Based on the code above, the variables used are temperature, radiation, and rainfall. The existing data is divided into two, namely training data and test data, with a percentage of 80% training data and 20% test data.

```
model = RandomForestClassifier(
    n_estimators=100,
    max_depth=10,
    max_features='sqrt',
    random_state=42
)
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

The code above is used to perform classification using the RandomForestClassifier function. Classification is based on the number of trees, maximum depth, and number of randomly selected features. A tree count of 100 is considered effective for most cases, and the depth is limited because unbounded search can lead to overfitting. Furthermore, the random state uses a common value of 42. After the decision tree is formed using the RandomForestClassifier function and the model is trained, the model can be used to make predictions. Based on the prediction results performed on the test data, an accuracy of 99.82% was obtained.

Akurasi: 0.9981965734896303

Classification Report:

	precision	recall	f1-score	support
Potensi Rendah/Tidak Potensial	1.00	1.00	1.00	520
Potensi Sedang	1.00	1.00	1.00	486
Potensi Tinggi	1.00	0.99	1.00	103
accuracy			1.00	1109
macro avg	1.00	1.00	1.00	1109
weighted avg	1.00	1.00	1.00	1109

Confusion Matrix:

```
[[520  0  0]
 [ 1 485  0]
 [ 0  1 102]]
```

Figure 6. Random Forest Model Evaluation

Based on Figure 6, it can be explained that the accuracy obtained was 99.82%. Previous research can also produce an accuracy of 99%, but the data used in the previous study was public data from Kaggle, while this study uses real-time and updated data. The results of the confusion matrix show that predictions in the low/no potential energy category can make large predictions, then the prediction of medium energy potential has only 1 error by predicting low/no potential energy potential, and prediction of high energy potential has 1 error also by predicting medium energy potential. The results of the model evaluation generally provide good accuracy with a low error rate, so the model can be used to make predictions on new data.

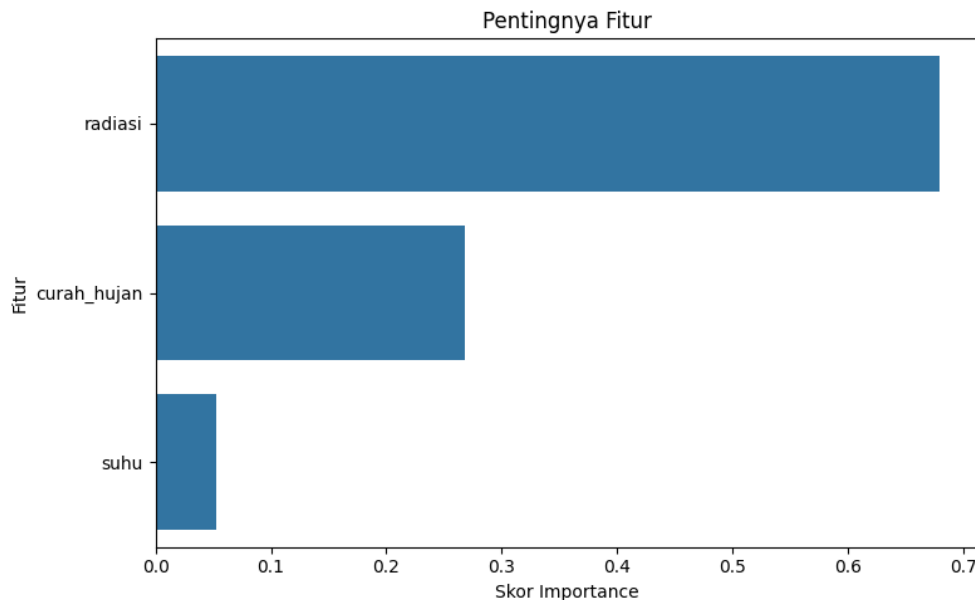


Figure 7. Feature Importance Determination Score

Figure 7 shows the feature importance levels based on the three variables used, specifically the radiation variable. Meanwhile, rainfall and temperature serve as supporting variables for optimally determining the potential for renewable solar energy. The next step to demonstrate the model's predictive capability is the code used to add new data and see the predicted energy potential.

```
new_data = pd.DataFrame([
    {'temperature':25,'radiation':4.3,'rainfall':300},
    {'temperature':25,'radiation':5.3,'rainfall':300},
    {'temperature':25,'radiation':7.3,'rainfall':300}
])
prediction = model.predict(new_data)

print("Energy potential category prediction:", prediction)
```

The code adds 1 new data with a temperature value of 25 degrees Celsius, radiation of 4.3 kwh/m², and the amount of rainfall per month of 300, so the results obtained based on the new data are:

```
Prediksi kategori potensi energi: ['Potensi Rendah/Tidak Potensial' 'Potensi Sedang' 'Potensi Tinggi']
```

Figure 8. New Data Prediction Results

Based on the prediction results in Figure 8, if referring to the training data, there are provisions based on the weighting of radiation values below 5 kwh/m², resulting in low/no potential energy

potential because the effective radiation value is above 5 kwh/m². Therefore, based on the new data tested, it has a suitability based on radiation values because radiation has the highest level of importance in the case of determining the potential for solar renewable energy. However, supporting variables still have weights that can support optimal decisions in the case of predicting the potential for solar renewable energy.

CONCLUSION

Based on the research results that have been conducted, the purpose of this study is to be able to produce a model to predict the potential of renewable solar energy using the Random Forest method that can be used with an accuracy percentage of 99.82%. The classification results can predict low/no potential energy correctly, medium energy potential with only 1 error, and high energy potential with only 1 error. The results of the study also answer the explanation related to previous research where the Random Forest method is reliable in classifying with the note that the data used must already have output labels because the Random Forest method is a supervised learning method. The Random Forest method also has reliability in producing patterns in determining variables that have a level of importance in classifying. Suggestions for further research include the addition of other variables, such as air humidity. Furthermore, it can be developed in combination with other methods, such as LSTM and SVM.

REFERENCES

- Abdelsamiea, A.T., & Abd El-Aal, M.F. (2025). Optimizing Energy Consumption Efficiency in Global Industrial Systems Using the Random Forest Algorithm. *International Journal of Energy Economics and Policy*, 15(3), 239–244. <https://doi.org/10.32479/ijeeep.18796>
- Annur, CM (2022). This is the Growth of Renewable Energy in Indonesia until 2021. Katadata Media Network. <https://databoks.katadata.co.id/datapublish/2022/11/23/ini-pertumbuhan-energi-terbarukan-di-indonesia-sampai-2021>
- Ardianto, Raharjo, AB, & Purwitasari, D. (2022). Random Forest Regression for Predicting Power Production of Solar Power Plants. *Briliant: Journal of Research and Conceptual Studies*, 7(4), 1058. <https://doi.org/10.28926/briliant.v7i4.1036>
- Directorate General of Electricity. (2021). Government Optimistic That 23% of New and Renewable Energy Will Be Achieved by 2025. Ministry of Energy and Mineral Resources. <https://www.esdm.go.id/id/berita-unit/direktorat-jenderal-ketenagalistrikan/pemerintah-optimistis-ebt-23-tahun-2025-tercapai>
- Efendi, MS, & Zyen, AK (2024). Application of the Random Forest Algorithm for Sales Prediction and Product Inventory Systems. 5(1), 12–20. <https://doi.org/10.30865/resolusi.v5i1.2149>
- Febriani Irma, M. (2024). High Temperature Rise Due to Increased Greenhouse Gas Emissions in Indonesia. *JSSIT: Journal of Science and Applied Science*, 2(1), 26–32. <https://doi.org/10.30631/jssit.v2i1.49>
- Febtiawan, EP, Syamsul, LA, Akbar, I., & Rachman, AS (2024). Forecasting Photovoltaic Energy Production Using the Random Forest Classification Algorithm. *Journal of Information System Research (JOSH)*, 5(4), 1053–1062. <https://doi.org/10.47065/josh.v5i4.5514>
- Horhoruw, LFM, Pratama, A., Asri, Y., & P, N. (2024). MACHINE LEARNING EXPLORATION WITH SCIKIT-LEARN: Machine Learning Strategy. *UWAIS Inspirasi Indonesia*.
- Liu, Z., Guo, H., Zhang, Y., & Zuo, Z. (2025). A Comprehensive Review of Wind Power Prediction Based on Machine Learning: Models, Applications, and Challenges. *Energies*, 18(2), 1–17. <https://doi.org/10.3390/en18020350>
- Mahmuda, S. (2024). Implementation of the Random Forest Method on YouTube Channel

- Content Categories. Jurnal Jendela Matematika, 2(01), 21–31.
<https://doi.org/10.57008/jjm.v2i01.633>
- Malakouti, S. M. (2023). Use machine learning algorithms to predict turbine power generation to replace renewable energy with fossil fuels. *Energy Exploration and Exploitation*, 41(2), 836–857. <https://doi.org/10.1177/01445987221138135>
- Mallala, B., Ahmed, AIU, Pamidi, SV, Faruque, MO, & M, RR (2025). Forecasting global sustainable energy from renewable sources using random forest algorithm. *Results in Engineering*, 25(November 2024), 103789. <https://doi.org/10.1016/j.rineng.2024.103789>
- Nguyen, HN, Tran, QT, Ngo, CT, Nguyen, DD, & Tran, VQ (2025). Solar energy prediction through machine learning models: A comparative analysis of regressor algorithms. *PLOS ONE*, 20(1), 1–23. <https://doi.org/10.1371/journal.pone.0315955>
- P, N., Gudiatto, C., Frigia, L., Horhoruw, M., & Asri, Y. (2025). *MACHINE LEARNING APPROACH: A Practical Guide to Big Data and Regression Analysis*. UWAIS Inspiration of Indonesia.
- Pratiwi, GE, & Nugroho, A. (2024). IMPLEMENTATION OF THE RANDOM FOREST METHOD FOR CLASSIFICATION OF BEST-SELLING SOAP PRODUCTS. 7(2), 531–540. <https://doi.org/10.37600/tekinkom.v7i2.1610>
- Putra, A., Veronica, D., & Bansa, Y.A. (2023). Revenue from Natural Resource Management and Its Impact on Non-Tax State Revenue in Indonesia. *Journal of Development*, 11(1), 29–37. <https://doi.org/10.53978/jd.v11i1.272>
- Sahin, G., Isik, G., & van Sark, W.G.J.M. (2023). Predictive modeling of PV solar power plant efficiency considering weather conditions: A comparative analysis of artificial neural networks and multiple linear regression. *Energy Reports*, 10(September), 2837–2849. <https://doi.org/10.1016/j.egyr.2023.09.097>
- Suci Amaliah, Nusrang, M., & Aswi, A. (2022). Application of the Random Forest Method for Classifying Coffee Drink Variants at the Konijiwa Bantaeng Coffee Shop. *VARIANSI: Journal of Statistics and Its Application on Teaching and Research*, 4(3), 121–127. <https://doi.org/10.35580/variansiunm31>
- Taneza, E., & Firdaus. (2025). SMART SYSTEM FOR MONITORING AND OPTIMIZING SOLAR POWER PLANTS PERFORMANCE. *Scientific Journal of Electrical Engineering*. <https://doi.org/10.14710/transmisi.27.1.20-32>
- Tjiwidjaja, H., & Salima, R. (2023). The Impact of Fossil Fuels on Climate Change and Green Energy-Based Solutions. 2(2), 166–172.
- Villegas-Mier, C.G., Rodriguez-Resendiz, J., Álvarez-Alvarado, J.M., Jiménez-Hernández, H., & Odry, Á. (2022). Optimized Random Forest for Solar Radiation Prediction Using Sunshine Hours. *Micromachines*, 13(9). <https://doi.org/10.3390/mi13091406>
- Yesğmur, A., Kayakuş, M., & Terzioğlu, M. (2023). Predicting renewable energy production by machine learning methods: The case of Turkey. *Environmental Progress & Sustainable Energy*, 42(3). <https://aiche.onlinelibrary.wiley.com/doi/10.1002/ep.14077>