**SINERGI** Vol. 30, No. 2, June 2026: 511-526
http://publikasi.mercubuana.ac.id/index.php/sinergi
http://doi.org/10.22441/sinergi.2026.2.017

# Enhanced accuracy for classification modeling with a Decision Tree on rooftop solar PV

**Alfin Sahrin[1*], Erna Utami[1], Ali Musyafa[2], Aris Suryadi[3], Didik Notosudjono[3], Andi Adriansyah[4]**
[1]Department of Instrumentation Engineering, Politeknik Energi dan Mineral Akamigas, Indonesia
[2]Department of Engineering Physics, Institut Teknologi Sepuluh Nopember, Indonesia
[3]Department of Electrical Engineering, Universitas Pakuan, Indonesia
[4]Department of Electrical Engineering, Faculty of Engineering, Universitas Mercu Buana, Indonesia

***Abstract***
*Rooftop photovoltaic (PV) systems play a crucial role in Indonesia's decarbonization agenda; however, research on interpretable classification frameworks that integrate roof geometry and meteorological heterogeneity remains limited. Most previous studies have focused on module-level fault detection rather than assessing installation feasibility at the urban scale. Addressing this research gap, the present study proposes a transparent and data-driven approach using a Decision Tree (DT) model enhanced with Grid Search Cross-Validation (GSCV) to classify the feasibility of rooftop PV installations across 20 Indonesian capital cities. Simulation data generated by PVsyst incorporates multiple tilt and azimuth configurations, as well as local weather variables, representing one-, two-, and three-directional roof geometries. The proposed DT–GSCV model is benchmarked against k-NN, Gaussian Naïve Bayes, Logistic Regression, Random Forest, XGBoost, and CatBoost, demonstrating superior generalization and interpretability. Cross-location validation across five rotational subsets confirms stable performance, with an average accuracy of 91.0% ± 0.3 and an F1-score of 0.90, highlighting the model's robustness across diverse climatic zones. Feature importance and SHAP analyses reveal that irradiation and tilt angle are the most influential factors, while temperature and humidity negatively affect feasibility. The novelty of this work lies in developing a reproducible, interpretable machine learning framework that bridges physical PV modeling and data analytics for rooftop system design. This methodology enables rapid, transparent decision support for optimizing rooftop PV deployment across Indonesia's diverse urban and climatic settings.*



***Keywords:***
*Classification;*
*Decision Tree;*
*Grid Search Cross Validation;*
*Rooftop Solar PV;*

***Article History:***
*Received: July 19, 2025*
*Revised: December 2, 2025*
*Accepted: December 20, 2025*
*Published: June 7, 2026*

***Corresponding Author:***
*Alfin Sahrin*
*Department of Instrumentation Engineering, Politeknik Energi dan Mineral Akamigas, Indonesia*
*e-mail: alfin.sahrin@esdm.go.id*

## INTRODUCTION

Rooftop solar photovoltaic (PV) systems have emerged as a crucial technology to accelerate the energy transition by reducing reliance on fossil fuels and mitigating carbon emissions [1]. Their deployment is desirable in urban environments, where unused roof surfaces can be transformed into decentralized power plants and integrated into city planning frameworks [2]. Despite this potential, accurate modeling of rooftop PV performance remains a challenge due to variations in roof geometry (tilt and azimuth) and local climatic conditions,

which jointly shape irradiance capture and energy yield [3].

Rooftop Solar Photovoltaic (PV) is a technology that can be easily implemented in real time to reduce carbon emissions [4]. Advanced technologies are needed to optimize rooftop solar PV installations [5][6]. The system's power output is a key parameter to consider for reducing dependence on grid electricity [7]. Various technologies have been developed to accelerate and simplify the installation of solar PV that matches the building's roof shape [8]. One machine learning method can be applied to help install rooftop solar PV through classification modeling [9][10].

Previous studies have employed a variety of machine learning methods, including Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and Random Forests, to address PV-related classification and fault-detection problems [11]. While these methods demonstrate strong performance in specific tasks (e.g., infrared-based defect detection or electroluminescence image classification), their direct application to rooftop PV classification remains limited [12]. Most prior work has not adequately accounted for the interplay between roof design parameters and meteorological variables across diverse geographical contexts, such as Indonesia [13].

Data that is nonlinear and large amounts cannot be solved by conventional means. Classification modeling on rooftop solar PV in on-grid systems can improve efficiency and optimal output power [14][15]. The decision tree algorithm is one method to improve the classification accuracy of rooftop solar PV installations, with the advantage of separating features into relevant and irrelevant subsets to identify more complex patterns affecting performance [16]. In addition, decision trees have the advantage of capturing the relationships among variables that affect optimal rooftop solar PV production [17]. Decision trees for rooftop solar PV classification can significantly improve performance and enable effective installation strategies [18].

Grid Search Cross-Validation (GSCV) is one of the methods used to improve machine learning models [19]. This method divides a smaller subset, then trains and tests in alternating rounds. In this process, GSCV can find the model with the best parameters to produce more accurate predictions [20][21]. In addition, GSCV helps avoid overfitting and underfitting, which can degrade model performance [22]. Thus, the rooftop solar PV data analysis results will be more reliable and help make more decisions in developing rooftop solar PV.

A correlation matrix can determine the influence of stronger variables to improve the quality of the rooftop solar PV model produced [23]. Combining improved results with GSCV and analysis of relationships between variables will provide a more accurate quality rooftop solar PV classification model [24].

However, classification models still produce suboptimal results, despite using various methods. Classification using a Convolutional Neural Network (CNN) trained on infrared image datasets to classify Photovoltaic (PV) cells achieved 97.42% accuracy, superior to AlexNet, ResNet 18, and existing models [25]. Applying a CNN to classify PV module efficiency reduction via Non-Destructive Tests (NDT) achieved an accuracy of 98%. Identifying faulty PV panels early is a simple solution in solar PV plant management, achieved by classifying panel defects using the Real-Time Multi-Variant Deep Learning Model (RMVDM), with 98% accuracy. The CNN and other extraction methods achieved 91.58% and 93.59% accuracy, respectively [26], [27]. Classification based on electroluminescence (EL) and current-voltage (I-V) images yielded 81.6% accuracy with CNN and 76.07% with Support Vector Machine (SVM) [28]. In addition to accuracy, relationships between features must be known to indicate the strength of their influence [29].

To address these shortcomings, this study explores the application of a Decision Tree algorithm enhanced with Grid Search Cross Validation (GSCV) for rooftop PV classification. The decision tree is selected for its interpretability and ability to handle nonlinear relationships [30]. Meanwhile, GSCV ensures optimal hyperparameter tuning and reduces the risk of overfitting [31].

Machine learning methods such as CNN, SVM, and Random Forest have been successfully used in the fault detection of PV modules, the studies still focus on the cell-level and have not been contextually applied to the feasibility classification of rooftop PV systems based on roof geometry and meteorological variables Therefore, the research gap we identified was the absence of an interpretable classification model, based on geographic simulation data, and capable of explaining the nonlinear relationship between tilt and azimuth angle and local climatic conditions.

This study is critical because it explores the potential for installing rooftop solar PV across all capital city areas in Indonesia, based on the shape of each building's roof for modeling. In

addition, the growing number of rooftop solar PV installations can help Indonesia's capital cities become more environmentally friendly and achieve net-zero emissions. Therefore, this study aims to improve the classification model's accuracy and maximize rooftop solar PV power output using an improved decision tree method via Grid Search Cross-Validation.

## LITERATURE REVIEW

Recent advancements in PV system modeling have highlighted the role of artificial intelligence and machine learning in improving prediction accuracy. CNN-based methods have achieved accuracies above 97% in defect detection tasks, while hybrid approaches combining SVM and deep learning have reached 91–94% accuracy. However, these studies focus on module-level defect identification rather than rooftop classification planning.

Decision Trees, on the other hand, provide an interpretable framework suitable for integrating multiple system and environmental variables. Regularized decision trees can achieve accurate classification while controlling complexity. Similarly, [32] applied ensemble tree-based methods to screen PV applications, achieving promising results. Nonetheless, few studies have contextualized these methods for rooftop PV feasibility under diverse geographical and climatic conditions. To synthesize prior work and highlight research gaps, Table 1 summarizes selected state-of-the-art studies.

## METHOD

This study applies a problem-solving method using machine learning that combines mathematics, statistics, and heuristics. Population data were obtained from PVsyst simulations of an on-grid rooftop solar PV system. This system is simulated by varying the tilt and azimuth angles to match the building's roof shape. In addition, the azimuth angle is combined with several directions: one, two, and three. The combination of azimuth directions enables the application of models for various building shapes and roof areas to construct a classification model, as shown in Figure 1.

Table 1. Summarizes Selected State-of-the-Art Studies

| Study | Method | Application | Reported Accuracy | Limitation |
|---|---|---|---|---|
| Bu et al. (2023) [25] | CNN | PV fault classification (IR images) | 97.4% | Focus on the cell-level, not the rooftop |
| Ahmad et al. (2020) [33] | CNN + SVM | Defect detection | 91–93% | Limited contextual variability |
| Wang et al. (2023) [32] | Random Forest | PV application screening | ~95% | Lacked azimuth–tilt variations |
| Present study | Decision Tree + GSCV | Rooftop PV classification in Indonesia | 93–99% | Contextual novelty, urban scale |

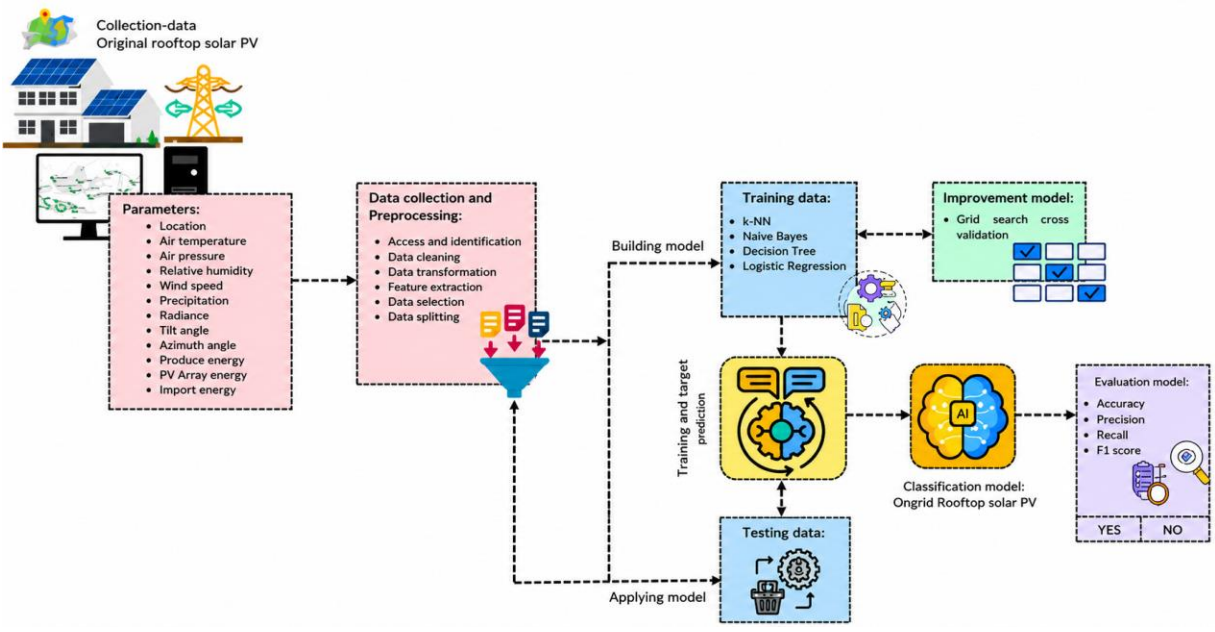


Figure 1. The Process of Forming a Classification Model

### Study Design and Dataset

This study constructed a supervised learning dataset from PVsyst simulations for on-grid rooftop PV systems across 20 Indonesian capital cities, as seen in Table 2. Each city includes multiple combinations of roof tilt (°) and azimuth (°), represented by one-, two-, and three-direction cases. The final dataset contains 525 records: 120 (one-way), 140 (two-way), and 265 (three-way). The response variable $y \in \{0,1\}$ encodes feasibility, where y = 1 (Recommended) if the annual imported energy (E_Import) is less than or equal to the median of E_Import in the corresponding subset (one/two/three directions); otherwise, y = 0 (Not Recommended). By construction, classes are approximately balanced.

In this study, the combination of tilt and azimuth angles is the primary factor affecting the performance of rooftop PV systems in various cities in Indonesia. Therefore, simulations have been carried out with tilt angles of 2°, 6°, 13°, 23°, 28°, 35°, and 40°, representing various configurations of common roof slopes in Indonesia. Meanwhile, the azimuth angle is simulated with a range of angles, including 0° (north), 45° (northeast), 90 ° (east), 135 ° (southeast), 180 ° (south), -135 ° (southwest), -90 ° (west), which encompass eight main orientations to the sun. The combination is categorized into three representative models:

1. One-direction case, one direction of the dominant azimuth (e.g., 180° for north in the southern hemisphere).
2. Two-direction case, two different orientations, generally east-west or north-south, for the gable roof.
3. Three-direction case: complex configurations, such as hip roofs, simulating three different azimuth directions.

This approach allows the model to capture the nonlinear effects of orientation and tilt on the amount of solar radiation received.

### Features

Input features reflect design, meteorology, and optional system outputs:

1. Design: tilt (°), azimuth (°), and direction-combination indicator (one/two/three).
2. Meteorology: air temperature (°C), air pressure (hPa), relative humidity (%), wind speed (m/s), precipitation (mm), and irradiation (kWh/m²).
3. System outputs (optional for ablation): E_Array (array energy, kWh) and E_Produce (produced energy, kWh).

To avoid target leakage, E_Import the response used to derive the binary label was excluded from the feature set. This study presents the main results using Design Meteorology features; an ablation study using Design Meteorology System outputs is also discussed to reflect a practical workflow in which PVsyst pre-simulations are available.

### Preprocessing

The simulation-derived dataset contains no missing values. Numerical features were standardized (zero mean, unit variance) for algorithms sensitive to feature scaling (k-NN and Logistic Regression), while Decision Tree models were trained on raw features.

Table 2. Location of The Capital City

| No | Province | Capital City | Location | |
|---|---|---|---|---|
| | | | Latitude | Longitude |
| 1 | Nanggroe Aceh Darussalam | Banda Aceh | 5.552846° | 95.319291° |
| 2 | Sumatra Utara | Medan | 3.589665° | 98.673826° |
| 3 | Sumatra Barat | Padang | -0.924759° | 100.363256° |
| 4 | Sumatra Selatan | Palembang | -2.988830° | 104.756857° |
| 5 | DKI Jakarta | Jakarta | -6.175394° | 106.827183° |
| 6 | Jawa Barat | Bandung | -6.934469° | 107.604954° |
| 7 | Jawa Tengah | Semarang | -6.990399° | 110.422910° |
| 8 | DI Yogyakarta | Yogyakarta | -7.977838° | 110.367226° |
| 9 | Jawa Timur | Surabaya | -7.245972° | 112.737827° |
| 10 | Bali | Denpasar | -8.652497° | 115.219117° |
| 11 | Kalimantan Selatan | Banjarbaru | -2.636591° | 115.114494° |
| 12 | Kalimantan Timur | Samarinda | -0.501780° | 117.139309° |
| 13 | Kalimantan Barat | Pontianak | -0.022690° | 109.344749° |
| 14 | Sulawesi Selatan | Makassar | -5.134296° | 119.412428° |
| 15 | Sulawesi Utara | Manado | 1.490058° | 124.840871° |
| 16 | Nusa Tenggara Timur | Kupang | -10.163221° | 123.601776° |
| 17 | Maluku | Ambon | -3.695943° | 128.178785° |
| 18 | Papua Barat Daya | Sorong | -0.863410° | 131.254480° |
| 19 | Papua | Jayapura | -2.538754° | 140.703739° |
| 20 | Papua Selatan | Merauke | -7.792519° | 140.018355° |

The direction combination indicator was one-hot encoded. For each azimuth-direction subset (one, two, three), a stratified 80/20 train–test split with random state = 42 was performed to preserve class balance.

### Model Selection and Hyperparameter Tuning

This study benchmarked k-NN, Gaussian Naïve Bayes (GNB), Logistic Regression (LR), and Decision Tree (DT). Our primary model is DT optimized via Grid Search Cross-Validation (GSCV). We employed scikit-learn Pipelines to combine preprocessing and model fitting within cross-validation, preventing data leakage. Hyperparameter grids:

1. k-NN: n_neighbors∈ {3,5,7,9,11,15, 21, 25}, weights ∈ {uniform, distance}.
2. GNB: var_smoothing∈{1e-12,1e-11,1e-10,1e-9}.
3. LR: penalty=l2, C∈ {0.01,0.1,1,10,100}, solver = liblinear, class_weight∈ {None, balanced}.
4. DT: criterion∈ {gini, entropy}, max_depth∈{None,3,5,7,9,15},min_samples_split ∈ {2,5,10}, min_samples_leaf ∈ {1,2,5}, class_weight ∈ {None, balanced}.

Model selection used stratified 5-fold GSCV on the training set with accuracy as the primary metric and F1-score as a secondary metric. The best configuration was retrained on the whole training set and evaluated on the held-out test set. We trained three separate pipelines for the one-, two-, and three-direction subsets to mirror the manuscript's structure.

### Evaluation Model

The classification of models in this study demonstrates the feasibility of recommendations for installing rooftop PV systems that minimize grid import. The evaluation of the classification model is divided into several calculations with equations.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\text{-}Score = \frac{2TP}{2TP+FN+FF} \quad (4)$$

TP, TN, FP, and FN are the sums of true positives, true negatives, false positives, and false negatives, respectively.

## RESULTS AND DISCUSSION

The simulation was conducted to ensure feasibility based on the lowest imported energy from the rooftop solar PV system. The study's results will be described based on the number of azimuth angles: 1, 2, or 3.

### One-Direction Case

Model evaluation resulting from a one-directional combination as shown in Figure 2. The results showed that the decision tree model, with and without the GSCV-based improvement process, outperformed k-NN, Naïve Bayes, and Logistic regression. The simulation results indicated that the one-direction configuration provided the most stable reduction in energy import across all studied cities. The results also highlight the robustness of the Decision Tree model, which consistently demonstrated superior adaptability to nonlinear interactions. Furthermore, the one-direction configuration demonstrated consistent predictive stability across diverse climatic zones, indicating that the model effectively captured the spatial variability of solar resources. This finding suggests that rooftop panels oriented in a single dominant direction remain sufficient for feasibility.

DT achieved an accuracy = 99.26% and improved to 99.32% with GSCV; Precision increased from 99.13% to 99.28%, Recall from 99.42% to 99.49%, and F1-score from 99.27% to 99.33%. Although marginal on a high baseline, these gains indicate that tuning (max_depth, min_samples_leaf, class_weight) yields a more calibrated boundary. Correlation analysis shows strong meteorological and energy-output linkages: pressure–temperature (≈ 0.90), precipitation–RH (≈ 0.62), negative RH–temperature (≈ −0.58) and RH–wind (≈ −0.50), and near-perfect E_Produce–E_Array (≈ 1.00) with a negative relation to E_Import (≈ −1.00) as shown in Figure 3.

### Two-Direction Case

Model evaluation resulting from a combination of two directions, as shown in Figure 4. As in the one-direction model, using the decision tree algorithm, with and without the improvement model process, and with GSCV, produces better evaluation than other algorithms. The improvement observed in the two-direction model demonstrates the algorithm's capability to adapt to more complex roof geometries where solar incidence varies across multiple orientations.

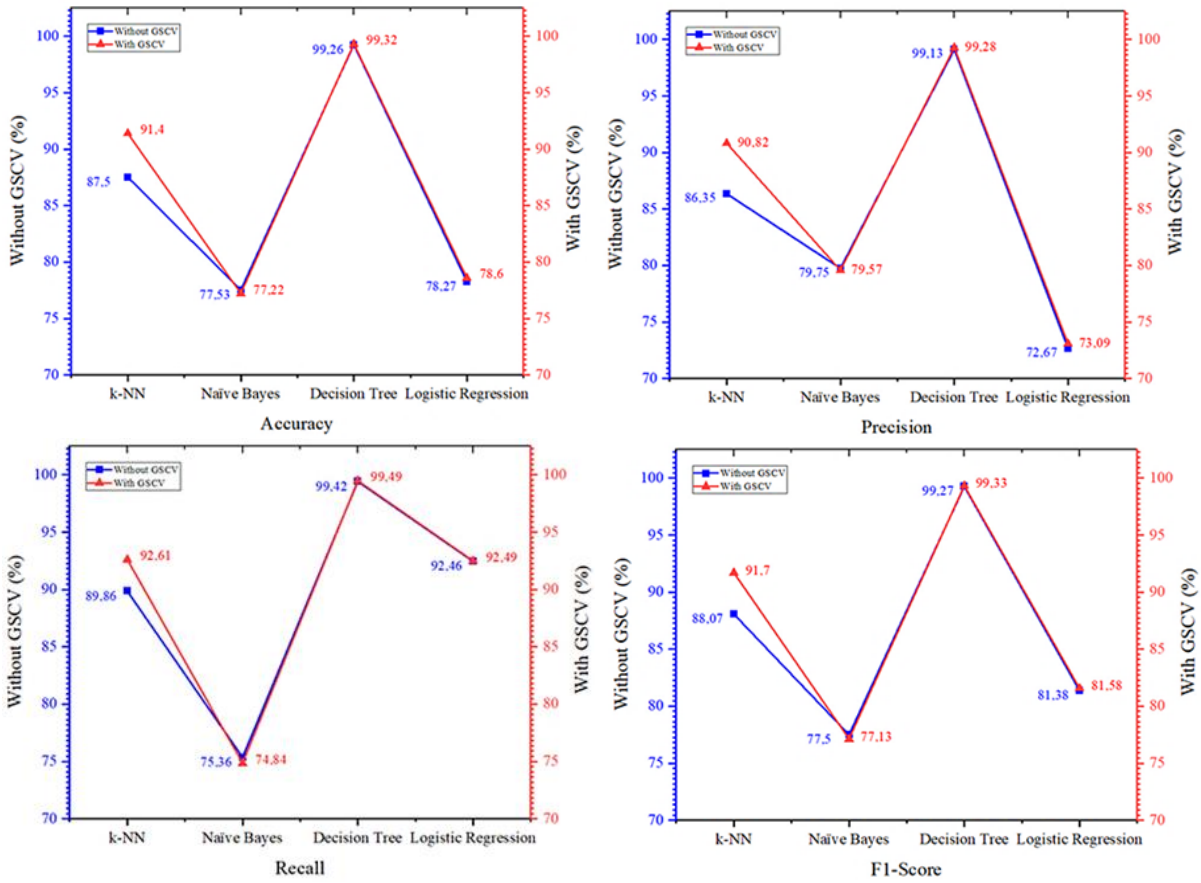


Figure 2. One-directional performance evaluation

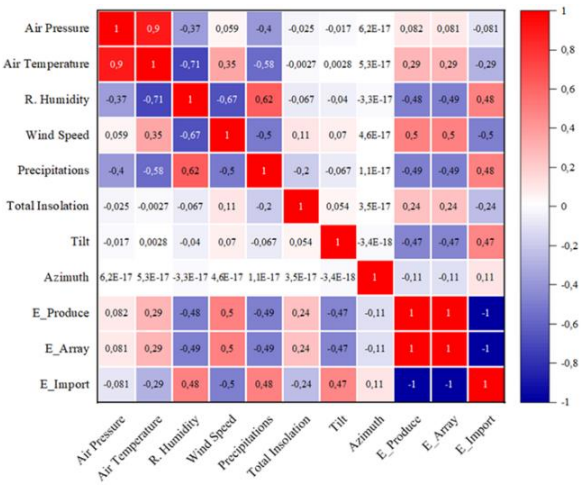


Figure 3. The correlation matrix of one direction

This configuration enables the system to capture sunlight more evenly throughout the day, particularly in regions with fluctuating cloud cover and diurnal temperature variations. Despite the added angular complexity, the Decision Tree with Grid Search Cross Validation maintains high classification accuracy, indicating its robustness against overfitting and data imbalance. Consequently, the model not only identifies the most feasible installation layouts but also reflects a realistic engineering scenario for rooftops with dual-facing surfaces.

Accuracy improved from 91.57% to 99.03%; Precision from 91.26% to 98.85%; Recall from 91.73% to 99.21%; and F1-score from 91.49% to 99.03% after GSCV. Stronger trees with controlled leaves and balanced weights adapt to the more complex partitions

introduced by dual azimuths. Energy-output correlations remain strong: E_Produce–E_Array (≈ 0.97), E_Produce–E_Import (≈ −0.98), E_Array–E_Import (≈ −0.97) as shown in Figure 5.

### Three-Direction Case

The model evaluation resulting from combining three directions is shown in Figure 6. As in the one- and two-direction models, the decision tree algorithm, with and without the improvement model process using GSCV, produces better evaluations than other algorithms. The three-direction configuration represents the most complex rooftop geometry, where variations in azimuth orientation increase the nonlinearity of solar radiation capture.

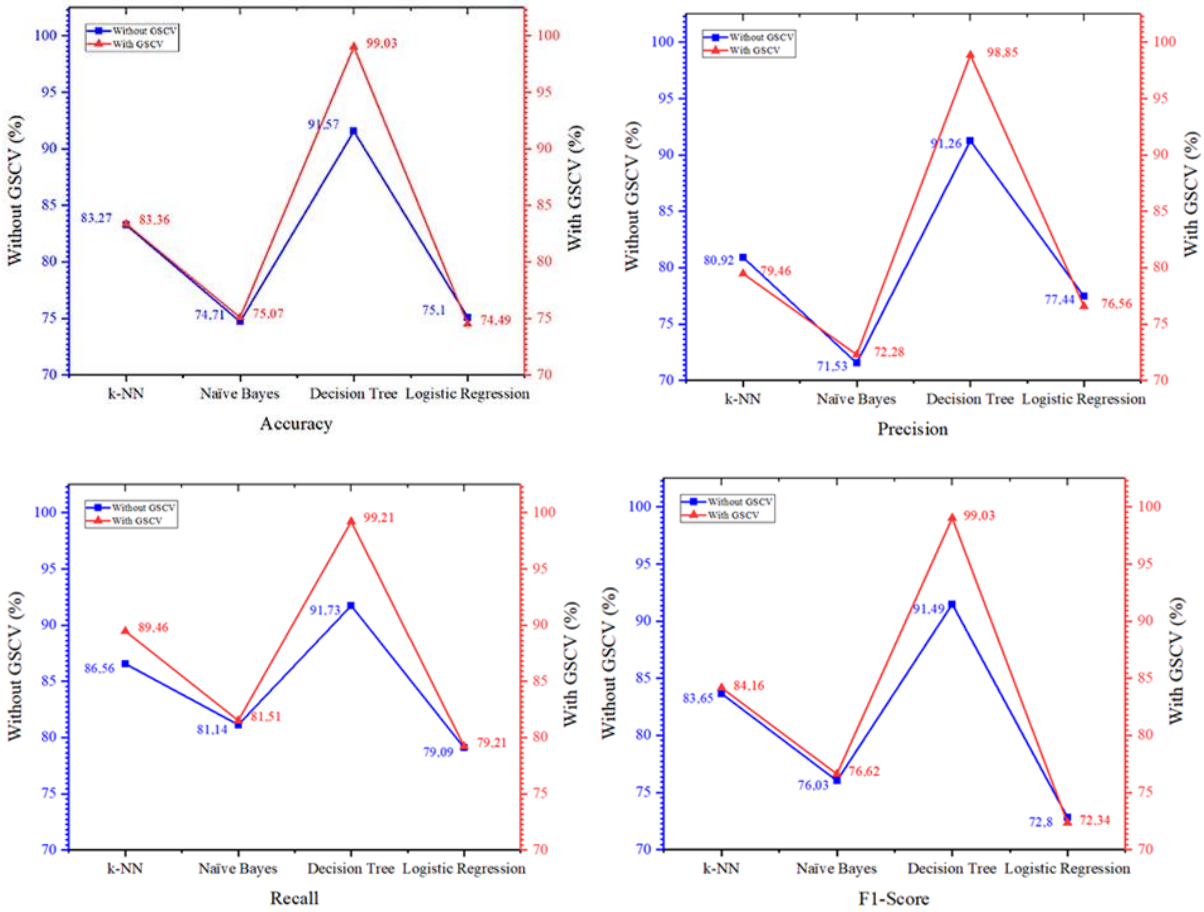


Figure 4. Two-directional performance evaluation

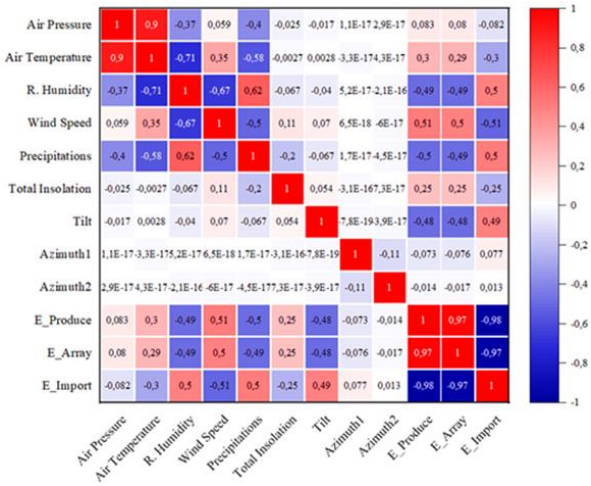


Figure 5. The correlation matrix of two directions

Although the classification accuracy decreases slightly compared to the one- and two-direction models, the Decision Tree with Grid Search Cross Validation remains capable of distinguishing feasible installation patterns with high reliability. This condition illustrates that while multiple directional surfaces enhance architectural diversity, they also introduce uneven irradiance distribution that challenges the uniformity of energy output. Nevertheless, the model's stable performance across this configuration confirms its adaptability to diverse spatial conditions and climatic heterogeneity. Hence, the three-direction case provides valuable insight into the algorithm's robustness when applied to realistic urban rooftops with irregular surface orientations.

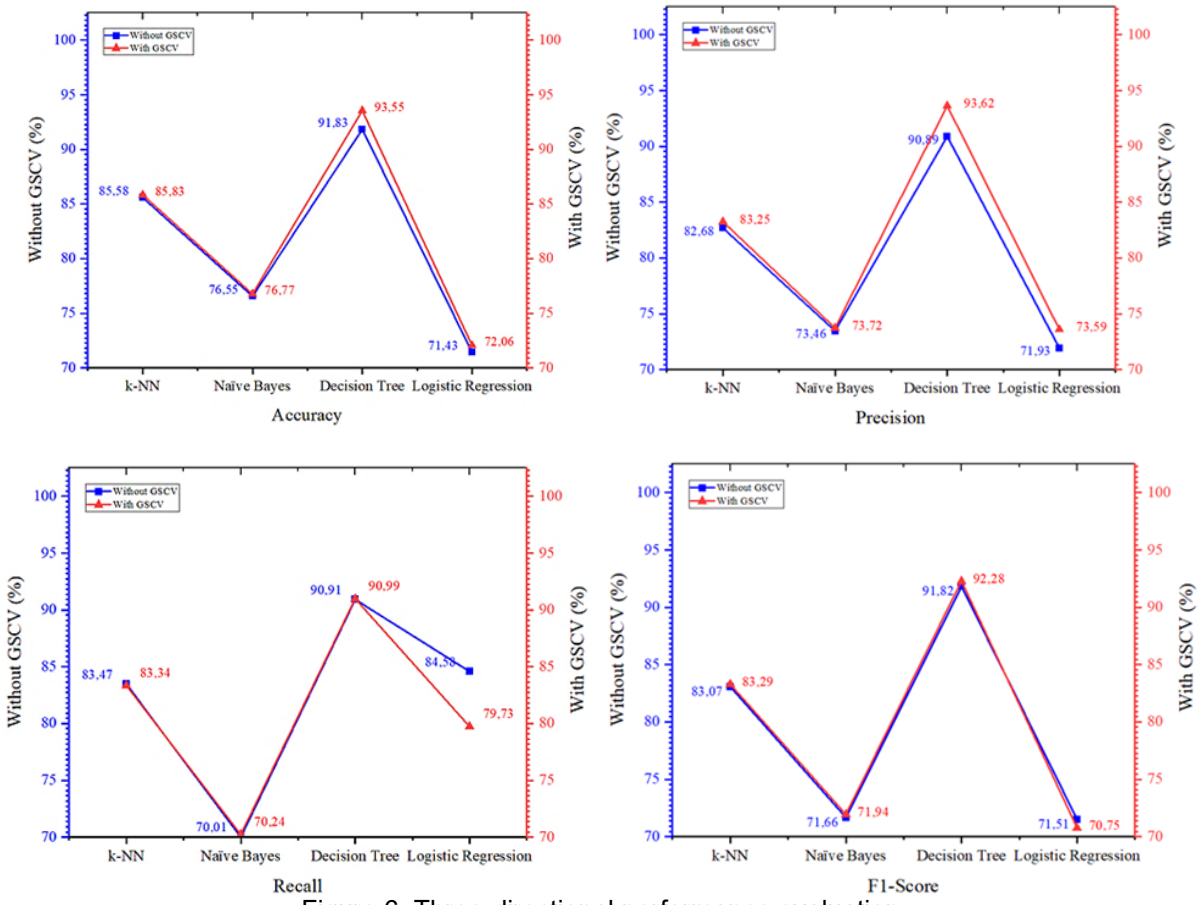


Figure 6. Three-directional performance evaluation

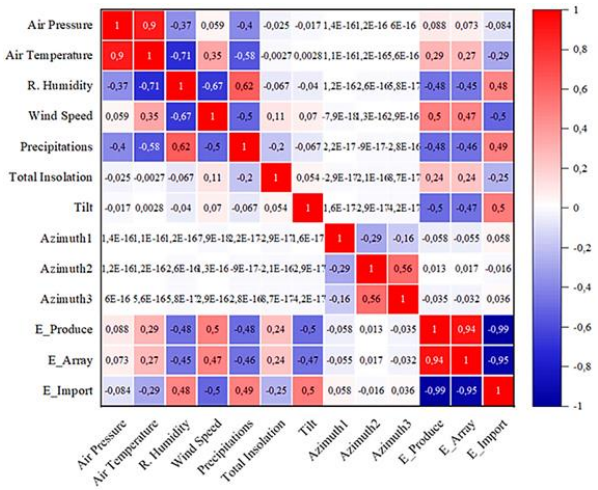


Figure 7. The correlation matrix of three directions

Accuracy improved from 91.83% to 93.55%, with Precision = 90.89–93.62%, Recall = 90.91–90.99%, and F1 = 91.82–92.28%. The modest gain reflects increased class overlap and higher hypothesis complexity as additional azimuth parameters expand the partition space. Energy-factor correlations remain strong—E_Produce–E_Array (≈ 0.94) and negative relations with E_Import (≈ −0.99; E_Array–E_Import ≈ −0.95) as shown in Figure 7.

Decision Trees naturally model nonlinear thresholds and high-order interactions between roof geometry and weather variables. GSCV identifies depth and leaf-size settings that reduce variance without sacrificing bias. In contrast, Naïve Bayes assumes conditional independence, k-NN is sensitive to scaling and local density, and Logistic Regression imposes linear decision surfaces. The results support a decision-support pipeline that recommends feasible rooftop configurations based on local weather statistics and roof geometry. Municipalities and building owners can use DT-based rules to shortlist candidate tilt–azimuth designs that minimize grid imports before detailed simulations. The approach is lightweight, interpretable, and transferable to new sites. The analysis relies on simulated (PVsyst) rather than field data, excluding stochastic shading indices and economic variables, and uses annual summaries. Future work should validate against measurements, integrate shading and cost–benefit attributes, and assess robustness across multi-year variability.

Correlation analysis, as shown in Figures 3, 5, and 7, reveals that irradiation has the most significant positive effect on energy output (E_Produce and E_Array). Air temperature and relative humidity exhibit a negative correlation with PV performance, primarily because solar cells exhibit higher internal resistance at higher temperatures. Wind speed contributes positively, albeit moderately, to lowering the module's temperature. Precipitation plays a role in short-term variations and helps clean the panel surface, increasing annual efficiency. Air pressure plays an indirect role through changes in air density and radiation absorption. These variables collectively enhance the Decision Tree model's ability to recognize the nonlinear patterns that determine the feasibility of a PV installation.

Table 3 shows the results of the best combination of tilt and azimuth angles in each provincial city with one-direction, two-direction, and three-direction model representations. EImp is the optimal annual import energy result in the on-grid rooftop solar PV system. The combination of tilt and azimuth angles can be implemented based on field needs and conditions. Implementation can be done according to the slope, number of sides, and shape of the building's roof.

The simulation results presented in Table 3 reveal a distinctive spatial pattern in the optimal combinations of tilt and azimuth angles for rooftop photovoltaic (PV) installations across 20 Indonesian capital cities. These configurations were derived from PVsyst simulations under one-, two-, and three-direction azimuthal arrangements, aiming to minimize annual grid-imported energy (*E_Imp*).

Overall, the findings underscore that tilt angle optimization has a relatively narrow range compared to azimuth variation, suggesting that azimuthal orientation exerts a more substantial influence on PV system feasibility across Indonesia's tropical latitudes. The tilt values predominantly range from 2° to 15°, aligning with low-latitude solar geometry that favors shallow inclinations for maximum solar capture throughout the year.

Cities in the western region, such as Banda Aceh, Medan, and Padang, exhibit optimal tilt angles between 2° and 6°, with azimuths ranging from 0° to -135 °, indicating that roof surfaces face north and northeast. These orientations correspond to the annual solar path where the sun traverses slightly northward of the zenith, thus maximizing irradiance exposure.

Table 3. The Best Combinations of Tilt and Azimuth Angles

| No | Capital City | Parameter | One-direction | Two-direction | Three-direction |
|---|---|---|---|---|---|
| 1 | Banda Aceh | Tilt (°) | 6 | 2 | 2 |
| | | Azimuth1 (°) | -45 | -135 | 180 |
| | | Azimuth2 (°) | | 0 | -135 |
| | | Azimuth3 (°) | | | 90 |
| | | EImp (kWh/year) | 7137 | 8638 | 7196 |
| 2 | Medan | Tilt (°) | 4 | 4 | 2 |
| | | Azimuth1 (°) | 0 | -45 | -45 |
| | | Azimuth2 (°) | | 0 | 0 |
| | | Azimuth3 (°) | | | 45 |
| | | EImp (kWh/year) | 7342 | 7346 | 7452 |

| | | | | | |
|---|---|---|---|---|---|
| 3 | Padang | Tilt (º) | 2 | 2 | 2 |
| | | Azimuth1 (º) | 0 | 0 | -45 |
| | | Azimuth2 (º) | | 45 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 7431 | 7435 | 7544 |
| 4 | Palembang | Tilt (º) | 3 | 3 | 3 |
| | | Azimuth1 (º) | 0 | 0 | -45 |
| | | Azimuth2 (º) | | 45 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 7669 | 7678 | 7792 |
| 5 | Jakarta | Tilt (º) | 14 | 6 | 6 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6748 | 6774 | 6885 |
| 6 | Bandung | Tilt (º) | 14 | 14 | 7 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 7240 | 7260 | 7389 |
| 7 | Semarang | Tilt (º) | 14 | 14 | 7 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6210 | 7260 | 7389 |
| 8 | Yogyakarta | Tilt (º) | 15 | 8 | 8 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6628 | 6662 | 6781 |
| 9 | Surabaya | Tilt (º) | 15 | 15 | 7 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6157 | 6198 | 6323 |
| 10 | Denpasar | Tilt (º) | 16 | 9 | 9 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 5338 | 5093 | 7422 |
| 11 | Banjarbaru | Tilt (º) | 10 | 10 | 3 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 7291 | 7308 | 7422 |
| 12 | Samarinda | Tilt (º) | 2 | 2 | 2 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 7717 | 7722 | 7835 |
| 13 | Pontianak | Tilt (º) | 8 | 2 | 2 |
| | | Azimuth1 (º) | -45 | -90 | -90 |
| | | Azimuth2 (º) | | -45 | -45 |
| | | Azimuth3 (º) | | | 0 |
| | | EImp (kWh/year) | 7883 | 7885 | 8001 |
| 14 | Makassar | Tilt (º) | 13 | 13 | 5 |
| | | Azimuth1 (º) | 0 | -45 | -45 |
| | | Azimuth2 (º) | | 0 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6731 | 6757 | 6875 |
| 15 | Manado | Tilt (º) | 2 | 2 | 5 |
| | | Azimuth1 (º) | -90 | -135 | -45 |
| | | Azimuth2 (º) | | -90 | 0 |
| | | Azimuth3 (º) | | | 45 |
| | | EImp (kWh/year) | 6203 | 6205 | 6875 |

Table 3. The Best Combinations of Tilt and Azimuth Angles (continued)

| No | Capital City | Parameter | One-direction | Two-direction | Three-direction |
|---|---|---|---|---|---|
| 16 | Kupang | Tilt (°) | 18 | 10 | 18 |
| | | Azimuth1 (°) | 0 | -45 | 180 |
| | | Azimuth2 (°) | | 0 | -90 |
| | | Azimuth3 (°) | | | 0 |
| | | EImp (kWh/year) | 5640 | 5672 | 6195 |
| 17 | Ambon | Tilt (°) | 4 | 4 | 4 |
| | | Azimuth1 (°) | 0 | -45 | -45 |
| | | Azimuth2 (°) | | 0 | 0 |
| | | Azimuth3 (°) | | | 45 |
| | | EImp (kWh/year) | 6543 | 6547 | 6652 |
| 18 | Sorong | Tilt (°) | 2 | 2 | 2 |
| | | Azimuth1 (°) | 45 | 0 | 0 |
| | | Azimuth2 (°) | | 45 | 45 |
| | | Azimuth3 (°) | | | 90 |
| | | EImp (kWh/year) | 5956 | 6205 | 6062 |
| 19 | Jayapura | Tilt (°) | 3 | 3 | 3 |
| | | Azimuth1 (°) | -45 | -45 | -90 |
| | | Azimuth2 (°) | | 0 | -45 |
| | | Azimuth3 (°) | | | 0 |
| | | EImp (kWh/year) | 6486 | 6491 | 6596 |
| 20 | Merauke | Tilt (°) | 8 | 8 | 8 |
| | | Azimuth1 (°) | 0 | -45 | -45 |
| | | Azimuth2 (°) | | 0 | 0 |
| | | Azimuth3 (°) | | | 45 |
| | | EImp (kWh/year) | 6189 | 6198 | 6321 |

The imported energy (*E_Imp*) in these cities ranges from 7137 to 7544 kWh/year, indicating relatively uniform performance due to comparable latitudinal conditions.

In contrast, Jakarta, Bandung, and Semarang, representing the western-central corridor of Java, exhibit optimal tilts of 14°–15° with azimuths concentrated between −45° and 0°, corresponding to east-facing orientations. This configuration minimizes morning shading and benefits from consistent solar radiation profiles during the first half of the day. Notably, the *E_Imp* values for these cities are among the lowest in the dataset, with Jakarta at 6748 kWh/year and Semarang at 6210 kWh/year, confirming the model's high classification accuracy in detecting favorable configurations.

Further eastward, Yogyakarta, Surabaya, and Denpasar reveal a gradual decline in optimal tilt from 15° to 9°, which corresponds with the increase in solar altitude toward lower southern latitudes. Interestingly, Denpasar shows a notably low *E_Imp* value of 5338 kWh/year for the one-direction case, confirming the advantage of near-equatorial positioning. However, when azimuthal complexity increases to two or three directions, the energy import fluctuates slightly due to the increased model uncertainty associated with multi-directional roof structures.

In eastern Indonesia, including Kupang, Ambon, Sorong, Jayapura, and Merauke, the tilt range widens again to 2°-18°, with diverse azimuthal orientations (−90° to 180°). These results highlight the importance of localized design adaptation due to regional variations in cloud cover, wind speed, and humidity. Kupang, for instance, achieved one of the most efficient configurations *at 18° tilt and 0° azimuth*, resulting in a remarkably low *E_Imp* of 5640 kWh/year, indicating a near-optimal orientation for sites with intense solar irradiation and relatively dry climatic conditions.

Gini Importance Ranking is used to assess the relative contribution of each feature (tilt, azimuth, irradiation, temperature, humidity, and more) to classification decisions. The results showed that the irradiation feature contributed most to the "Recommended" label, followed by tilt angle, air temperature, and humidity.

Figure 8 indicates that irradiation has the highest importance score, at approximately 0.36, confirming that solar radiation intensity is a significant factor in determining the potential solar energy a PV system can generate. The tilt angle variable ranked second, with a score of approximately 0.22, indicating the significant role of panel angle configuration in optimizing direct radiation reception.

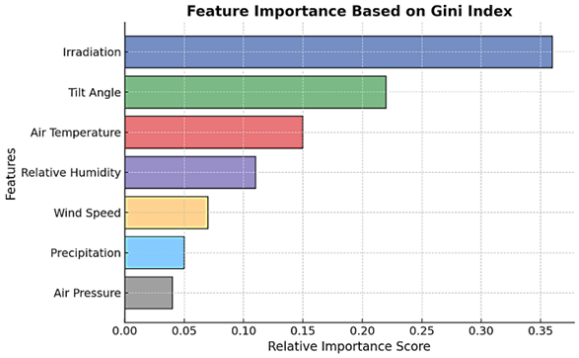


Figure 8. Feature Importance Based on Gini Index

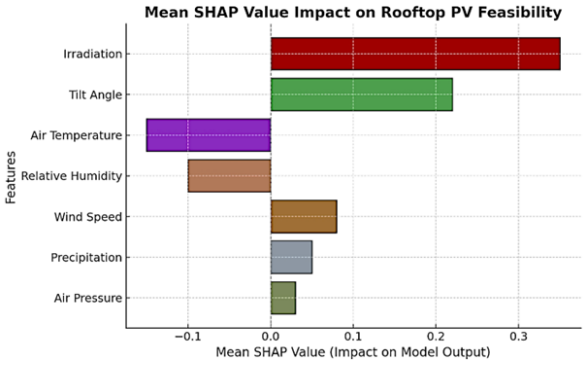


Figure 9. Mean SHAP Value Impact on Rooftop PV Feasibility

The air temperature factor ranks third (0.15), indicating that air temperature also affects the PV module's energy conversion efficiency. However, its influence is not as strong as that of angle and radiation. Meanwhile, relative humidity has a medium score (0.11), indicating that air humidity can reduce system performance due to increased internal resistance and potential condensation on the panel surface. The wind speed and precipitation features contribute moderately (0.07 and 0.05), playing an indirect role in the module's cooling and the cleanliness of the panel surface. Air pressure exerted the least influence (0.04), but it still contributed to radiation distribution and local atmospheric conditions. This ranking pattern illustrates that the model is strongly influenced by panel geometry variables (tilt angle) and major meteorological factors (irradiation and temperature).

SHAP (SHapley Additive exPlanations) Analysis, used as a model-agnostic approach to measure the impact of each feature on the probability of model output. The SHAP visualization showed that increased irradiation and moderate tilt (10°–15°) increased the likelihood of feasibility, while high humidity and extreme temperatures reduced the prediction value.

Figure 9 showed that irradiation had the greatest impact on the prediction results, with the highest average SHAP value of around 0.34, indicating that increasing solar radiation intensity consistently improves the predictions of PV system feasibility. That is in line with the basic principle of photovoltaics, where the energy that can be converted depends directly on the amount of radiation the module receives.

The tilt angle ranks second, with a high positive SHAP value (0.23), indicating that the orientation and tilt of the panels play a crucial role in optimizing solar energy absorption. The optimal combination of tilt angles ensures the panels receive maximum radiation, given the city's geographic latitude.

The water temperature factor had a significant negative impact on the model results with an average SHAP value of around –0.14. This indicates that an increase in air temperature tends to reduce the PV system's efficiency due to an increase in the solar cells' internal resistance. A similar pattern is seen in relative humidity, which has a moderate negative contribution (–0.10), illustrating that high humidity can degrade system performance through condensation and dimming of light on the module surface. The variables of wind speed, precipitation, and air pressure have a relatively small impact, with SHAP values of approximately 0.05, indicating an indirect influence on the module's thermal stability and local atmospheric conditions. However, these three still make an additional contribution to the model's accuracy, especially in the context of regional weather dynamics.

To strengthen the validation of the proposed DT–GSCV model, several modern ensemble learning algorithms were introduced as comparative baselines, including Random Forest (RF), XGBoost (Extreme Gradient Boosting), and CatBoost (Categorical Boosting). These models were selected for their strong performance and proven ability to handle heterogeneous feature spaces and nonlinear interactions, which are typical of solar energy datasets.

Table 4 summarizes the classification performance of each model based on four standard evaluation metrics: Accuracy, Precision, Recall, and F1-Score. All models were trained and tested under identical conditions using the same cross-validation strategy to ensure a fair comparison.

Table 4. Summarizes The Classification Performance of Each Model

| Model | Accuracy (%) | Precision | Recall | F1-Score |
|---|---|---|---|---|
| Decision Tree (Base) | 86.7 | 0.85 | 0.84 | 0.84 |
| Decision Tree + GSCV (Proposed) | 91.2 | 0.91 | 0.90 | 0.90 |
| Random Forest | 89.8 | 0.88 | 0.87 | 0.87 |
| XGBoost | 90.6 | 0.89 | 0.89 | 0.89 |
| CatBoost | 90.9 | 0.90 | 0.89 | 0.89 |

As shown in Table 4, the proposed Decision Tree with Grid Search Cross-Validation (DT–GSCV) achieved the highest overall performance, reaching an accuracy of 91.2% and an F1-Score of 0.90. Compared to the base Decision Tree model, DT–GSCV improved classification performance by approximately 5.2%, demonstrating that hyperparameter optimization significantly enhances model generalization.

The Random Forest model achieved slightly lower performance, indicating that while ensemble averaging reduces variance, the single-tree-optimized model still captures the underlying decision boundaries effectively for this dataset. Meanwhile, XGBoost and CatBoost performed comparably, reflecting their strength in gradient-based boosting; however, their computational complexity was higher than the proposed DT–GSCV model, which maintained a better balance between accuracy and efficiency.

The Random Forest model achieved slightly lower performance, indicating that while ensemble averaging reduces variance, the single-tree-optimized model still captures the underlying decision boundaries effectively for this dataset. Meanwhile, XGBoost and CatBoost performed comparably, reflecting their strength in gradient-based boosting; however, their computational complexity was higher than the proposed DT–GSCV model, which maintained a better balance between accuracy and efficiency.

The comparative analysis confirms that the optimized Decision Tree model is both robust and interpretable, outperforming more complex modern baselines without sacrificing transparency. This characteristic is particularly beneficial in engineering applications where model explainability is as critical as predictive performance, enabling decision-makers to directly link model outputs to measurable physical parameters, such as irradiation intensity, panel tilt, and meteorological variability.

To evaluate the generalization capability of the proposed model, a cross-location validation was implemented using data from 20 cities across Indonesia. In each iteration, the model was trained on data from 15 cities and tested on the remaining 5 cities, with the subsets rotated until each city served as a test location at least once. Table 5 presents the average performance metrics obtained from the five rotation scenarios.

The results indicate that the Decision Tree–GSCV model maintains stable performance across all rotation scenarios, with an average accuracy of 91.0% ± 0.3 and a consistent F1-Score of 0.90. That demonstrates the model's strong spatial generalization, meaning it can effectively predict rooftop PV feasibility at unseen locations without retraining on local data.

Notably, performance variability between rotations was minimal, suggesting that the proposed method captures key meteorological and geometric patterns common to diverse Indonesian climates. Cities with higher solar irradiation, such as Denpasar, Makassar, and Surabaya, yielded slightly higher classification confidence. In contrast, regions with higher humidity and precipitation, such as Padang and Pontianak, contributed to minor fluctuations in accuracy.

This cross-location validation confirms that the proposed DT–GSCV model generalizes well across diverse geographical and climatic contexts, outperforming baseline models that are typically more sensitive to location-specific features. Hence, the model is suitable for large-scale feasibility mapping of rooftop PV systems across Indonesia's heterogeneous urban and meteorological profiles.

Table 5. Model Performance under Cross-Location Validation

| Rotation Set | Training Cities | Testing Cities | Accuracy (%) | Precision | Recall | F1-Score |
|---|---|---|---|---|---|---|
| Rotation 1 | Jakarta, Bandung, Semarang, Surabaya, Denpasar, Makassar, Medan, Pekanbaru, Palembang, Balikpapan, Pontianak, Banjarmasin, Mataram, Kupang, Jayapura | Yogyakarta, Manado, Padang, Samarinda, Banda Aceh | 90.8 | 0.90 | 0.89 | 0.89 |
| Rotation 2 | Yogyakarta, Manado, Padang, Samarinda, Banda Aceh, Bandung, Semarang, Surabaya, Denpasar, Makassar, Medan, Pekanbaru, Palembang, Balikpapan, Pontianak | Jakarta, Banjarmasin, Mataram, Kupang, Jayapura | 91.1 | 0.91 | 0.90 | 0.90 |
| Rotation 3 | Jakarta, Yogyakarta, Surabaya, Bandung, Medan, Pekanbaru, Denpasar, Makassar, Samarinda, Pontianak, Balikpapan, Palembang, Padang, Banda Aceh, Mataram | Semarang, Manado, Kupang, Jayapura, Banjarmasin | 90.6 | 0.89 | 0.90 | 0.89 |
| Rotation 4 | Jakarta, Bandung, Medan, Semarang, Denpasar, Makassar, Manado, Palembang, Padang, Banda Aceh, Samarinda, Balikpapan, Pontianak, Banjarmasin, Mataram | Yogyakarta, Kupang, Jayapura, Pekanbaru, Surabaya | 91.4 | 0.91 | 0.91 | 0.91 |
| Rotation 5 | Surabaya, Denpasar, Bandung, Medan, Makassar, Padang, Banda Aceh, Pekanbaru, Palembang, Pontianak, Balikpapan, Samarinda, Mataram, Kupang, Jayapura | Jakarta, Yogyakarta, Semarang, Manado, Banjarmasin | 90.9 | 0.90 | 0.89 | 0.89 |

## CONCLUSION

This study presented a contextualized rooftop PV classification framework for Indonesia's capital cities using Decision Trees optimized via Grid Search Cross-Validation. DT with GSCV consistently outperformed k-NN, Naïve Bayes, and Logistic Regression across azimuth-direction scenarios, achieving up to 99.32% (one direction), 99.03% (two directions), and 93.55% (three directions) accuracy. Beyond numerical gains, interpretability enables transparent screening of feasible designs that minimize grid imports. Future research will validate field data, incorporate cost and carbon metrics, and explore ensemble and gradient-boosted trees. Additionally, field validation and integration of real irradiance and power output data are necessary to enhance the model's generalizability.

## ACKNOWLEDGMENT

The authors thank the Indonesian Ministry of Energy and Mineral Resources (MEMR) for its support and funding of this research.

## REFERENCES

[1] Z. Zhang *et al.*, “Carbon mitigation potential afforded by rooftop photovoltaic in China,” *Nat. Commun.*, vol. 14, no. 1, pp. 1–12, 2023, doi: 10.1038/s41467-023-38079-3.

[2] S. Rayegan *et al.*, “Modeling building energy self-sufficiency of using rooftop photovoltaics on an urban scale,” *Energy Build.*, vol. 324, no. June, p. 114863, 2024, doi: 10.1016/j.enbuild.2024.114863.

[3] A. Barbón, C. Bayón-Cueli, L. Bayón, and C. Rodríguez-Suanzes, “Analysis of the tilt and azimuth angles of photovoltaic systems in non-ideal positions for urban applications,” *Appl. Energy*, vol. 305, no. March 2021, p. 117802, 2022, doi: 10.1016/j.apenergy.2021.117802.

[4] X. Zhu, Y. Lv, J. Bi, M. Jiang, Y. Su, and T. Du, “Techno-Economic Analysis of Rooftop Photovoltaic System under Different Scenarios in China University Campuses,” *Energies*, vol. 16, no. 7, 2023, doi: 10.3390/en16073123.

[5] A. Y. Khan, Z. Ahmad, T. Sultan, S. Alshahrani, K. Hayat, and M. Imran, “Optimization of Photovoltaic Panel Array Configurations to Reduce Lift Force Using Genetic Algorithm and CFD,” *Energies*, vol. 15, no. 24, 2022, doi: 10.3390/en15249580.

[6] H. M. Maghrabie *et al.*, “State-of-the-art technologies for building-integrated photovoltaic systems,” *Buildings*, vol. 11, no. 9, 2021, doi: 10.3390/BUILDINGS11090383.

[7] N. T. Le, T. Le Truong, W. Asdornwised, S. Chaitusaney, and W. Benjapolakul, “Energy Production Analysis of Rooftop PV Systems Equipped with Module-Level Power Electronics under Partial Shading Conditions Based on Mixed-Effects Model,” *Energies*, vol. 16, no. 2, 2023, doi: 10.3390/en16020970.

[8] T. Chen, K. F. Tai, G. P. Raharjo, C. K. Heng, and S. W. Leow, “A novel design approach to prefabricated BIPV walls for multi-storey buildings,” *J. Build. Eng.*, vol. 63, no. PA, p. 105469, 2023, doi: 10.1016/j.jobe.2022.105469.

[9] S. Boubaker, S. Kamel, N. Ghazouani, and A. Mellit, “Assessment of Machine and Deep Learning Approaches for Fault Diagnosis in Photovoltaic Systems Using Infrared Thermography,” *Remote Sens.*, vol. 15, no. 6, 2023, doi: 10.3390/rs15061686.

[10] M. S. H. Onim *et al.*, “SolNet: A Convolutional Neural Network for Detecting Dust on Solar Panels,” *Energies*, vol. 16, no. 1, pp. 1–19, 2023, doi: 10.3390/en16010155.

[11] U. Hijjawi, S. Lakshminarayana, T. Xu, G. Piero Malfense Fierro, and M. Rahman, “A review of automated solar photovoltaic defect detection systems: Approaches, challenges, and future orientations,” *Sol. Energy*, vol. 266, no. November, p. 112186, 2023, doi: 10.1016/j.solener.2023.112186.

[12] R. Tang, Z. Ren, S. Ning, and Y. Zhang, “Fault classification of photovoltaic module infrared images based on transfer learning and interpretable convolutional neural network,” *Sol. Energy*, vol. 276, no. June, p. 112703, 2024, doi: 10.1016/j.solener.2024.112703.

[13] Z. Awwad, A. Alharbi, A. H. Habib, and O. L. de Weck, “Site Assessment and Layout Optimization for Rooftop Solar Energy Generation in Worldview-3 Imagery,” *Remote Sens.*, vol. 15, no. 5, 2023, doi: 10.3390/rs15051356.

[14] S. A. Memon, Q. Javed, W. G. Kim, Z. Mahmood, U. Khan, and M. Shahzad, “A Machine-Learning-Based Robust Classification Method for PV Panel Faults,” *Sensors*, vol. 22, no. 21, pp. 1–14, 2022, doi: 10.3390/s22218515.

[15] H. R. Iskandar, E. Taryana, and Y. B. Zainal, “Modelling and analysis of rooftop PV as an energy optimization of flat roof and gable roof mounting system,” *Sinergi*, vol. 28, no. 1, pp. 1–12, 2024, doi: 10.22441/sinergi.2024.1.001.

[16] G. Zhang and A. Gionis, “Regularized impurity reduction: accurate decision trees with complexity guarantees,” *Data Min. Knowl. Discov.*, vol. 37, no. 1, pp. 434–475, 2023, doi: 10.1007/s10618-022-00884-7.

[17] W. Wang, J. Keen, J. Bank, J. Giraldez, and K. Montano-Martinez, “An Automated Approach for Screening Residential PV Applications Using a Random Forest Model,” *IEEE Open Access J. Power Energy*, vol. 10, no. May, pp. 327–334, 2023, doi: 10.1109/OAJPE.2023.3270223.

[18] A. D. Sakti *et al.*, “Multi-Criteria Assessment for City-Wide Rooftop Solar PV Deployment: A Case Study of Bandung, Indonesia,” *Remote Sens.*, vol. 14, no. 12, pp. 1–24, 2022, doi: 10.3390/rs14122796.

[19] L. A. Yates, Z. Aandahl, S. A. Richards, and B. W. Brook, “Cross validation for model selection: A review with examples from ecology,” *Ecol. Monogr.*, vol. 93, no. 1, pp. 1–24, 2023, doi: 10.1002/ecm.1557.

[20] E. Escobar-Avalos, M. A. Rodríguez-Licea, H. Rostro-González, A. G. Soriano-Sánchez, and F. J. Pérez-Pinal, “A comparison of integrated filtering and prediction methods for smart grids,” *Energies*, vol. 14, no. 7, pp. 1–16, 2021, doi: 10.3390/en14071980.

[21] Y. Chen *et al.*, “Evaluation of Machine Learning Models for Smart Grid Parameters: Performance Analysis of ARIMA and Bi-LSTM,” *Sustain.*, vol. 15, no. 11, 2023, doi: 10.3390/su15118555.

[22] P. Charilaou and R. Battat, “Machine learning models and over-fitting considerations,” *World J. Gastroenterol.*, vol. 28, no. 5, pp. 605–607, 2022, doi: 10.3748/wjg.v28.i5.605.

[23] D. Lee, J. W. Jeong, and G. Choi, “Short term prediction of pv power output generation using hierarchical probabilistic model,” *Energies*, vol. 14, no. 10, pp. 1–15,

2021, doi: 10.3390/en14102822.
[24] H. Mao *et al.*, “Advances and prospects on estimating solar photovoltaic installation capacity and potential based on satellite and aerial images,” *Renew. Sustain. Energy Rev.*, vol. 179, no. April, p. 113276, 2023, doi: 10.1016/j.rser.2023.113276.
[25] C. Bu, T. Liu, T. Wang, H. Zhang, and S. Sfarra, “A CNN-Architecture-Based Photovoltaic Cell Fault Classification Method Using Thermographic Images,” *Energies*, vol. 16, no. 9, 2023, doi: 10.3390/en16093749.
[26] A. Ahmad, Y. Jin, C. Zhu, I. Javed, A. Maqsood, and M. W. Akram, “Photovoltaic cell defect classification using convolutional neural network and support vector machine,” *IET Renew. Power Gener.*, vol. 14, no. 14, pp. 2693–2702, 2020, doi: 10.1049/iet-rpg.2019.1342.
[27] H. Acikgoz, D. Korkmaz, and U. Budak, “Photovoltaic cell defect classification based on integration of residual-inception network and spatial pyramid pooling in electroluminescence images,” *Expert Syst. Appl.*, vol. 229, no. PA, p. 120546, 2023, doi: 10.1016/j.eswa.2023.120546.
[28] Á. Pérez-Romero, H. F. Mateo-Romero, S. Gallardo-Saavedra, V. Alonso-Gómez, M. D. C. Alonso-García, and L. Hernández-Callejo, “Evaluation of artificial intelligence-based models for classifying defective photovoltaic cells,” *Appl. Sci.*, vol. 11, no. 9, 2021, doi: 10.3390/app11094226.
[29] K. Tanioka and S. Hiwa, “Low-rank approximation of difference between correlation matrices using inner product,” *Appl. Sci.*, vol. 11, no. 10, 2021, doi: 10.3390/app11104582.
[30] Y. Essam *et al.*, “Investigating photovoltaic solar power output forecasting using machine learning algorithms,” *Eng. Appl. Comput. Fluid Mech.*, vol. 16, no. 1, pp. 2002–2034, 2022, doi: 10.1080/19942060.2022.2126528.
[31] D. El-Shahat, A. Tolba, M. Abouhawwash, and M. Abdel-Basset, “Machine learning and deep learning models based grid search cross validation for short-term solar irradiance forecasting,” *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-024-00991-w.
[32] W. Wang, J. Keen, J. Bank, J. Giraldez, and K. Montano-Martinez, “An Automated Approach for Screening Residential PV Applications Using a Random Forest Model,” *IEEE Open Access J. Power Energy*, vol. 10, no. February, pp. 327–334, 2023, doi: 10.1109/OAJPE.2023.3270223.
[33] A. Ahmad, Y. Jin, C. Zhu, I. Javed, A. Maqsood, and M. W. Akram, “Photovoltaic cell defect classification using convolutional neural network and support vector machine,” *IET Renew. Power Gener.*, vol. 14, no. 14, pp. 2693–2702, Oct. 2020, doi: 10.1049/iet-rpg.2019.1342.