SHM PUBLISHER
SURYA HIJAU MANFAAT

Journal of Information System Exploration and Research

https://shmpublisher.com/index.php/joiser

p-ISSN 2964-1160 | e-ISSN 2963-6361

# Analysis of the Stacking Ensemble Learning Model of Categorical Boosting and Naïve Bayes Algorithms for Crop Selection Based on Soil Characteristics

**Ilham Maulana[1*], Budi Prasetiyo[2]**
*[1,2] Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Semarang, Indonesia*

DOI: https://doi.org/10.52465/joiser.v3i2.604
Received 22 August 2025; Accepted 04 September 2025; Available online 30 July 2025

**Article Info**

***Keywords:***
Crop Recommendation;
Classification;
CatBoost;
Naïve Bayes;
Stacking

**Abstract**

This study aims to develop a machine learning model for selecting crop types based on soil characteristics, using the Categorical Boosting and Naïve Bayes algorithms as base learners. Next, an ensemble learning technique using a stacking approach was applied to improve the performance of the base model that was built. This was done to analyze and compare the performance results of each ensemble model that was carried out. Model performance was evaluated using evaluation metrics including precision, recall, f1-score, and accuracy. The results of this study indicate that the stacking ensemble model with Random Forest as the meta learner can provide better performance compared to other ensemble models. This model achieved a precision of 98.85337%, a recall of 99.84848%, an F1-score of 99.84844%, an accuracy of 99.84848%, and a model training time of 78.61110 seconds. Based on these results, this study is expected to provide tangible contributions and new knowledge in plant selection classification based on soil characteristics, thereby aiding in the precise and efficient determination of suitable plant types.



## 1. Introduction

Based on the decision of the Ministry of Agriculture of the Republic of Indonesia No. 484/KPTS/RC.020/M/8/2021, Indonesia has a land area of approximately 191.1 million ha, which is divided into two types of land, namely 43.6 million ha of wet land and 144.5 million ha of dry land [1]. There are 15.9 million ha of the total land area that has the potential to be used as agricultural land [1]. Agriculture is the main foundation of life throughout the world and the heart of a country's economy as it provides food and raw materials. However, in land cultivation, many farmers still do not understand the importance of selecting the right type of crop to plant [2].The accuracy of crop selection

* ***Corresponding Author:***

Ilham Maulana,
Computer Science Department, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Semarang,
Semarang, Indonesia.
Email: ilhammaulana05303005@students.unnes.ac.id

is crucial because each crop requires different levels of nitrogen, phosphorus, potassium, temperature, humidity, pH, and rainfall to achieve optimal yields [3]. Therefore, a technological approach is needed that can automatically and accurately process information from these factors in order to provide farmers with the right crop selection. With the development of technology, machine learning-based approaches are beginning to be used in the agricultural sector to assist in crop selection based on data. The machine learning approach enables the creation of automated systems that can learn from historical data and recognize certain patterns, resulting in more precise crop selection. Machine learning generally has several types of approaches, one of which is supervised learning.

Supervised learning is a learning process that uses labeled datasets, meaning that the input data has a known target [4]. The objective of this approach is to create an algorithmic model that can predict output results based on new input data that has not been previously trained [5]. This approach is suitable for classification and regression problems, such as predicting crop yields, plant diseases, and selecting crop types based on soil and weather characteristics. Examples of algorithms classified as supervised learning include Naïve Bayes and Categorical Boosting, each with their own characteristics and advantages in processing classification data. Naïve Bayes itself has advantages in training and prediction speed, efficiency, and is effective in handling noise, making it suitable for handling categorical data [6]. However, Naïve Bayes has a drawback in terms of feature independence, which often does not align with real-world data, leading to a decline in model performance [7].Meanwhile, Categorical Boosting (CatBoost) can process categorical features directly without encoding, reducing the risk of overfitting, and performing better on multi-class classification data [8], [9]. However, CatBoost has high model complexity, resulting in longer average training times and average performance measurements compared to other boosting models [10].Therefore, to address model shortcomings and achieve stable and accurate model performance, one approach is to use ensemble learning techniques. This ensemble learning technique combines several base learners to form a model with better performance. The ensemble learning technique used is stacking, which has a way of improving model performance by combining several classification models [11]. Stacking is a concept of combining several base models to build a stronger meta-model [11].

There have been several previous studies that have conducted machine learning-based research on the topic of plant selection based on soil characteristics. Dahiphale et al. [12] conducted research using several machine learning models, namely Logistic Regression, Decision Tree, Random Forest, K-Nearest Neighbors, Naïve Bayes, SVM, and Neural Network. Each model achieved an accuracy of 95.955%; 98.682%; 99.5%; 98.045%; 99.5%; 98.682%; and 97.73%. Based on these results, it is evident that the Naïve Bayes and Random Forest models achieved the highest accuracy of 99.5%, outperforming the other models in the study. Furthermore, the research by Sultana et al. [13] highlights that the application of ensemble learning techniques using the stacking method on base models can improve accuracy performance. The base models used were Naïve Bayes and Random Forest. After applying stacking to both models, the accuracy results showed an improvement in accuracy compared to each model individually. The Naïve Bayes model achieved an accuracy of 82%, while the Random Forest model achieved an accuracy of 92%. After combining them using the stacking method, the accuracy increased to 97%, indicating that applying the stacking method enhances classification prediction performance.

Based on the background description above, this proposed research focuses on analyzing the evaluation of the CatBoost and Naïve Bayes models using several ensemble learning techniques. The ensemble learning technique used in this research is stacking. The application of these techniques aims to obtain the model with the best performance. The model with the best evaluation results will then be tested through a website to determine its predictions for new data that has not been trained before.

## 2. Literature Review

This study was conducted with reference to a number of relevant literature reviews to ensure the correct direction and focus on the object and methods used. The literature used as a basis consists of the results of previous studies that specifically discuss plant classification based on soil characteristics, thereby providing a strong theoretical foundation and enriching the analysis in this study.

There are several studies that utilize classical machine learning for plant classification based on soil characteristics using tabular data. One such study was conducted by There are several studies that utilize classical machine learning for plant classification based on soil characteristics using tabular data.

One study conducted by Vyapari et al. [14] began exploring the combination of soil characteristic data with additional feature engineering such as pH, temperature, humidity, and NPK content, as well as plant type. The data underwent feature engineering to improve data quality in order to produce a good machine learning model. This study used several algorithm models, namely Logistic Regression, Naïve Bayes, Random Forest, Decision Tree, Support Vector Machine, and Gradient Boosting. The model performance evaluation results showed that the Random Forest and Gradient Boosting models achieved higher accuracy compared to the others, at 98.56% and 98% respectively. This study supports a new approach in plant classification research that begins to leverage the potential of boosting models to achieve higher prediction accuracy beyond traditional methods.

Furthermore, the use of derivative models of Gradient Boosting such as XGBoost, LightGBM, and CatBoost has begun to be adopted in studies to produce high accuracy. Research conducted by P. Mahesh et al. [15] using gradient-based machine learning models helped in the selection of plants based on soil characteristics. The main algorithm models used include eXtreme Gradient Boosting, Light Gradient-Boosting Machine, and Categorical Boosting (CatBoost). These models were trained using data on pesticides, rainfall, and average temperature, as well as historical harvest data from 10 crops commonly consumed worldwide. Based on the model performance evaluation results, the CatBoost model demonstrated superior performance compared to other boosting algorithms. CatBoost achieved a prediction accuracy of 99.123% and an R2 value of 0.964, indicating that the model has more reliable performance. Although the CatBoost method requires relatively longer model training time, it has proven to be the most effective in making predictions.

In addition, machine learning technology is advancing rapidly and the need for predictive model accuracy is increasing. The approach using classification models no longer relies solely on individual model training. Although a single model often provides good accuracy during training, in some cases, combining several models can provide more stable and accurate performance. This has led to the creation of the ensemble learning technique, which combines the probability results of several machine learning algorithms to improve prediction accuracy and reduce the risk of overfitting. Similarly, Afzal et al. [16] aimed to classify soil types to select the right crops, irrigation systems, and appropriate fertilization. The research results utilized the stacking ensemble learning technique, which combines several classification models—KNN, Random Forest, and XGBoost as base learners—and the AdaBoost algorithm as a meta-learner. The model evaluation results showed that the stacking ensemble learning model achieved higher accuracy, at 88.7935%, compared to models trained individually. This ensemble technique demonstrates that combining multiple classification models can enhance efficiency and achieve higher accuracy compared to models trained individually.

### 2.1 Categorical Boosting

Categorical Boosting, commonly referred to as CatBoost, is an algorithm model that focuses on categorical features and is a Gradient Boosting Decision Tree algorithm introduced by Prokhorenkova et al. and Dorogush et al [17]. CatBoost itself uses a complex ensemble learning technique, which is built based on a series of Decision Trees. Each subsequent tree is trained from the previous tree to improve model performance and reduce loss [18]. The CatBoost algorithm model differs from other Gradient Boosting trees in two key features: ordered boosting and efficient handling of categorical features [19].

Categorical Boosting builds predictions incrementally by adding weak learners at iteration t, which can be calculated using equation 2.1 [20].

$$F_t(x) = F_{t-1}(x) + \alpha \cdot h_m(x) \tag{2.1}$$

Description,
$F_t(x)$ = Model with iteration-t,
$F_{t-1}(x)$ = Previous model,
$h_m(x)$ = Newly added weak learner,
$\alpha$ = Learning rate to adjust the contribution of $h_m(x)$.
Next, the final prediction result $\widehat{y}_i$ is obtained from the sum of all Decision Trees, then the final predictor is calculated using equation 2.2 [21].

$$\widehat{y}_i = \sum_{k=1}^{K} f_k\,(x_i), f_k \in F \tag{2.2}$$

Description,
$\hat{y}_i$ = Prediction result for data-*i*,
$K$ = Total number of Decision Trees,
$x_i$ = Input feature-i,
$f_k$ = Base learner-k.

Next, an illustration of the Categorical Boosting algorithm is shown in Figure 1. The illustration shows that Categorical Boosting works by combining several weak learners to form a strong and accurate prediction model, especially for handling categorical data.

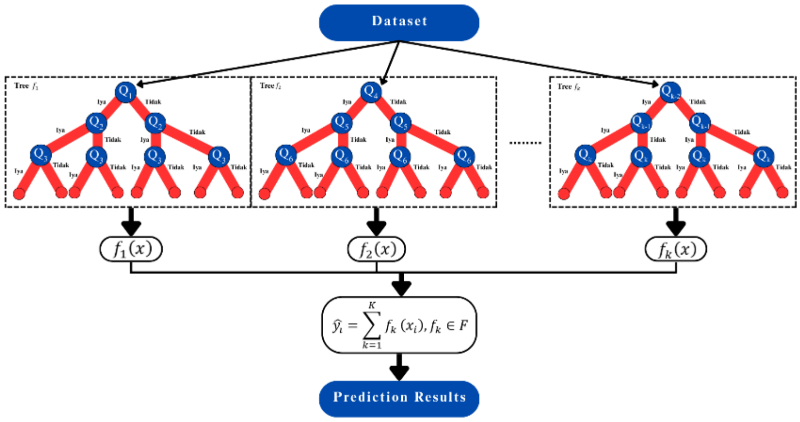


Figure 1. Illustration of how Categorical Boosting works [21]

## 2.2 Naïve Bayes

Naïve Bayes is one of the most commonly used classification algorithms due to its simple structure and high speed [22]. This algorithm is often used as a benchmark for comparison when designing new predictive approaches [23]. Naïve Bayes has the same network structure, where all attributes $x_1 \cdots x_n$ are considered conditionally independent based on the class value $C$. As shown in Figure 2, this illustrates an example of the structure obtained when there are four attributes.

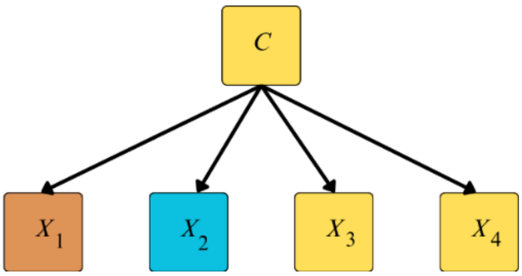


Figure 2. Structure of the Naïve Bayes algorithm [24]

The main objective of the Naïve Bayes algorithm is to predict that test sample x (with m class attributes) belongs to class $C_i$ if and only if the posterior probability $P(C_i|X)$ is higher than the posterior probabilities of other classes. Mathematically, this can be seen in equation 2.3 [22].

$$P(C_i|X) = Max\{P(C_1|X), P(C_2|X), \cdots P(C_n|X)\} \quad (2.3)$$

Description,
$C_i$ = Target class
$X$ = Input feature to be classified
$P(C_i|X)$ = Probability that input data $x$ belongs to class $\mathrm{C_i}$
$Max\{\cdot\}$ = Select the maximum value from all $P(C_i|X)$
$C_1, C_2, \cdots\cdots, C_n$ = All possible classes (total of $n$ classes)

## 2.3 Stacking

Stacking is a technique that combines several machine learning models to build a single, more powerful machine learning model. This technique aims to overcome the weaknesses of each individual model and improve the generalization ability and overall strength of the model [25], [26]. This technique uses a layered learning approach. In a study [25] it is mentioned that stacking has two layers, namely:

1. Layer-1 (base learners) is a set of basic learning models that are trained independently with training data. These basic learning models can be homogeneous or heterogeneous, which means that several types of machine learning algorithms can be combined. The purpose of using different types of models is to capture different patterns in the data.
2. Layer-2 (base learner) is an advanced model trained using the prediction results from base learners as input. This meta learner is tasked with optimally combining the predictions of base learners by minimizing errors and maximizing accuracy.

In the stacking training process, training is first performed on base learners using training data and test data to train the base learner model. The training process often utilizes the K-fold cross validation technique to avoid overfitting and data leakage [25], [26]. An illustration of the stacking technique process can be seen in Figure 3.

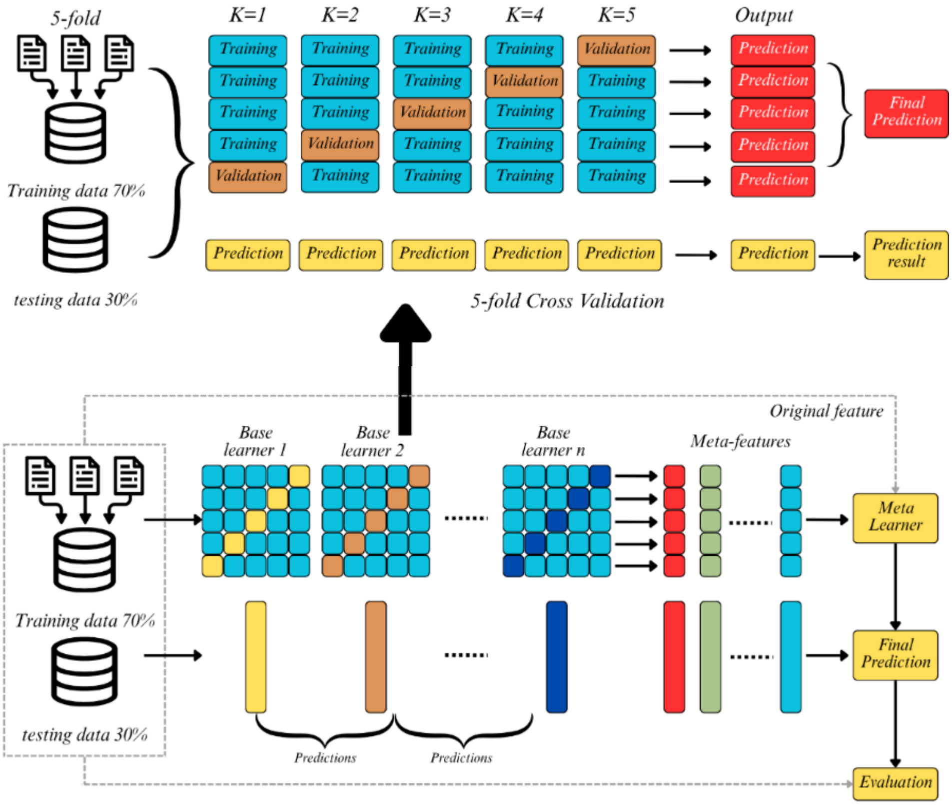


Figure 1. Stacking and fold CV model structure [25]

## 3. Method

This research approach uses a quantitative approach because the dataset obtained will be analyzed statistically to determine the performance and accuracy of the algorithm model. The quantitative approach is used because it is in line with the research objectives of measuring and evaluating the performance of the classification model that has been built. This approach is carried out in every process, starting from data preprocessing, building a classification model, to the model evaluation process. Model evaluation will be analyzed using statistical methods such as accuracy, precision, recall, F1-score, and confusion matrix calculations. This approach provides objective and measurable results in assessing model performance and is expected to yield potential improvements in the future. The stages carried out in this research can be seen in the flowchart in Figure 4.

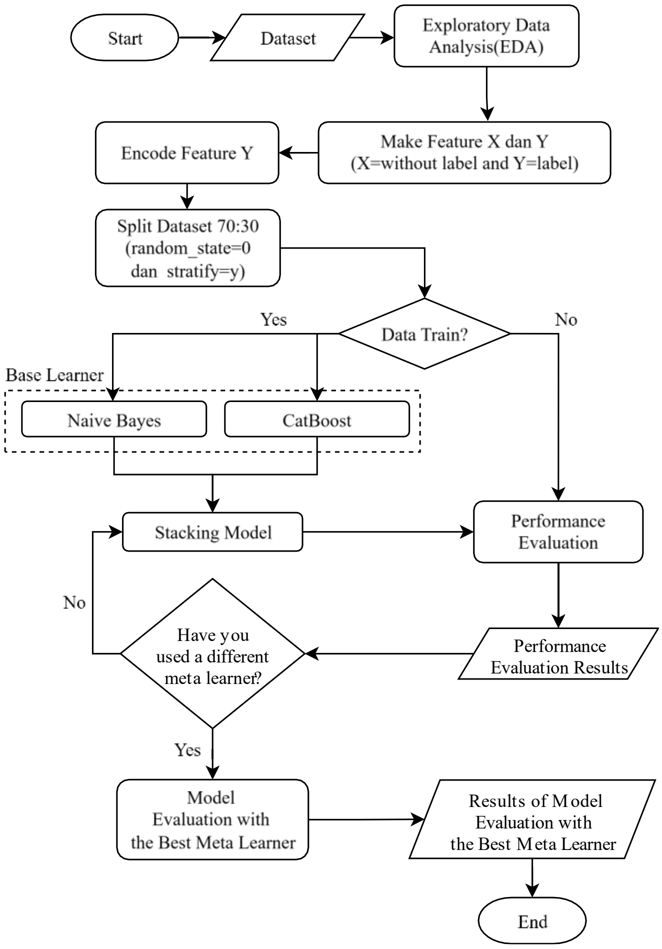


Figure 4. Research method flowchart

### 3.1 Dataset

The data used in this study was obtained from the Kaggle platform under the title Crop Recommendation Dataset [27]. The dataset has undergone augmentation by its author, combining various data sources such as rainfall, climate information, and soil characteristics from the Indian region. Overall, this dataset contains 2,200 data points on soil conditions and consists of eight attributes. The main focus of this dataset is to classify the most suitable crop types for the available soil conditions. Further details about the dataset can be found in Table 1.

Table 1. Crop recommendation dataset attributes

| Atribut | Deskripsi | Tipe Data |
|---|---|---|
| N (Nitrogen) | Nitrogen content in soil | int64 |
| P (phosphorus) | Phosphorus content in soil | int64 |
| K (potassium) | Potassium content in soil | int64 |
| Temperature | Soil temperature in degrees | float64 |
| Humidity | Moisture level in soil in degrees | float64 |
| pH | Acidity and alkalinity levels in soil | float64 |
| Rainfall | Rainfall levels in soil | float64 |
| Label | Type of plant | object |

### 3.2 Data Preprocessing

Data preprocessing is the stage of preparing data so that it is ready to be used in the training model. There are several stages of data preprocessing used in this study, namely exploratory data analysis, feature separation, feature encoding, and dataset splitting. The following is a detailed explanation of each stage.

### 3.3.1 Exploratory Data Analysis

Based on Figure 5, the relationship between input attributes shows that the nutrients phosphorus (P) and potassium (K) have a fairly strong positive correlation of 0.74, indicating a correlation between their distribution. Meanwhile, the correlation between nutrient elements and climate parameters such as temperature, humidity, pH, and rainfall is weak, with a value of <0.2, meaning that climate factors only have a limited influence on the concentration of these nutrient elements. Therefore, the features indicate that each attribute carries unique information that does not overlap with other attributes, making all attributes suitable for use in the model-building stage. The low correlation between features reduces the risk of multicollinearity in the classification model and has the potential to improve classification quality because each feature contributes different information.

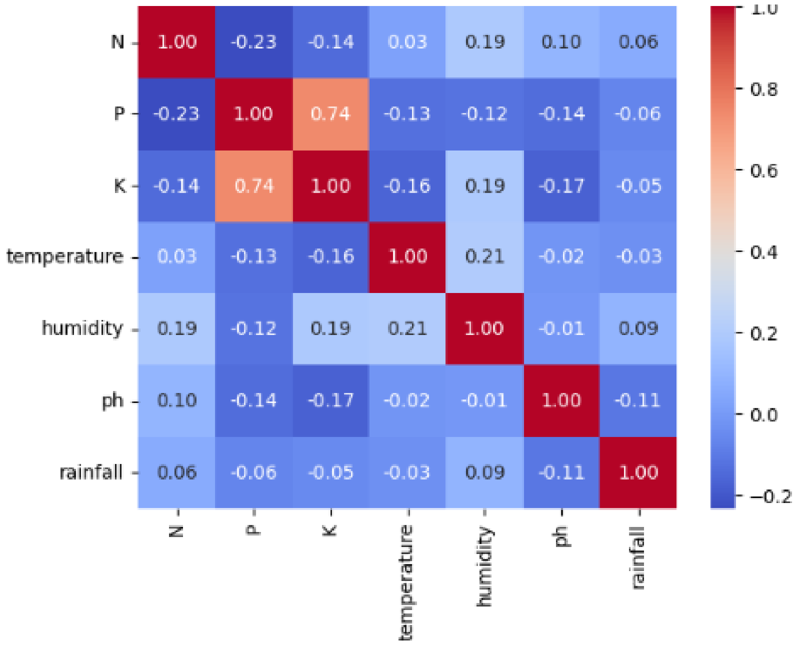


Figure 5. Feature Correlation Visualization

### 3.3.2 Features Used

The next step is to separate the features into two parts, namely input features and output features. Input features are independent attributes that serve as model inputs in the model training process. Meanwhile, output features are target attributes or outputs generated by the prediction model based on the information provided by the input features. This separation aims to facilitate the model training process in finding patterns between input features and targets. The list of attributes used for input and output features can be seen in Table 2.

Table 2. Separation of input and output features

| Features Input | Features Output |
|---|---|
| N (Nitrogen) | |
| P (phosphorus) | |
| K (potassium) | |
| Temperature | Label (Type of plant) |
| Humidity | |
| pH | |
| Rainfall | |

### 3.3.3 Encoding Label

Encoding is the process of converting categorical data into a numerical format so that it can be used in the model training stage of the algorithm. This stage is very important because some machine learning algorithms can only handle data in numerical form. This stage is performed on feature output, which is a string data type, so it is converted into a numerical format. The feature encoding process uses the LabelEncoder function from the scikit-learn library. Details of the output feature encoding can be seen in Table 3.

Table 3. Output feature encoding

| Encoding Feature Ouput | Uncoding Feature Output |
|---|---|
| 0 | apple |
| 1 | banana |
| 2 | blackgram |
| 3 | chickpea |
| 4 | coconut |
| 5 | coffee |
| 6 | cotton |
| 7 | grapes |
| 8 | jute |
| 9 | kidneybeans |
| 10 | lentil |
| 11 | maize |
| 12 | mango |
| 13 | mungbean |
| 14 | mothbeans |
| 15 | muskmelon |
| 16 | orange |
| 17 | papaya |
| 18 | pigeonpeas |
| 19 | pomegranate |
| 20 | rice |
| 21 | watermelon |

**3.3.4 Split Dataset**

The next step is to separate the data into two main parts, namely training data and test data, which aims to ensure that the model can perform well with data that has never been trained before or training data. Training data is used to train the algorithm model so that it can learn patterns between features, while test data is used to evaluate the performance of the model that has been trained on data that has never been trained before. This data separation is performed using the train_test_split function, with 70% of the data used for training the model and 30% used for testing. Additionally, the parameter random_state=0 is applied to ensure consistent data separation results each time the process is run. Furthermore, the parameter stratify = (encoded label) is applied to maintain a balanced class distribution between the training and testing data.

**3.4 Variations Algorithm Model**

After the data goes through the preprocessing stage, it then enters the algorithm model building stage. The model is built using several different model variations, namely the CatBoost model, the Naïve Bayes model, and the stacking model. Details of the model variations built can be seen in Table 4. All model variations use default parameters and only add the parameter random_state = 0.

Table 4. Variations in the Algorithm Models Used

| Metode | Ensemble | Final Estimator | Model |
|---|---|---|---|
| CatBoost | - | - | Model #1 |
| Naïve Bayes | - | - | Model #2 |
| CatBoost-Naïve Bayes | Stacking | Logistic Regression | Model #3 |
| CatBoost-Naïve Bayes | Stacking | Random Forest | Model #4 |

### 3.5 Evaluation Model

Algorithm model evaluation is a stage for assessing how well a model performs in making predictions using several evaluation metrics [28]. Common evaluation metrics include the confusion matrix, precision, recall, F1-score, and accuracy [28], [29], [30]. A confusion matrix is a table showing the accuracy of the algorithm model and the precision of the algorithm model in making predictions [28], [31]. The confusion matrix consists of four cells: true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The confusion matrix is shown in Table 5. The mathematical formulas for the evaluation matrix can be seen in Equations 3.1, 3.2, 3.3, and 3.4.

Table 5. Confusion Matrix

| | | Predicted Class | |
|---|---|---|---|
| | | Positive | Negative |
| True Class | Positive | **TP** | **FP** |
| | Negative | **FN** | **TN** |

$$Precision = \frac{TP}{TP + FP} \quad (3.1)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.2)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.3)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.4)$$

## 4. Results and Discussion

### 4.1 Model Evaluation Results

This study focuses on classifying plant types based on soil characteristics using the Categorical Boosting and Naïve Bayes classification algorithms. In addition to using individual models, this study also implements ensemble learning techniques, such as soft voting, blending, and stacking, to improve model accuracy and stability against data variations. To assess the performance of each model built, the models were evaluated based on accuracy, precision, recall, F1-score, and confusion matrix metrics. The evaluation results were then compared comprehensively to determine which model provided the best performance in classification. A complete comparison of the performance evaluation of each model can be seen in Table 6, which displays the accuracy results of each model built in this study.

Table 6. Comparison of Model Performance Evaluation Results

| Metode | Ensemble | Final Estimator | Accuracy | Model |
|---|---|---|---|---|
| CatBoost | - | - | 99,545% | Model #1 |
| Naïve Bayes | - | - | 99,545% | Model #2 |
| CatBoost-Naïve Bayes | Stacking | Logistic Regression | 97,273% | Model #3 |
| **CatBoost-Naïve Bayes** | **Stacking** | **Random Forest** | **99,848%** | **Model #4** |

Based on Table 6, which compares the evaluation results of the seven classification model variations tested in this study, Model #4, which is a stacking model using Random Forest as the base learner and Categorical Boosting and Naïve Bayes as the base learners, shows superior performance compared to other model variations. This model achieved the highest accuracy of 99.848%, making it the best-performing classification model variation among the others. Model #4 uses Random Forest as the base learner, unlike the stacking model in Model #4, which uses Logistic Regression as the base learner and only achieves an accuracy of 97.273%. This indicates that Random Forest provides a more complex non-linear approach, enabling it to capture more predictive patterns and achieve higher and more stable accuracy.

### 4.2 Discussion and Comparation Model

After analyzing the performance of the constructed model, Table 7 shows a comparison of the model's performance with previous studies.

Table 7. Comparison of Model Performance with Previous Researchers

| Previous Research | Method | Dataset | Results Accuracy |
|---|---|---|---|
| [32] | SVM | Crop recommendation dataset | 87,38% |
| | XGBoost | | 95,62% |
| | Random Forest | | 97,18% |
| | Decision Tree | | 86,64% |
| | KNN | | 96,36% |
| [33] | Multilayer Perceptron | Crop recommendation dataset | 98,2273% |
| | Decision Table | | 88,5909% |
| | Repeated Incremental Pruning to Produce Error Reduction (Jrip) | | 96,2273% |
| [34] | SVM (rbf) | Indian Chamber of Food and Agriculture | 96,36% |
| | SVM (linear) | | 97,57% |
| | SVM (poly) | | 97,87% |
| | Decision Tree | | 98% |
| [35] | LightGBM + Decision Tree | Crop recommendation dataset | 98, 64% |
| [36] | Support Vector Machine Classifier | Crop recommendation dataset | 98,6% |
| | Decision Tree Classifier | | 97,7% |
| | Random Forest Classifier | | 99,3% |
| | GaussianNB | | 99% |
| | Kneighbors Classifier | | 97,7% |
| | AdaBoost Classifier | | 10,4% |
| | Gradient Boosting Classifier | | 97,9% |
| **Proposed Method** | **Stacking CatBoost-Naïve Bayes with Final Estimator Random Forest** | **Crop recommendation dataset** | **99,848%** |

Based on Table 7, it can be seen that the proposed model, namely the Categorical Boosting and Naïve Bayes stacking model as the base learner, and Random Forest as the base learner, achieved an accuracy of 99.848% on the crop recommendation dataset. The accuracy obtained is superior to the research conducted by [32] which showed that the model with the highest accuracy was Random Forest, with an accuracy of 97.18%. These results indicate that the proposed model combination in this study performs better in handling the characteristics of the crop recommendation dataset used. Furthermore, the study by [33] showed the highest accuracy using the Multilayer Perceptron approach at 98.2273%. This accuracy is still relatively low compared to the proposed model. The lower performance is due to the model's limitations in handling complexity and distributing diverse data, as well as the lack of application of the ensemble learning stacking technique, which can reduce the shortcomings and strengthen each model.

The study conducted by [34] also used several variations of the Support Vector Machine and Decision Tree models. The accuracy results show that the Decision Tree model achieved a higher accuracy of 98% compared to other model variations. However, this accuracy is still below that of the proposed model in this study. This indicates that while Decision Trees are effective in classifying data, the stacking approach with the appropriate combination of base learners has proven to be superior in learning complex patterns in the dataset.Furthermore, the study conducted by [35] which applied a combination of LightGBM and Decision Tree models, achieved an accuracy of 98.64%. Although this result is relatively high, the stacking model proposed in this study still demonstrates better performance. This advantage indicates that the use of stacking techniques by combining two base learners with different characteristics, namely Categorical Boosting and Naïve Bayes, can provide more accurate predictions, especially when combined with non-linear base learners such as Random Forest.

Furthermore, [36] tested several classification algorithms, such as Random Forest, Gaussian Naïve Bayes, K-Nearest Neighbors, AdaBoost, and Gradient Boosting, on a crop recommendation dataset, recording the highest accuracy of 99.3% for Random Forest and 99% for GaussianNB. Although both results provide accuracy close to the proposed model, the proposed model still has more optimal performance, achieving 99.848% from the stacking model in this study. This reinforces the finding that combining algorithm models can complement each other by utilizing stacking techniques, which enhance generalization and classification model performance compared to using a single model.

Overall, it can be concluded that the proposed stacking model in this study excels in accuracy and in leveraging the strengths of the algorithms used. This advantage demonstrates the significant potential of the model for plant selection based on soil characteristics, with prediction accuracy superior to models applied in previous studies.

## 5. Conclusion

The evaluation results of the stacking ensemble model with Categorical Boosting and Naïve Bayes as base learners, and Random Forest as the base learner, produced better evaluation results than the other three model variations. The evaluation was conducted using precision, recall, f1-score, and accuracy. The evaluation results of the stacking model variation with the Random Forest base learner achieved a precision of 99.85337%; recall of 99.84848%; F1-score of 99.84844%; and accuracy of 99.84848%. These results show that the stacking ensemble learning model with the Random Forest base learner is capable of more accurate classification in plant selection based on soil characteristics.

## References


[1] K. P. R. Indonesia, “Keputusan Menteri Pertanian Republik Indonesia Nomor 484/KPTS/RC.020/M/8/2021 Tentang Perubahan Kedua Atas Keputusan Menteri Pertanian Nomor 259/KTPS/RC.020/M/05/2020 Tentang Rencana Strategis Kementrian Pertanian Tahun 2020-2024,” Jakarta, 2021. [Online]. Available: https://bpmsph.ditjenpkh.pertanian.go.id/wp-content/uploads/2023/10/RENSTRA-KEMENTAN-2020-2024-REVISI-2-26-Agt-2021.pdf

[2] A. F. Negarawan, M. U. Siregar, A. Fatwanto, and M. D. R. Wahyudi, “An Implementation of Profile Matching Method to Determine Agricultural Crops that Suit the Land,” *Proc. Int. Conf. Sci. Eng. (ICSE-UIN-SUKA 2021)*, vol. 211, pp. 124–129, 2022, doi: 10.2991/aer.k.211222.020.

[3] Mahmudi, I. Sasli, and T. H. Ramadhan, “Growth and yield of rice from mycorrhizal enrichment seedlings on different soil water content,” *Indones. J. Agron.*, vol. 51, no. 2, pp. 173–180, 2023, doi: 10.24831/ija.v51i2.46201.

[4] B. Mahesh, “Machine Learning Algorithms - A Review,” *Int. J. Sci. Res.*, vol. 9, no. 1, pp. 381–386, 2020, doi: 10.21275/art20203995.

[5] Q. Liu and Y. Wu, “Encyclopedia of the Sciences of Learning,” *Encycl. Sci. Learn.*, no. January 2012, 2012, doi: 10.1007/978-1-4419-1428-6.

[6] P. J. B. Pajila, B. G. Sheena, A. Gayathri, J. Aswini, M. Nalini, and R. Siva Subramanian, “A Comprehensive Survey on Naive Bayes Algorithm: Advantages, Limitations and Applications,” *Proc. 4th Int. Conf. Smart Electron. Commun. ICOSEC 2023*, no. September 2023, pp. 1228–1234, 2023, doi: 10.1109/ICOSEC58147.2023.10276274.

[7] R. Blanquero, E. Carrizosa, P. Ramírez-Cobo, and M. R. Sillero-Denamiel, “Variable selection for Naïve Bayes classification,” *Comput. Oper. Res.*, vol. 135, p. 105456, 2021, doi: 10.1016/j.cor.2021.105456.

[8] W. Chang, X. Wang, J. Yang, and T. Qin, “An Improved CatBoost-Based Classification Model for Ecological Suitability of Blueberries,” *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23041811.

[9] C. Yan, X. Zhang, and J. Shen, “Credit Score Classification Using Advanced Machine Learning : A

Comprehensive Approach," vol. 18, no. 3, pp. 98–112, 2025, doi: 10.4236/jsea.2025.183007.
[10] A. A. Ibrahim, R. L. Ridwan, M. M. Muhammed, R. O. Abdulaziz, and G. A. Saheed, "Comparison of the CatBoost Classifier with other Machine Learning Methods," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 11, pp. 738–748, 2020, doi: 10.14569/IJACSA.2020.0111190.
[11] M. R. P. Putra and E. Utami, "Comparative Analysis of Hybrid Model Performance Using Stacking and Blending Techniques for Student Drop Out Prediction In MOOC," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 8, no. 3, pp. 346–354, 2024, doi: 10.29207/resti.v8i3.5760.
[12] D. Dahiphale, P. Shinde, K. Patil, and V. Dahiphale, "Smart Farming: Crop Recommendation using Machine Learning with Challenges and Future Ideas," 2023, doi: 10.36227/techrxiv.23504496.v1.
[13] S. Sultana, T. Hasan, F. I. Tanjil, S. Islam, M. O. Rahman, and N. R. Chakraborty, "An IoT and Hybrid Machine Learning Based Real-time Web Framework for Optimizing Crop Recommendation," *CAMA Cent. Appl. Macroecon. Anal.*, vol. 13, no. January, pp. 909–920, 2023, doi: 10.2139/ssrn.5132627.
[14] C. Vyapari, P. Bhosale, and C. Vyapari Prathamesh Bhosale Ameya Parkar, "Comparative Analysis of Machine Learning Algorithms for Crop Recommendations," pp. 0–4, 2023, [Online]. Available: https://doi.org/10.21203/rs.3.rs-3181672/v1
[15] P. Mahesh and R. Soundrapandiyan, "Yield prediction for crops by gradient-based algorithms," *PLoS One*, vol. 19, no. 8, pp. 1–20, 2024, doi: 10.1371/journal.pone.0291928.
[16] H. Afzal *et al.*, "Incorporating soil information with machine learning for crop recommendation to improve agricultural output," *Sci. Rep.*, vol. 15, no. 1, pp. 1–15, 2025, doi: 10.1038/s41598-025-88676-z.
[17] H. Wang and L. Cheng, "C AT B OOST MODEL WITH SYNTHETIC FEATURES IN APPLICATION TO LOAN RISK ASSESSMENT OF SMALL," 2021.
[18] S. Hussain, M. Wazir, T. A. Jumani, and S. Khan, "A novel feature engineered-CatBoost-based supervised machine learning framework for electricity theft detection," *Energy Reports*, vol. 7, pp. 4425–4436, 2021, doi: 10.1016/j.egyr.2021.07.008.
[19] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost : unbiased boosting with categorical features," no. Section 4, pp. 1–23, 2019.
[20] R. Taherdangkoo *et al.*, "Modeling unsaturated hydraulic conductivity of compacted bentonite using a constrained CatBoost with bootstrap analysis," *Appl. Clay Sci.*, vol. 260, no. March, p. 107530, 2024, doi: 10.1016/j.clay.2024.107530.
[21] X. Huang, W. Liu, Q. Guo, and J. Tan, "Prediction method for the dynamic response of expressway lateritic soil subgrades on the basis of Bayesian optimization CatBoost," *Soil Dyn. Earthq. Eng.*, vol. 186, no. September, p. 108943, 2024, doi: 10.1016/j.soildyn.2024.108943.
[22] H. Chen, S. Hu, R. Hua, and X. Zhao, "Improved naive Bayes classification algorithm for traffic risk management," *EURASIP J. Adv. Signal Process.*, vol. 2021, no. 1, 2021, doi: 10.1186/s13634-021-00742-6.
[23] J. S. A. Ruiz, C. Romero, and A. Cicconardi, "XNB: Explainable Class-Specific NaIve-Bayes Classifier," 2024, [Online]. Available: http://arxiv.org/abs/2411.01203
[24] T. Miquélez, E. Bengoetxea, and P. Larrañaga, "Evolutionary computation based on Bayesian classifiers," *Int. J. Appl. Math. Comput. Sci.*, vol. 14, no. 3, pp. 335–349, 2004, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.110.3623&rep=rep1&type=pdf
[25] L. Wu *et al.*, "Prediction model for the compressive strength of rock based on stacking ensemble learning and shapley additive explanations," vol. 123, 2024, doi: 10.1007/s10064-024-03896-3.
[26] Q. Li and X. Wang, "Bayesian optimization of stacking ensemble learning model for HPC compressive strength prediction," *Expert Syst. Appl.*, vol. 288, no. May, p. 128281, 2025, doi: 10.1016/j.eswa.2025.128281.
[27] A. Ingle, "Crop Recommendation Dataset." Accessed: Jan. 13, 2025. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset
[28] M. Vakili, M. Ghamsari, and M. Rezaei, "Performance Analysis and Comparison of Machine and Deep Learning Algorithms for IoT Data Classification," no. January, 2020, doi: 10.48550/arXiv.2001.09636.
[29] A. Mohammed and R. Kora, "A comprehensive review on ensemble deep learning: Opportunities and challenges," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 35, no. 2, pp. 757–774, 2023, doi: 10.1016/j.jksuci.2023.01.014.

[30] D. M. W. Powers, “Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,” pp. 37–63, 2020, [Online]. Available: http://arxiv.org/abs/2010.16061

[31] I. Markoulidakis and G. Markoulidakis, “Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis,” *Technologies*, vol. 12, no. 7, 2024, doi: 10.3390/technologies12070113.

[32] N. N. Thilakarathne, M. S. A. Bakar, P. E. Abas, and H. Yassin, “A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming,” 2022. doi: 10.3390/s22166299.

[33] K. Bakthavatchalam *et al.*, “IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms,” *Technologies*, vol. 10, no. 1, 2022, doi: 10.3390/technologies10010013.

[34] S. Chunduri, H. Raj, and V. G. Narendra, “Intelligent Systems for Crop Recommendation using Machine Learning,” *WSEAS Trans. Comput.*, vol. 24, pp. 14–19, 2025, doi: 10.37394/23205.2025.24.2.

[35] R. K. Banoth and R. B. V. Murthy, “Automatic Crop Recommendation System Using LightGBM and Decision Tree Machine Learning Models,” *J. Mach. Comput.*, vol. 5, no. 1, pp. 343–355, 2025, doi: 10.53759/7669/jmc202505026.

[36] M. Bouni, B. Hssina, K. Douzi, and S. Douzi, “Interpretable Machine Learning Techniques for an Advanced Crop Recommendation Model,” *J. Electr. Comput. Eng.*, vol. 2024, 2024, doi: 10.1155/2024/7405217.