BAREKENG: Journal of Mathematics and Its Applications
September 2025 Volume 19 Issue 3 Page 1853-1864
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol19iss3pp1853-1864

COMPARATIVE ANALYSIS OF MACHINE LEARNING METHODS IN CLASSIFYING THE QUALITY OF PALU SHALLOTS

Desy Lusiyanti[1], Selvy Musdalifah[2*], Agusman Sahari[3], Iman Al Fajri[4]

[1,2,3,4] *Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Tadulako*
Jln. Soekarno Hatta No.KM. 9, Tondo, Mantikulore Sub-District, Palu, Central Sulawesi, 94148, Indonesia

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Brawijaya*
Jln. Veteran, Malang, East Java, 65145, Indonesia

*Corresponding author's e-mail: *selvymusdalifah@gmail.com*

Article History:

Received: 5th November 2024
Revised: 7th February 2025
Accepted: 8th April 2025
Published: 1st July 2025

Keywords:

Classification;
Machine Learning;
Palu Shallots;
Random Forest;
SVM.

ABSTRACT

This study conducts a comparative analysis of various machine learning methods for classifying the quality of Palu shallots based on the Indonesian National Standard (SNI). The dataset consists of 1,500 samples of Palu shallots, each characterized by 10 key features, including size, color, texture, and moisture content. Five machine learning models—Naïve Bayes, Decision Tree, Random Forest, Support Vector Machine (SVM), and Logistic Regression—were evaluated using accuracy, precision, recall, and F1 score as performance metrics. The results indicate that Random Forest achieved the best performance with an accuracy of 95.4%, followed by Decision Tree (90.7%) and SVM (90.2%). Random Forest also excelled in precision (93.6%) and F1 Score (93.5%), making it the most reliable model for shallot quality classification. Meanwhile, SVM demonstrated a good balance between recall and precision, making it a strong alternative. Implementing machine learning models has the potential to enhance the efficiency and accuracy of agricultural product quality assurance. The findings of this study provide valuable insights for farmers, agribusiness practitioners, and researchers adopting artificial intelligence technology for more precise and efficient agricultural quality assessment.

How to cite this article:

D. Lusiyanti, S. Musdalifah, A. Sahari and I. Al Fajri., "COMPARATIVE ANALYSIS OF MACHINE LEARNING METHODS IN CLASSIFYING THE QUALITY OF PALU SHALLOTS," *BAREKENG: J. Math. & App.,* vol. 19, no. 3, pp. 1853-1864, September, 2025.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Shallots (*Allium Ascalonium L.*) are a horticultural commodity that plays an important role in the agricultural sector in Indonesia [1]. Shallots are not only a staple ingredient in Indonesian cuisine but also have high economic value due to their increasing demand. In Palu City, Central Sulawesi, shallots are one of the leading agricultural products that significantly contribute to the local economy [2].

However, the quality of the shallots produced often varies. This quality variation can be attributed to various factors, including cultivation techniques, soil conditions, weather, and post-harvest practices. Inconsistent shallot quality can affect selling prices and consumer satisfaction. Therefore, classifying the quality of shallots is crucial to ensure that the products meet the standards desired by the market.

In several previous studies, methods such as linear regression, artificial neural networks, and pattern recognition have been used to classify the quality of shallots. For instance, a Pradana (2023) study focused on detecting shallot quality using circularity image processing [3]. Another study by Mutia (2020) used regression analysis to predict shallot quality based on environmental factors [4]. Although these approaches are quite effective, they have some drawbacks, such as the need for large datasets and complexity in implementation. On the other hand, advancements in machine learning technology offer more adaptive and efficient solutions. Machine learning enables faster and more accurate data processing by utilizing algorithms that can learn from data and improve their performance over time.

Various machine learning methods, such as k-nearest neighbors (KNN), support vector machines (SVM), and neural networks, have been successfully applied in various classification applications in agriculture [5], [6], [7]. However, each method has advantages and disadvantages depending on the data characteristics and classification objectives. Therefore, it is essential to conduct a comparative analysis of different machine learning methods to determine the most effective method for classifying the quality of Palu shallots.

This research aims to fill the gap in the literature by providing a comprehensive comparative analysis of several machine learning methods for classifying the quality of Palu shallots. The study also seeks to identify key factors affecting the performance of each method and to offer practical recommendations for farmers and agribusiness entrepreneurs in choosing the most suitable method for their needs.

One of the strengths of this research is the application of more advanced and adaptive machine learning techniques compared to traditional approaches. Additionally, the research utilizes a larger and more diverse dataset, which is expected to yield more accurate and representative results. Thus, this study not only contributes to improving the quality of Palu shallots but also provides new insights into the application of machine learning technology in agriculture.

The novelty of this research lies in the in-depth analysis of the performance of various machine learning methods in the specific context of classifying the quality of Palu shallots. The study not only compares existing methods but also explores new combinations and hybrid approaches that have not been extensively researched before. The results of this study are expected to serve as an important reference for the development of quality classification technology for agricultural products in the future.

2. RESEARCH METHODS

2.1. Data Collection

The data used in this research consists of Palu shallot samples collected from various locations in Sigi Regency, specifically from the villages of Maku, Maranata, Sidondo, Sidera, Oloboju, Soulowe, and Watubula. Subsequently, quality measurements of the samples were conducted in the laboratory by agronomists using standard measuring instruments. Each sample was tested based on the parameters established by the Government for shallot commodities, namely the Indonesian National Standard (SNI 01-3159-1992), as follows:

Table 1. Quality Standard of Red Onions

Characteristic	Requirements		Testing Method
	Quality I	Quality II	
Similarity of Varieties	Alike	Alike	Organoleptic
Ages	Old	Quite Old	Organoleptic
Texture	Solid	Quite Solid	Organoleptic
Diameter (cm) minimum	>2,5	1,5-2,5	SP-SMP-309-1981
Damage, % weight/maximum weight	5	8	SP-SMP-310-1981
Rotten, % weight/maximum weight	1	2	SP-SMP-311-1981
Dirt, % weight/maximum weight	No	No	SP-SMP-313-1981
Dryness	Dry	Dry	Organoleptic
Water Content (%)	80-85	75-80	SP-SMP-313-1981

**Data source*: Badan Standardisasi Nasional, 1992

2.2. Data Preprocessing

Before training the machine learning models, the dataset undergoes several preprocessing steps to ensure its suitability for analysis and to enhance model performance. The main preprocessing steps include data normalization and labeling, critical to standardizing the input data and providing meaningful target values for classification.

a. Data Normalization

Data normalization is performed to eliminate differences in scale among the features extracted from the samples [8]. This step ensures that features with larger numerical ranges do not dominate those with smaller ranges during model training [9]. For example, size or weight may have values in different units or magnitudes compared to colour intensity or texture metrics. Normalizing the data rescales all features to a standard range, typically between 0 and 1, or to a standard normal distribution with a mean of 0 and a standard deviation of 1. This step improves model convergence and ensures that algorithms sensitive to feature scales, such as Support Vector Machines (SVM) and Logistic Regression, perform optimally.

b. Labeling

After normalization, each sample is assigned a quality label based on SNI standards (Indonesian National Standards) and expert assessments. The labeling process involves classifying the shallot samples into predefined quality categories, such as "high quality," "medium quality," or "low quality." These categories are determined by specific criteria outlined in the SNI standards, which consider attributes such as size, color, texture, and overall condition of the shallots. Additionally, expert evaluations are incorporated to refine the labeling process and ensure accuracy. The labeled dataset serves as the target variable for supervised machine learning, enabling the models to learn the patterns and characteristics associated with each quality category.

By performing these preprocessing steps, the dataset is transformed into a standardized and structured format, ensuring that the machine learning models can effectively learn from the data and produce accurate classifications.

2.3. Model Training and Evaluation

Several machine learning methods are applied and evaluated to determine the most effective method for classifying the quality of Palu shallots:

2.3.1 Support Vector Machine

Support Vector Machines (SVM) are a class of machine learning algorithms that are widely used for classification tasks by creating a maximal margin hyperplane to separate different classes in the data [10]. SVMs have gained popularity due to their ability to handle high-dimensional data and their robustness in dealing with complex datasets [11]. These algorithms are known for their solid mathematical foundations, including properties like margin maximization and the use of kernels for nonlinear classification [12].

Given a dataset $\{(\boldsymbol{x_i}, \boldsymbol{y_i})\}$ where x_i represents feature vectors and $y_i \in \{-1,1\}$ represents class labels, the objective of SVM is to find a hyperplane that best separates the two classes. The decision function is given by:

$$f(x) = w^T x + b \quad (1)$$

where w is the weight vector and b is the bias term. The optimization problem for SVM can be formulated as:

$$\text{Min } \frac{1}{2}\|w\|^2 \quad (2)$$
$$y_i(x_i.w + b) \geq 1, for\ \forall i$$

To handle non-linearly separable data, SVM employs the kernel trick to map input data into a higher-dimensional space using kernel functions [13] . Common kernels include:

a. Linear Kernel: $K(x_i,x_j) = x_i^T x_j$

b. Polynomial Kernel: $K(x_i,x_j) = \left(x_i^T x_j + c\right)^d$

c. Radial Basis Function (RBF) Kernel: $K(x_i,x_j) = e^{-\gamma\|x_i - x_j\|^2}$

The SVM algorithm operates as follows [14] :

a. Kernel Selection: The user selects an appropriate kernel based on the type of data and the nature of the problem to be solved (e.g., linear, polynomial, RBF).

b. Model Training: During the training process, the SVM identifies support vectors, which are the data points lying around the maximal margin. These are the critical points that define the position of the hyperplane.

c. Prediction: After training is complete, the SVM can be used to make new predictions. New data are classified based on their relative position to the hyperplane learned during training. The distance of new data points from the hyperplane is used to determine the predicted class.

d. Model Evaluation: The performance of the SVM model is evaluated using appropriate metrics for the classification task (e.g., accuracy, precision, recall).

e. Parameter Tuning: In some cases, parameters such as the C parameter (which controls the trade-off between margin size and classification error) and gamma (for nonlinear kernels like RBF) need to be adjusted to improve model performance.

2.3.2 Decision Tree

Decision trees are a popular classification method [15]. The Decision Tree method is a technique in machine learning that utilizes a tree structure where each internal node represents a test on an attribute, each branch signifies the result of that test, and each leaf node represents a class label or output value. A decision tree splits the dataset into subsets based on attributes, aiming to maximize the separation between different classes or outputs [16].

The Decision Tree algorithm begins with attribute selection, where the best attribute is chosen to split the dataset. The selection criteria are based on measures such as Entropy and Information Gain in the ID3 algorithm or the Gini Index in C4.5 and CART [17]. The entropy of a dataset is defined as:

$$H(S) = -\sum p_i log_2 p_i \ , \quad (3)$$

here represents the proportion of instances belonging to class (Mitchell, 1997). The Information Gain for a given attribute is calculated as:

$$IG(S,A) = H(S) - \sum \frac{|S_v|}{|S|} H(S_v) \quad (4)$$

where represents the subset of after splitting on attribute. A higher Information Gain indicates a better attribute for splitting [18].

The Decision Tree algorithm operates as follows:

a. Attribute Selection: The algorithm selects the best attribute to split the dataset. The selection criteria are usually based on a measure that evaluates the quality of the split, such as Entropy and Information Gain in ID3, or the Gini Index in C4.5 and CART.

b. Dataset Splitting: The dataset is divided into subsets based on the values of the selected attribute. This process continues recursively on each subset, creating a tree structure.

c. Node Creation: Each node of the tree represents a selected attribute, and each branch from the node represents the value or range of values of that attribute.

d. Stopping: The splitting process continues until one of the stopping conditions is met:

   i. All examples in the subset have the same label.

   ii. No attributes remain for further splitting.

   iii. The tree reaches a predefined maximum depth.

e. Pruning: After the tree is built, pruning can be performed to reduce complexity and prevent overfitting. Pruning removes branches that contribute little to the model's accuracy.

2.3.3 Random Forest

Random Forests (RF) are a combination of several decision trees that can solve classification and regression problems, making it an ensemble learning method that contains several basic algorithms, namely decision trees [19]. The algorithm used is as follows:

a. For $b = 1$ until B do bootstrap sample $Z^*$of size N from train data

   i. Build a random forest tree T_b on the bootstrapped data, by repeating the following steps for each terminal node in the tree, until the minimum node size is reached. Choose m variable randomly on p variable

   ii. Take the best variable/splitting point between m

b. The output of the ensemble tree is $\{T_b\}_1^B$.

The RF algorithm operates as follows [20]:

a. Building Decision Trees:

   i. Data Sampling: From the original dataset with N samples, perform random sampling with replacement to create multiple bootstrap samples.

   ii. Feature Selection: For each node in the tree, select a random subset of the available features.

   iii. Node Splitting: From the selected subset of features, determine the feature and split point that provide the best separation based on metrics such as Gini impurity or information gain.

   iv. Build Trees: Continue this process until the tree is fully grown (or meets stopping conditions such as maximum depth or minimum number of samples in a node).

b. Combining Results (Ensembling):

   i. Classification: For classification problems, each tree in the forest provides a class prediction. RF combines these predictions using majority voting.

   ii. Regression: For regression problems, the predictions from each tree are averaged to provide the final value.

c. Prediction: Once the Random Forest model is built, use it to make new predictions. Input new samples into each tree in the forest, and combine the prediction results from each tree to provide the final prediction.

2.3.4 Logistic Regression

The logistic regression algorithm is a widely used classification algorithm in machine learning [21]. It is commonly applied in various fields, such as fraud detection, medical diagnosis, and recommendation systems [22]. Logistic regression is particularly suitable for binary classification tasks where the target

variable is binary, and the features are either continuous or categorical [23]. Logistic regression can be extended to handle multiclass classification problems using one-vs-rest (OvR) or softmax regression [24].

The principle of logistic regression lies in transforming a linear model into a probabilistic output using the logistic (sigmoid) function. The logistic function is defined as:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{5}$$

where z represents the linear combination of input features and weights:

$$z = w_0 + w_1 x_1 + w_2 x_2 + \cdots + w_n x_n$$

where, w_0 is the intercept (bias), w_i are the weights assigned to each feature x_1, and n is the number of features.

The output of the sigmoid function, $\sigma(z)$, produces a probability value between 0 and 1, which represents the likelihood that a given input belongs to the positive class. A classification decision is then made based on a predefined threshold, typically set at 0.5. If the probability is greater than or equal to 0.5, the sample is classified as belonging to the positive class; otherwise, it is classified as belonging to the negative class.

The principle of logistic regression:

a. Logistic model. Logistic regression uses the logistic (sigmoid) function to transform the linear output of regression into probabilities.

b. Probability output. The output of the sigmoid function is a value between 0 and 1, which is interpreted as the probability that a sample belongs to the positive class (e.g., class 1).

c. Classification decision. Based on the generated probability, a sample is classified into the positive class if its probability is above a certain threshold (typically 0.5) and into the negative class if it is below the threshold.

2.3.5 Naïve Bayes

Naive Bayes is a widely used algorithm in various domains, known for its simplicity and effectiveness in classification tasks. It is a probabilistic classifier that applies Bayes' theorem with the assumption of feature independence within a given category [25]. The algorithm calculates probabilities by summing frequencies and combinations of values from a dataset, making it straightforward and efficient for classification [26]. The Naïve Bayes algorithm is based on Bayes' Theorem, which is the theoretical foundation of probabilistic classification. Bayes' Theorem is formulated as follows:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \tag{6}$$

where: $P(A|B)$ is the posterior probability of the class (A, for example, the target class) given the predictor (B, for example, the feature), $P(B|A)$ is the likelihood, which is the probability of the predictor given the class, $P(A)$ is the prior probability of the class, and $P(B)$ is the prior probability of the predictor.

2.4 Performance Evaluation

The performance measurement of the classification algorithm in this research is conducted using a confusion matrix. Confusion matrix is a very popular measure used while solving classification problems [27]. It can be applied to binary classification as well as for multi-class classification problems. An example of a confusion matrix for binary classification is shown in Table 2.

Table 2. Confusion Matrix for Binary Classification

		Predicted	
		Negative	**Positive**
Actual	**Negative**	TN	FP
	Positive	FN	TP

Confusion matrices represent counts from predicted and actual values. The output "TN" stands for True Negative, which accurately shows the number of negative examples classified. Similarly, "TP" stands for Predicted Actual for True Positive, which indicates the number of positive examples classified accurately. The term "FP" shows a False Positive value, i.e., the number of actual negative examples classified as positive, and "FN" means a False Negative value, which is the number of actual positive examples classified as negative. One of the most commonly used metrics while performing classification is accuracy. Accuracy is the ratio of correctly predicted performance observations from the total observation [28], using Equation (7):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \tag{7}$$

Precision and recall are widely used and popular metrics for classification. Precision shows how accurate the model is for predicting positive values. Thus, it measures the accuracy of an expected positive outcome. It is also known as the Positive Predictive Value. Recall is helpful to measure the strength of a model to predict positive outcomes, and it is also known as the Sensitivity of a model. Both measures provide valuable information, but the objective is to improve recall without affecting the precision. Precision and Recall values can be calculated in Python using the "precision_score()" and "recall_score()" functions. Both of these functions can be imported from "sklearn.metrics". The formulas for calculating precision and recall are given below.

$$Precision = \frac{TP}{TP+FP} \tag{8}$$

$$\boldsymbol{Recall = \frac{TP}{TP+FN}} \tag{9}$$

3. RESULTS AND DISCUSSION

The dataset used in this study consists of 150 samples of Palu shallots, each characterized by 10 features representing key attributes such as size, color, texture, and moisture content. The dataset was divided into training and testing sets with an 80:20 ratio to evaluate the performance of the machine learning models. The training set was used to train the models, while the testing set provided an unbiased evaluation of their performance on unseen data. To ensure robustness and minimize the potential variability caused by data splitting, 5-fold Cross Validation was employed. This method splits the training data into five subsets, using four subsets for training and the remaining one for validation, rotating the validation subset in each iteration. This approach ensures that all data points are used for training and validation, providing a comprehensive assessment of model performance.

The study evaluated the effectiveness of five machine learning models: Naive Bayes, Decision Tree, Random Forest (RF), Support Vector Machine (SVM), and Logistic Regression, in classifying the quality of Palu shallots. These models were chosen due to their broad applicability in classification tasks and their differing underlying algorithms, offering diverse data handling approaches. Each model's performance was assessed using four critical evaluation metrics: Accuracy, Precision, Recall, and F1 Score. Accuracy measures the overall correctness of the model's predictions, Precision evaluates the ability to avoid false positives, Recall measures the ability to identify true positives, and F1 Score balances Precision and Recall, making it particularly useful for imbalanced datasets.

Hyperparameter tuning was conducted to optimize the models' effectiveness and optimize their performance. For example, parameters such as the number of estimators (n_estimators) and the maximum tree depth (max_depth) were tuned for Random Forest. In contrast, the kernel type and regularization parameter (C) were adjusted for SVM. This process ensured that each model operated at its optimal configuration, enabling fair comparisons.

The results, summarized in Table 3, reveal that while all models performed relatively well, their effectiveness varied depending on the metric evaluated. Random Forest emerged as the most reliable model, achieving the highest scores across all metrics, reflecting its robustness in handling complex data and its ability to capture intricate feature interactions. SVM also demonstrated strong performance, particularly in Recall and Precision, making it another viable choice for this classification task. In contrast, Naive Bayes,

while computationally efficient, showed lower performance due to its assumption of feature independence, which may not hold for the dataset used in this study.

Figure 1 illustrates the comparative performance of the models across the four evaluation metrics, providing a clear visualization of their strengths and weaknesses. Given their consistently high performance across multiple metrics, this analysis highlights the suitability of Random Forest and SVM for classifying Palu shallots. These findings emphasize the importance of selecting models that align with the specific requirements and characteristics of the dataset to achieve optimal classification outcomes.

Table 3. Performance Metrics and Parameters of Machine Learning Models

Model	Parameter	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Naive Bayes	-	85.3	89.8	94.7	93.2
Decision Tree	*max_depth* : 10	90.7	91.2	93.3	92.4
Random Forest	*n_estimators*: 100, *max_depth*: 15	95.4	93.6	94.9	93.5
SVM	*C*: 1.0, *kernel*: RBF	90.2	91.4	92.5	91.7
Logistic Regression	*solver*: liblinear	90.1	91.3	92.1	91.6

Performance analysis by metric:

a. Accuracy

Random Forest demonstrated the highest accuracy at 95.4%, outperforming all other models. This reflects its capability to effectively handle the complex and interdependent features of the dataset, such as size, texture, and color of shallots. Decision Tree and SVM followed closely with accuracy values of 90.7% and 90.2%, respectively, while Logistic Regression achieved 90.1%. In contrast, Naive Bayes had the lowest accuracy at 85.3%, indicating its limitation in datasets where feature independence assumptions do not hold. **Figure 1** (a) illustrates the comparison of accuracy values across models, emphasizing the superior performance of Random Forest.

b. Precision

Precision measures the proportion of correctly classified positive samples among all predicted positives. Random Forest achieved the highest precision (93.6%), demonstrating its ability to minimize false positives, which is crucial in applications where misclassification of lower-quality shallots as high-quality could lead to economic losses. Logistic Regression and SVM followed closely with precision values of 91.3% and 91.4%, respectively, while Decision Tree achieved 91.2%. Naive Bayes had the lowest precision at 89.8%, reflecting its difficulty in avoiding false positives in this classification task. **Figure 1** (b) shows the precision values of each model, highlighting Random Forest's advantage in minimizing false positives.

c. Recall

Recall reflects the proportion of correctly identified positive samples among all actual positives. Both Random Forest (94.9%) and Naive Bayes (94.7%) demonstrated strong recall, indicating their ability to correctly identify high-quality shallots across the dataset. Logistic Regression and SVM exhibited recall values of 92.1% and 92.5%, respectively, while Decision Tree recorded 93.3%. These results suggest that Random Forest and Naive Bayes are particularly effective in tasks where identifying true positives is critical. **Figure 1** (c) provides a visual comparison of recall values, showing that Random Forest and Naive Bayes excel in this metric.

d. F1 Score

F1 Score, the harmonic mean of precision and recall, balances these two metrics to provide an overall measure of model performance. Random Forest achieved the highest F1 Score (93.5%), showcasing its balanced performance across all metrics. Naive Bayes followed closely with an F1 Score of 93.2%, while Decision Tree, SVM, and Logistic Regression scored 92.4%, 91.7%, and 91.6%, respectively. The high F1 Score of Random Forest highlights its robustness and reliability in classification tasks requiring balanced precision and recall. **Figure 1** (d) depicts the F1 Score of each model, reinforcing Random Forest's leading performance.

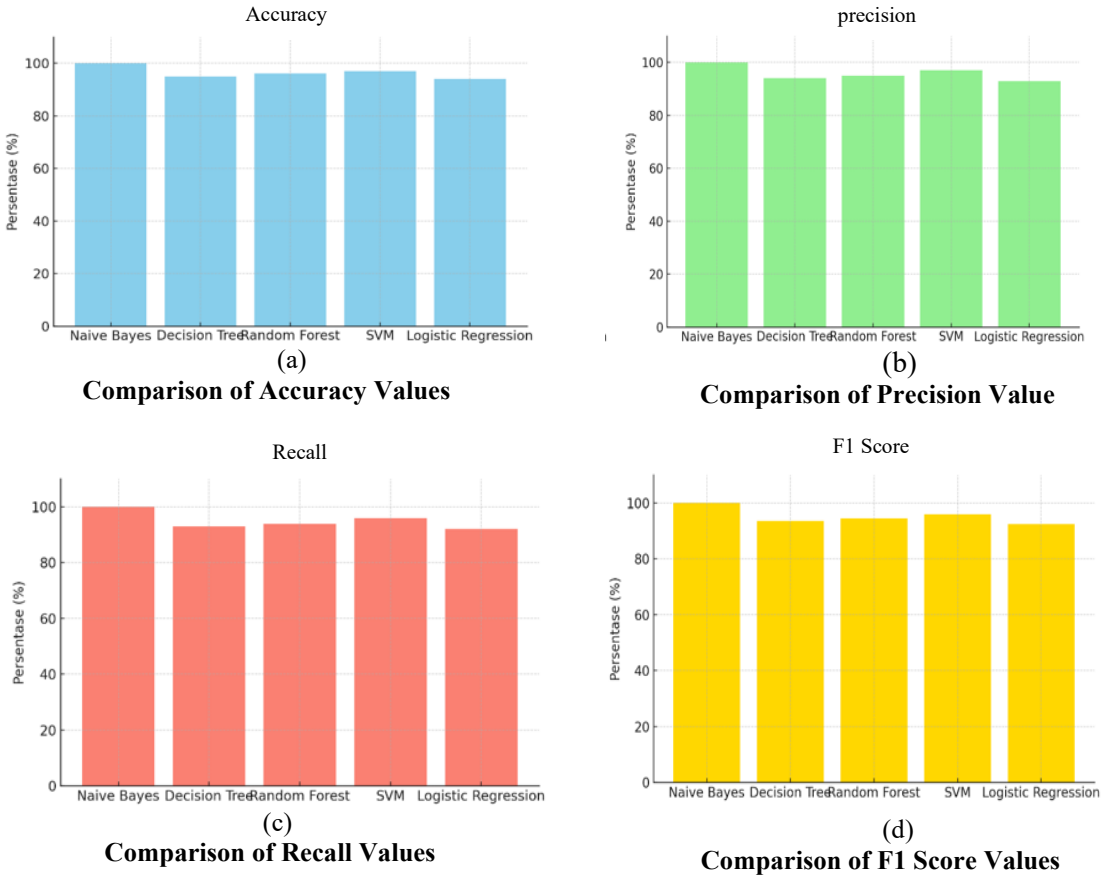

Figure 1. Performance Comparison of Machine Learning Models in Classifying Shallot Quality (a) Comparison of Accuracy Values – Evaluates the Overall Correctness of Predictions. (b) Comparison of Precision Values – Measures the Proportion of Correctly Classified Positive Samples. (c) Comparison of Recall Values – Assesses the Model's Ability to Identify True Positive Cases. (d) Comparison of F1 Score Values – Balances Precision and Recall for Overall Performance Evaluation.

The varying performances of the models reflect the different underlying algorithms used. As an ensemble model, RF combines multiple decision trees to provide a stable and accurate classification that excels in all evaluated metrics. The high accuracy, precision, and F1 Score of RF underscore its ability to handle complex feature interactions, which benefits the nuanced task of shallot quality classification. Similarly, SVM demonstrates competitive performance due to its ability to create distinct margins between classes, making it particularly effective in scenarios with clear decision boundaries. However, Naive Bayes shows relatively lower performance, likely due to its simplistic assumption of feature independence; in the context of shallot quality, where features such as color, texture, and size may be interdependent, Naive Bayes struggles to capture these complexities, leading to lower accuracy and precision.

These findings suggest that RF could be implemented to ensure high accuracy and balanced classification in practical settings, such as quality control within agricultural supply chains. For instance, RF's high F1 Score indicates its reliability in identifying high-quality shallots while minimizing the risk of misclassifying lower-quality products as high quality, thereby reducing sorting errors and enhancing consumer trust in product quality. On the other hand, SVM's high recall makes it suitable for applications where it is crucial to identify as many high-quality products as possible, even if it occasionally includes borderline cases. For future research, hybrid models or further feature engineering could be explored to improve classification accuracy. Techniques such as hyperparameter tuning or deep learning approaches, like convolutional neural networks for image-based quality assessment, may yield even better results, especially if visual quality features (e.g., colour and texture) are included. Additionally, expanding the dataset to incorporate more samples and varied environmental conditions, such as lighting and storage effects, would enhance the generalizability of the models.

4. CONCLUSIONS

This study compared five machine learning models for classifying the quality of Palu shallots. The results show that Random Forest (RF) and Support Vector Machine (SVM) performed best, with RF excelling in accuracy, precision, and F1 Score, while SVM was strong in recall. Naïve Bayes performed the weakest due to its assumption of feature independence. RF is recommended for reliable quality control, while SVM is suited for applications prioritizing recall. Future research should explore hybrid models, deep learning, and larger datasets to enhance performance. This study highlights the potential of machine learning for efficient agricultural quality classification.

ACKNOWLEDGMENT

We would like to thank the Faculty of Mathematics and Natural Sciences, Universitas Tadulako, for funding this research with Rector's Decree of Universitas Tadulako Number: 1413/UN28.16/AL.04/2024, dated May 29, 2024.

REFERENCES

[1] S. Musdalifah, D. Lusiyanti, and A. Sahari, "CLUSTERING THE QUALITY OF PALU SHALLOT (ALLIUM ASCALONIUM L.) USING THE KOHONEN ALGORITHM," 2023, doi: https://doi.org/10.1063/5.0133333.

[2] D. Lusiyanti, S. Musdalifah, A. Sahari, and Y. Darmawanti, "RANCANG BANGUN SISTEM CLUSTERING KUALITAS BAWANG MERAH PALU (ALLIUMASCALONIUM L.) MENGGUNAKAN ALGORITMA KOHONEN," *J. Ilm. Mat. Dan Terap.*, vol. 19, no. 1, pp. 103–110, 2022, doi: https://doi.org/10.22487/2540766X.2022.v19.i1.15940.

[3] Y. A. Pradana, L. P. Dewi, S. Muthiah, Y. Setyawati, F. R. S. Prakoeswa, and I. Untari, "DETEKSI KUALITAS BAWANG MERAH DENGAN CIRCULARITY IMAGE PRO-CESSING," *J. Keilmuan Dan Keislam.*, pp. 19–28, 2023, doi: https://doi.org/10.23917/jkk.v3i1.175.

[4] A. K. Mutia, Y. A. Purwanto, and L. Pujantoro, "PERUBAHAN KUALITAS BAWANG MERAH (Allium Ascalonicum L.) SELAMA PENYIMPANAN PADA TINGKAT KADAR AIR DAN SUHU YANG BERBEDA ((Allium ascalonicum l.) During Storage at Different Temperature and Water Content)," *J. Penelit. Pascapanen Pertan.*, vol. 11, p. 108, Jan. 2017, doi: https://doi.org/10.21082/jpasca.v11n2.2014.108-115.

[5] E. Hossain, M. F. Hossain, and M. A. Rahaman, "A COLOR AND TEXTURE BASED APPROACH FOR THE DETECTION AND CLASSIFICATION OF PLANT LEAF DISEASE USING KNN CLASSIFIER," in *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2019, pp. 1–6. doi: https://doi.org/10.1109/ECACE.2019.8679247.

[6] Z. H. Kok, A. R. Mohamed Shariff, M. S. M. Alfatni, and S. Khairunniza-Bejo, "SUPPORT VECTOR MACHINE IN PRECISION AGRICULTURE: A REVIEW," *Comput. Electron. Agric.*, vol. 191, p. 106546, 2021, doi: https://doi.org/10.1016/j.compag.2021.106546.

[7] N. Lestari, I. Indahwati, E. Erfiani, and E. Julianti, "A COMPARISON OF ARTIFICIAL NEURAL NETWORK AND NAIVE BAYES CLASSIFICATION USING UNBALANCED DATA HANDLING," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 17, no. 3 SE-Articles, Sep. 2023, doi: https://doi.org/10.30598/barekengvol17iss3pp1585-1594.

[8] D. Singh and B. Singh, "FEATURE WISE NORMALIZATION: AN EFFECTIVE WAY OF NORMALIZING DATA," *Pattern Recognit.*, vol. 122, p. 108307, 2022, doi: https://doi.org/10.1016/j.patcog.2021.108307.

[9] J. M. Johnson and T. M. Khoshgoftaar, "SURVEY ON DEEP LEARNING WITH CLASS IMBALANCE," *J. Big Data*, vol. 6, no. 1, p. 27, 2019, doi: https://doi.org/10.1186/s40537-019-0192-5

[10] D. Mustafa Abdullah and A. Mohsin Abdulazeez, "MACHINE LEARNING APPLICATIONS BASED ON SVM CLASSIFICATION A REVIEW," *Qubahan Acad. J.*, vol. 1, no. 2 SE-Articles, pp. 81–90, Apr. 2021, doi: https://doi.org/10.48161/qaj.v1n2a50.

[11] B. Pes, "Ensemble Feature Selection For High-Dimensional Data: A Stability Analysis Across Multiple Domains," *Neural Comput. Appl.*, vol. 32, no. 10, pp. 5951–5973, 2020, doi: https://doi.org/10.1007/s00521-019-04082-3.

[12] Y. Aksu, D. J. Miller, G. Kesidis, and Q. X. Yang, "MARGIN-MAXIMIZING FEATURE ELIMINATION METHODS FOR LINEAR AND NONLINEAR KERNEL-BASED DISCRIMINANT FUNCTIONS," *IEEE Trans. Neural Networks*, vol. 21, no. 5, pp. 701–717, 2010, doi: https://doi.org/10.1109/TNN.2010.2041069.

[13] C. Oganis, S. Musdalifah, and D. Lusiyanti, "KLASIFIKASI STATUS GIZI IBU HAMIL UNTUK MENGIDENTIFIKASI BAYI BERAT LAHIR RENDAH (BBLR) MENGGUNAKAN METODE SUPPORT VECTOR MACHINE (SVM) (STUDI KASUS DI PUSKESMAS LABUAN)," *J. Ilm. Mat. Dan Terap.*, vol. 14, no. 2, pp. 144–151, 2017, doi: https://doi.org/10.22487/2540766X.2017.v14.i2.9017.

[14] H. Li and Y. Ping, "RECENT ADVANCES IN SUPPORT VECTOR CLUSTERING: THEORY AND APPLICATIONS," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 29, Oct. 2014, doi: https://doi.org/10.1142/S0218001415500020.

[15] T. Naing and A. N. Oo, "DECISION TREE MODELS FOR MEDICAL DIAGNOSIS," *Int. J. Trend Sci. Res. Dev.*, vol. Volume-3, no. Issue-3, pp. 1697–1699, 2019, doi: https://doi.org/10.31142/ijtsrd23510.

[16] B. P. Battula, K. Krishna, and T. Kim, "An Efficient Approach for Knowledge Discovery in Decision Trees Using Inter Quartile Range Transform," *Int. J. Control Autom.*, vol. 8, no. 7, pp. 325–334, 2015, doi: 10.14257/ijca.2015.8.7.32.

[17] A. Purwanto, B. Sartono, and K. A. Notodiputro, "A COMPARISON OF RANDOM FOREST AND DOUBLE RANDOM

FOREST: DROPOUT RATES OF MADRASAH STUDENTS IN INDONESIA," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 19, no. 1 SE-Articles, Jan. 2025, doi: https://doi.org/10.30598/barekengvol19iss1pp227-236.

[18] S. Tangirala, "EVALUATING THE IMPACT OF GINI INDEX AND INFORMATION GAIN ON CLASSIFICATION USING DECISION TREE CLASSIFIER ALGORITHM*," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 2, 2020, doi: https://doi.org/10.14569/IJACSA.2020.0110277.

[19] A. Nugroho and A. Abdullah, "PERFORMANCE ANALYSIS OF RANDOM FOREST USING ATTRIBUTE NORMALIZATION," *Sistemasi*, vol. 11, no. 1, p. 186, 2022, doi: https://doi.org/10.32520/stmsi.v11i1.16811.

[20] S. Wang, "DIABETES PREDICTION USING RANDOM FOREST IN HEALTHCARE," *Highlights Sci. Eng. Technol.*, vol. 92, pp. 210–217, 2024, doi: https://doi.org/10.54097/5ndh9a05

[21] "MATHEMATICAL JUSTIFICATION ON THE ORIGIN OF THE SIGMOID IN LOGISTIC REGRESSION," 2022, doi: 10.57030/23364890.cemj.30.4.135.

[22] J. Liu, J. Cui, and C. Chen, "ONLINE EFFICIENT SECURE LOGISTIC REGRESSION BASED ON FUNCTION SECRET SHARING," 2023, doi: https://doi.org/10.1145/3583780.3614998.

[23] A. Al Imran, A. Rahman, H. Kabir, and S. Rahim, "THE IMPACT OF FEATURE SELECTION TECHNIQUES ON THE PERFORMANCE OF PREDICTING PARKINSON'S DISEASE," *Int. J. Inf. Technol. Comput. Sci.*, vol. 10, no. 11, pp. 14–29, 2018, doi: https://doi.org/10.5815/ijitcs.2018.11.02.

[24] Y. Wang, "ANALYSIS OF THE EVOLUTION OF MODERN CHINESE HISTORY BASED ON DATA MINING," *Appl. Math. Nonlinear Sci.*, vol. 9, no. 1, 2023, doi: https://doi.org/10.2478/amns.2023.2.00428.

[25] A. Wahid, F. Azim, and F. Firdausi, "APPLICATION OF DATA MINING TO CLASSIFY RECEIVING SOCIAL ASSISTANCE USING THE NAÏVE BAYES METHOD," vol. 1, no. 2, pp. 62–66, 2023, doi: https://doi.org/10.31967/inside.v1i2.881.

[26] S. Shang, "AN IMPROVED TEXT SENTIMENT ANALYSIS ALGORITHM BASED ON TF-GINI," 2018, doi: https://doi.org/10.23940/ijpe.18.09.p8.20082014.

[27] A. Luque, A. Carrasco, A. Martín, and A. de las Heras, "THE IMPACT OF CLASS IMBALANCE IN CLASSIFICATION PERFORMANCE METRICS BASED ON THE BINARY CONFUSION MATRIX," *Pattern Recognit.*, vol. 91, pp. 216–231, 2019, doi: https://doi.org/10.1016/j.patcog.2019.02.023.

[28] D. Ballabio, F. Grisoni, F. Grisoni, and R. Todeschini, "MULTIVARIATE COMPARISON OF CLASSIFICATION PERFORMANCE MEASURES," *Chemom. Intell. Lab. Syst.*, vol. 174, pp. 33–44, 2017, doi: https://doi.org/10.1016/j.chemolab.2017.12.004.