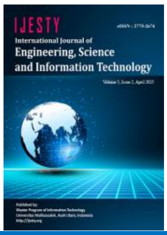
International Journal of Engineering, Science and Information Technology
Volume 5 No. 2 (2025) pp. 448-454
ISSN 2775-2674 (online)
Website: http://ijesty.org/index.php/ijesty
DOI: https://doi.org/10.52088/ijesty.v5i2.862
Research Paper, Short Communication, Review, Technical Paper



Comparison of the Results of the Weighted Moving Average Method and the Least Absolute Shrinkage and Selection Operator Method for Predicting Total Palm Oil Production at PT. Mora Niaga Jaya

Sakha Ardiansyah*, Rozzi Kesuma Dinata, Ar Razi

Department of Informatics, Faculty of Engineering, Universitas Malikussaleh, Aceh, Indonesia

**Corresponding author Email: sakha.200170176@mhs.unimal.ac.id*

The manuscript was received on 18 November 2024, revised on 28 December 2024, and accepted on 28 March 2025, date of publication 19 April 2025

Abstract

This study compares two prediction methods, Weighted Moving Average (WMA) and Least Absolute Shrinkage and Selection Operator (LASSO), in forecasting the total palm oil production at PT. Mora Niaga Jaya. Accurate forecasting is essential in the palm oil industry to support decision-making, optimize production planning, and manage supply chains efficiently. The WMA method produced more realistic prediction results, with a Mean Absolute Error (MAE) of 114,854 tons and a Mean Absolute Percentage Error (MAPE) of 220.45%, despite still having a considerable margin of error. These values suggest that while WMA is not perfectly accurate, it performs moderately well, given the complexity and variability inherent in agricultural production data. On the other hand, the LASSO method yielded significantly worse results, with an extremely high and unrealistic MAE and a MAPE of 291,456.000%, indicating that this approach is unsuitable for palm oil production forecasting in this specific case. The underperformance of the LASSO method may be due to the nature of the data used, which may not meet the assumptions required for LASSO to function optimally, such as linear relationships and minimal noise. This highlights the importance of aligning forecasting methods with the dataset's characteristics. Based on the comparison, it can be concluded that the WMA method is more appropriate for predicting palm oil production than LASSO. However, further steps such as parameter optimization, data normalization, and outlier removal should be undertaken to achieve better predictive accuracy. This research provides valuable insights into the importance of selecting the correct predictive method and ensuring data quality in forecasting. Ultimately, careful model selection and data preprocessing support effective operational and strategic decisions in the palm oil industry.

Keywords: *Data Mining, Palm Prediction, Comparison Method, WMA, Lasso.*

1. Introduction

The agricultural sector in Indonesia is divided into three types: plantations, rice fields, and fields. Among these three sectors, the plantation sector is more favored because it has a high selling value, can be cultivated on a large scale, and its appeal continues to increase. Plantation crops in Indonesia are dominated by palm oil, cocoa, rubber, sugarcane, and coffee. Among these crops, oil palm is the most profitable. The production results of palm oil significantly impact the demand for processed products, processing output targets, the budget for harvest workers' wages, and the production budget. Therefore, the company must conduct production calculations or predictions before determining the next production quantity. This prediction aims to understand the policies that must be implemented [1] [2] [3].

Prediction (forecasting) is an effort to estimate future events based on past events. This prediction is carried out by collecting historical data, which is then analyzed and calculated to identify specific patterns. These patterns are used to generate predictive data (forecasts) for the future[4]. Time series data prediction has many benefits that are powerfully felt in the present. Government agencies and private companies often use predictive results as a reference for making accurate decisions [5] [6] [7].

Selecting prediction techniques is a crucial step that must be taken by considering high accuracy results and low deviation between the predicted results and reality. In addition to accuracy, prediction techniques can also be chosen based on their responsiveness to data changes. One of the methods frequently used for making predictions is the Weighted Moving Average (WMA) method. This method applies a technique of assigning different weights to the available data, with the most recent data considered the most relevant for forecasting, thus given a higher weight. In applying this weighted moving average method, management or data analysts must first determine the weights (weight factors) of the existing data. The determination of these weights is subjective and depends on the experience and opinion of the data analyst [8].

Besides the WMA method, the Lasso method (Least Absolute Shrinkage and Selection Operator) is also one of the techniques used in prediction, especially in the context of regression. The Lasso method is a regularization technique that reduces model complexity by adding a penalty to the absolute value of the regression coefficients [9]. This technique effectively addresses the problem of overfitting that often occurs when the model has too many predictor variables.

PT. Mora Niaga Jaya, an oil palm plantation company, faces challenges in determining the optimal production quantity because production results significantly impact the demand for processed products and the company's budget. To assist in making more accurate decisions, predicting the total daily production of palm oil becomes crucial. The Weighted Moving Average (WMA) method is proposed because it gives greater weight to the most recent data, making it responsive to trend changes without assuming a specific data distribution [10]. In addition, the Lasso method (Least Absolute Shrinkage and Selection Operator) can also simplify prediction models by reducing insignificant variables, thereby increasing accuracy and avoiding overfitting [11]. The combination of these two methods is expected to produce more accurate predictions of daily palm oil production and support the company's strategic decisions more effectively.

2. Literature Review

2.1. Palm Oil

Oil palm is one of Indonesia's leading commodities, and it is growing rapidly. To achieve optimal results in oil palm cultivation, three essential factors must be considered: the environment, the land's physical properties, and the soil's or soil fertility's chemical properties. In commercial plantations, oil palm can grow well at temperatures between 24 and 28 degrees Celsius. To achieve optimal results, it is crucial to pay attention to the physical and chemical properties of the soil, including good soil structure and adequate drainage [12].

Oil palm plantations play a significant role in the country's economy and provide many jobs for the community. In October 2023, Indonesia's palm oil production reached 4.95 million tons, an increase of 9.17% from the previous month of 4.54 million tons. Oil palm plays an essential role in producing crude palm oil (CPO) and palm kernel oil (PKO) through the processing of Fresh Fruit Bunches (FFB), making it one of the leading commodities in Indonesia's agricultural sector [13].

2.2. Knowledge Discovery In Database

Knowledge Discovery in a Database (KDD) is a method for discovering unknown knowledge and information from a database.KDD is often a synonym for data mining, although the two terms have different concepts. However, both are interconnected, where data mining is the core of the entire KDD process. Data mining is a technique for discovering, searching, and extracting new information and insights from enormous datasets by integrating or combining it with other fields of science, such as statistics, artificial intelligence, and machine learning, producing valuable information [14].

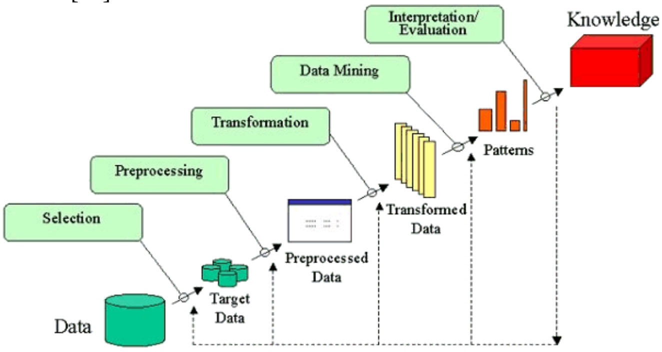


Fig 1. Knowledge Discovery In Database

2.3. Data Mining

Data mining is discovering or extracting new patterns from large datasets using methods from statistics and artificial intelligence [15]. This technique allows the discovery of desired patterns to extract useful information from large, complex databases. These patterns are identified using tools to provide valuable and in-depth data analysis. The results of this process can be further utilized with other decision support tools to generate more comprehensive insights and assist in better decision-making.

Data mining is one of the most common techniques in knowledge discovery in a database (KDD), but it is also essential. This is because Data Mining focuses on finding meaningful patterns in large data sets that may not be visible with traditional data analysis methods. This technique integrates various disciplines, such as statistics, artificial intelligence, and machine learning, to extract information from complex data. By using Data Mining, organizations can identify trends, anomalies, and relationships in data that can be used for various purposes, including marketing, fraud detection, risk management, and much more.

In business, Data Mining helps companies understand customer behavior, identify new market opportunities, and improve operational efficiency. In the scientific field, this technique is used to find patterns and relationships in research data that can lead to discoveries and innovations. Therefore, Data Mining is a tool for data analysis and a critical component of modern data-driven decision-making strategies [16].

2.4. Predictions

Prediction is estimating future events, which includes various aspects such as time, quality, quantity, and location to meet the demand for goods or services. Predictions can be classified based on the time frame into short-term, medium-term, and long-term, with short-term covering a few days to a few months ahead. The function of prediction is to analyze past data behavior to provide systematic solutions and increase confidence in the prediction results. Prediction methods are divided into two types: qualitative and quantitative. Qualitative predictions are based on opinions, surveys, and expert opinions, while quantitative predictions are based on historical data and have two main models: time series and causal. This prediction technique is widely used to estimate the optimal inventory of raw materials and products to meet consumer demand and achieve optimal production targets, thereby maximizing profits [17]

2.5. Time Series

Time series is a forecasting system that collects patterns from historical data to predict future data. The advantage lies in its ease of use and development without requiring complex systems like genetic algorithms or neural networks. In forecasting, time series models are generally used more frequently. This forecasting method utilizes two main theories, namely smoothing and decomposition. Based on the principle of Averaging Smoothing past error, Smoothing calculates the forecast value by adding previous forecast values and calculating the percentage error between the actual value and the forecast value. Meanwhile, the decomposition method divides time data into several main components, such as Trend, Cyclical, Seasonal, and Random Effects, which are combined to obtain more accurate predictions from each element, except for the random effects that are difficult to predict. This approach provides a more detailed structure for analyzing and predicting time series data patterns [18].

2.6. Weighted Moving Average

The Weighted Moving Average (WMA) method is a forecasting technique that considers the average value of a series of historical data, giving more weight to more recent data. This method is very effective when recent trend changes are more relevant and must be emphasized more than older data. In the WMA method, each historical data value is given a different weight, determining how much influence the data value has on the prediction. The main principle of the WMA method is to multiply each historical data value by a certain weight and then add up the results of these multiplications to get the predicted value. More weight is given to more recent data to have a greater influence on the forecasting results. Here is the general formula used in WMA:

$$WMA = \frac{\sum_{i=1}^{n}(w_i \times x_i)}{\sum_{i=1}^{n} w_i} \quad \text{.......................(1)}$$

Information :

x_i : Data value in period i

w_i : Weight is given to the i-th period data

n : Total number of data periods used

$\sum_{i=1}^{n}(w_i \times x_i)$: The sum of multiplying each data value by its weight.

$\sum_{i=1}^{n} w_i$: Total amount of weight given

The steps for using the Weighted Moving Average (WMA) method can be seen in the following explanation:

1. Determine the weight for each data period: The first step in the WMA method is determining the weight for each historical data. For example, if we have data from three periods (n3), we might assign a weight of 3 to the most recent period, 2 to the period before that, and 1 to the most distant period. These weights are assigned based on the importance of the data in determining the prediction. The closer the data is to the prediction time, the greater the weight.
2. Multiply each data value by the specified weight: After determining the weight, each historical data value is multiplied by its weight. For example, if historical data has weights, then we calculate them.
3. Adding the multiplication results: Add all the multiplication results from the second step to get the total weighted value. For example, $(w_1 \times x_1) + (w_2 \times x_2) + w_3 \times x_3$.
4. Adding all weights: Add all the weights determined to get the total weight. For example, $w_1 + w_2 + w_3$.
5. Generating predictions: The predicted value is obtained by dividing the total weighted value by the total weight: $WMA = \frac{(w_1 \times x_1)+(w_2 \times x_2)+(w_3 \times x_3)}{w_1+w_2+w_3}$

2.7. Least Absolute Shrinkage and Selection Operator

The Lasso (Least Absolute Shrinkage and Selection Operator) method is a regression technique used for prediction and variable selection in linear regression models. This method is very effective when there are many predictor variables, and only a few are relevant, and when there is a risk of overfitting due to too many variables. The main principle of the Lasso method is to add a penalty to the absolute sum of the regression coefficients during the model fitting process, which results in some regression coefficients being zero. This means that Lasso performs parameter estimation and variable selection automatically, retaining only the most significant variables in the model. Here is the general formula used in Lasso Regression:

$$Minimize = (\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij})^2 + \lambda \sum_{j=1}^{p} |\beta_j|) \quad (2)$$

Information :

y_i : The value of the data observed in period i

x_{ij} : The predictor j value for the data i

β_0 : *Intercept* or constant in the model

β_j : Regression coefficient for the j predictor variable

p : Total number of predictor variables

λ : Penalty parameter that controls the magnitude of the penalty on the absolute number of regression coefficients

$\sum_{j=1}^{p} |\beta_j|$: Absolute sum of all regression coefficients

The steps for using the Lasso method in making predictions are as follows:

1. Preparing Data and Predictor Variables: The first step is to prepare the dataset with relevant variables that may influence the target variable to be predicted. For example, in the case of predicting house prices, the predictor variables could include land area, building area, number of bedrooms, etc.
2. Performing Data Normalization and Standardization: Data must be normalized or standardized to ensure all variables are on the same scale. This is important because Lasso regression is sensitive to scale differences between predictor variables.
3. Choosing the Penalty Parameter (λ): Choosing the optimal value of λ is essential in Lasso. This λ value controls the strength of the penalty applied to the coefficients. This parameter is usually determined through cross-validation to find the value that produces the best model performance.
4. Applying Lasso Regression to the Data: After the value of λ is determined, Lasso Regression is applied to the data. This method will estimate the regression coefficients while applying a penalty to the absolute sum of those coefficients, thereby forcing some insignificant coefficients to become zero.
5. Model Evaluation and Result Interpretation: The final step is to evaluate the model's performance using metrics such as Mean Squared Error (MSE), R-squared, or other relevant metrics. A good Lasso model will balance prediction accuracy and model simplicity, retaining only the most significant variables.
   By following these steps, the Lasso method can be used to build an effective predictive model, helping to reduce the risk of overfitting and improve model interpretability with optimal variable selection.

2.8. Evaluation Metrics

Evaluation metrics are measures or parameters used to evaluate the performance or accuracy of a model or algorithm in predicting or classifying data [19]. These metrics help us understand how well or poorly the model handles a specific task. By using evaluation metrics, we can get an idea of how close the model's predictions are to the actual values of the data, allowing us to assess the reliability and accuracy of the model.

Some common examples of evaluation metrics are Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).MAE measures the average absolute error between the values predicted by the model and the actual values, providing a clearer picture of the average error without giving more weight to larger errors like RMSE. Meanwhile, MAPE measures the average absolute error as a percentage of the actual value, providing a perspective on how significant the error is relative to the predicted value [20].

Penggunaan metrik evaluasi yang tepat sangat penting dalam memastikan bahwa model yang dikembangkan memiliki kinerja yang memadai dan dapat diandalkan untuk tugas yang diinginkan. Selain MAE dan MAPE, terdapat berbagai metrik evaluasi lainnya yang dapat dipilih berdasarkan karakteristik dan kebutuhan spesifik dari tugas atau data yang sedang dianalisis.

3. Research methods

3.1. Place and Time of Research

The location and timing of this research are significant to ensure the smoothness and success of the study. This research will be conducted at PT. Mora Niaga Jaya is in Rantau District, Aceh Tamiang Regency, Aceh Province. This location was chosen because it is the main activity center for the total palm oil production data that will be investigated. The implementation of this research is planned to start from November 2024 until completion. The selection of this time considers the most representative and relevant conditions to obtain accurate and valuable data to understand the company's daily daily palm oil production movement.

3.2. System Schema

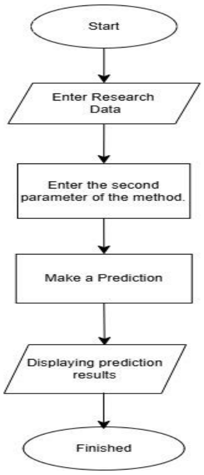


Fig 2. System Schema

Description of the system schema:
The system scheme illustrates the workflow carried out in the process of predicting palm oil production using two methods, namely Weighted Moving Average (WMA) and Least Absolute Shrinkage and Selection Operator (LASSO). The process begins with inputting the research data, namely the historical data on palm oil production that will be used for prediction. After the data is entered, the next step is to input the parameters of the two methods. Users need to specify the relevant parameters for each prediction method, such as the WMA period and the regularization value in LASSO.
Context: Users need to determine the relevant parameters for each prediction method, such as the WMA period and the regularization value in LASSO. After the parameters are specified, the system will make predictions by performing calculations using the two selected methods. Text to translate: After the parameters are determined, the system will make predictions by performing calculations using both chosen methods. The results of this prediction will then be displayed in the form of graphs and evaluation metrics, such as MAE (Mean Absolute Error) and MAPE (Mean Absolute Percentage Error), which indicate how accurate the predictions are compared to the actual data. This process ends with the completion step, which suggests that the prediction has been completed and the results are ready to be analyzed or used in palm oil production planning.

4. Results and Discussion

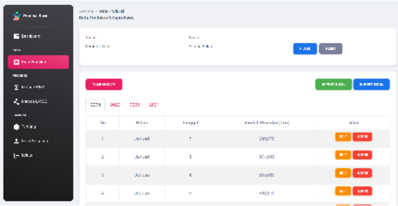
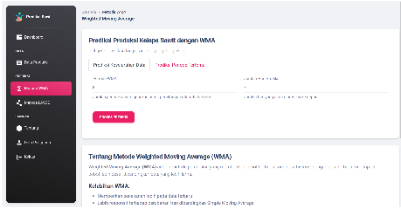
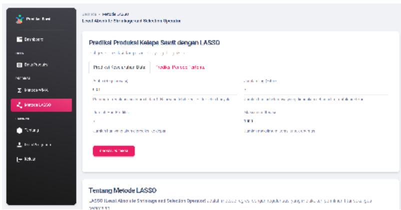
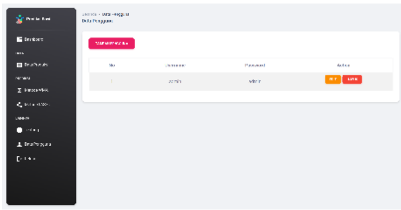
This study compares the Weighted Moving Average (WMA) method and the Least Absolute Shrinkage and Selection Operator (LASSO) in predicting the total palm oil production at PT. Mora Niaga Jaya. The analysis used production data over four years, namely 2021, 2022, 2023, and 2024. The prediction results are evaluated by measuring the accuracy level using Mean Absolute Percentage Error (MAPE) and Mean Absolute Error (MAE). The results show each method's performance in generating accurate predictions of total palm oil production. With this comparison, it is hoped that the most optimal method can be identified to support decision-making in the company's production planning. In addition, this research also provides an overview of the effectiveness of both approaches in handling the production data patterns used.

4.1. White Box Testing

At this stage, the White Box Testing method ensures that each component in the system runs according to the previously determined design. This testing focuses on testing the program's internal structure and logic flow, including checking conditions, loops, and code paths to detect possible errors or anomalies not detected by other testing methods. This approach thoroughly tests each function and subroutine to ensure no mistakes can affect the system's overall performance. Thus, White Box Testing helps ensure the developed system operates efficiently and according to the expected specifications.

Table 1. White Box Testing

No	Action	Testing	Output	Results
1	Opening the System	Displaying the Login page	Valid	

No	Action	Testing	Output	Results
2	Pressing the Production Data Menu	Displaying the production data page	Valid	
3	Pressing the WMA Process Menu	Displaying the WMA process page	Valid	
4	Pressing the Lasso Process Menu	Displaying the process pageLASSO	Valid	
5	Pressing the About Menu	Displaying the page about	Valid	
6	Pressing the User Data Menu	Displaying user data page	Valid	

5. Conclusion

This study compares two methods of predicting palm oil production: the Weighted Moving Average (WMA) and LASSO. The results show that WMA provides more realistic predictions than LASSO, although it still has a relatively high error rate. Meanwhile, LASSO produces very inaccurate predictions, making it unsuitable for use in this context. The simultaneous use of both methods is considered ineffective because their approaches and results are very different and do not complement each other. The data quality also plays a vital role in the accuracy of predictions, especially in the LASSO method, which is sensitive to unclean data or data that does not meet the model's assumptions. Overall, WMA is more suitable for predicting palm oil production because it can better adapt to trend changes. However, it still requires refinement of its parameters to make the predictions more accurate.

References

[1] S. J. Arditna and R. Kurniawan, "Prediksi Harga Tandan Buah Segar dengan Algoritma K-Nearest Neighbor," *J. Sist. Komput. dan Inform.*, vol. 5, no. 1, pp. 92–101, 2023, doi: 10.30865/json.v5i1.6818.

[2] V. Wineka Nirmala, D. Harjadi, and R. Awaluddin, "Sales Forecasting by Using Exponential Smoothing Method and Trend Method to Optimize Product Sales in PT. Zamrud Bumi Indonesia During the Covid-19 Pandemic," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 4, 2021, doi: 10.52088/ijesty.v1i4.169.

[3] S. Kumari and A. Harikrishnan, "Importance of Financial literacy For Sustainable Future Environment: A Research Among People In Rural Areas With Special Reference To Mandi District,Himachal Pradesh," *Int. J. Eng. Sci. Inf. Technol.*, vol. 1, no. 1, 2021, doi: 10.52088/ijesty.v1i1.36.

[4] C. A. Rahayu, "Prediksi Penderita Diabetes Menggunakan Metode Naive Bayes," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3, 2023, doi: 10.23960/jitet.v11i3.3055.

[5] R. Akbar, R. Santoso, and B. Warsito, "Prediksi Tingkat Temperatur Kota Semarang Menggunakan Metode Long Short-Term Memory (Lstm)," *J. Gaussian*, vol. 11, no. 4, pp. 572–579, 2023, doi: 10.14710/j.gauss.11.4.572-579.

[6] A. E. Syaputra and Y. S. Eirlangga, "Prediksi Tingkat Kunjungan Pasien dengan Menggunakan Metode Monte Carlo," *J. Inf. dan Teknol.*, vol. 4, no. 2, pp. 1–5, 2022, doi: 10.37034/jidt.v4i2.202.
[7] D. Abdullah, M. Farhan Aulia Barus, and M. Riansyah, "Forecasting Palawija Harvest Results In North Aceh Using Multiple Linear Regression Method," *Int. J. Artif. Intelegence Res.*, vol. 6, no. 1, 2022.
[8] S. Nurhayati and A. Syafiq, "Sistem Prediksi Jumlah Produksi Baju Menggunakan Weighted Moving Average," *J. Manaj. Inform.*, vol. 12, no. 1, pp. 14–24, 2022, doi: 10.34010/jamika.v12i1.6680.
[9] H. P. Stroke, "ANALISIS REGRESI LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR (LASSO) TERHADAP WAKTU TAHAN," vol. 01, no. 1, pp. 21–30, 2024.
[10] Y. J. Pradana, A. H. Mujianto, and T. Kistofer, "Implementasi Metode Weighted Moving Average Untuk Prediksi Pendapatan Di Ika Laundry Berbasi Web 53 IMPLEMENTASI METODE WEIGHTED MOVING AVERAGE UNTUK PREDIKSI PENDAPATAN DI IKA LAUNDRY BERBASI WEB," pp. 53–62, 2021.
[11] N. D. Ovalingga, N. Amalita, Y. Kurniawati, and Z. Martha, "Regularized Ordinal Regression with LASSO : Identifying Factors in Students ' Public Speaking Anxiety at Universitas Negeri Padang," vol. 2, no. 1999, pp. 475–482, 2024.
[12] A. N. Chansa *et al.*, "Upaya Pemanfaatan Limbah Kelapa Sawit," *J. Sains Student Res.*, vol. 2, no. 2, pp. 148–153, 2024.
[13] T. Pipit Muliyah, Dyah Aminatun, Sukma Septian Nasution, Tommy Hastomo, Setiana Sri Wahyuni Sitepu, "済無No Title No Title No Title," *J. GEEJ*, vol. 7, no. 2, pp. 317–323, 2020.
[14] S. Pujiono, R. Astuti, and F. Muhamad Basysyar, "Implementasi Data Mining Untuk Menentukan Pola Penjualan Produk Menggunakan Algoritma K-Means Clustering," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 1, pp. 615–620, 2024, doi: 10.36040/jati.v8i1.8360.
[15] Ar Razi, "Klasifikasi Penerima Beasiswa Aceh Carong (Aceh Pintar) Di Universitas Malikussaleh Menggunakan Algoritma Knn (K-Nearest Neighbors)," *J. Tika*, vol. 7, no. 1, pp. 79–84, 2022.
[16] F. M. Almufqi and A. Voutama, "Perbandingan Metode Data Mining Untuk Memprediksi Prestasi Akademik Siswa," *J. Tek.*, vol. 15, no. 1, pp. 61–66, 2023, doi: 10.30736/jt.v15i1.929.
[17] V. A. Lestari, A. Y. Ananta, and P. Basudewa, "Sistem Informasi Prediksi Persediaan Obat Di Apotek Naylun Farma Menggunakan Holt-Winters," *J. Inform. Polinema*, vol. 9, no. 2, pp. 229–236, 2023, doi: 10.33795/jip.v9i2.1289.
[18] Fungki Wahyu and Billy Hendrik, "Perbandingan Algoritma Time Series Dan Fuzzy Inference System Dalam Analisis Data Deret Waktu," *J. Penelit. Teknol. Inf. dan Sains*, vol. 1, no. 3, pp. 16–24, 2023, doi: 10.54066/jptis.v1i3.711.
[19] E. Erlin, Y. Desnelita, N. Nasution, L. Suryati, and F. Zoromi, "Dampak SMOTE terhadap Kinerja Random Forest Classifier berdasarkan Data Tidak seimbang," *MATRIK J. Manajemen, Tek. Inform. dan Rekayasa Komput.*, vol. 21, no. 3, pp. 677–690, 2022, doi: 10.30812/matrik.v21i3.1726.
[20] A. A. Suryanto, "Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi," *Saintekbu*, vol. 11, no. 1, pp. 78–83, 2019, doi: 10.32764/saintekbu.v11i1.298.