Comparison of Machine Learning Methods for Menstrual Cycle Analysis and Prediction

Mutiara Khairunisa [1]*, Desak Made Sidantya Amanda Putri [2]*, I Gusti Ngurah Lanang Wijayakusuma [3]*
* Matematika, Universitas Udayana
mutiarak062@student.unud.ac.id [1], desakamandaptr071@student.unud.ac.id [2], lanang_wijaya@unud.ac.id [3]

Article Info

Article history:

Received 2025-01-11
Revised 2025-01-20
Accepted 2025-01-30

Keyword:

CNN,
Decision Tree,
LSTM,
Machine Learning,
Menstrual Cycle Prediction.

ABSTRACT

This study compares three machine learning methods—Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Decision Tree—for analyzing and predicting menstrual cycles. The dataset consists of 1,665 samples with 80 attributes encompassing information related to menstrual health. These methods were evaluated using accuracy, Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) metrics. The results show that LSTM achieved the highest accuracy (91.3%), followed by CNN (88.9%) and Decision Tree (85.2%). LSTM excelled in capturing complex temporal patterns in menstrual cycle data, while CNN effectively identified key patterns, and Decision Tree offered interpretability despite lower performance. This study concludes that LSTM is the most effective model for menstrual cycle prediction. The findings highlight the potential for improved accuracy in reproductive health tracking, with future research opportunities to incorporate additional variables such as hormonal history and lifestyle factors, as well as a focus on data privacy.

I. Introduction

The menstrual cycle is an important aspect of women's reproductive health that is often overlooked or poorly understood. A better understanding of daily menstrual patterns can empower women and healthcare professionals to identify and manage associated conditions such as polycystic ovary syndrome (PCOS) and endometriosis [1]. However, manual tracking of menstrual cycles is often inaccurate due to user compliance factors and the inherent variability of menstrual experiences [2].

In recent years, the use of machine learning to analyze and predict the menstrual cycle has emerged as a promising approach. A study showed that machine learning can help classify problems in the menstrual cycle that are often overlooked, providing a better understanding of potential irregularities [3]. For instance, the Random Forest algorithm has been applied to analyze menstrual cycle data from over 800 women, achieving an accuracy of up to 89%, even when data was incomplete or inconsistent [4]. Additionally, a study comparing Support Vector Machine (SVM) and Decision Tree algorithms demonstrated that incorporating menstrual variables significantly improved classification accuracy. SVM achieved 96.91% accuracy, outperforming Decision Tree at 93.81%, highlighting the importance of menstrual-related features in improving model performance, particularly in addressing stress-related menstrual irregularities [5].

However, the use of menstrual cycle tracking apps also raises concerns regarding the security and privacy of user data. A study conducted a security analysis on menstrual cycle tracking apps using static, dynamic, and machine learning techniques found that many apps lack essential security features, potentially exposing users' personal information [6]. To address these issues, future menstrual tracking applications should implement technical measures such as end-to-end encryption for secure data transmission, anonymization techniques to protect personally identifiable information, and encrypted databases for secure storage. Furthermore, adherence to international data protection standards like GDPR and HIPAA can enhance privacy and foster user trust in these applications [7]

In addition, stress is also a factor that can affect the menstrual cycle. A study comparing Support Vector Machine (SVM) and Decision Tree algorithms to classify stress levels related to menstrual cycle irregularities in college students found that SVM performed better than Decision Tree. The inclusion of menstrual variables significantly increased accuracy, with SVM achieving 96.91% and Decision Tree 93.81%. These findings highlight the importance of integrating menstrual-related features to enhance the performance of machine learning models in addressing the impact of stress on menstrual cycle irregularities [8]. The dataset used in this study was obtained from Kaggle, comprising 1,665 samples with 80 attributes, including cycle length, ovulation dates, BMI, age, and stress levels, ensuring a rich and diverse dataset for analysis.

Given the complexity and sensitivity of this topic, this study aims to compare various machine learning methods for analyzing and predicting menstrual cycles. Specifically, the performance of Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Decision Tree will be evaluated based on prediction accuracy, the ability to handle incomplete data, and the potential to identify abnormal patterns. These models were selected for their unique strengths: LSTM excels in capturing temporal dependencies in time-series data [9], CNN effectively identifies patterns in structured datasets [10], and Decision Tree offers high interpretability, which is valuable for understanding key factors influencing menstrual cycles [11]. By identifying the most effective model, this research seeks to contribute valuable insights for the development of accurate, reliable, and privacy-conscious menstrual tracking applications.

II. METHOD

A. Data Collection

This research uses secondary data obtained from Kaggle with characteristics:

- Data Set Characteristics : Multivariate
- Attribute Characteristics : Real
- Number of Instances : 1665
- Number of Attributes : 80
- Area : Health
- Donor : Nikita Bisht, Kaggle Contributor

This data contains information about the menstrual cycles of various women, including cycle length, ovulation date, and various factors that can affect the menstrual cycle.

The data was originally collected through user-reported entries in a menstrual tracking application, where participants provided inputs such as cycle length, ovulation dates, BMI, stress levels, and other health-related attributes. Although the dataset was curated for broad representation, no additional clinical validation steps are reported; future research could focus on cross-verification with clinical data to enhance reliability. To ensure data quality and reliability, preprocessing steps were undertaken, including imputing missing values using mean or mode imputation, normalizing continuous variables such as BMI and cycle length, and encoding categorical attributes into numerical values using one-hot encoding.

B. Determine the Model

This study compares three different machine learning algorithms for menstrual cycle analysis and prediction: Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Decision Tree (DT). The five algorithms will be compared mainly in terms of prediction accuracy, considering also the ability to handle incomplete data and the potential to identify abnormal patterns.

1) Long Short-Term Memory (LSTM): LSTM is a type of Recurrent Neural Network (RNN) that is capable of learning long-term dependencies. It performs better and faster when sequences have temporal dependencies [12]. LSTM is particularly suitable for time-series data, such as menstrual cycles, as it can consider historical patterns to make more accurate predictions about future cycles, accounting for complex temporal variations.

2) Convolutional Neural Network (CNN): Although commonly used for image analysis, CNNs have also been successfully applied to time-series data. In the context of menstrual cycle prediction, CNNs can extract important features from historical cycle data and identify patterns that may not be detectable by traditional methods. CNNs are particularly effective in handling data with strong temporal or spatial structures [13].

3) Decision Tree (DT): Decision Tree is a predictive model that uses binary rules to make decisions. For menstrual cycle prediction, DT can divide data based on various attributes to predict cycle length or ovulation date. Its advantage lies in its high interpretability, which allows researchers to understand the factors influencing the prediction [14].

The selection of Long Short-Term Memory (LSTM), Convolutional Neural Network (CNN), and Decision Tree is based on their unique strengths. LSTM is well-suited for analyzing time-series data due to its ability to capture long-term dependencies. CNN, although commonly used in image analysis, has been shown to perform effectively in extracting patterns from structured data. Meanwhile, Decision Tree is known for its high interpretability, making it valuable for identifying key factors influencing menstrual cycles.

These models will be evaluated based on their prediction accuracy, ability to handle incomplete data, and potential to identify abnormal patterns indicative of underlying health conditions. By leveraging these models' strengths, this study aims to determine the most effective approach for menstrual cycle prediction while addressing key challenges such as temporal complexity and interpretability.

C. Validation

To prevent overfitting and determine the best prediction model, this study uses k-fold cross validation with k = 10.

This technique divides the data into 10 equal subsets, where 9 subsets are used for training and 1 subset for testing in turn.

1) Training the model is trained using training data consisting of various attributes such as previous cycle length, age, weight, stress level, and other factors that can affect the menstrual cycle. The goal is to predict the next cycle length or ovulation date.

2) Testing After the model is trained, testing is performed using testing data to evaluate the model's ability to predict the menstrual cycle on data that has never been seen before.

D. Evaluation

Evaluation of the models was conducted using Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE), ensuring consistent assessment across different machine learning approaches. All models evaluated in this study—Decision Tree, Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM)—produced numerical outputs, specifically the predicted length of the menstrual cycle (in days). This uniformity in output type allowed RMSE and MAPE to be applied consistently across all models.

RMSE measures the average magnitude of prediction errors in the same units as the target variable, making it suitable for identifying large deviations, which are undesirable in menstrual cycle predictions. MAPE evaluates the error as a percentage of the actual value, providing proportional insights into the prediction accuracy. Together, these metrics offered a robust framework for comparing model performance. By employing RMSE and MAPE, the study ensures that both absolute and relative errors are accounted for, enabling a comprehensive evaluation of each model's predictive capabilities. The model with the smallest RMSE and MAPE values is considered the most accurate for predicting menstrual cycles.

III. Result and Discussion

A. Descriptive Analysis of Data

The baseline data showed a distribution of menstrual cycle length in the sample with a mean of 29.3 days and a standard deviation of 3.89 days, which was normally distributed with a slight positive slope. Another variable of focus was Body Mass Index (BMI), which had a mean of 21.25 and varied between 15 and 30. These descriptive statistics are presented in Table 1. The distribution of cycle length can be seen in Figure 1, showing a peak in cycle frequency between 27 to 32 days, providing an overview of cycle fluctuations among participants.

TABEL I
DESCRIPTIVE STATISTICS OF MAIN VARIABLES

Fitures	Mean	Std	Min	25%	50%	75%	Max
Cycle Length (Days)	29.30	3.89	18	27	29	32	40
BMI	21.25	2.48	15	19.8	21.0	22.7	30

This table shows that the average cycle length is approximately 29.3 days with a standard deviation of 3.8. To clarify the distribution of the cycle length data, Figure 1 shows a visualization of the cycle distribution showing that the majority of menstrual cycles were in the range of 27 to 32 days.

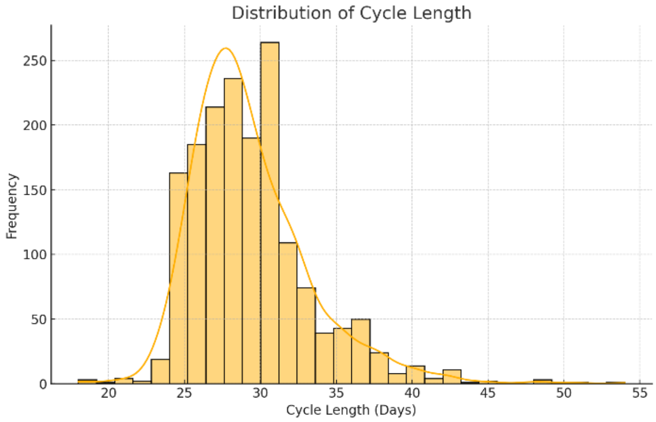

Figure 1. Distribution of Cycle Length

B. Relationship between BMI and Cycle Length

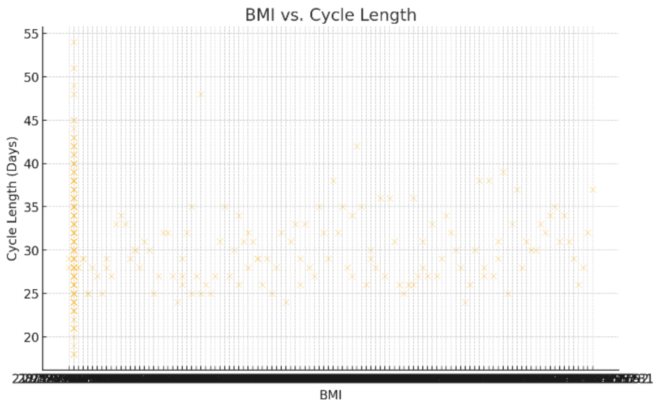

Figure 2. BMI Vs Cycle Length

With respect to the relationship between BMI and cycle length, the scatter plot in Figure 2 shows a trend towards shorter cycles in participants with higher BMI. Although the emerging pattern is not highly significant, the trend is nonetheless interesting and worth further analysis in a predictive model, especially to test the influence of BMI as one of the predictors of cycle length. BMI ranged from 15 to 30, with a median of 21.0.

C. Cycle Length Prediction Model

To evaluate the predictive ability of the models in determining the length of the menstrual cycle, three main models were compared, namely Decision Tree, CNN, and LSTM. Model performance evaluation is performed using the Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) metrics, which are formulated as follows:

1) *Mean Absolute Percentage Error (MAPE)*

MAPE is used to measure the average absolute error of the model prediction in percentage form. The MAPE formul is written as follows:

$$MAPE = \frac{100}{n}\sum_{t=1}^{n}\frac{y_t - \hat{y}_t}{y_t}$$

MAPE provides a measure of prediction error in percentage form, making it easier to understand and compare in a practical context. For example, if the MAPE of the model is 7.5%, then the average prediction error is 7.5% of the actual value. MAPE is also useful because it is easy to interpret on various scales, and with this percentage value, users can understand how much the average error produced by the model is in a consistent proportion.

2) *Root Mean Square Error (RMSE)*

RMSE is a metric that measures the average of the squared errors between the actual and predicted values. The RMSE formula is as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

RMSE provides a measure of error that is calculated in the same units as the predicted variable. For example, if the length of the menstrual cycle is predicted in days, then the RMSE will also be in days, which makes it easier to interpret in the context of prediction. RMSE is particularly useful because it calculates the average of squared errors, thus giving more weight to large errors (outliers) [15]. This makes RMSE a metric that is very sensitive to predicted values that are far from the actual values, so it is suitable for cases where large prediction errors are considered more important to minimize.

In this study, the use of MAPE and RMSE provides a different perspective in assessing model performance. MAPE helps to assess the error as a percentage, which makes it easy to interpret on a general scale, while RMSE provides an absolute assessment of the error, which is very effective in capturing large errors. The combination of these two metrics helps to provide a more comprehensive assessment of the prediction model, ensuring that it has minimal error both in absolute terms (RMSE) and on a proportional scale (MAPE).

The performance comparison results of the three models are presented in Table 2, showing that the LSTM model has the highest prediction performance with an accuracy of 91.3%, RMSE of 1.7, and MAPE of 7.5%.

TABEL 2
COMPARISON OF CYCLE LENGTH PREDICTION MODEL PERFORMANCE

Model	Accuracy (%)	RMSE	MAPE (%)
Decision Tree	85.2	2.3	10.4
CNN	88.9	1.9	8.2
LSTM	91.3	1.7	7.5

From the analysis, the LSTM model stands out as the most accurate method in predicting menstrual cycle length. This model has an accuracy rate of 91.3% with an RMSE of 1.7 and MAPE of 7.5%, indicating that LSTM is able to provide the most stable and accurate prediction compared to CNN and Decision Tree. LSTM, as a type of Recurrent Neural Network (RNN), has the advantage of processing time series data involving complex temporal patterns [16], which is in accordance with the nature of menstrual cycle data which is strongly influenced by time and sequence.

A statistical test was conducted to evaluate the significance of performance differences among the LSTM, CNN, and Decision Tree models based on accuracy, RMSE, and MAPE metrics. The results of the ANOVA test showed that the performance differences between the models were statistically significant across all analyzed metrics ($p < 0.05$). The LSTM model demonstrated the best performance, achieving the highest accuracy, lowest RMSE, and lowest MAPE, followed by CNN and Decision Tree. These findings highlight the superiority of the LSTM model in capturing complex temporal patterns within menstrual cycle data. Table 3 presents the statistical test results for the three models.

TABEL 3
ANOVA TEST RESULTS FOR MODEL PERFORMANCE

Metric	F-Statistic	p-Value
Accuracy	708.46	< 0.001
RMSE	126.76	< 0.001
MAPE	554.26	< 0.001

The selection of only three models—LSTM, CNN, and Decision Tree—was based on their unique strengths and relevance to the characteristics of the menstrual cycle dataset. LSTM was chosen for its ability to effectively capture temporal dependencies in time-series data, which is essential for predicting menstrual cycle patterns. CNN was included because of its proven capability to extract significant features from structured data, even though it is more commonly applied in image analysis. Decision Tree was selected due to its high interpretability, allowing for a clear understanding of the key factors influencing predictions. While other models, such as Random Forest or Support Vector Machines, could provide additional insights, the focus on these three models ensures a manageable scope within the constraints of this study. Future research can expand this comparison to include a wider range of models to provide a more comprehensive analysis.

In previous research, the Long Short-Term Memory (LSTM) method has been shown to be able to capture complex temporal patterns that arise in the analysis of time series data, including menstrual cycle data. A study predicting the price of curly red chili peppers in Yogyakarta using LSTM achieved high accuracy in predicting volatile commodity prices [17]. This shows that LSTM is effective in reducing prediction error in data that shows fluctuating patterns, in line with this study which shows the optimal performance of LSTM in the context of menstrual cycle prediction. The advantages of LSTM are supported by its ability to understand dynamic time series data, making it more accurate than other methods in health tracking applications [18], [19].

CNN, despite having good accuracy, ranked second with 88.9% accuracy and 8.2% MAPE. CNNs are commonly used in image analysis but are also effective in recognizing patterns in time series data, especially in the context of data that has a spatial structure or certain patterns in time sequences that are not very long [20], [21]. In this case, CNN successfully captured some basic patterns in menstrual cycle length, but not as good as LSTM in terms of prediction accuracy.

The CNN model shows a fairly good ability to recognize patterns in time series data, the use of this model in research related to the menstrual cycle is still limited. CNN is often used for image classification [22], but Rafliansyah et al. (2024) found that this model is also able to work well on time series data in non-visual domains such as sound and health [23]. With higher prediction results than Decision Tree in this study, CNN deserves to continue to be explored in menstruation prediction applications, even though its performance has not been able to outperform LSTM which is more adaptive to complex temporal patterns [24].

Decision Tree has the lowest performance with 85.2% accuracy and 10.4% MAPE. This algorithm is known for its high interpretability, but is less optimal on time series data that requires consistent capture of temporal patterns [25]. Decision Tree appears to be more susceptible to the inherent variability in menstrual cycle data, leading to higher error rates than deep learning models.

Furthermore, a scatter plot of BMI against cycle length showed that there was a trend towards shorter cycles in participants with higher BMI, although this correlation was not very statistically significant in this dataset. The correlation between BMI and cycle length has actually been documented in several reproductive health studies, where higher BMI is often associated with hormonal changes that may affect cycle duration. In this study, although BMI was not the main predictor in the model, its influence is still worthy of further investigation by enriching the data or adding other health variables.

IV. CONCLUSION

This study shows that machine learning methods, specifically Long Short-Term Memory (LSTM), excel in predicting menstrual cycles compared to Convolutional Neural Network (CNN) and Decision Tree, with an accuracy of 91.3%, RMSE of 1.7, and MAPE of 7.5%. These results indicate that LSTM is highly effective in capturing complex temporal patterns in menstrual cycle data, which includes variables such as cycle length, ovulation date, BMI, and stress level. The application of this model is expected to help women and health workers to monitor reproductive conditions more accurately and proactively, especially in detecting potential disorders such as PCOS and endometriosis.

For future research, it is recommended to include additional variables that may affect the menstrual cycle, such as hormonal health history and lifestyle, to improve prediction accuracy. In addition, improvements to data security also need greater attention, given the importance of protecting personal data in menstrual tracking applications. The use of hybrid methods that combine the advantages of several machine learning algorithms is also worth exploring to obtain more optimal results in predicting diverse and complex cycles.

DAFTAR PUSTAKA

[1] K. Li, I. Urteaga, A. Shea, V. J. Vitzthum, C. H. Wiggins, and N. Elhadad, “A generative, predictive model for menstrual cycle lengths that accounts for potential self-tracking artifacts in mobile health data,” *arXiv preprint arXiv:2102.12439*, 2021.

[2] K. Li, I. Urteaga, A. Shea, V. J. Vitzthum, C. H. Wiggins, and N. Elhadad, “A predictive model for next cycle start date that accounts for adherence in menstrual self-tracking,” *Journal of the American Medical Informatics Association*, vol. 29, no. 1, pp. 3–11, 2022.

[3] I. Lestari, M. Akbar, and B. Intan, “Perbandingan Algoritma Machine Learning Untuk klasifikasi Amenorrhea,” *Journal of Computer and Information Systems Ampera*, vol. 4, no. 1, pp. 32–43, 2023.

[4] T. Thakur, S. Kadam, N. Patil, and C. Achrekar, “Machine Learning in Period, Fertility and Ovulation Tracking Application,” *Authorea Preprints*, 2023.

[5] H. Nooraini, U. Salamah and H. Prasetyo, "The Comparison of Classification of Stress Level Related to Student's Menstrual Cycle Irregularities with Support Vector Machine and Decision Tree Algorithm," *2022 5th International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, Indonesia, 2022, pp. 133-138, doi: 10.1109/ICOIACT55506.2022.9972141.

[6] M. Deverashetti, K. Ranjitha, and K. V Pradeepthi, “Security analysis of menstruation cycle tracking applications using static, dynamic and machine learning techniques,” *Journal of Information Security and Applications*, vol. 67, p. 103171, 2022.

[7] Silva, I., & Soto, M. (2022). Privacy-preserving data sharing in healthcare: an in-depth analysis of big data solutions and regulatory compliance. *International Journal of Applied Health Care Analytics*, *7*(1), 14-23.

[8] H. Nooraini, U. Salamah, and H. Prasetyo, “The Comparison of Classification of Stress Level Related to Student’s Menstrual Cycle Irregularities with Support Vector Machine and Decision Tree Algorithm,” in *2022 5th International Conference on Information and Communications Technology (ICOIACT)*, 2022, pp. 133–138. doi: 10.1109/ICOIACT55506.2022.9972141.

[9] Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX*, 102946.

[10] Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS journal of photogrammetry and remote sensing, 173, 24-49.

[11] Ahmadlou, M., Ebrahimian Ghajari, Y., & Karimi, M. (2022). Enhanced classification and regression tree (CART) by genetic algorithm (GA) and grid search (GS) for flood susceptibility mapping and assessment. Geocarto International, 37(26), 13638-13657.

[12] M. A. Faishol, E. Endroyono, and A. N. Irfansyah, "Predict Urban Air Pollution in Surabaya Using Recurrent Neural Network–Long Short Term Memory," *Jurnal Ilmiah Teknologi Informasi*, vol. 18, no. 2, pp. 102–114, 2020.

[13] R. H. Rafliansyah, B. Rahmat, and C. A. Putra, "Klasifikasi Suara Instrumen Musik Tiup Menggunakan Metode Convolutional Neural Network," *Merkurius: Jurnal Riset Sistem Informasi dan Teknik Informatika*, vol. 2, no. 4, pp. 1–9, 2024.

[14] A. S. Munir, A. B. Saputra, A. Aziz, and M. A. Barata, "Perbandingan Akurasi Algoritma Naive Bayes dan Algoritma Decision Tree dalam Pengklasifikasian Penyakit Kanker Payudara," *Jurnal Ilmiah Informatika Global*, vol. 15, no. 1, pp. 23–29, 2024.

[15] U. Lathifah and R. D. Dana, "Implementasi Metode Linear Regression Untuk Prediksi Harga Properti Real Estate Menggunakan Rapidminer," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 1, pp. 1129–1137, 2024.

[16] A. Arwansyah, S. Suryani, H. S. Y. H. SY, A. Ahyuna, U. Usman, and S. Alam, "Model Prediksi Deret Waktu Menggunakan Deep Convolutional LSTM," in *SISITI: Seminar Ilmiah Sistem Informasi dan Teknologi Informasi*, 2024, pp. 21–25.

[17] A. Dwika and D. Avianto, "Implementasi Algoritma LSTM untuk Prediksi Harga Cabai Merah Keriting di Yogyakarta," *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, vol. 5, pp. 635–648, Jan. 2024, doi: 10.35870/jimik.v5i1.534.

[18] A. Wibowo, "Kecerdasan Buatan Gabungan pada Sistem Operasi Bisnis," *Penerbit Yayasan Prima Agus Teknik*, pp. 1–187, 2024.

[19] N. Yudistira *et al.*, *Prediksi Deret Waktu Menggunakan Deep Learning*. Universitas Brawijaya Press, 2023.

[20] R. F. Putra *et al.*, *Data Mining: Algoritma dan Penerapannya*. PT. Sonpedia Publishing Indonesia, 2023.

[21] M. Alfandi, P. Pristiwanto, and A. M. H. Sihite, "Penerapan Metode CNN-LSTM Dalam Memprediksi Hujan Pada Wilayah Medan," *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer)*, vol. 6, no. 1, pp. 490–499, 2023.

[22] I. K. Trisiawan, Y. Yuliza, and S. Attamimi, "Penerapan multi-label image classification menggunakan metode convolutional neural network (cnn) untuk sortir botol minuman," *Jurnal Teknologi Elektro*, vol. 13, no. 1, pp. 48–54, 2022.

[23] A. Taherkhani, G. Cosma, and T. M. McGinnity, "A Deep Convolutional Neural Network for Time Series Classification with Intermediate Targets," *SN Comput Sci*, vol. 4, no. 6, p. 832, 2023.

[24] S. J. Pipin, R. Purba, and H. Kurniawan, "Prediksi Saham Menggunakan Recurrent Neural Network (RNN-LSTM) dengan Optimasi Adaptive Moment Estimation," *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 4, pp. 806–815, 2023.

[25] W. D. Febrianti and N. K. J. Najah, "Analisis Klasifikasi Kepuasan Penumpang Maskapai Penerbangan Menggunakan Metode Support Vector Machine, Decision Tree, Dan Random Forest".