Journal of Computation Physics and Earth Science
ISSN: 2776-2521 (online)
Volume 5, Number 1, April 2025, Page 74 - 79
https://journal.physan.org/index.php/jocpes/index

Comparison of Random Forest and LSTM Methods for Temperature Prediction

Muhammad Asyril Ar Rosyad[1], Aviv Maghridlo[2]
[1,2]State of Meteorology Climatology and Geophysics Agency, Tangerang, Banten, Indonesia

Article Info

Article history:

Received March 11, 2025
Revised March 16, 2025
Accepted March 17, 2025

Keywords:

Long Short-Term Memory (LSTM)
Random Forest (RF)
Temperature Prediction
RMSE
MAE
R^2 Score
Machine Learning
Time Series Forecasting
Temperature Trends
Model Evaluation

A B S T R A C T

This study compares the performance of Long Short-Term Memory (LSTM) and Random Forest (RF) models in predicting temperature data from Tanjung Priok, Indonesia, using evaluation metrics such as RMSE, MAE, and R^2 Score. The LSTM model demonstrated its ability to capture temporal dependencies and temperature trends, achieving an R^2 score of 0.4493 and an MAE of 0.5863. In contrast, the RF model performed better in minimizing prediction errors, with a lower RMSE of 0.6498 and an R^2 score of 0.4066. While the LSTM model excelled in explaining variance in the temperature data, the RF model was more effective in stable periods, exhibiting lower prediction errors. The results highlight that both models have distinct advantages, with LSTM better suited for capturing long-term temperature trends and RF performing well during periods of stability. Future research could explore hybrid models or further optimization of these techniques to improve prediction accuracy, particularly for dynamic and extreme temperature fluctuations.

Corresponden Author:

Muhammad Asyril Ar Rosyad
State of Meteorology Climatology and Geophysics Agency
Tangerang City, Banten
Email: asyrilrosyad@gmail.com

1. INTRODUCTION

This journal discusses the comparison of the use of RF model and LSTM machine learning methods for predicting temperatures with a focus on accuracy assessment using RMSE, MAE and R2 Score. Temperature prediction plays a critical role in sectors such as agriculture, energy management, and disaster preparedness, particularly in regions like Indonesia, where climate patterns can vary significantly.

Because of the unpredictable and chaotic character of atmospheric conditions and our limited knowledge of the several atmospheric processes, weather forecasting is a difficult undertaking [1]. Temperature is one of the weather factors. On a numerical scale, temperature can be a physical quantity that represents the degree of hotness or coldness. It is measured using a thermometer, which can be calibrated in any of the various temperature scales, and may represent a degree of the warm liveliness per molecule of matter or radiation [2]. Predicting air temperature is one of the most important parts of studying the climate. When making decisions to solve ecological, industrial, or environmental issues, accurate temperature prediction can offer vital direction [3]. In this study, the data used is the daily mean temperature, which provides an average value of temperatures recorded throughout the day, offering a comprehensive representation of the temperature trends.

Machine learning can be used to estimate air temperature [4][5]. The logical consideration of computations and quantifiable models that computer systems use to carry out certain tasks without requiring explicit modification is known as machine learning (ML)[6]. LSTM and Random Forest are two of the several machine learning techniques [7][8][9].

One advancement in neural networks that can learn long-term dependencies is LSTM [10]. Recurrent neural networks are extended by LSTM networks, which overcome the vanishing gradient problem that plagues RNNs and enable us to recall and account for historical data in memory [11]. By highlighting contextually relevant information in textual data, LSTM models in conjunction with attention processes have the potential to greatly improve sentiment classification model performance [12].

When modeling the data, the Random Forest method is employed by combining multiple independent decision trees [13]. The algorithm splits the data into several random subsets and builds decision trees for each subset. This process is known as an ensemble method [13][14]. In the context of time series forecasting, Random Forest can capture non-linear patterns and complex relationships between variables affecting the data [15].

To determine which machine learning model is most suited for predicting air temperature, accuracy evaluation is applied to the models [16][17][5]. MAE, the distances between the estimated and observed values for N are averaged in this measure, which is an error statistic. The average squared discrepancy between the expected and observed temperature data is known as the MSE. For N, the root mean square error RMSE is as follows: This metric represents the standard deviation of the discrepancy between the actual observed data and the estimation. The sensitivity of this measure to large prediction mistakes is higher [5]. Additionally, the R^2 score is used to evaluate the proportion of variance in the observed data that is predictable from the independent variables [18]. A higher R^2 value indicates a better fit of the model to the data, with 1 being a perfect prediction and 0 indicating no predictive power. The R^2 score provides an additional layer of assessment in determining the overall accuracy of the temperature prediction model [18][19].

The main objective of this journal is to evaluate how well two time series prediction techniques—RF and LSTM—predict temperature. The goal of the study is to illustrate the advantages and disadvantages of each approach: LSTM's ability to identify complex non-linear patterns, particularly in irregular datasets, and RF ability to handle non-linear relationships, handle unstructured data, and be robust to overfitting, making it suitable for assessing feature importance. The journal also highlights the potential of hybrid models, which combine the strengths of both strategies to improve accuracy, as well as the approach of accuracy evaluation metrics to determine the most suitable method for temperature prediction in the study are.

2. RESEARCH METHOD

The method used in this study was based on LSTM and RF techniques, with accuracy evaluation using RMSE, MAE, and R^2 Score. The temperature data from Tanjung Priok, spanning from January 1, 2013 to December 31, 2023, was processed for this analysis. The data was obtained from the Climate Data Online website, created by NOAA. The collected data was used to compare the performance of LSTM and Random Forest models in predicting temperature, with a focus on evaluating the models' accuracy using the mentioned metrics. This study aims to determine which model provides more accurate predictions for temperature in the Tanjung Priok region, offering insights into the effectiveness of these machine learning approaches.

3. RESULTS

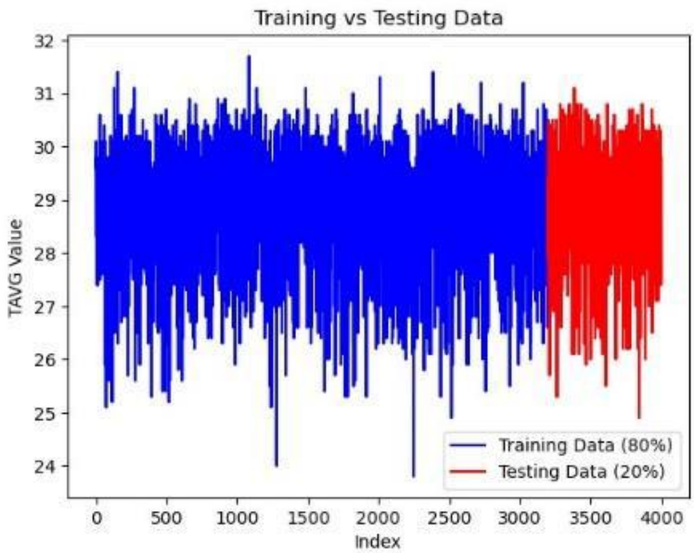

Fig. 3.1 Data Splitting

The temperature data from Tanjung Priok, spanning January 1, 2013, to December 31, 2023, was sourced from the Climate Data Online platform provided by NOAA. For the analysis, the dataset was divided into 80% for training and 20% for testing. The study compared the performance of LSTM and RF models in temperature prediction, with model accuracy evaluated using selected metrics.

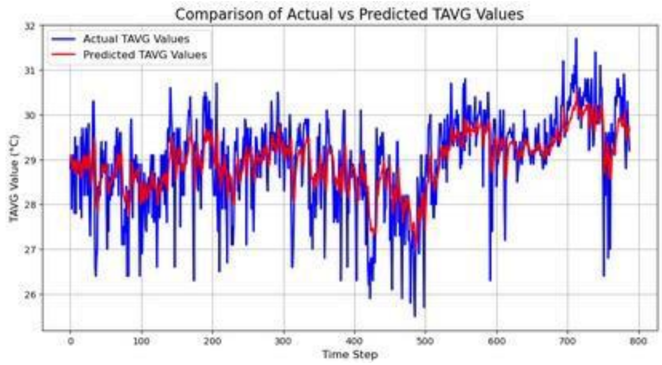

Fig. 3.2 Actual and Predicted TAVG using LSTM

The graph illustrates the comparison between actual and predicted average temperature (TAVG) values generated using the LSTM model. The blue line represents the actual TAVG values, while the red line indicates the predicted values obtained from the LSTM model. The results show a close alignment between the two lines, highlighting the LSTM model's ability to capture the temperature trends over the observed time steps accurately. Despite slight deviations at certain time intervals, the overall pattern suggests that the LSTM model effectively learns the temporal dependencies in the dataset and provides reliable predictions.

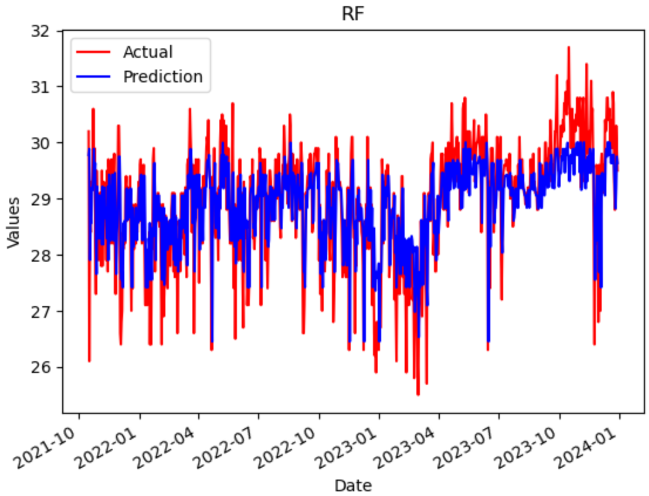

Fig. 3.3 Actual and Predicted TAVG using RF

The graph illustrates the comparison between actual and predicted temperature values generated using the RF model over the period from October 2021 to January 2024. The results show that the RF model successfully captures the overall trend and seasonal patterns of the temperature data, as indicated by the close alignment between the actual values (red line) and predicted values (blue line). However, discrepancies are observed during periods of extreme variations, where the model struggles to accurately predict sudden spikes or drops in temperature. Despite these deviations, the RF model performs well during relatively stable periods, demonstrating its capability to model the general behavior of the temperature data. These findings suggest that while the Random Forest model is effective for temperature prediction, further improvements, such as incorporating additional meteorological variables or optimizing model parameters, could enhance its performance in capturing dynamic and extreme temperature fluctuations.

Table 1. Evaluation Metrics

Model	RMSE	MAE	R^2 Score
LSTM	0,7742774935015089	0.5863260356302794	0,4493465040506345
RF	0,6498641163085757	0,6122487388374491	0,4066512320396348

The table presents the performance comparison between the LSTM and Random Forest models based on three evaluation metrics: RMSE, MAE, and R^2 Score. The LSTM model achieved an RMSE of 0.7743, an MAE of 0.5863, and an R^2 Score of 0.4493, while the RF model obtained a lower RMSE of 0.6498 and a slightly higher MAE of 0.6122, with an R^2 Score of 0.4066. These results indicate that while the RF model produces lower prediction errors (RMSE), the LSTM model explains a slightly higher proportion of the data variance (R^2 Score) and yields a lower MAE. Overall, both models show competitive performance, but LSTM demonstrates a marginally better ability to capture the underlying temperature trends.

4. DISCUSSION

The results from this study highlight the performance comparison between the LSTM and RF models in predicting temperature data. Both models have shown competitive results, though with subtle differences in terms of predictive accuracy.

In terms of RMSE, the LSTM model had a value of 0.7743, while the RF model had a lower RMSE of 0.6498. RMSE measures the average magnitude of error between predicted and actual values, with lower values indicating better model performance. The lower RMSE of the RF model suggests that it is more effective at minimizing the difference between actual and predicted temperatures. However, RMSE does not necessarily reflect the model's ability to capture the variance in the data, which makes other metrics like MAE and R-squared (R^2) crucial for further evaluation [20].

When comparing MAE, the LSTM model achieved a slightly better result with an MAE of 0.5863, compared to the Random Forest model's MAE of 0.6122. MAE measures the average of the absolute differences between predicted and actual values, providing a simpler view of the prediction accuracy [21]. The lower MAE of the LSTM model indicates that it produced slightly more accurate predictions in terms of the average magnitude of error, making it more reliable when absolute errors are a key concern.

Regarding the R-squared (R^2) value, which measures how well the model explains the variance in the data [17][22], the LSTM model outperformed the Random Forest model with an R^2 score of 0.4493, compared to 0.4066 for RF. The LSTM model explains around 44.93% of the variance in the temperature data, capturing more of the underlying temperature trends, especially in terms of temporal dependencies. In contrast, the RF model explains about 40.66% of the variance, which suggests that while it performs well overall, it may not capture as much of the temporal patterns present in the data.

Further analysis revealed that the LSTM model had an advantage in handling complex, sequential patterns in temperature data, allowing it to better capture long-term dependencies and fluctuations over time. This makes the LSTM model particularly valuable for scenarios where temperature forecasting requires consideration of historical trends and seasonality. On the other hand, the Random Forest model, which operates through ensemble learning and decision trees, excelled in its ability to minimize overall errors (as indicated by the RMSE) and performed consistently well for stable, less volatile periods in the dataset. However, its limitations in capturing sudden temperature spikes or drops suggest the need for further enhancements, such as incorporating additional meteorological features like humidity, wind speed, or atmospheric pressure.

Both models demonstrated competitive performance, but each excelled in different aspects. The LSTM model's ability to capture temporal dependencies and trends in the data gives it a slight edge in learning from time-series data, which makes it suitable for long-term temperature forecasting. On the other hand, the RF model performed better in minimizing prediction error (lower RMSE), making it effective for stable periods in temperature data. However, it struggled to accurately predict sudden spikes or drops in temperature, which could be improved by incorporating additional meteorological variables or optimizing its parameters.

In conclusion, this study highlights that model selection depends on the specific forecasting goals. The LSTM model is advantageous for scenarios requiring time-aware analysis and long-term trend detection, while the Random Forest model is more suitable for cases prioritizing overall error minimization and short-term stability. Future studies could explore hybrid approaches that combine the strengths of both models, or incorporate other deep learning architectures, such as GRU or CNN, to further improve predictive performance. Additionally, enriching the dataset with external factors like climate anomalies, El Niño events, or urbanization trends could enhance the robustness of temperature predictions for regions like Tanjung Priok.

5. CONCLUSION

This study compared the performance of LSTM and RF models for predicting temperature data from Tanjung Priok, using three evaluation metrics: RMSE, MAE, and R^2 Score. The LSTM model demonstrated superior performance in capturing temporal dependencies and trends in the data, with a higher R^2 score and lower MAE, which suggests it was more effective in understanding the underlying temperature patterns. This makes the LSTM model particularly suitable for long-term temperature forecasting, where learning from past temperature trends and making predictions based on these temporal dependencies is crucial. However, the LSTM model exhibited a higher RMSE, reflecting slightly larger prediction errors compared to the RF model.

On the other hand, the Random Forest model had a lower RMSE, indicating that it performed better in minimizing prediction errors, particularly in stable temperature periods. While RF was effective at handling non-linear relationships and avoiding overfitting, it struggled with sudden temperature spikes and extreme fluctuations, which led to discrepancies in its predictions. Despite this, the lower RMSE suggests RF is more robust in stable conditions. Both models demonstrated competitive strengths, with LSTM excelling in capturing trends and RF performing well in minimizing error. Future improvements could include hybrid models combining the strengths of both approaches, as well as incorporating additional meteorological variables or optimizing model parameters to better handle dynamic and extreme weather conditions.

REFERENCE

[1] E. De Saa and L. Ranathunga, “Comparison between ARIMA and Deep Learning Models for Temperature Forecasting,” 2020, [Online]. Available: http://arxiv.org/abs/2011.04452

[2] M. Asamoah-Boaheng, “Using SARIMA to Forecast Monthly Mean Surface Air Temperature in the Ashanti Region of Ghana Using Seasonal Autoregressive Integrated Moving Average (SARIMA) model, the study determined an adequate forecasting model for the mean temperature of Ashanti Reg,” Int. J. Stat. Appl., vol. 4, no. 6, pp. 292–299, 2014, doi: 10.5923/j.statistics.20140406.06.

[3] M. Yu, F. Xu, W. Hu, J. Sun, and G. Cervone, “Using Long Short-Term Memory (LSTM) and Internet of Things (IoT) for Localized Surface Temperature Forecasting in an Urban Environment,” IEEE Access, vol. 9, pp. 137406–137418, 2021, doi: 10.1109/ACCESS.2021.3116809.

[4] D. Fister, J. Pérez-Aracil, C. Peláez-Rodríguez, J. Del Ser, and S. Salcedo-Sanz, “Accurate long-term air temperature prediction with Machine Learning models and data reduction techniques,” Appl. Soft Comput., vol. 136, p. 110118, 2023, doi: 10.1016/j.asoc.2023.110118.

[5] J. Cifuentes, G. Marulanda, A. Bello, and J. Reneses, “Air temperature forecasting using machine learning techniques: A review,” Energies, vol. 13, no. 6, pp. 1–28, 2020, doi: 10.3390/en13164215.

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

[7] D. Gente, “P-issn 2244-4432 e-issn 2984-7125,” vol. 14, pp. 73–83, 2023.

[8] A. A. Akbar, Y. Darmawan, A. Wibowo, and H. K. Rahmat, “Accuracy Assessment of Monthly Rainfall Predictions using Seasonal ARIMA and Long Short-Term Memory (LSTM),” vol. 5, no. 2, pp. 99–114, 2024.

[9] M. Elsaraiti and A. Merabet, “A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed,” Energies, vol. 14, no. 20, 2021, doi: 10.3390/en14206782.

[10] J. Lu, G. Li, S. S. R. Moustafa, and S. S. Khodairy, “A comparison of SARIMA and LSTM in forecasting dengue hemorrhagic fever incidence in Jambi , Indonesia A comparison of SARIMA and LSTM in forecasting dengue hemorrhagic fever incidence in Jambi , Indonesia”, doi: 10.1088/1742-6596/1566/1/012054.

[11] S. Khedhiri, “Comparison of SARFIMA and LSTM methods to model and to forecast Canadian temperature,” Reg. Stat., vol. 12, no. 2, pp. 177–194, 2022, doi: 10.15196/RS120204.

[12] A. K. Wardhana, Y. Riwanto, and B. W. Rauf, “Performance Comparison Analysis on Weather Prediction using LSTM and TKAN,” vol. 04, pp. 3–10, 2024, doi: 10.31763/iota.v4i3.808.

[13] R. Meenal, P. A. Michael, D. Pamela, and E. Rajasekaran, “Weather prediction using random forest machine learning model,” Indones. J. Electr. Eng. Comput. Sci., vol. 22, no. 2, pp. 1208–1215, 2021, doi: 10.11591/ijeecs.v22.i2.pp1208-1215.

[14] W. Y. N. Naing and Z. Z. Htike, “Forecasting of monthly temperature variations using random forests,” ARPN J. Eng. Appl. Sci., vol. 10, no. 21, pp. 10109–10112, 2015.

[15] X. Wang, “Ladle furnace temperature prediction model based on large-scale data with random forest,” IEEE/CAA J. Autom. Sin., vol. 4, no. 4, pp. 770–774, 2017, doi: 10.1109/JAS.2016.7510247.

[16] S. Kendzierski, B. Czernecki, L. Kolendowicz, and A. Jaczewski, “Air temperature forecasts’ accuracy of selected short-term and long-term numerical weather prediction models over Poland,” Geofizika, vol. 35, no. 1, pp. 19–37, 2018, doi: 10.15233/gfz.2018.35.5.

[17] B. Azari, K. Hassan, J. Pierce, and S. Ebrahimi, “Evaluation of Machine Learning Methods Application in Temperature Prediction,” Comput. Res. Prog. Appl. Sci. Eng., vol. 8, no. 1, pp. 1–12, 2022, doi: 10.52547/crpase.8.1.2747.

[18] D. Chicco, M. J. Warrens, and G. Jurman, “The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation,” PeerJ Comput. Sci., vol. 7, pp. 1–24, 2021, doi: 10.7717/PEERJ-CS.623.

[19] A. V Tatachar, “Comparative assessment of regression models based On model evaluation metrics,” Int. Res. J. Eng. Technol., vol. 8, no. 9, pp. 853–860, 2021, [Online]. Available: https://d1wqtxts1xzle7.cloudfront.net/73250877/IRJET_V8I9127-libre.pdf

[20] Doreswamy, K. S. Harishkumar, Y. Km, and I. Gad, “Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models,” Procedia Comput. Sci., vol. 171, no. 2019, pp. 2057–2066, 2020, doi: 10.1016/j.procs.2020.04.221.

[21] A. Ahmad, K. A. Ostrowski, M. Maślak, F. Farooq, I. Mehmood, and A. Nafees, “Comparative study of supervised machine learning algorithms for predicting the compressive strength of concrete at high temperature,” Materials (Basel)., vol. 14, no. 15, 2021, doi: 10.3390/ma14154222.

[22] R. M. Adnan et al., “Air temperature prediction using different machine learning models,” Indones. J. Electr. Eng. Comput. Sci., vol. 22, no. 1, pp. 534–541, 2021, doi: 10.11591/ijeecs.v22.i1.pp534-541.