Journal of Technomaterial Physics
Journal homepage: https://talenta.usu.ac.id/JoTP

Comparison of Artificial Neural Network Models for Rainfall Prediction in Palu City

Arya Zaki Ramadhan*, Febby Debora Abigael, and Muhammad Fany Nurwibowo
State College of Meteorology Climatology and Geophysics, Tangerang, 15119, Banten, Indonesia
*Corresponding Author: aryazakiramadhan@gmail.com

ARTICLE INFO

Article history:
Received 1 November 2024
Revised 4 February 2025
Accepted 7 February 2025
Available online 28 February 2025

E-ISSN: 2656-0755
P-ISSN: 2656-0747

How to cite:
A. Z. Ramadhan, F. D. Abigael and M. F. Nurwibowo, " Comparison of Artificial Neural Network Models for Rainfall Prediction in Palu City," Journal of Technomaterial Physics, vol. 07, no. 01 pp., pp. 22-27, Feb. 2025, doi: 10.32734/jotp.v7i1.19757.

ABSTRACT

Rainfall prediction is crucial to support natural disaster mitigation and water resource management, especially in areas like Palu City with dynamic rainfall patterns. This study evaluated the performance of three Artificial Neural Network (ANN) models with different architectures to identify the most accurate model in predicting rainfall in 2023. To obtain the model, the historical data of nine meteorological parameters in Palu City from 2018 to 2022 was processed using the Python programming language through pre-processing, processing, post-processing, and verification stages. All three models obtained are designed with hidden layers and different nodes. The best model obtained was Model A with one hidden layer, 8 nodes, and a MAPE value of 9.42%, putting it in the excellent category. Meanwhile, Model B and Model C are in a suitable category with MAPE values of 14.43% and 10.23%. The challenge of using the ANN method in predicting rainfall is its tendency to equalize extreme rain. Therefore, complete data is needed to improve ANN performance.

Keywords: Artificial Neural Network, Meteorological Parameters, Palu City, Python, Rainfall Prediction

ABSTRAK

Prediksi curah hujan sangat penting untuk mendukung mitigasi bencana alam dan pengelolaan sumber daya air, khusussnya di wilayah seperti Kota Palu yang memiliki pola curah hujan yang dinamis. Penelitian ini mengevaluasi kinerja tiga model Artificial Neural Network (ANN) dengan arsitektur yang berbeda untuk mengidentifikasi model yang paling akurat dalam memprediksi curah hujan tahun 2023. Untuk mendapatkan model, data historis sembilan parameter meteorologi Kota Palu tahun 2018 hingga 2022 diolah menggunakan bahasa pemrograman Python melalui tahapan pra-pemrosesan, pemrosesan, pasca-pemrosesan, dan verifikasi. Ketiga model yang diperoleh dirancang dengan lapisan tersembunyi dan node yang berbeda. Model terbaik yang diperoleh adalah Model A dengan satu lapisan tersembunyi, 8 node, dan nilai MAPE sebesar 9,42% yang menjadikannya pada kategori sangat akurat. Sedangkan Model B dan Model C berada pada kategori baik dengan nilai MAPE sebesar 14,43% dan 10,23%. Tantangan penggunaan metode ANN dalam memprediksi curah hujan adalah kecenderungannya dalam memeratakan nilai ekstrem. Oleh karena itu, data yang lengkap diperlukan untuk meningkatkan kinerja ANN.

Kata kunci: Artificial Neural Network, Parameter Meteorologi, Kota Palu, Python, Prediksi Curah Hujan

1. Introduction

One of the aspects to consider in decision-making during daily activities is the weather and climate conditions [1]. One of the weather parameters that has a significant impact is rainfall [2]. There are many challenges in predicting rainfall due to various factors influencing it, such as geographic location, dynamic atmospheric conditions, and maritime factors [3]. Therefore, a method that can accurately predict rainfall is required. The Artificial Neural Network (ANN) is an effective method for predicting rainfall.

Previous research has demonstrated the success of the ANN method in predicting rainfall in Bangkok with a lead time of 1-3 hours [4]. Utilizing historical weather data, ANN has also proven effective in predicting rainfall in Indonesia [2]. Furthermore, applying a three-layer ANN model on daily weather data at the Kemayoran Meteorological Station has shown increased accuracy [5]. Nevertheless, the results from previous studies remain varied due to the quantity and types of data that have yet to be optimized.

Based on data from BPS (Badan Pusat Statistik), Palu City has daily rainfall in 2023, reaching 587.55 mm with a monthly average of 48.95 mm, where monthly rainfall tends to be dynamic with varying differences. This shows the need to predict rainfall in Palu City. This study aims to maximize the existing data in predicting rainfall in Palu City using various weather parameters. The comparison of the ANN model in this study is intended to obtain accurate results so that it can contribute to mitigating the impact of natural disasters and support better water resources management planning [6].

2. Methods

2.1. Data

The data used consists of two types, namely training data and testing data. The training data includes data on rainfall, humidity, air pressure, duration of solar irradiation, air temperature (maximum, minimum, and average), as well as wind speed and direction [2], [5] in the period 2018 to 2022. Meanwhile, the testing data includes rainfall data in 2023. Both data types were taken from The Mutiara SIS Al-Jufri Class II Meteorological Station in Palu City, Central Sulawesi, as shown in Figure 1, through the BMKGSoft and Ogimet portals.

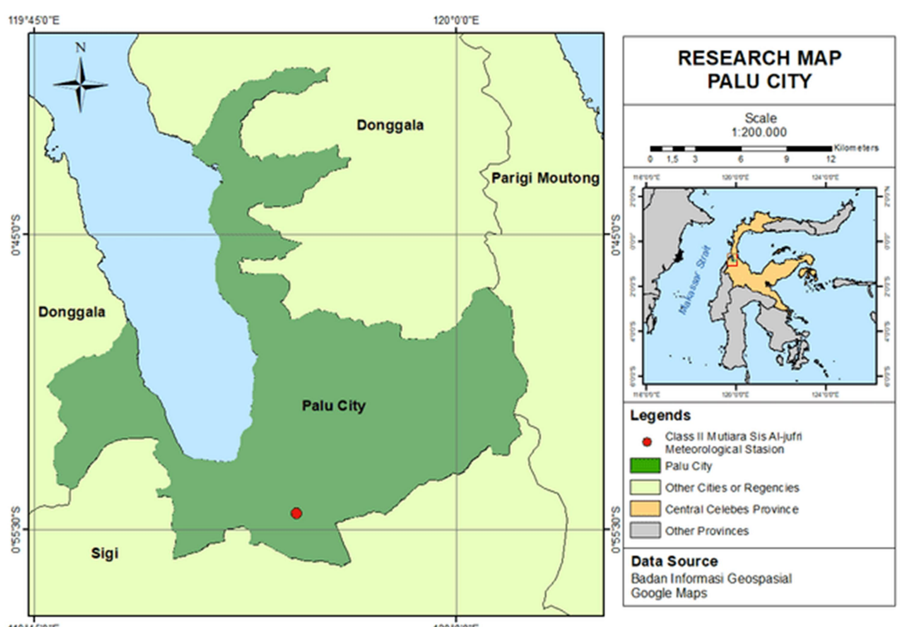

Figure 1. Research map

2.2. Research procedures

Data processing uses Google Colab with Python as the programming language. Data processing goes through the following stages:

1) Pre-processing: cleans data collected in .xlx format, where blank data with 0 is filled in because the value of 0 is considered neutral and does not make a significant positive or negative contribution to calculating the weight and activation of neurons. Then, normalize the data to the range [0,1, 0,9] because the output value does not reach 0 or 1 if using the sigmoid activation function, so it does not

use the range [0, 1] [7]. Then, divide the data into 80% training data and 20% test data [8].

2) Processing: using a Sequential Model with the number of layers built with the sigmoid activation function [9], configuring the data with Adam Optimizer, monitoring the performance of the model with MAE (Mean Absolute Error), MSE (Mean Squared Error), and RMSE (Root Mean Squared Error) [10] in Equations (1) to (3) [11], [12], [13], and training the data for 500 epochs with 10% validation data.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \tag{1}$$

$$MSE = \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n} \tag{2}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}} \tag{3}$$

Where n is the total number of data, x_i is the actual value of the i^{th} data, and y_i is the predicted value of the i^{th} data.

3) Post-processing: using the model.predict function to predict based on testing data.
4) Denormalization: returns data from the normalized scale to its original scale through Equation (4) [14].

$$vi = \frac{(vi' - new_{minA})(max_A - min_A)}{(new_{maxA} - new_{minA})} + min_A \tag{4}$$

Where v_i is the data of the i^{th} denormalization result, v_i' is the data of the i^{th} normalization result, max_A is the maximum value of the original data, min_A is the minimum value of the original data, new_{maxA} is the maximum value on the scale used for normalization (0.9), and new_{minA} is the minimum value on the scale used for normalization (0.1).

2.3. Data Verification

The data is verified by analyzing the MAPE (Mean Absolute Percentage Error) value, which can be seen in Equation (5), where a smaller value indicates a more accurate prediction [15].

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i}\right| \times 100\% \tag{5}$$

The obtained MAPE value has an accuracy classification that can be seen in Table 1.

Table 1. Classification of MAPE values [16], [17]

MAPE Values	Prediction Categories
<10%	Excellent
10% - 20%	Good
20% - 50%	Reasonable
>50%	Inaccurate

3. Result and Discussion

3.1. Comparison of Models

The accuracy of the rainfall forecast for 2023 was evaluated for three ANN models (Model A to C in Figure 2) during the initial testing phase. In this experiment, three models with different hidden layer structures were trained, each for 500 epochs, to predict rainfall. The models were evaluated using metrics such as MAE, MSE, and RMSE, allowing for a comprehensive comparison of their predictive capabilities. The results of this evaluation are summarized in Table 2.

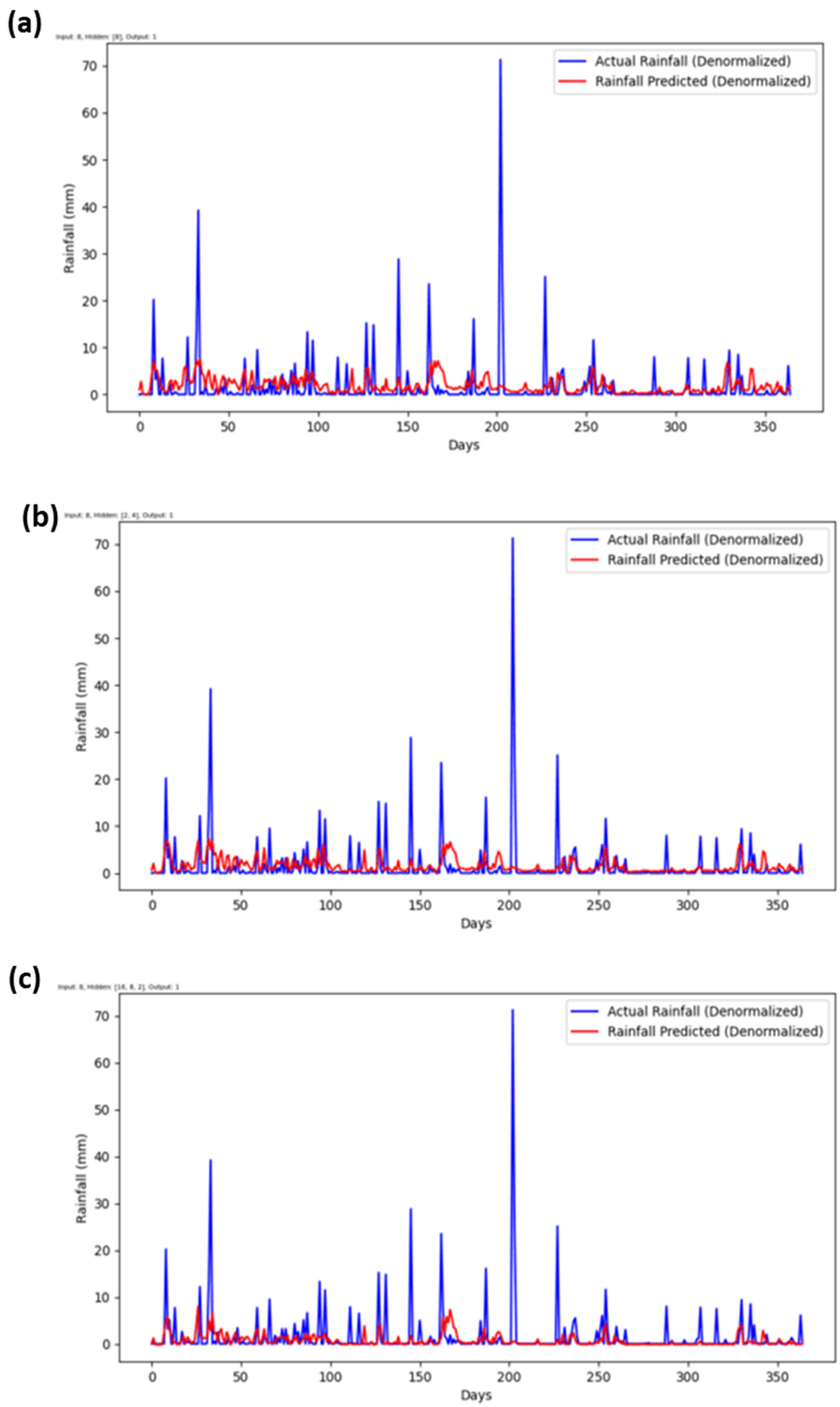

Figure 2. Actual and prediction rainfall of (a) Model A with 1 hidden layer, (b) Model B with 2 hidden layers, and (c) Model C with 3 hidden layers

Table 2. Results and evaluation of three models

Sample	Architecture	Learning Rate	Epoch	MAE	MSE	RMSE
Model a	[8] [8] [1]	0.01	500	1.92	33.09	57.53
Model b	[8] [2 4] [1]	0.01	500	2.45	32.68	57.16
Model c	[8] [16 8 2] [1]	0.01	500	2.05	32.83	57.30

From Figure 2 and Table 2, the first model (Model A), consisting of a single hidden layer with 8 nodes, achieved an MAE of 1.92, an MSE of 33.09, and an RMSE of 57.53. This relatively simple architecture demonstrates solid performance, indicating its ability to capture key patterns in the data while maintaining computational efficiency. This is because the model is more likely to generalize well to unseen data rather than memorizing noise in the training dataset. However, its simplicity might limit its capacity to learn more complex relationships, potentially leaving room for improvement.

The second model (Model B) featured two hidden layers with 2 and 4 nodes, respectively. Despite its deeper architecture, this model achieved a slightly higher MAE of 2.45, with an MSE of 32.68 and an RMSE of 57.16. These results suggest that while the two-layer structure may offer enhanced flexibility, the low number of nodes in each layer may have constrained its capacity to generalize effectively. Additionally, the increased depth did not significantly improve the simpler model, highlighting the importance of selecting appropriate layer sizes.

The third model (Model C) utilized three hidden layers with 16, 8, and 2 nodes, respectively. It achieved an MAE of 2.05, an MSE of 32.83, and an RMSE of 57.30. This deeper and more complex architecture demonstrates good performance but does not outperform the single-layer model significantly. While the increased number of nodes in the first two layers might have improved its ability to capture intricate patterns, adding a third layer with only 2 nodes may have limited its effectiveness. Furthermore, the added complexity likely increased the computational cost without delivering a proportional gain in accuracy.

In all three models, the predicted rainfall pattern generally follows the overall shape of the actual data. However, there are striking differences in amplitude, especially at certain peaks. One example is the model's tendency to underestimate extreme spikes in precipitation, which results in flattening effects. In addition, the large amount of blank data on the rainfall parameters used as training data is also a factor in this striking difference. Nonetheless, all models achieved low error values in the MAE, MSE, and RMSE metrics, thus still showing good potential for rainfall prediction.

3.2. Models Performance Analysis

When evaluating the performance of predictive models, MAPE provides valuable insights into the accuracy of predictions relative to actual values. However, even though MAPE shows mean percentage errors, MAPE does not fully capture the nuances of model performance, especially when the model fails to represent critical patterns in the data adequately. In this experiment, three neural network models with varying architectures were compared. The MAPE values of the three models can be seen in Table 3.

Table 3. MAPE values of three models

Sample	Architecture	MAPE (%)
Model a	[8] [8] [1]	9.42
Model b	[8] [2 4] [1]	14.43
Model c	[8] [16 8 2] [1]	10.23

Based on the classification of MAPE values, the MAPE value in the Model A of 9.42% is in the excellent category. Meanwhile, the MAPE values of Model B and Model C of 14.43% and 10.23% are in a suitable category [16], [17]. Although Model A is considered highly accurate, the graphical analysis in Figure 2 indicates that Model A struggles to capture rainfall spikes. This can be observed from the model's tendency to flatten extreme peaks, resulting in the loss of detail in significant events within the dataset. The flattening of extreme peaks may be caused by the inherent complexity of the data and the model's insufficient capacity to learn highly complex patterns. Model B, which has the lowest accuracy compared to the other two models, shows its inability to generalize effectively. The limited number of nodes in each layer likely limits the model's capacity to learn the complex relationships needed to predict rainfall patterns accurately.

In contrast, Model C performs better than Model B, although it is still worse than Model A. Despite having a deeper architecture and greater nodes in the first two layers, Model C also struggles to capture highly localized precipitation surges, as seen in Figure 2. In this experiment, adding more layers and nodes provided only marginal performance gains, indicating that increased complexity did not significantly improve generalization. The addition of complexity does not improve the model's ability to capture important details, and even the bottlenecks caused by the small number of nodes in the third layer may further limit its effectiveness.

Alternative approaches can be explored to address the challenges in rainfall prediction, particularly in capturing extreme variations. Ensemble learning techniques, such as Random Forest or XGBoost, can be employed to combine multiple models, reducing bias and variance while improving overall accuracy. Additionally, replacing traditional activation functions like ReLU with advanced alternatives such as Swish or ELU may enhance the model's ability to respond to sudden shifts in rainfall patterns. Given the sequential nature of rainfall data, time-series models like LSTM or RNN could be integrated to capture temporal dependencies better. Hybrid models, such as combining LSTM with ARIMA, offer a balanced approach by leveraging the strengths of both deep learning and statistical methods to model linear and non-linear trends.

4. Conclusion

Based on the results and discussion, Model A is the best model that predicts rainfall in Palu City in 2023 with one hidden layer and 8 nodes. Model A is in the excellent category with a MAPE value of 9.42%, but its tendency to flatten extreme peaks makes some estimates underestimated when predicting rainfall. This can be due to the inherent complexity of the data and the insufficient capacity of the model to learn highly complex patterns. The ANN method can be an effective tool for predicting rainfall if the model is designed to be simple, efficient, and has complete training data.

References

[1] S. P. Cipta, "Penerapan Algoritma Evolving Neural Network Untuk Prediksi Curah Hujan," *Jurnal Teknologi Informasi Universitas Lambung Mangkurat*, vol. 1, no. 1, pp. 1–8, 2016.

[2] D. A. H. Panggabean, F. M. Sihombing, and N. M. Aruan, "Prediksi Tinggi Curah Hujan Dan Kecepatan Angin Berdasarkan Data Cuaca Dengan Penerapan Algoritma Artificial Neural Network (ANN)," *Seminastika*, vol. 3, no. 1, pp. 1–7, 2021, doi: 10.47002/seminastika.v3i1.237.

[3] N. Mishra, H. K. Soni, S. Sharma, and A. K. Upadhyay, "Development and Analysis of Artificial Neural Network Models for Rainfall Prediction by Using Time-Series Data," *IJISA*, vol. 10, no. 1, pp. 16–23, 2018, doi: 10.5815/ijisa.2018.01.03.

[4] N. Q. Hung, M. S. Babel, S. Weesakul, and N. K. Tripathi, "An artificial neural network model for rainfall forecasting in Bangkok, Thailand," *Hydrol. Earth Syst. Sci.*, vol. 13, no. 8, pp. 1413–1425, 2009, doi: 10.5194/hess-13-1413-2009.

[5] R. M. Putra and N. Anjar Rani, "Prediksi Curah Hujan Harian di Stasiun Meteorologi Kemayoran Menggunakan Artificial Neural Network (ANN)," *BGB*, vol. 1, no. 2, pp. 101–108, 2020, doi: 10.31172/bgb.v1i2.35.

[6] N. Kamal and Y. Ramadhani, "Optimasi Algoritma Neural Network berbasis Fitur Seleksi Menggunakan Algoritma Genetika Untuk Prediksi Curah Hujan," in *E-Prosiding Teknik Informatika*, vol. 4, pp. 269–282, Universitas Adhijarasa Resawara Sanjaya, 2023.

[7] D. Jauhari, A. Himawan, and C. Dewi, "Prediksi Distribusi Air PDAM menggunakan Metode Jaringan Syaraf Tiruan Backpropagation di PDAM Kota Malang," *JTIIK*, vol. 3, no. 2, pp. 83–87, 2016.

[8] R. C. Nabila, A. Arnita, A. Fitria, and N. Suryani, "Application Of The Backpropagation Method To Predict Rainfall In North Sumatra Province," *BAREKENG: J. Math. & App.*, vol. 17, no. 1, pp. 0449–0456, 2023, doi: 10.30598/barekengvol17iss1pp0449-0456.

[9] W. Winarnie, "Pemanfaatan Sequential Convolutional Neural Networks pada Deteksi Malware Android," *Jurnal Ekonomi dan Teknik Informatika*, vol. 12, no. 1, pp. 65–70, 2024, doi: 10.37601/jneti.v12i1.257.

[10] E. G. Dada, H. J. Yakubu, and D. O. Oyewola, "Artificial Neural Network Models for Rainfall Prediction," *EJECE*, vol. 5, no. 2, pp. 30–35, 2021, doi: 10.24018/ejece.2021.5.2.313.

[11] D. I. Mulayana and Marjuki, "Optimasi Prediksi Harga Udang Vaname dengan Metode RMSE dan MAE Dalam Algoritma Regresi Linier," *Jurnal Ilmiah Betrik*, vol. 13, no. 1, pp. 50–58, 2022.

[12] M. R. M, R. D. Putri, M. R. R, S. Amin, and M. Akli, "Pengaplikasian Artificial Neural Network (ANN) dalam Memprediksi Curah Hujan Menggunakan Python," in *SinarFE7*, vol. 4, pp. 369–373, 2021.

[13] H. Ihsan, I. Irwan, and A. I. E. Nensi, "Implementation of Backpropagation and Hybrid Arima-NN Methods in Predicting Accuracy Levels of Rainfall in Makassar City," *BAREKENG: J. Math. & App.*, vol. 18, no. 4, pp. 2435–2448, 2024, doi: 10.30598/barekengvol18iss4pp2435-2448.

[14] M. D. Ariani, "Prediksi Curah Hujan Kecamatan di Jawa Timur Menggunakan Algoritma Backpropagation Neural Network Berdasarkan Faktor Penentu Curah Hujan," Universitas Islam Negeri Maulana Malik Ibrahim, Malang, 2022

[15] T. G. Lasijan, R. Santoso, and A. R. Hakim, "Prediksi Harga Emas Dunia Menggunakan Metode Long-Short Term Memory," *J.Gauss*, vol. 12, no. 2, pp. 287–295, 2023, doi: 10.14710/j.gauss.12.2.287-295.

[16] L. Hakim, Anik Vega Vitianingsih, Gita Indah Marthasari, Kresna Arief Nugraha, and Anastasia Lidya Maukar, "Analisis Spasial untuk Klasifikasi Pengembangan Tempat Penampungan Sementara Menggunakan Metode Jaringan Syaraf Tiruan," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 6, no. 1, pp. 108–114, 2022, doi: 10.29207/resti.v6i1.3713.

[17] U. I. Arfianti, D. C. R. Novitasari, N. Widodo, Moh. Hafiyusholeh, and W. D. Utami, "Sunspot Number Prediction Using Gated Recurrent Unit (GRU) Algorithm," *Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 2, p. 141, 2021, doi: 10.22146/ijccs.63676.