JOURNAL OF DATA SCIENCE | Vol.2023:04
eISSN:2805-5160

Rainfall Prediction in Palembang City Using the GRU and LSTM Methods

Surta Wijaya[1], Tri Basuki Kurniawan Kurniawan[1*], Edi Surya Negara[2], Yesi Novaria Kunang[1]

[1] Magister of Information Technology, University of Bina Darma, Palembang, Indonesia
[2] Computer Science, University of Bina Darma, Palembang, Indonesia

Email: tribasukikurniawan@binadarma.ac.id

Abstract

Rainfall is one of the weather elements that are very important for the survival of an area. Palembang City, as one of the big cities in Indonesia, is heavily influenced by the level of rainfall that occurs every month. Variations in precipitation can affect various aspects of people's lives, such as agriculture, industry, tourism, etc. Accurate rainfall predictions can assist in preparing for multiple activities and making the right decisions. Therefore, it is crucial to research predicting rainfall in Palembang. It is expected to simplify the prediction process and produce more accurate results. This research uses the Gated Recurrent Unit (GRU) and Long-Shorted Term Memory (LSTM) methods to make daily rainfall predictions for the next month using weather element data for ten (10) years in Palembang, utilizing the deep learning method. The hyperparameter model tuning experiment was conducted to obtain the best prediction results. From the research results, it can be concluded that the LSTM model is overall better than the GRU model in predicting daily rainfall in Palembang City. GRU has RMSE 9.33 and R^2 0.54, while the LSTM Model has RMSE 7.45 and R^2 0.70.

Keywords

Rainfall, Prediction, Gated Recurrent Unit (GRU), Long-Shorted Term Memory (LSTM)

Introduction

Rainfall is the main factor that can influence the occurrence of hydro-meteorological disasters, such as causing droughts and floods. Deficiency occurs due to low rainfall over a long period, limiting the water supply for human, agricultural and industrial needs. High rainfall in a short time can cause flooding, while rainfall that occurs in the long term can cause landslides (Collier, 2016). An example of a hydrometeorological disaster in Palembang is that there were four (4) floods in October 2022 due to the intensity of heavy rainfall in one day (BMKG, 2022).

Prediction models are made as part of disaster mitigation measures. Deep learning is a method that can predict rainfall to overcome the problems that arise from high rain (Aswin et al., 2018). This research uses deep learning Gated Recurrent Unit (GRU) and Long-Shorted Term

Submission: 28 February 2023; **Acceptance:** 21 March 2023

Memory (LSTM). GRU and LSTM were chosen because of their excellent performance in predicting time series data (Rajagukguk et al., 2020). Previous studies applied the GRU and LSTM methods to predict stock prices using a time series model. The result was that GRU had the best performance compared to linear regression and LSTM (Sofi et al., 2021).

Research on using the LSTM model in predicting rainfall has previously been carried out. LSTM is used as a method to predict precipitation in the city of Malang. The result is that LSTM can work optimally with fairly good predictive results (Rizki et al., 2020). Another study using LSTM was also carried out to predict rainfall in Bandung. Using epoch 50 and batch size 1 produces the highest accuracy with an RMSE train value of 12.24 and a test of 8.86 (Firdaus & Paputungan, 2022).

This research aims to create a daily rainfall prediction model that can predict the city of Palembang for the next one (1) month using the GRU and LSTM methods. This study also tries to compare the ability of the GRU and LSTM plans to predict rainfall using weather element data for ten (10) years.

Methodology

The method used in this study consisted of 8 stages, starting from the data collection stage to the evaluation of the prediction results by the model made. The research flowchart can be seen in Figure 1.

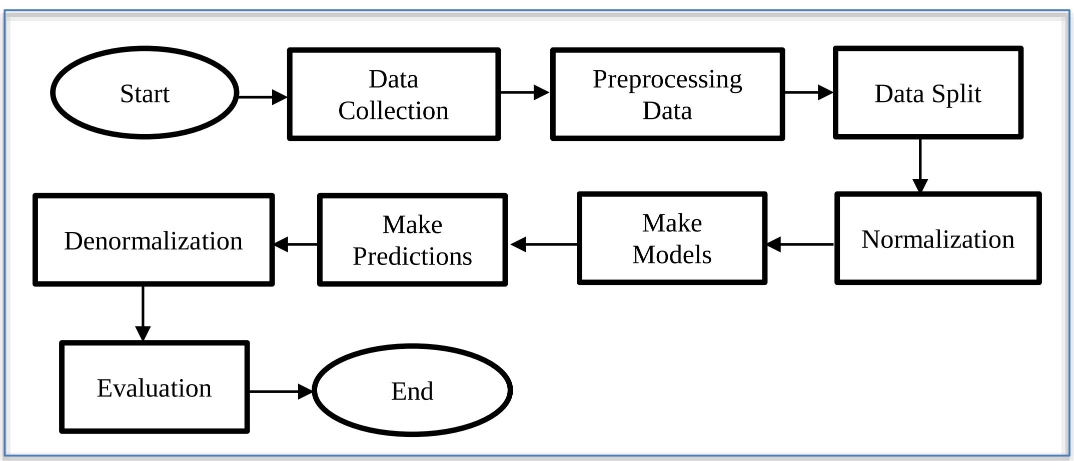

Figure 1. Research Flowchart

Data collection

This study uses a dataset of several weather elements per day from 2011 to 2021, originating from the Palembang Climatology Station. The weather elements used are rainfall (RR) average temperature (T_AVG), the temperature at 07 and 13 local times (T_07, T_13), average humidity (RH_AVG), humidity at 07, 13 and 18 (RH_07, RH_13, RH_18), evaporation (EEE), sunshine duration (SS), and average wind speed (WS_AVG). The data was chosen because it correlates well with the rainfall data. Rainfall data serves as target data, and other data functions as predictor data. An example of the dataset used for modeling can be seen in Table 1.

Table 1. Sample Research Dataset

	Date	EEE	RH_AVG	SS	T_AVG	T_07	T_13	T_Max	T_Min	RH_07	RH_13	RH_18	WS_AVG	RR
0	1/1/2011	3.7	83	40	27.4	24.9	30.8	32.4	24.2	95	67	75	3	0
1	1/2/2011	4.4	81	49	27.6	25	32.8	33.5	24.8	92	56	85	4	0
2	1/3/2011	5.2	84	0	26.6	23.6	30.2	30.3	22.8	92	71	81	3	1
3	1/4/2011	3.1	82	69	27	23.4	31.4	32.4	23.1	95	64	73	4	3
4	1/5/2011	5	93	4	24.7	22.2	27.2	29.9	22	98	86	89	2	26
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3829	6/26/2021	3.9	83	60	28.1	25.6	32	33.4	25.4	95	65	76	5	0
3830	6/27/2021	3.7	83	64	27.2	24.4	31.8	32.2	24	93	67	77	4	0
3831	6/28/2021	3.4	86	35	26.9	24.5	30.2	31.7	24.2	94	73	82	4	2
3832	6/29/2021	3.3	83	81	27.3	24.8	32.2	33.2	24	95	61	80	7	0
3833	6/30/2021	2.1	86	19	26.6	24.3	29.6	30.2	24.2	94	71	85	6	0

Preprocessing Data

Preprocessing data in deep learning is the stage of the data preparation process before model training. This stage is crucial so that the data used for model training is of good quality and appropriate to the characteristics of the model you want to use. Proper data preprocessing can improve the performance and accuracy of deep learning models. The dataset has been created by selecting 13 variables which will later be used for the modelling process.

After creating the dataset, the next step is to divide the dataset into training data, evaluation data and testing data. Training and evaluation data use daily data from 2011 to 2020 with a ratio of 90:10. Then, data testing uses data for 2021. The training and evaluation data distribution is used to train and make prediction models. In contrast, data testing will be used to test the models created.

The next stage is the data normalization process. This process is carried out to change the data scale for each feature or variable in the dataset to the same scale. This research uses the MinMax Scaler method. The minimum value will reduce each value in the data; then, the result will be divided by the difference between the maximum and minimum values. This process will generate new data with a range between 0 and 1. The formula used is as follows:

$$X_i' = \frac{X_i - X_{min}}{X_{max} - X_{min}} \quad \text{Eq. (1)}$$

Where X_i' is the new normalized data, X_i is the value in the data to be normalized, X_{min} is the minimum value of the data, and X_{max} is the maximum value of the data. This process aims to facilitate the model training process and improve model performance. The denormalization process is carried out to restore normalized data. The formula used is as follows.

$$X_i = (X_i' \times (X_{max} - X_{min})) + X_{min} \quad \text{Eq. (2)}$$

Where, X_i is the data to be returned before normalization, X_i' is data that has been normalized, X_{max} is the maximum value of the data before normalization, and X_{min} is the minimum value before normalization. After the normalization process, the data is converted into a format suitable for the input of the deep learning model.

Deep Learning Models

At this stage, the model's design is done through hyperparameter tuning. Hyperparameter tuning is a process for optimizing the values in the parameters used during the deep learning model creation process. Determining parameter values is done manually by trying several different weights and comparing the model's performance to the testing data. Incorrect parameter values can cause the model to be overfitting or underfitting. This process is carried out to improve the performance of making deep learning models better to make the prediction results more accurate.

Evaluation

The model evaluation stage is a process to evaluate how well the model has been made and produces predictive values. This study uses MSE (Mean Squared Error), RMSE (Root Mean Squared Error), and R^2 (R-Squared) evaluations. During the data training process, MSE calculates Loss and Validation Loss values. This process aims to monitor training data and evaluation data when building models. MSE measures the average of the squared difference between the actual value and the predicted value (Salman et al., 2017). The formula for obtaining MSE is as follows.

$$\text{MSE} = \sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n} \qquad \text{Eq. (3)}$$

Where, $\hat{y}_i$ is the predicted data on the ith data, and y_i is the actual data on the *i*th data. The lower the MSE value, the better the performance of the model that has been made. Furthermore, after the model is created and produces predictive values, evaluation is carried out using RMSE and R^2. RMSE measures the average square root of the difference between predicted and actual data. (Chai & Draxler, 2014) The formula for obtaining RMSE is as follows.

$$\text{RMSE} = \sqrt{\sum_{i}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \qquad \text{Eq. (4)}$$

Where, $\hat{y}_i$ is the predicted value in data *i* and y_i is the actual value in data *i*. The lower the RMSE value, the closer the predicted value is to the actual value. R^2 is used to measure how well the model predicts the data by comparing the variation in the predicted value produced by the model with the actual value. (Lu et al., 2020) The formula for obtaining R^2 is as follows.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad \text{Eq. (5)}$$

Where, $\hat{y}_i$ is the predicted value in data *i*, y_i is the actual value in data i and y_i is the actual average value. R^2 values range from 0 to 1. The closer to 1, the better the prediction accuracy.

Rainfall

Rainfall is a quantitative measurement of the amount of water that falls to the earth's surface as rain. Rainfall has a complex relationship with other weather elements such as temperature, humidity, solar radiation, atmospheric pressure, wind, and evaporation. (Ahrens, 2014) In the hydrologic cycle, evaporation that occurs due to sunlight and high air temperatures causes the formation of clouds and rain.

The higher the humidity, the more likely it is to rain. The high water evaporation rate will increase the water vapor in the atmosphere, thereby affecting air pressure and wind. Wind conditions and air pressure will affect water exchange between regions and the atmosphere, rainfall, and water vapour spread to other areas.

As rainfall increases, air pressure decreases, and winds will become gentler. Increased rainfall will also lower temperatures and increase humidity. Conversely, when rain decreases, air pressure increases, winds strengthen, and temperature and humidity drop. (Donald Ahrens & Henson, 2015).

Time Series data

Time series is a series of data measured at regular intervals. The data is usually taken from the past and can be used to predict or forecast future values. An example of a time series is weather data that is measured every month or every year. Time series data prediction is the process of predicting the future value of a time series based on patterns and trends found in historical data. This process can assist in making better business decisions, such as production planning, demand forecasting, or stock management (Prabhakaran, 2019).

There is a technique called sliding windows in making deep learning models that use time series data. The sliding window is a technique of dividing time series data into several parts or windows of the same size and overlapping each other at certain time intervals. The purpose of this technique is to prepare data that can be used as input to machine learning models to make predictions or classifications based on time series data.

In general, sliding windows in machine learning time series retrieve a group of values or observations from time series data stored at a specific time range. Then, these values or statements are placed into a window with a specified size. (Yahmed et al., 2015) Furthermore, the window will be shifted over time with a certain distance or step size so that new values or observations will be entered, and the oldest values or observations will be removed.

At each iteration, the values or observations in each window can be used as input to the machine learning model. The output from the model can be used to predict values or categories in the following time series data (Hota et al., 2017).

GRU - LSTM

Gated Recurrent Unit (GRU) and Long-Short Term Memory (LSTM) are two types of Recurrent Neural Networks (RNN) that can process sequential data, such as time series. They can study the relationship between previous data and future predictions, making it suitable for predicting time series. LSTM and GRU are part of the RNN, which has been modified to overcome the problems of exploding and vanishing gradients when retaining memory and processing long sequential data. (Manaswi et al., 2018) These problems can reduce the accuracy of the prediction results in the RNN model (Zhao et al., 2017)

GRU is a Deep Learning method similar to LSTM but has a more straightforward structure and faster training process (Cho et al., 2014). GRU's gating mechanism allows the model to select and discard relevant information. GRU performs equivalent to LSTM in predicting time series but has the advantage of a faster and more efficient training process (Chollet F., 2017). Each unit in the GRU has two types of gates, namely, the update gate and the reset gate. The update gate determines the extent to which the team must update its current state with new information, while the reset gate determines the size to which the unit must forget information from the past.

The LSTM architecture has three (3) gates: input, forget and output (Chung & Shin, 2018). Each gate has a function and task in collecting and processing data. Forget gate has the task of forgetting various unnecessary and inappropriate information by a system so that the LSTM can display different information that is complete but still actual according to what is needed. The input gate's job is to send valuable data so that correct data may be produced. The input gate is also in charge of entering data that the forget gate previously chose. The output gate creates accurate and comprehensive data information, as shown in Figure 2.

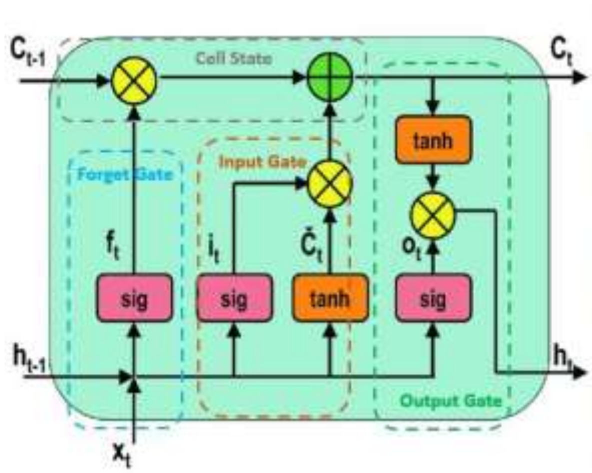

Figure 2. Gate Structure on LSTM (Singhal Gaurav, 2020)

Results and Discussion

This study will make predictions for the next one (1) month using 13 variables or features to create GRU and LSTM models with targeted rainfall data. The testing data used to evaluate the prediction results is January 2021, with 31 data. The model uses four (4) hidden layers of 2 GRU/LSTM layers with eight (8) units/neurons in each layer and two dropout layers of 0.2. Then the model uses the early stop feature, which stops the training process when the validation loss value does not decrease. Determining the initial values of other deep learning parameters can be seen in Table 2.

Table 2. Default setting parameter Deep Learning

Parameters	Value
Number of Hidden Layer	4
Number of units layer GRU/LSTM	8
Dropout rate	0.2
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Loss function	MSE
Validation Split	0.1
Epoch Max	200

Determine the Total Window Size

Determining the number of Window Sizes (WS) is a sliding window technique on time series data used to predict future values in a time series. This technique separates time series data into several overlapping time windows according to the number of WS. In this study, the initial number of WS was determined based on monthly rainfall trends. The selection of the number of WS in each model can be seen in the following Tables dan Figures.

Table 3. Selection of the Number of WS on the GRU model

WS	Total Time	Epoch	Loss	Val Loss	RMSE	R^2
30	122400 ms	200	0.0069	0.0063	15.02	0.42
60	103632 ms	127	0.0070	0.0062	11.2	0.66
90	153318 ms	138	0.0068	0.0063	11.8	0.63
120	260000 ms	167	0.0068	0.0062	11.46	0.66
150	247104 ms	156	0.0066	0.0065	13.11	0.55
180	256956 ms	138	0.0068	0.0065	11.46	0.66

Table 4. Selection of the Number of WS in the LSTM model

WS	Total Time	Epoch	Loss	Val Loss	RMSE	R^2
30	51000 ms	100	0.0072	0.0063	14.95	0.42
60	101898 ms	111	0.0072	0.0062	10.35	0.72
90	130290 ms	129	0.0068	0.0063	8.30	0.80
120	170300 ms	131	0.0068	0.0068	12.38	0.60
150	194832 ms	123	0.0066	0.0065	9.82	0.75
180	183456 ms	117	0.0068	0.0064	9.49	0.76

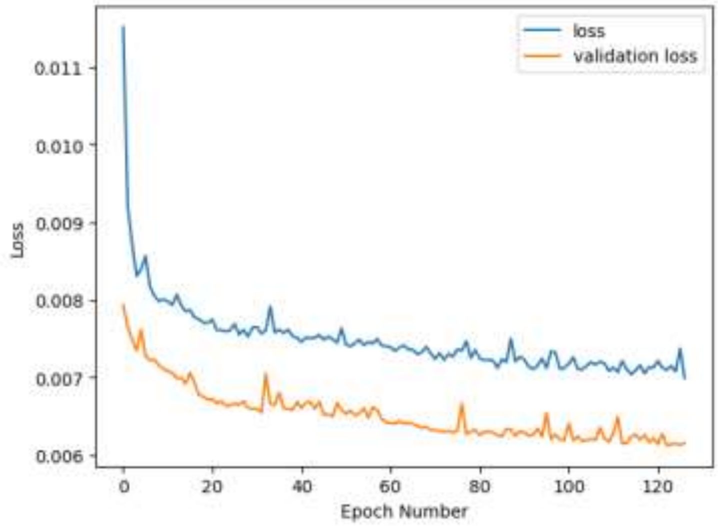

Figure 3. GRU model training plots

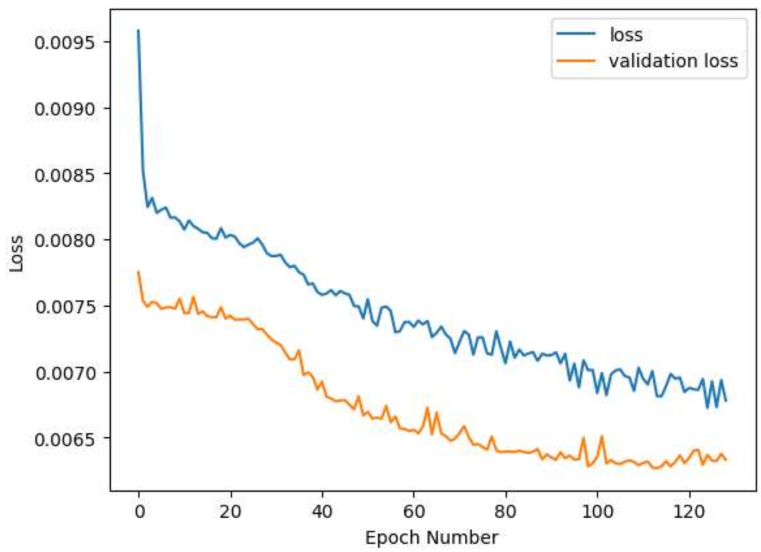

Figure 4. LSTM model training plots

In Figure 3 and Figure 4, based on the model evaluation results in tables and tables, the more WS is used, the longer the training process will take. In the GRU model, WS with several of 60 has the lowest RMSE value of 11.20 and the highest R2 value of 0.66, while in the LSTM model, WS with some 90 has the lowest RMSE value of 8.30 and the highest R2 value of 0.80. The training process for both models can be seen in the figure. From the evaluation results of comparing the two models, the temporary LSTM model produces a better predictive than the GRU model.

Determine the Number of Units

After determining the number of WS in each model, the next hyperparameter tuning determines the number of units or neurons in each layer. This process is carried out to maximize the creation and training of deep learning models to obtain more accurate predictive results. The GRU Model uses a WS 60, and the LSTM model uses WS 90. Each model was carried out eight (8) times with a combination of different numbers of units in each experiment. The number of units in each layer and the evaluation results can be seen in Table 5 and Table 6.

Table 5. Selection of the number of units in the GRU layer

Layer 1 GRU	Layer 2 Dropout	Layer 3 GRU	Layer 4 Dropout	Epoch	RMSE	R^2
16	0.2	8	0.2	112	12.3	0.6
32	0.2	8	0.2	107	10.8	0.69
64	0.2	8	0.2	91	10.94	0.69
128	0.2	8	0.2	68	11.86	0.64
8	0.2	16	0.2	150	11.10	0.68
8	0.2	32	0.2	121	11.02	0.68
8	0.2	64	0.2	89	10.52	0.71
8	0.2	128	0.2	81	11.29	0.67

Table 6. Selection of the number of units in the LSTM layer

Layer 1 LSTM	Layer 2 Dropout	Layer 3 LSTM	Layer 4 Dropout	Epoch	RMSE	R^2
16	0.2	8	0.2	93	9.10	0.78
32	0.2	8	0.2	94	7.86	0.84
64	0.2	8	0.2	76	4.23	0.95
128	0.2	8	0.2	78	9.91	0.74
8	0.2	16	0.2	133	7.55	0.85
8	0.2	32	0.2	110	8.60	0.81
8	0.2	64	0.2	101	10.18	0.73
8	0.2	128	0.2	97	9.27	0.78

Based on the evaluation results in tables 5 and 6, the more units used in each layer, the fewer the number of epochs needed, so the time required during the training process is faster. The GRU model at layer 1 has eight (8) units, and layer 3 has 64 units; Epoch 89 has the best value with the lowest RMSE, which is 10.52 and the highest R^2 value, which is 0.71. A comparison of predicted data results with actual data can be seen in Figure 5.

The LSTM model at layer 1 with 64 units, layer 3 with eight (8) units, epoch 76 has the best value with the lowest RMSE, which is 4.23 and the highest R^2 value is 0.95. A comparison of predicted data with actual data can be seen in Figure 6. Therefore it can be concluded that the best model for predicting daily rainfall for the next one (1) month is the LSTM model, as shown in Figure 6.

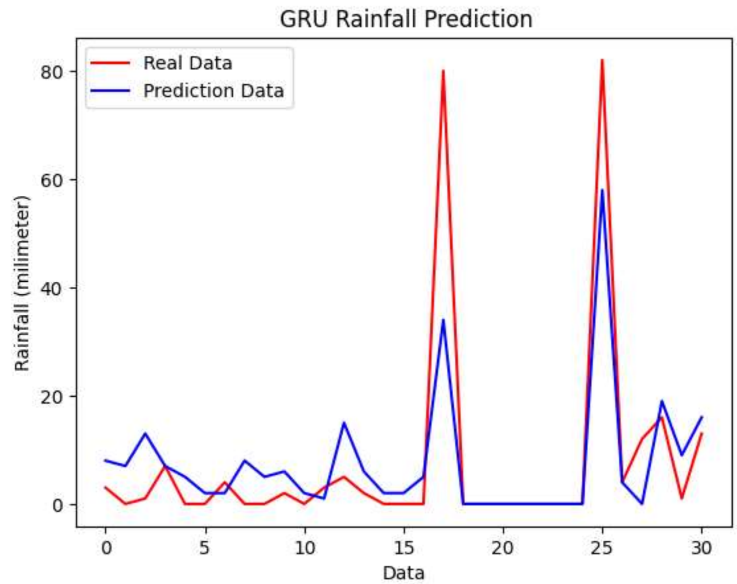

Figure 5. The plot of GRU Prediction Results

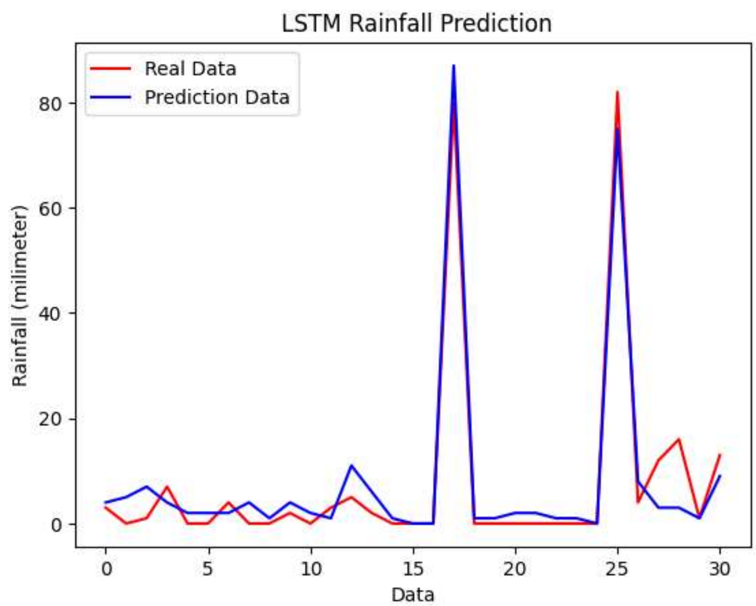

Figure 6. The plot of LSTM prediction results

Model Testing

After the model successfully predicts rainfall for the next one (1) month in January. The next step is to test the model to make predictions for one month in another month to see whether the model that has been made can be applied in other months and whether it can get the exact best results as in January. Tests are carried out every month until June 2021. The model evaluation results from every month can be seen in Table 7 and Table 8.

Table 7. GRU model evaluation results

Month	Epoch	Loss	Val Loss	RMSE	R^2
Jan	89	0.0071	0.0063	10.52	0.71
Feb	101	0.0070	0.0064	7.43	0.26
Mar	84	0.0071	0.0052	12.47	0.51
Apr	59	0.0075	0.0058	7.98	0.20
Mei	63	0.0077	0.0050	9.88	0.36
June	116	0.0070	0.0043	5.83	0.57

Table 8. Evaluation results of the LSTM model

Month	Epoch	Loss	Val Loss	RMSE	R^2
Jan	76	0.0067	0.0063	4.23	0.95
Feb	88	0.0067	0.0063	7.58	0.23
Mar	90	0.0066	0.0050	9.94	0.69
Apr	40	0.0072	0.0056	6.97	0.39
Mei	56	0.0070	0.0043	9.66	0.39
June	93	0.0065	0.0046	4.38	0.76

The GRU and LSTM models can adequately carry out the training process based on the validation loss value. The GRU model has the highest RMSE value in March with a value of 12.47, the lowest value in June with a value of 5.83, while the lowest R^2 value is in April with a value of 0.20, the highest in January with a value of 0.71. The LSTM model has the highest RMSE value in March with a value of 9.94; the lowest value is still in January, while the lowest R^2 value is in February with a value of 0.23, and the highest value in January is 0.95.

The total test evaluation results for six months, namely the GRU model, have RMSE = 9.33 and R^2 = 0.54. The LSTM model has RMSE = 7.45 and R^2 = 0.70; from these results, the LSTM model is overall better than the GRU model. The plot of the results of the prediction data for the two models can be seen in Figure 7.

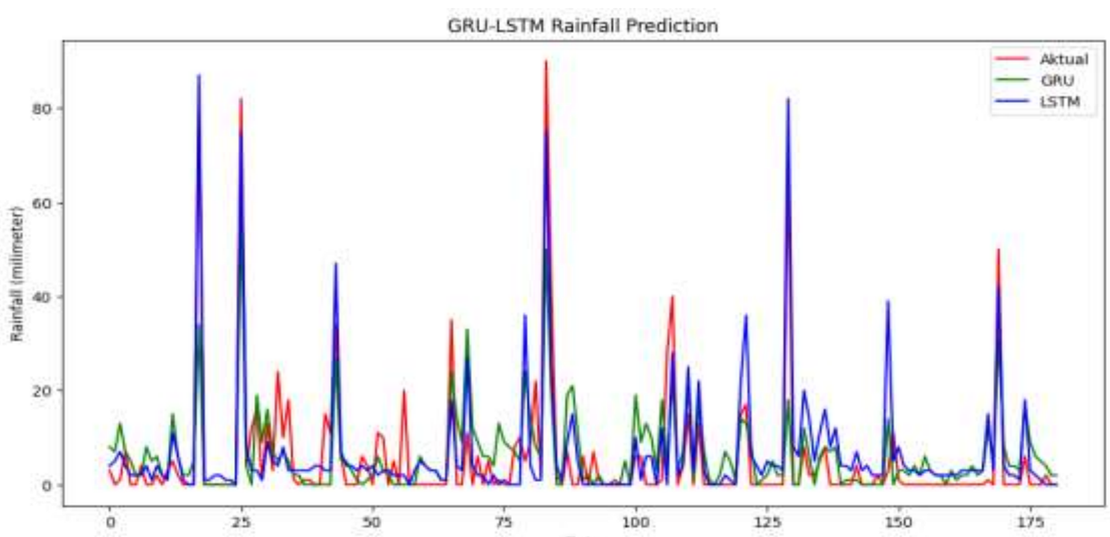

Figure 7. The plot of GRU and LSTM Prediction Results for six months

Conclusion

Based on the evaluation of the predicted and actual values, it can be concluded that the LSTM model is better than the GRU model in predicting daily rainfall in Palembang City. The GRU model has RMSE 9.33 and R^2 0.54, while the LSTM model has a value of RMSE 7.45 and R^2 0.70. When predicting rainfall for the next one month, the GRU model has the best predictions in January with an RMSE value of 10.52 and R^2 of 0.71. The LSTM model has the best prediction as GRU in January with RMSE 4.23 and R^2 0.95. The worst prediction of the GRU and LSTM model is when making predictions in February. GRU has an RMSE value of 7.43 and R^2 0.26, while LSTM has an RMSE value of 7.58 and R^2 0.23.

References

Ahrens, C. D. (2014). *Essentials of Meteorology: An Invitation to the Atmosphere*. Cengage Learning. https://books.google.co.id/books?id=heHKAgAAQBAJ

Aswin, S., Geetha, P., & Vinayakumar, R. (2018). Deep learning models for the prediction of rainfall. *2018 International Conference on Communication and Signal Processing (ICCSP)*, 657–661.

BMKG Stasiun Klimatologi Sumatera Selatan. (2022, November). *Buletin Ikim Edisi November 2022*.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Cho, K., van Merrienboer, B., Bahdanau, D., & Bengio, Y. (2014). On the Properties of Neural Machine Translation: Encoder–Decoder Approaches. *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, 103–111. https://doi.org/10.3115/v1/W14-4012

Chollet Francois. (2017). *Deep Learning with Python*. Manning Publications.

Chung, H., & Shin, K. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability, 10*(10), 3765.

Collier, C. G. (2016). *Hydrometeorology*. John Wiley & Sons.

Donald Ahrens, C., & Henson, R. (2015). *Meteorology today: an introduction to weather, climate and the environment*.

Firdaus, R. F., & Paputungan, I. V. (2022). Prediksi Curah Hujan di Kota Bandung Menggunakan Metode Long Short Term Memory. *Jurnal Penelitian Inovatif*, *2*(3), 453–460.

Hota, H. S., Handa, R., & Shrivas, A. K. (2017). Time series data prediction using sliding window based RBF neural network. *International Journal of Computational Intelligence Research*, *13*(5), 1145–1156.

Lu, W., Li, J., Li, Y., Sun, A., & Wang, J. (2020). A CNN-LSTM-based model to forecast stock prices. *Complexity*, *2020*, 1–10.

Manaswi, N. K., Manaswi, N. K., & John, S. (2018). *Deep learning with applications using python*. Springer.

Prabhakaran, S. (2019). Arima model–complete guide to time series forecasting in python. *Machine Learning Plus*, *18*.

Rajagukguk, R. A., Ramadhan, R. A. A., & Lee, H.-J. (2020). A review on deep learning models for forecasting time series data of solar irradiance and photovoltaic power. *Energies*, *13*(24), 6623.

Rizki, M., Basuki, S., & Azhar, Y. (2020). Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory (LSTM) Untuk Prediksi Curah Hujan Kota Malang. *Jurnal Repositor*, *2*(3), 331–338.

Salman, M. S., Kukrer, O., & Hocanin, A. (2017). Recursive inverse algorithm: Mean-square-error analysis. *Digital Signal Processing*, *66*, 10–17. https://doi.org/10.1016/j.dsp.2017.04.001

Singhal Gaurav. (2020, September 9). *Introduction to LSTM Units in RNN*. Https://Www.Pluralsight.Com/Guides/Introduction-to-Lstm-Units-in-Rnn. Diakses pada tanggal 1 Februari 2023.

Sofi, K., Sunge, A. S., Riady, S. R., & Kamalia, A. Z. (2021). Perbandingan algoritma linear regression, LSTM, dan GRU dalam memprediksi harga saham dengan model time series. *PROSIDING SEMINASTIKA*, *3*(1), 39–46.

Yahmed, Y. ben, Bakar, A. A., RazakHamdan, A., Ahmed, A., & Syed Abdullah, S. M. (2015). ADAPTIVE SLIDING WINDOW ALGORITHM FOR WEATHER DATA SEGMENTATION. *Journal of Theoretical & Applied Information Technology*, *80*(2).

Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y., & Liu, J. (2017). LSTM network: a deep learning approach for short-term traffic forecast. *IET Intelligent Transport Systems*, *11*(2), 68–75.