e-ISSN : 2722-4368

Timeseries forecasting for Local Average Temperature in Northern Sumatera Using Long Short-Term Memory Model

Marzuki Sinambela[1], Maman Sudarisman[2], Munawar[3]
[1]Undergraduate Program in Applied of Instrumentation Meteorology, Climatology Geophysics, STMKG, Indonesia
[2,3]Undergraduate Program in Applied of Climatology, STMKG, Indonesia
Correspondence should be addressed to Marzuki Sinambela: sinambela.m@gmail.com

***Abstrak**. For better management and planning of water resources in a basin, it is important to understand trends and predict average temperature as one of the parameters of weather and climate data. The study of weather trends using normal and local annual average temperature, comparison and observation. In this study, we will analyse the local and normal average temperature data in the city of Medan, based on the observation station in situ. The main objective of this study is to compare the normal temperature with the local station and to predict the temperature data in the city of Medan, North Sumatra by using the long term short term memory model. Based on the result of normal data science of exploring temperature with local temperature correlation, we got the display of training curve, residual plot and the scatter plot are shown using these codes. The good performance of Kualanamu and better than Deliserdang station had MSE value 0.01 and R2 value 0.98, close to zero represents better prediction quality.*

Keywords: temperature, trends, forecasting, lstm, timeseries

INTRODUCTION

Predicting the weather is an interesting research problem related to air navigation, environment, climate and agriculture [1], [2]. An important weather data in climate issue is temperature. Temperature is an important factor in all phases of the climate issue. It is the most important weather factor influencing fire behaviour. Higher temperatures mean that heat waves are likely to occur more frequently and to last for a longer period. Higher temperatures can also cause a chain reaction of other changes around the world [3], [4]. That's because an increase in air temperature also has an effect on the oceans, weather patterns, snow and ice, and plants and animals. In this study, we focused on the evaluation of the time series of the average temperature which was collected from 2 stations in Medan, North Sumatra. The two station are Tuntungan and Kualanamo station. The main objective of this study is to use Long Short Term Memory (LSTM) model to analyse intermediate variabels, evaluate and predict the timeseries of temperature average from 2008 to 2020 in Medan, North Sumatra. The models have been tested using weather time series data at each station, where the Root Mean Square Error (RMSE) obtained by the LTSM model has been collected and compared. The results of this experiment show that the LSTM model, with or without the intermediate variable, performs better when the time series are available.

In general, the results of this study will illustrate how to validate and predict the average temperature as a parameter in climate data in Medan area based on each station and can be used as a reference in further analysis of climate data sources that can be monitored by network climate in Medan area based on local observation that is the source. Using the clustering approach, this study aims to identify and predict the distribution of temperature based on segment activity.

DATA AND METHODS

For this study, we used temperature data from the climate catalogue of Meteorology, Climatology and Geophysics Regional I. The recorded temperature data from 2008 to 2020, with the category of average temperature on Medan City. The selection of temperature series data provides a solid basis for the evaluation of the performance of the observed, which have been installed and reviewed in the previous reports. A list of temperature recorded in each station can be seen in Table 1.

Table 1. List of the Station in Medan Area

No	Latitude	Longitude	Code Station
1	3.41	98.47	KNO
2	3.5	98.56	TSI

The time series of temperature data recorded at each of the stations around Medan were analysed and the time series were trained, tested and predicted with the help of the LSTMs [5]–[7].
In Figure 2, as a brief explanation, the input to the LSTM cell is a time-series set of data x that goes through a number of sigmoid activation gates σ. Each gate calculates a function. Each gate calculates a specific function to compute the cell states. What we have provided is only a very brief explanation of the way in which LSTM works. To learn more about LSTM, it's still much better to take a course on deep learning to help understand the concepts. The LSTM cell consists of an input gate - which controls the flow of input activations into the memory cell, an output gate - which controls the output flow of the cell activation, and a forget gate - which filters the information from the input and the previous output and decides which should be remembered or forgotten and discarded. In addition to the three gates, the LSTM cell contains a cell update, which is usually a tan-h layer to be part of the cell state. Three variables, the current input x_t, the previous output h_t-1 and the previous cell state c_t-1, enter the cell in each LSTM cell. Conversely, two variables come out of each LSTM cell: the current output ht and the current cell state c_t.

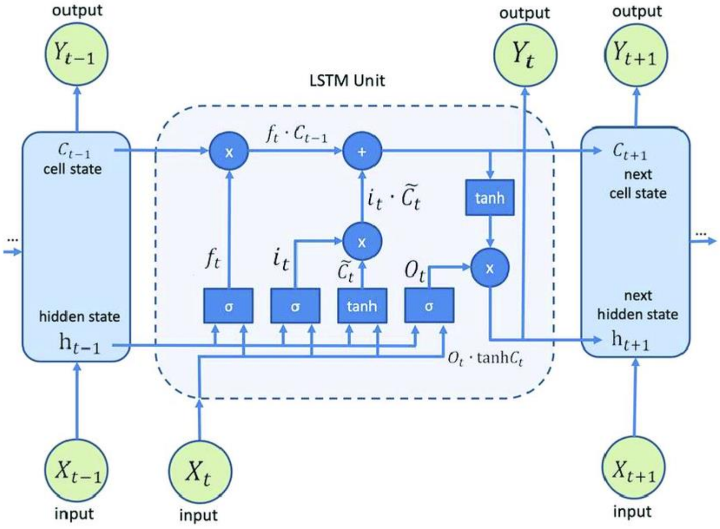

Figure 1. The LSTM Model [8]

The LSTM cell can be define as neural network where the input vector x = (x_1, x_2, x_3, … x_t) in time t, maps to the output vector y=(y_1, y_2,… ,y_m), through the calculation of the following layers [9]:

$$F_t = \sigma(W_f * [h_{t-1} * x_t] + b_f \quad (1)$$

Using the previous output h_t-1, the input vector xt and the matrix of weights from the Forget layer W_f with the addition of the corresponding bias bi, the Forget gate sigmoid layer for the time t, f_t is calculated:

$$i_t = \sigma(W_i * [h_{t-1}.x_t + b_i \quad (2)$$

It is calculated using the previous output h_t-1, the input vector x_t and the matrix of weights from the input layer W_i, adding the corresponding bias bi:

$$i_t = \sigma(W_i * \left[h_{t-1,}, x_t\right] + \mathrm{b}_i \quad (3)$$

From the forget gate ft and the previous cell state Ct-1, the cell state at time t, Ct, is calculated. The result is summed with the input gate it and the cell update state {\widetilde{c}}t, i.e. the layer

e-ISSN : 2722-4368
calculated from the previous output h_t-1, the input vector x_t and the weight matrix for the cell with addition of the appropriate bias bi:

$$O_t = \sigma(W_0 * [h_{t-1}, x_t] + b_0 \quad (4)$$

The previous output h_t-1, the input vector x_t and the matrix of weights from the output layer Wo, with the addition of the corresponding bias bi, are used to calculate the output gate sigmoid layer for the time t, o_t:

$$h_t = o_t \otimes \tanh(C_t) \quad (5)$$

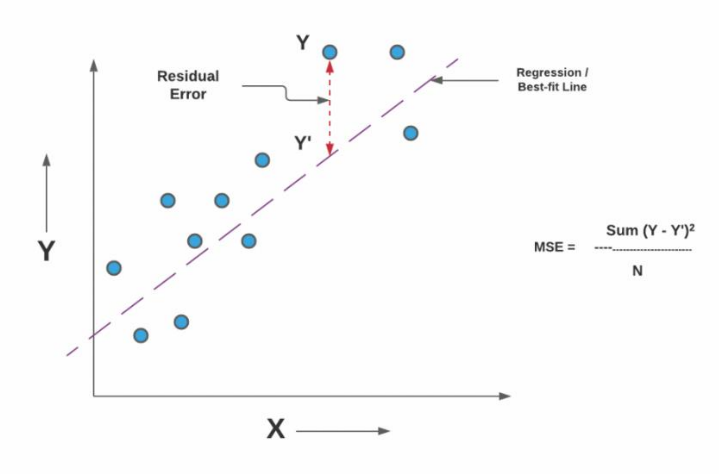

Figure 2. Mean Squared Error Representation [10]

The calculation of the current output h_t is the last stage of the LSTM cell. The multiplication operation \otimes between the output gate layer and the tanh layer of the current cell state C_t is used to calculate the current output. The current output h_t has passed through the network as the previous state for the next LSTM cell, or as the input for the output layer of the neural network.
In this section you will learn about the concepts of mean squared error and R-squared. The mean square error (MSE) is the average of the sum of the squares of the differences between the actual value and the predicted or estimated value. It is also known as the average squared deviation (MSD). The mathematical representation is as follows [10]:

The value of MSE is always positive or greater than zero. A value close to zero will represent better quality of the estimator / predictor (regression model). An MSE of zero (0) represents the fact that the predictor is a perfect predictor. When you take a square root of MSE value, it becomes root mean squared error (RMSE). In the above equation, Y represents the actual value and the Y' is predicted value. Here is the diagrammatic representation of MSE [10].

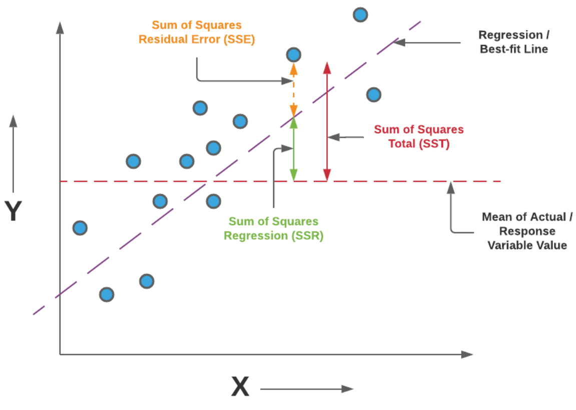

Figure 3 Diagrammatic representation for understanding R-Squared [10]

e-ISSN : 2722-4368
R-Squared can also be represented using the following formula. The diagram and note that greater the value of SSR, more is the variance covered by the regression / best fit line out of total variance (SST). R-Squared can also be represented using the following formula in eq. 6.

$$R^2 = \frac{SST}{SST} = \frac{\sum(\hat{y} - \bar{y})^2}{\sum(y_i - \bar{y})^2} \quad (6)$$

Based on the diagram and note that smaller the value of SSE, smaller is the value of (SSE/SST) and hence greater will be value of R-Squared.

R-Squared = 1 – (SSE/SST) (7)

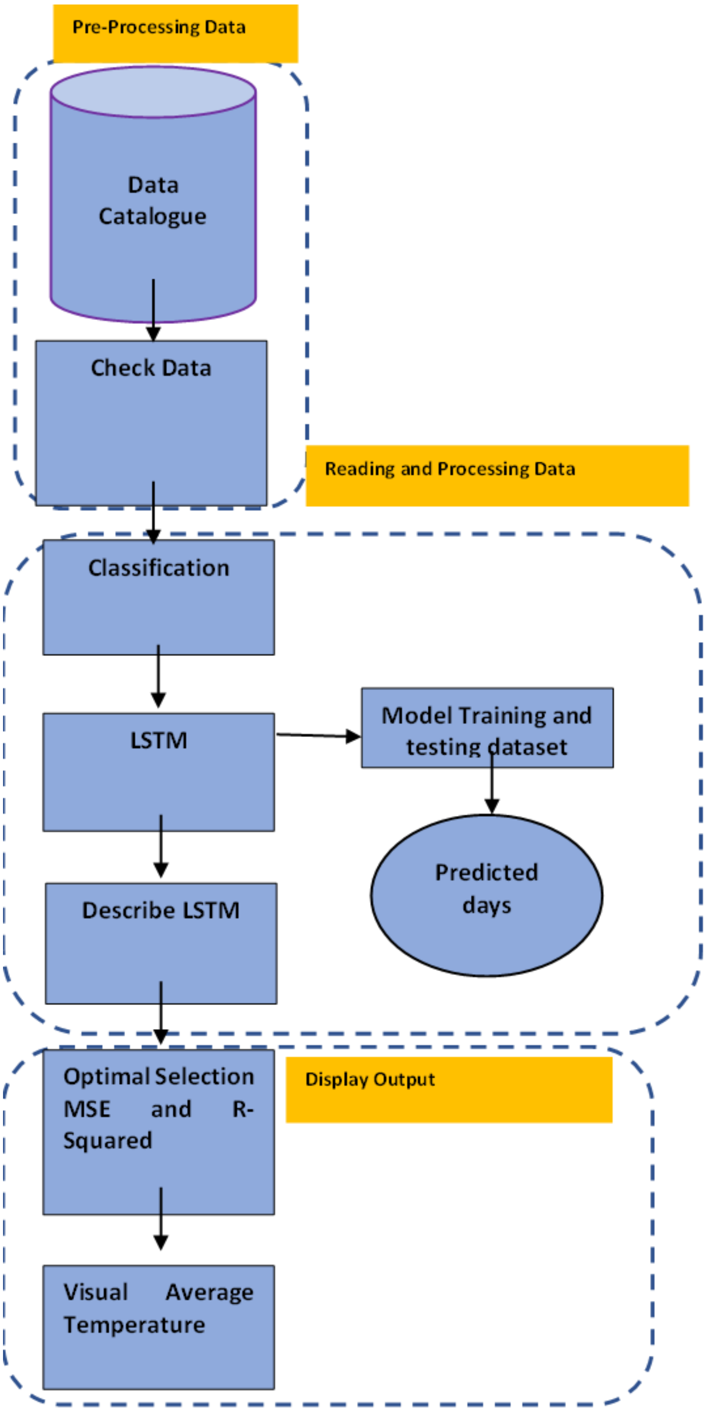

Figure 4. The flowchart of Forecasting Average Temperature

RESULT and DISCUSSION

The original data of average temperature had been computed to see that our model prediction was successful. However, it can be observed from the predicted (n days) that the errors are usually from the unexpected rise or decline in the data such as in days 350-360. But, based on the first 75 days, the model can properly follow the pattern of the data. The visual of all data of synoptic Balai Besar (96041) can be shown in fig.2. Based on the series data of temperature of 96041 station, the MSE value that is 0.01 and R^2 value that is 0.98. Based on the series data of temperature of 96035 station in figure 7, the MSE value that is 0.01 and R^2 value that is 0.98. in this study, the result of 96035 and 96041 show that the LSTM model yields with same value.

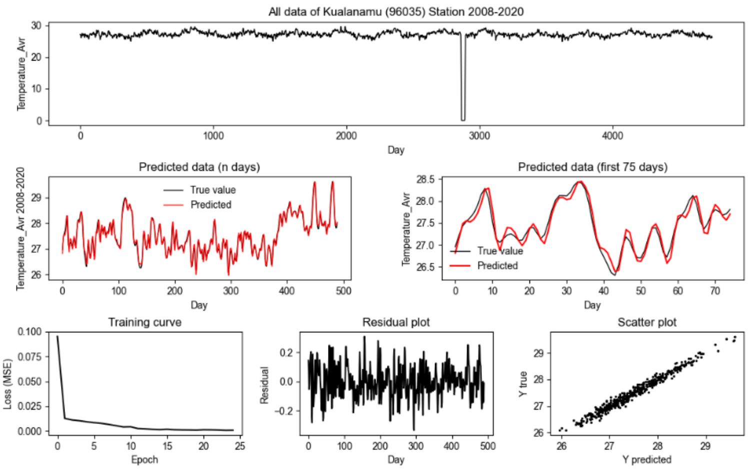

Figure 5 Stacked LSTM prediction results in 96035 Station

With the LSTM approach, the characteristics of the temperature in the 2008-2020 period are divided into MSE and R^2 value. From the results of figure 9, the MSE value of 96037 value that is 0.83 and R2 value is 0.83.

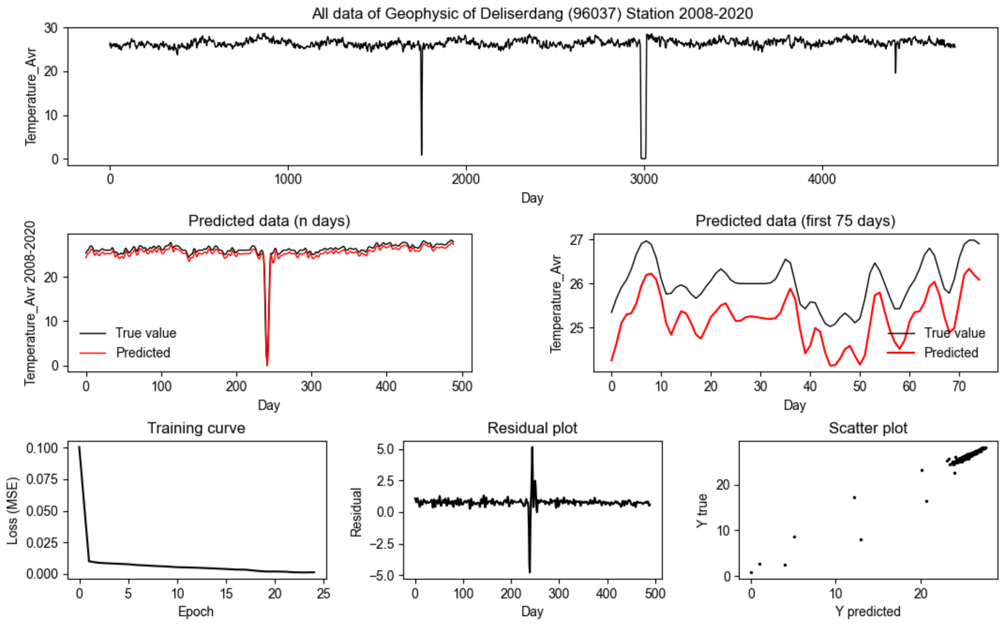

Figure 6 Stacked LSTM prediction results in 96037 Station

Based on the table 2, the represented of each station show the value close to zero will represent better quality of the estimator / predictor (regression model). It can be seen from the table that the Stacked LSTM performed best compared to the other types of each station. We can also see that the single-cell worked great using 100 input days but we found that this kind of set up was too computationally expensive. The bidirectional LSTM also performed worse with more input days.

Table 2. The value of MSE and R-Squared of each station

No	Stations	MSE	R-Squared
1.	Tuntungan	0.83	0.86
2.	Kualanamu	0.01	0.98

CONCLUSIONS

In this study, we found that LSTM is a good tool for predicting average temperature data. Based on the the representation of each station shown Kualanamu and Balai Station had value close to zero will represent better quality of the estimator / predictor (regression model). However, we can see from here that there are several things to take home as lessons in using LSTM. First, more input days does not really mean that the model will be more accurate. Other than that, temperature data conditioning may help in making the model more accurate. Lastly, even though we haven't shown, LSTM needs a certain amount of data to be applied. From there, we can imagine that LSTM can be used for predicting weather and trends of average temperature.

Data Availability

The data in this article the data was put to the public service so that reviewers could verify the reliability of the results. So, if reviewers need the raw data in this study, send me your request.

Acknowledgments

The author is grateful to Indonesia Meteorology, Climatology and Geophysics Agency and STMKG for providing and observed the climate regions stations and data collection.

References

[1] B. Choubin, G. Zehtabian, A. Azareh, E. Rafiei-Sardooi, F. Sajedi-Hosseini, and Ö. Kişi, "Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches," *Environ. Earth Sci.*, vol. 77, no. 8, pp. 1–13, Apr. 2018, doi: 10.1007/s12665-018-7498-z.

[2] F. S.-H. & Ö. K. Bahram Choubin, Gholamreza Zehtabian, Ali Azareh, Elham Rafiei-Sardooi, "Precipitation forecasting using classification and regression trees (CART) model: a comparative study of different approaches | SpringerLink," *https://link.springer.com/article/10.1007/s12665-018-7498-z?shared-article-renderer*. https://link.springer.com/article/10.1007/s12665-018-7498-z?shared-article-renderer (accessed Mar. 10, 2021).

[3] Wikepedia, "Effects of climate change - Wikipedia," *https://en.wikipedia.org/wiki/Effects_of_climate_change*, 2021. https://en.wikipedia.org/wiki/Effects_of_climate_change (accessed Mar. 10, 2021).

[4] NASA, "Global Warming," *https://earthobservatory.nasa.gov/features/GlobalWarming*, 2010. https://earthobservatory.nasa.gov/features/GlobalWarming (accessed Mar. 10, 2021).

[5] S. Afshin, H. Fahmi, A. Alizadeh, H. Sedghi, and F. Kaveh, "Long term rainfall forecasting by integrated artificial neural network-fuzzy logic-wavelet model in karoon basin," *Sci. Res. Essays*, vol. 6, no. 6, pp. 1200–1208, 2011, doi: 10.5897/SRE10.448.

[6] M. I. Hutapea, Y. Y. Pratiwi, I. M. Sarkis, I. K. Jaya, and M. Sinambela, "Prediction of relative humidity based on long short-term memory network," *AIP Conf. Proc.*, vol. 2221, no. March,

2020, doi: 10.1063/5.0003171.
[7] A. G. Salman, Y. Heryadi, E. Abdurahman, and W. Suparta, "Weather forecasting using merged Long Short-Term Memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model," *J. Comput. Sci.*, vol. 14, no. 7, pp. 930–938, 2018, doi: 10.3844/jcssp.2018.930.938.
[8] D. Zhou, X. Zuo, and Z. Zhao, "Constructing a Large-Scale Urban Land Subsidence Prediction Method Based on Neural Network Algorithm from the Perspective of Multiple Factors," pp. 0–3, 2022.
[9] Stanford, "Understanding LSTM Networks," *https://web.stanford.edu/class/cs379c/archive/2018/class_messages_listing/content/Artificial_Neural_Network_Technology_Tutorials/OlahLSTM-NEURAL-NETWORK-TUTORIAL-15.pdf*, pp. 1–13, 2015, [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/.
[10] A. Kumar, "Mean Squared Error or R-Squared, Data Analytics," *https://vitalflux.com/mean-square-error-r-squared-which-one-to-use/*, 2020. https://vitalflux.com/mean-square-error-r-squared-which-one-to-use/ (accessed Mar. 10, 2021).