Journal of Computers Vol. 31 No. 4, 2020, pp. 42-53
doi:10.3966/199115992020083104004

Anomaly Detection in Climate Data Using Stacked and Densely Connected Long Short-Term Memory Model

Teguh Wahyono[1,2*], Yaya Heryadi[1], Haryono Soeparno[3], Bahtiar Saleh Abbas[4]

[1] Doctor of Computer Science, Binus Graduate Program, Bina Nusantara University, Jakarta, Indonesia
teguhsalatiga@gmail.com, yayaheryadi@binus.edu

[2] Faculty of Information Technology, Satya Wacana Christian University, Salatiga, Indonesia
teguh.wahyono@uksw.edu

[3] Bioinformatics and Data Science Research Center, Bina Nusantara University, Jakarta, Indonesia
haryono@binus.edu

[4] Department of Industrial Engineering, Bina Nusantara University, Jakarta, Indonesia
bahtiar@binus.edu

Received 18 December 2018; Revised 28 May 2019; Accepted 31 May 2019

Abstract. Climate anomalies are considered as an important factor closely related to many disasters causing many human losses, such as airline crash, wildfires, drought and flooding in many areas. Many researchers have projected that the rising global temperature will increase the draught, especially in the mid-latitude areas. Taking those problems into account, studies on anomaly detection in climate are crucial. While climate prediction aims to analyze and model regular pattern of climate, climate anomaly studies aim to model climate deviation from its previous general patterns. Long Short Term Memory (LSTM) is a method employed in this research because it has been proven to work effectively in several anomaly detection studies, especially for data with similar characteristics. This paper presents empirical results of using Basic LSTM, Densely Connected (DC) LSTM, and Stacked-DC LSTM models to detect temperature anomaly as a manifestation of climate change in Semarang City. The results show that Stacked DC LSTM produced higher accuracy in detecting anomaly than the other two methods

Keywords: anomaly detection, deep learning, Denseley Connected, Stacked-DC LSTM

1 Introduction

Research on anomaly detection is extremely important for people's lives [1-3]. Anomaly detection has been used to solve crucial life problems. Many research topics in the field have been studied, for examples financial fraud detection [4-5], network security [6-7], medical and health [8], and machinery errors detection [9-10].

One of the most important research topics for humans' life is climate anomaly detection [11]. The escalated level of global climate change in recent years has increased the number of disasters internationally. Intergovernmental Panel on Climate Change (IPCC) has projected since 2007 that the rising global temperature will increase draught, especially in mid-latitude areas. The water level has also been increasing in many parts of the world. There have been many damages caused by storms [12-13]. As an archipelagic country located on the equator line, Indonesia is prone to natural disasters due to climate change. One of the most severe impacts caused by various climate anomalies is the irregularities of seasons in Indonesia. For example, El Nino phenomenon caused the delay of rainy season and reduced the rainfall to 60% from the normal condition. The shift between dry and rainy seasons can no longer be

* Corresponding Author

predicted. The air temperature escalates dramatically, followed by extreme occurrences, such as tornados, El-Nino, La-Nina, flood, and draughts.

Considering the above problems, studies on anomaly detection in climate are crucial. However, there were only few studies in the field [1-2]. Research on anomaly detection in climate faces some obstacles, such as the availability of data [2, 11], especially labeled data, which are usually expensive and difficult to obtain. The climate data processing also experiences some challenges as the data have strong seasonality components [11] and complex structures due to the influence of spatial and temporal autocorrelation [11, 14-15].

This research aims to develop a model of accurate climate anomaly events. The data were obtained from the climate data, especially temperature in Semarang, a city located in Central Java Province of Indonesia. The city was chosen as it is vulnerable to climate change [16-17]. The city has several areas that are prone to disasters, like lowlands which are often affected by tidal flood, houses on river banks often affected by flood, hilly areas hit by gale or strong wind, and housing areas in the suburbs that are far from reservoirs [18].

To perform anomaly detection analysis of climate, a method that is able to analyze spatial and temporal time-series data is needed. Long Short Term Memory (LSTM) is a method chosen for this research because it has been proven to work effectively in several anomaly detection studies, especially for data with similar characteristics [19-21]. LSTM was also chosen because it could save long term memory by using memory cell units which were self-connected to accumulate external signals. To improve the score of detection accuracy, this research employs Stacked LSTM by applying densely connected.

2 Literature Review

2.1 Climate Anomaly and Related Research

Climate is the average condition of weather for a long period of time in a specific location [14]. The climate features that make the atmosphere conditions vary are sunrays radiation, rains, humidity, air temperature, air pressure, wind, gasses and solid particles composition in the atmosphere. In other words, rainfall and temperature condition in an area indicate the climate of the area [22].

Climate change generally happens seasonally and can be affected by humans' activities which increase perceived climate diversity over a long period of time [14, 22]. The normal climate change refers to changes in the variation of average climate conditions in an area over a long period of time; it is shown statistically to real variability. In other words, climate change should be a long seasonal change in terms of temperature, humidity, wind, and rainfall [13, 23].

Irregular climate change causes climate anomaly. Some experts define anomaly as deviation of observed data that is inconsistent to other data sets [24]. Data anomaly has patterns in the data that does not match with normal data behaviors [2, 25]. Therefore, climate anomaly can be defined as an irregularity in climate patterns that happens in a certain area within a certain period of time [11]. The peculiarity of the climate or weather situation deviates from the previous patterns so that variations and inconsistancies occur. Anomaly calculation is very important as it is the fundamental issue in a climate science since most of climate data analyses depend on the anomaly calculation as the first step [11].

Although most studies in the climate domain face the problem of data set unavailability, the lack of labeled data, as well as expensive and scarce data [21, 26], some experts have conducted valuable studies. Celik [15] compared the implementation of several statistical methods by using Density-Based Spatial Clustering of Appllications with Noise (DBSCAN) algorithm to detect anomaly in the monthly temperature data. The result of the study showed that the statistic method could detect the anomaly points located above and below the threshold (extreme), although it experienced difficulties in detecting the presence of data that rarely happened. The anomaly points in this case did not only show the extreme points, but also the emergence of data that rarely happened; DBSCAN algorithm could find both types of anomalies. However, this study also had a limitation in which the researcher had not analyzed the performance of the compared methods.

Das [14] used Distance-Based Outlier Detection and Neighborhood-based Outlier Detection to detect the anomaly of global air temperature and precipitation gathered from heterogenic sensors within a period of fifty years (1950-1999). The same method had also been used by Adam [27] to detect anomaly in high dimensional spatiotemporal sensor datasets. Both studies showed that the proposed methods could detect the anomaly behaviors in the global climate, including sudden changes and extreme occurrence. Das [14] also clarified the results of the research by opening the historical records of climate occurrence in the respective areas and the results matched temporally and spatially. The limitation of those studies was the absence of comparison and performance test of the used methods.

Auto-regressive non-stationer (AR) model was used by Wakaura [28] to detect anomaly in the average daily surface temperature within a period of 40 years (1961-2000). This model was applied in high-pass filtered data to investigate the relationship between seasonal structures and high frequency variabilities of the anomaly to help find the influence of climate on surface temperature anomaly. The applied model was much better compared to the usual AR model for normal datasets of the surface air temperature obtained by almost all stations in Japan; it showed significant seasonal structures in their automatic correlations. Further, an illustration of the climate inclination value was presented to achieve better non-stationer model performance than the usual AR model. The limitation of the study was that Auto Regressive (AR) was highly accurate for short term predictions. However, for long term predictions, the implementation was not really good; it tended to be flat and constant.

The development of deep learning approach in the past ten years has motivated many researchers to apply deep learning models such as Long Short-Term Memory (LSTM) to analyze climate and detect climate anomaly. For example, a study conducted by Singh [21] analyzed the temperature data from Numenta Machine and Power Demand Dataset. Apart from that, Al Dosari [20] analyzed the sinusoidal time series signal data, power demand, and electro-cardiogram. Some of those study reports showed evidences that LSTM can accurately detect anomaly than some prominent methods. Singh compared it with Deed Forward Neural Network. Meanwhile, Dosari compared it to DBSCAN and Nearest Neighbor [21]. Dosari also concluded that the application of LSTM in anomaly detection needed a parameter optimization for an optimal performance, so that a well-established and thorough experiment was needed.

This research tries to answer the challenges presented in the previous researches, including the availablility of dataset [21, 26], strong seasonality component, and complex structures in the climate data [11, 14-15], as well as the inavailability of performance analysis from the applied methods [11, 15]. The application of Stack and Densely Connected into LSTM in this research is a recommendation from Aldosari's research to modify the model and optimize the parameter to improve the accuracy. The addition of dense layer is used to reduce gradient vanishing so that it enables the error signal to get even further back propagation in the neural network. Meanwhile, Stack LSTM is a machine learning model with some hidden LSTM layers application in it to reconnect the representations learnt in the previous layers and create new representations in the higher abstraction level. By the application of those two things, it is expected that the accuracy of detection in this research can improve significantly compare to the ordinary LSTM model.

2.2 Long Short Term Memory (LSTM) Method

Long Short-Term Memory (LSTM) is a deep learning model which can be viewed as an improvement of the previous Recurrent Neural Network (RNN) model, which was designed for sequence and pattern learning. In contrast to neural network model whose architecture was used as its base, the RNN architecture has a special function that makes it able to save past information output by a node fed to the respected node as its additional input. In many studies, RNN showed high accuracy in solving many tasks using sequential data as input. However, the algorithm has a weakness of not being able to store data for a long time. For this reason, Long Short-Term Memory (LSTM) was proposed by Hochreiter and Schmidhuber [29]. LSTM has been proven to be more effective than general RNN, especially when multiple layers are used in every time step.

LSTM improved RNN by providing a mechanism to store past information from a long period of time by using memory cell units connected to accumulate the external signals. The LSTM structures can be seen in the Figure 1 below. The mechanism to store past information was instrumented by 4 gates namely: input node, input gate, forget gate, and output gate [21, 29-30].

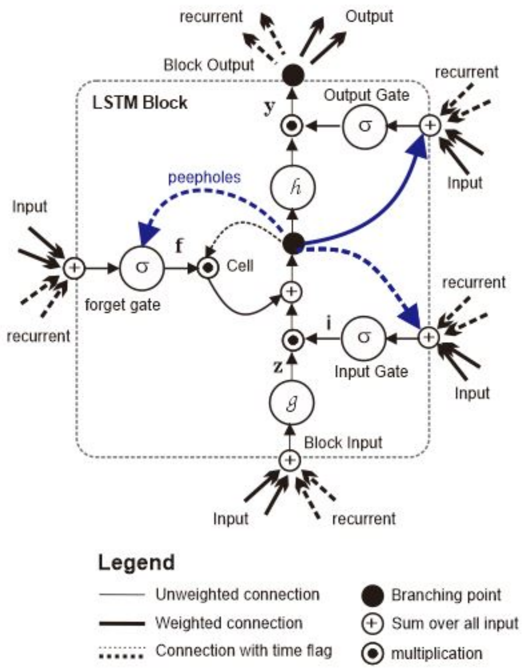

Fig. 1. LSTM with input, output and forget gate [21, 30]

LSTM was started by running sigmoid layer that cut the parts of which cell state are created [29-30]. Then, forget gate decides the information that should be thrown away from cell state with the following formula:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \tag{1}$$

This forget gate decides whether or not information from input X1 and output ht-1 is allowed to pass the gate. The output value close to 1 means the output is allowed to pass, whereas the output close to 0 will be ignored.

The input gate will also decide what kind of new information will be saved in the cell state. The gate has two similar layers with the following equations.

$$i_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_i). \tag{2}$$

$$C_t' = \tan h(W_C \cdot [h_{t-1}, x_t] + b_C). \tag{3}$$

Equation 2 is a sigmoid layer that decides which values will be renewed. Meanwhile, equation 3 is tanh layer that creates value vectors for new candidates that will be added into the state. The next step is updating the old cell state Ct-1 into a new cell state Ct with the following equation:

$$C_t = f_t * C_{t-1} + i_t * C_t'. \tag{4}$$

In equation 4, the old state of Ct-1 times ft is used to forget something that has been decided. After that, adding it with * Ct as the values of new candidates obtained from how we decide to renew every state value. Next, output gate will decide what values to be created. The output is based on cell state to be the desired version (filtered version).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o). \tag{5}$$

$$h_t = o_t * \tan h(C_r). \tag{6}$$

By training the four gates using the training dataset, the job delegation is done well where input gate decides what to be stored, cell state stores the information, and output gate decides when the information is used. Therefore, the outstanding events or occurrence can be saved or stored longer in the memory.

3 Methodology

3.1 Dataset and Research Framework

The LSTM based model was tested using datasets of air temperature in Semarang City from 2011-2016. The dataset was obtained from Meteorology, Climatology, and Geophysics Agency (BMKG) of Semarang City, Central Java Province, Indonesia. The observation was conducted from 2011-2016 with a total of 1980 data records.

Research Framework of this study is shown in Fig. 2. From the diagram, it can be seen that after obtaining the datasets, preprocessing step was implemented. It aimed to ensure that the data were ready to be used in both the training and testing processes with the model developed for anomaly detection. Deep learning model proposed for this study was Stacked-Densely Connected LSTM.

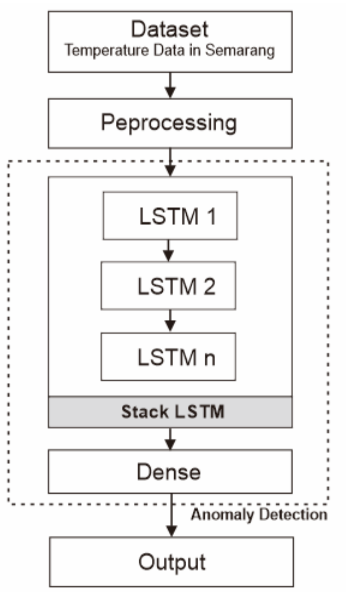

Fig. 2. The research framework

3.2 Model Performance and Finding the Anomaly

The performance of the machine learning model in this research was measured by Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) metrics for the following reasons. First, MSE gave understandings of how consistent the developed model was. By minimizing the MSE value, the variant of the model was also minimized. The model with less variant was able to give more consistent results for the entire data input. Below is the MSE calculation.

$$MSE = \frac{1}{n}\sum_{t=1}^{n} e_t^2 , \tag{7}$$

Second, RMSE was used to evaluate the prediction technique used especially in calculating the accuracy level of a prediction model. RMSE is the mean of the total errors' square that can describe how big the error made by a prediction model is. The low RMSE value showed that the variation value obtained by a prediction model was close to the observed value variation. Below is the RMSE calculation.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} e_t^2} . \tag{8}$$

In the context of machine learning, anomaly is basically a deviation of previous general patterns that have been trained by the model. This can be seen from the distance between data observation and data prediction. If the deviation is too far, then anomaly can potentially happen. RMSE from the time series data is often used in adaptive thresholding model to detect anomaly. Several literatures stated that an anomaly point is a point located more than two standard deviations of means [31]. So, for the standardized normal dataset with the mean of 0 and standard deviation of 1, all the data points between -2 and 2 are considered normal. However the data points more than 2.5 will be regarded as anomalies.

4 Result and Discussion

4.1 Data Exploration

The LSTM-based model was tested using datasets of air temperature in Semarang City within the period of 2011-2016. The dataset was obtained from Meteorology, Climatology, and Geophysics Agency (BMKG) of Semarang City, Central Java Province, Indonesia. The observation data were from the period of 2011-2016 with a total of 1980 data records.

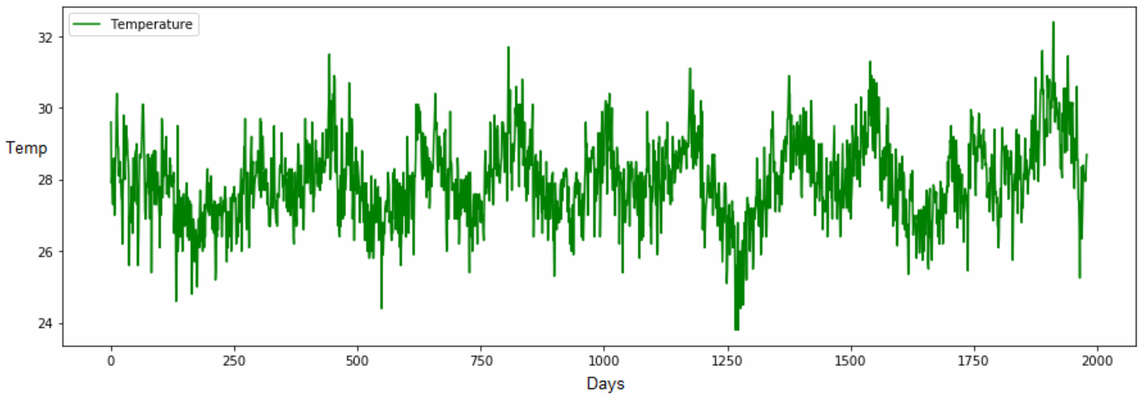

Fig. 3. Sequence temperature data visualization

Some descriptive statistic of the input dataset can be summarized in Table 1.

Table 1. Statistic data

Descriptive Statistic	Value
Mean	28.0
Standard Deviation	1.1
Max	32.4
Min	23.8

Data normalization was implemented through several phases; it started from deleting unnecessary features from the obtained data, such as dates, maximum and minimum temperatures. Improvements of data sequences were done on August 1, 2011 to December, 31 2016. Further, data normalization process was done by using MinMAxScaler, a helpful function of Python Scikit-Learn. Next, to prepare the process by using LSTM model, the dataset was separated into two parts, which were data training and data testing. In this research, the separation was done by classifying the 4 year data for data training and the last 1 year data for data testing.

4.2 Anomaly Detection Using Basic LSTM

In the first experiment, the training model was treated using Basic LSTM, with the following model.

Table 2. Model sumary of Basic LSTM

Layer (Type)	Output Shape	Param #
LSTM	(None, 1)	12
Dropout	(None, 1)	0
Total params		12
Trainable params		12
Non trainable params		0

The dropout layer was used in the Basic LSTM architecture to avoid overfitting. Based on the study by Sriyastaya [30], dropout between 0.5 for the hidden unit and 0.2 for the input can run well in several tasks.

In the LSTM model, dropout (0.2) was used, resulting in a training process with the accuracy of Mean Square Error (MSE) value of 0.7124 and Root Mean Square Error (RMSE) value of 0.8440. The training duration was 81.904 seconds. Meanwhile, the process of testing model produced a graphic presented in Figure 4 below. The accuracy values obtained from the process were 0.6644 for the MSE and 0.8151 for the RMSE. The graphic of prediction testing is shown below.

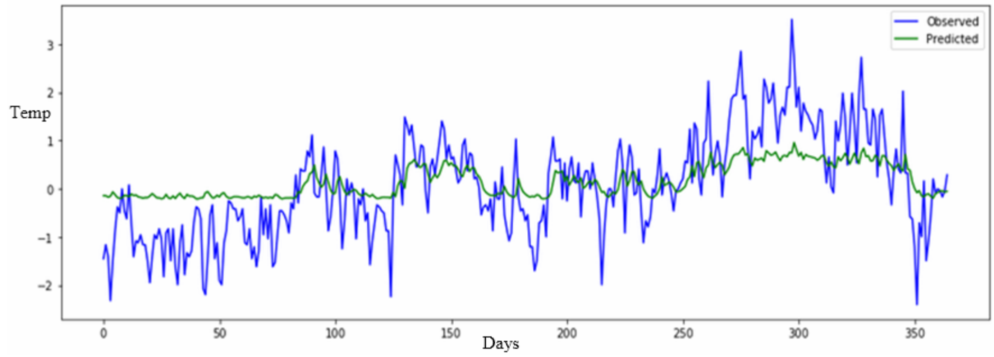

Fig. 4. The graphic of prediction testing

Anomaly detection was done by comparing the observed data with the prediction data. The prediction data are the data made based on the generalized data from the model, so that anomaly can be seen when there is a high deviation between the prediction data and the real observed data. To see the deviation, RMSE graphic is used, shown in Fig. 5.

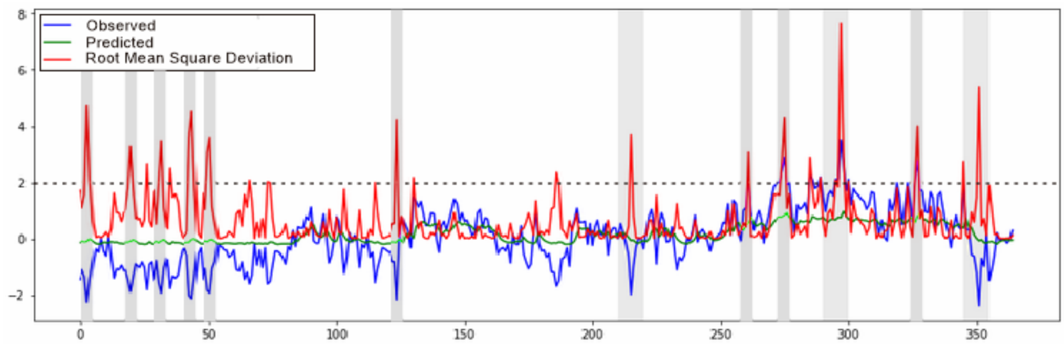

Fig. 5. The results of anomaly detection

As Fig. 5 indicates, the result of the detection testing found 12 dots or points in 1 year predicted as anomaly dots because the deviation between the predicted and observed data was high. The test values of the model were 0.6644 for MSE and 0.8151 for RMSE.

4.3 Anomaly Detection Using Densely Connected LSTM

In the second experiment, densely-connected output layer was added with linear activation function. The architecture can be seen in Fig. 6.

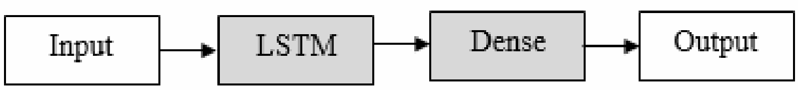

Fig. 1. The results of anomaly detection using Basic LSTM

Dense was an additional layer that is inserted to reduce network complexity. The use of dense layer in the above model was aimed to avoid gradient vanishing through more direct connections between layers.

Dense connection made it possible for the error signal to get back further propagation in the deep neural network, so that the gradient vanishing became smaller. The application of dense layer was able to outperform the residual network that had previously been a superior model.

Table 3. Model summary of DC-LSTM

Layer (Type)	Output Shape	Param #
LSTM	(None, 32)	4352
Dropout	(None, 32)	0
Dense	(None, 1)	33
Activation	(None, 1)	0
Total params		4,385
Trainable params		4,385

Table 3 shows the summary of the Densely Connected LSTM Model. The model produced training process with MSE result of 0.6086 and RMSE result of 0.7801. The duration of the training process was 363.123 seconds. Meanwhile, the testing model process produced the graphic presented in Fig. 7. The accuracy values produced from the process were 0.5734 for MSE and 0.7572 for the RMSE. The graphic of prediction testing is shown below.

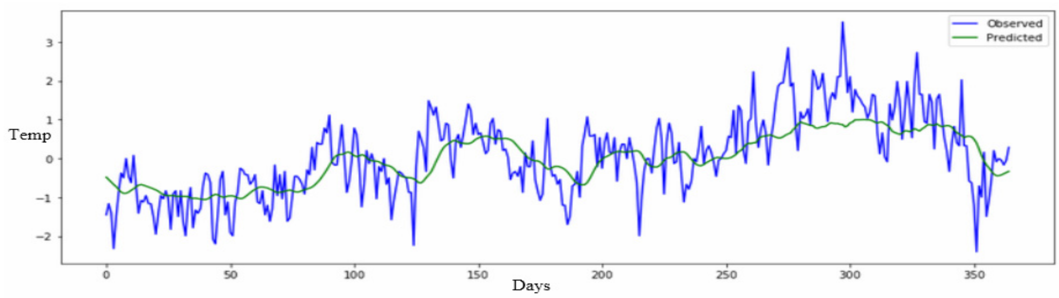

Fig. 7. Prediction testing graphic with densely connected LSTM

From the picture, we can observe that in the detection result, there were 9 points or dots in 1 year, which were predicted as anomaly dots because of the high deviation between the predicted and observed data. The result of the second experiment showed that the addition of dense layer in LSTM could reduce the MSE and RMSE values, which means that it improved the value of the accuracy of the model. In the training process, MSE decreased from 0.7124 for the Basic LSTM to 0.6086 for DC-LSTM, which meant it improved the training accuracy value of ± 15%. The same thing happened with the MSE value that changed from 0.6644 for the Basic LSTM to 0.5734 for DC-LSTM, which meant it improved the accuracy value of ± 14%.

4.4 Anomaly Detection Using Stack-DC LSTM

The third experiment applied Stacked-DC LSTM with the architecture as shown in Figure 8. Stacked LSTM is a learning machine that applies some hidden LSTM layers, in which every layer contains multiple memory cells. Having several hidden LSTM layers makes this model deeper, so that it is called deep learning that can gather object descriptions more accurately [32].

Fig. 8. The architecture of Stacked DC-LSTM

The success of deep neural network is generally associated with hierarchies introduced in various layers. Every layer processes some parts of the tasks that will be completed and continues them to the next tasks.

Table 4. Model summary of Stack DC-LSTM

Layer (Type)	Output Shape	Param #
LSTM 1	(None, 99, 32)	4,352
Dropout 1	(None, 99, 32)	0.00
LSTM 2	(None, 99, 128)	82,432
Dropout 2	(None, 99, 128)	0.00
LSTM 3	(None, 100)	91,600
Dropout 3	(None, 100)	0.00
Dense	(None, 1)	101
Activation	(None, 1)	0.00
Total params		178,485
Trainable params		178,485

The model produced training process with MSE score of 0.5985 and RMSE score of 0.7736. The duration of the training process was 18.0329 seconds. Meanwhile, the model testing process produced graphic shown in Fig. 9.

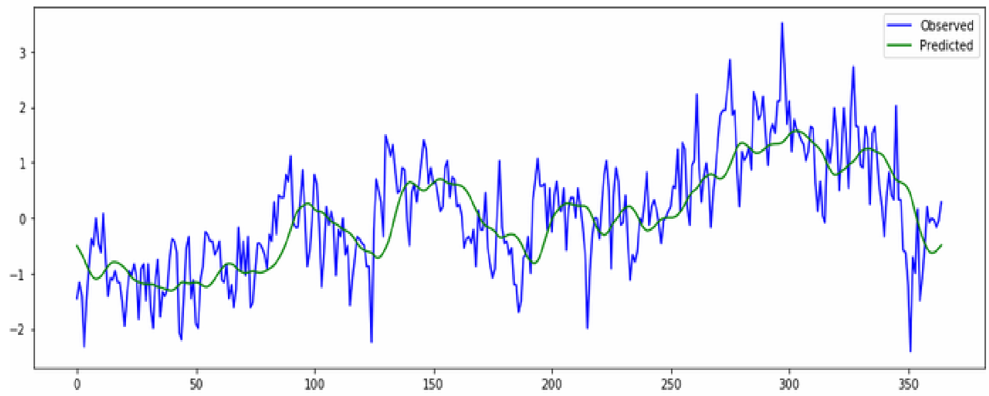

Fig. 9. Prediction testing graphic with DC-LSTM

The accuracy values produced for the prediction testing on Fig. 10 were 0.5350 for MSE and 0.7320 for RMSE. Further, from the anomaly detection analysis, 7 anomaly dots were found in 1 year, as per seen in Fig. 10.

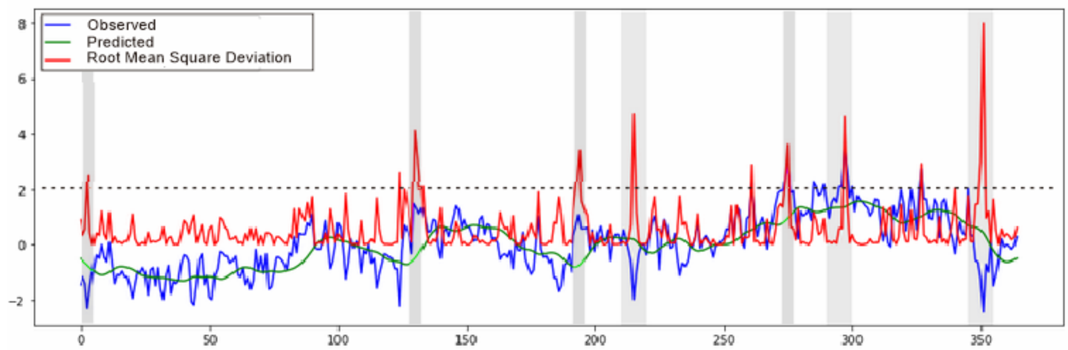

Fig. 10. Prediction testing graphic with DC-LSTM

The result of the third experiment showed that the application of stacked LSTM could improve the value of the model's accuracy even higher. In the training process, the MSE value decreased from 0.6086 for DC-LSTM to 0.5985 for Stack LSTM. The same thing happened in the testing process, in which the MSE value changed from 0.5734 fot DC-LSTM to 0.5350. The accuracy level improved ± 5% from the previous model.

4.5 Improving LSTM Model

The next experiment was tuning the parameters of the three models. Table 5 shows the summary of the training and testing results of those experiments. From the table, it can be seen that the accuracy level of Stack and DC-LSTM was better compared to DC-LSTM and Basic LSTM, in both training and testing processes.

Table 5. Summary of the experiments' Results

Model	E	Training			Testing		N
		T	M	R	M	R	
Basic LSTM	1	21.90	0.71	0.84	0.66	0.81	12
Basic LSTM	50	56.79	0.58	0.76	0.53	0.73	10
Basic LSTM	100	108.61	0.53	0.73	0.45	0.67	10
DC-LSTM	1	63.12	0.61	0.78	0.57	0.76	9
DC-LSTM	50	59.89	0.45	0.67	0.35	0.60	7
DC-LSTM	100	120.73	0.44	0.66	0.34	0.58	6
Stack & DC-LSTM	1	18.03	0.59	0.77	0.53	0.73	7
Stack & DC-LSTM	50	619.37	0.45	0.67	0.34	0.58	4
Stack & DC-LSTM	100	828.27	0.43	0.65	0.34	0.50	4

Notes: E: number of epoch; M: MSE
T: running time (sec); R: RMSE
N: Number of detected anomaly

The increased e-poch in every model also reduced the MSE and RMSE values, which means it increased the accuracy of the model. From the results of anomaly detection, it can also be seen that Stack and DC LSTM had higher accuracy level by finding real anomaly values, unlike other previously tested models that still contained bias in finding the anomaly data.

4 Conclusions

Climate anomaly detection is extremely crucial to be done in Indonesia due to the geographical condition that is prone to disasters related to climate change. To perform anomaly detection analysis from the climate data, a method that can analyze time series, spatial, and temporal data is needed. Long Short-Term Memory method is used in this study since it has been proven to work effectively in several previous anomaly detection studies, especially for data with similar characteristic. The application of Dense in Basic LSTM could reduce network complexity and improve the detection accuracy. The research output showed that the addition of dense layers in the LSTM could improve the training and the testing accuracies of ± 15%. Dense layers could reduce gradient vanishing through more direct connections between layers so that error signals could get propagation back further in deep neural network.

Representational optimization in machine learning model in this research was done by applying Stacked LSTM. Stack LSTM was a machine learning model that applied several hidden LSTM layers in it. The additional hidden LSTM layers were used to reconnect learnt representations from the previous layers and create new representations in higher abstraction levels. The research output showed that the application of Stack and Densely Connected LSTM could improve the training and the testing accuracies of ± 20% compared to the Basic LSTM model. The stack LSTM in this model was proven to make the model deeper and increase the accuracy level. Furthermore, an experiment by increasing e-poch in every model could also de-crease the MSE and RMSE values, increasing the accuracy of the model.

The RMSE graphic in the testing data showed the differences between the predicted and the observed data so that it could be used to detect the presence of anomalies in temperature data in one year. Further studies should verify the anomaly data by using ground truth data, which indicate the real anomaly occurrences in Semarang City. Therefore, it can be proven whether or not the results of the anomaly detection are similar to the real conditions.

References

[1] A. Patcha, J.M. Park, An Overview of Anomaly Detection Techniques: Existing Solutions and Latest Technological Trends, Computer Networks 51(12)(2007) 3448-3470.

[2] V. Chandola, A. Banerjee, V. Kumar, Anomaly detection: a survey, ACM Computing Surveys 41(3)(2009).

[3] S. Omar, A. Ngadi, H.H. Jebur, Machine Learning Techniques for Anomaly Detection: An Overview, International Journal of Computer Applications 79(2)(2013) 0975-8887.

[4] C. Guo, L. Xu, H. Liu, L. Wang, X. Yu and B. Han, The Financial Data of Anomaly Detection Research Based on Time Series, in: Proc. 2015 International Conference on Computer Science and Applications, 2015.

[5] M. Ahmed, N. Choudhury, S. Uddin, Anomaly Detection on Big Data in Financial Markets, in: Proc. 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, 2017.

[6] C.F. Tsai, C.Y. Lin, A triangle area based nearest neighbors approach to intrusion detection, Pattern Recognition 43(1)(2010) 222-229.

[7] P. Casas, P. Fiadino, A. D'Alconzo, Machine-Learning Based Approaches for Anomaly Detection and Classification in Cellular Networks, in: Proc. 2016 The Traffic Monitoring and Analysis workshop, 2016.

[8] Srinivasan, Anomalies Detection in Healthcare Services, IT Professional 16(6)(2014) 12-15.

[9] L. S. Lopes, L.M. Camarinha-Matos, A machine learning approach to error detection and recovery in assembly, in: Proc. 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems, 1995.

[10] N. Even-Chen, Stavisky, J. C. Kao, S. I. Ryu, K. V. Shenoy, Auto-deleting brain machine interface: Error detection using spiking neural activity in the motor cortex, in: Proc. 2015 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015.

[11] J. Kawale, S. Chatterjee, A. Kumar, S. Liess, M. Steinbach, V. Kumar, Anomaly Construction in Climate Data : Issues and Challenges, in: Proc. 2011 NASA Conference on Intelligent Data Understanding CIDU, 2011.

[12] Contribution of Working group II to the Fourth Assessment Report, Intergovernmental Panel on Climate Change (IPCC) Summary for Policymakers, 2007.

[13] Climate Change 2014: Mitigation of Climate Change, Intergovernmental Panel on Climate Change (IPCC) Assessment Report, 2014.

[14] M. Das, S. Parthasarathy, Anomaly detection and spatio-temporal analysis of global climate system, in: Proc. 3rd International Workshop on Knowledge Discovery from Sensor Data, 2009.

[15] M. Celik, F. Dadaser-Celik, A. Dokuz, Anomaly detection in temperature data using dbscan algorithm, in: Proc. 2011 Innovations in Intelligent Systems and Applications (INISTA) International Symposium, 2011.

[16] United Nations Population Fund (UNFPA), Climate vulnerability and adaptation in the Semarang Metropolitan Area: a spatial and demographic analysis. <https://www.preventionweb.net/publications/view/35324>, 2015 (accessed 20.08.18).

[17] Urbanisation, Demographics and Adaptation to Climate Change in Semarang, Indonesia, Urbanization and Emerging Population Issues Working Paper, IIED, 2016.

[18] Wijaya, Empat Alasan Semarang Rentan Bencana Perubahan Iklim, Buletin Ekuatorial 10(2)(2016).

[19] P. Maholtra, L. Vig, G. Shroff, P. Agarwal, Long Short Term Memory Networks for Anomaly Detection in Time Series, in: Proc. 23rd European Symposium on Artificial Neural Networks-Computational Intelligence and Machine Learning, 2015.

[20] M. S. AlDosari, Unsupervised Anomaly Detection in Sequences Using LSTM-RNN, [thesis], George Mason University, 2016.

[21] A. Singh, Anomaly Detection for Temporal Data using Long Short-Term Memory (LSTM), [thesis], KTH Information and Communication Technology, 2017.

[22] Tondokusumo, Memahami La Nina, Dampaknya bagi Produksi Tanaman Pangan, dan Adaptasinya, Agricultural Bulletin 12(1)(2016).

[23] NASA Website, What is Climate Change <http://www.nasa.gov/audience/forstudents/k-4/stories/nasa-knows/what-is-climate-change-k4.html>, 2014 (accessed 13.08.18).

[24] Z. Yang, Meratnia, P. Havinga, Outlier Detection Techniques For Wireless Sensor Networks: A Survey, Department of Computer Science University of Twente, 2007.

[25] A. Shrutti, S. Janpreet, Outlier Detection using k-Mean and Hybrid Distance Technique on Multidimensional Data Set, International Journal of Computer Applications 122(8)(2013) 975-8887.

[26] G. Campos, A. Zimek, J. Sander, R. J. G. B. Campello, B. Micenková, E. Schubert, I. Assent, M. E. Houle, On the evaluation of unsupervised outlier detection: measures, datasets, and an empirical study, Data Mining and Knowledge Discovery, 30(4)(2016) 891-927.

[27] N. R. Adam, V. P. Janeja, V. Atluri, Neighborhood based detection of anomalies in high dimensional spatio-temporal sensor datasets, in: Proc. 2004 ACM symposium on Applied computing, 2004.

[28] M. Wakaura and Y. Ogata, A time series analysis on the seasonality of air temperature anomalies, Meteorological Applications, 14(4)(2007) 425-434.

[29] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural computation, 9(8)(1997) 1735-1780.

[30] K. Greff, R. K. Srivastava, J. KoutnÃk, B. R. Steunebrink, J. Schmidhuber, LSTM: A Search Space Odyssey, in: Proc. 2017 IEEE Transactions on Neural Networks and Learning Systems, 2017.

[31] C. Chio, D. Freeman, Machine Learning and Security, O'Reilly Media, Gravenstein Highway North, 2017.

[32] J. Brownlee, Long Short-Term Memory Networks With Python: Develop Sequence Prediction Models With Deep Learning, Machine Learning Mastery, 2017.