Monthly Prediction of Rainfall in Nickel Mine Area with Artificial Neural Network

Tedy Agung Cahyadi[1, a)], Herlina Jayadianti[2, b)], Nur Ali Amri[1, c)], Muhammad Fathurrahman Pitayandanu[3, d)], Abu Ashar[4, e)]

[1]Lecturer, Mining Engineering, Faculty of Mineral Technology, Universitas Pembangunan Nasional "Veteran" Yogyakarta, 55283 Sleman, Indonesia

[2]Lecturer, Informatics Engineering, Faculty of Industrial Engineering, Universitas Pembangunan Nasional "Veteran" Yogyakarta, 55283 Sleman, Indonesia

[3] Undergraduate Student, Informatics Engineering, Faculty of Industrial Engineering, Universitas Pembangunan Nasional "Veteran" Yogyakarta, 55283 Sleman, Indonesia

[4] Director of Enviroment and Energy, PT. Vale Indonesia, Sorowako, Luwu Timur, 92984 South Sulawesi, Indonesia

[a)]Corresponding author: tedyagungc@upnyk.ac.id
[b)]herlinajayadianti@gmail.com
[c)]nuraliamri@upnyk.ac.id
[d)]fathurdanu@gmail.com
[e)]Abu.Ashar@vale.com

Abstract. Rainfall prediction in the mining area was needed to assist the process of mine drainage, and monitoring the availability of water in the reservoir, which is a source of hydroelectric power. Various ANN architectures were examined in this study to make predictions with monthly rainfall parameters t-3, t-2 and t-1. It included supporting parameters such as average exposure time, humidity, temperature, wind speed, and finally predicts rainfall in the month of occurrence. The ANN architecture contains a hidden layer which was examined by the optimal number of neurons and epochs. Hidden neurons were tried from seven to fourteen. The results of experiment showed that the architecture [7-8-1, 500 epochs] concluded that ANN gave good resultts of MSE which were 0.05865 for training and 0.08725 for testing. Furthermore, the ANN algorithm has provided to predict rainfall with a good model.

INTRODUCTION

Weather forecasting is an action to predict the conditions of the atmosphere for a given location and time with applicating science and technology. It is done by collecting the current state data of the atmosphere at a given place. The data is then analyzed, processed, and used to support various research needs. Predictive analytics or modelling is an area of statistics that deals with extracting information from data and using it to predict trends and behavior patterns. One of the applications is in rainfall prediction based on data collected through observation. The prediction model adopted should be accurate. For example, Rachmawati predicted the dengue disease outbreak using the C-Support Vector Classification (2016), which was proven to have high level of accuracy. Furthermore, it can be used in mining activities and infrastructure system engineering [1].

It is difficult to accurately predict rainfall because weather is based on an infinitely complex and constantly changing system. Various studies conducted were unable to predict rainfall with certainty. Research by Wahyuni, (2018) had predicted rainfall using the Tsukamoto Fuzzy Inference System in Tengger, Indonesia. The research successfully concluded the optimization of fuzzy membership with the recommended Meta-Heuristic algorithm [2]. Furthermore, Nourani (2016) presented a new generation of Artificial Intelligence-based models for daily rainfall-runoff modeling of the watershed, which was the first hydrological implementation of the Emotional Artificial Neural Network (EANN). Improving the network learning process is a goal of why it used An EANN with simulated emotional parameters [3].

Tan et al., (2018) recommended the use of the SAR Model during dry seasons, and the use of AEEMD-ANN in rainy seasons. This was shown in the research, which examined the locations of Ertan, Cuntan and Yichang Stations, China [4]. Also, the research on modeling rainfall runoff using Artificial Neural Network (ANN) Back-propagation was conducted by Hadihardaja et al., (2005). It was concluded that ANN Back-propagation gave relatively good results, but the results were not accurate [5]. Xiang et al., (2018) succeeded in predicting rainfall with the conclusion that the E-SVR-ANN model was recommended for further research [6]. The Artificial Neural Network method had also been used in modeling rainfall runoff in watersheds on the island of Bali, which was conducted [7]. The results show that Back-propagation is an effective tool for improving the accuracy of predictions.

Susilokari et al., (2015) conducted research in the irrigation areas of Curugagung, Cileuleuy, Cinangka, and Pangsor to make comparisons using the methods namely: the Fast Fourier Transformation (FFT), Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Network. The ANN model was recommended because it provided the best results compared to FFT and ARIMA [8]. Famesa et al., (2015) predicted rainfall with Hybrid Neural Network and Evolutionary Programming algorithms. The result recommended Hybrid Artificial Neural Network Algorithm models in Evolutionary Programming Algorithms with architecture 3-1-1 on the Simple Moving Average 3 method -MA and architecture 5-2-1 at the Simple Moving Average 5-MA in the Soreang Region of Bandung Regency [9].

Furthermore, Sutikno et al., (2015) recommended the ASTAR method, especially EX-ANTE at Juanda Station Surabaya to estimate the weather, using the ARIMA, ANN and ASTAR methods [10]. Rachmawati (2015) also conducted research on rainfall prediction in Pontianak on a daily and monthly scale. The result showed was better recognized by ANN than the basic and daily scales [1]. Winarti (2018) conducted research to estimate the incidence using the ANN-Fuzzy approach in Pontianak, Indonesia. Based on the result, Winarti recommended using ANN-Fuzzy for rain predictions [11]. Gumbel (1941) conducted rainfall modelling using an empirical model. Furthermore, the results of Gautama's (2012) estimation of rainfall used the Gumbel approach to design the mine drainage system facilities [12]. He objective of this research was to predict monthly rainfall based on humidity, temperature, wind and irradiation data.

MATERIALS AND METHODS

This chapter explains about the material and the method that needed for build the model. Furthermore, these divided in two subsections specifically dataset and Artificial Neural Network. ANN is recommended for rainfall forecast [13]

Dataset

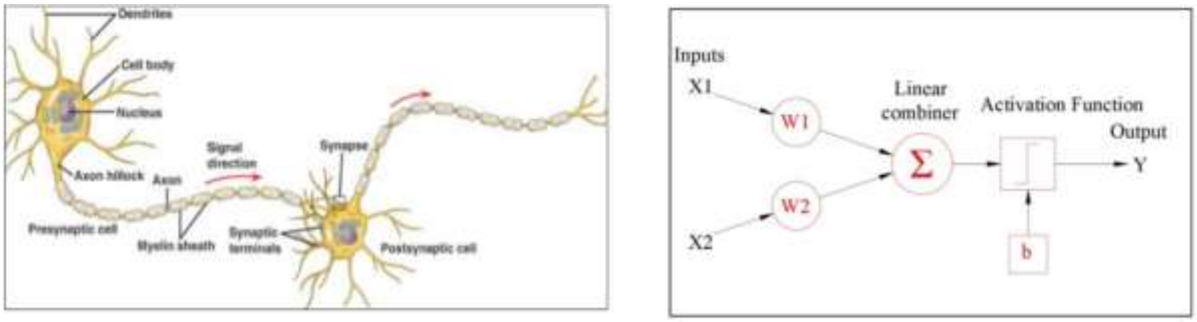
The dataset used for the rainfall modelling process was obtained daily at PT. Vale Indonesia. This was processed into monthly from 2008 to 2017 by adding up the rainfall data. It was also combined with supporting parameters which include: average irradiation time, average humidity, average temperature, and average wind speed. This data collection was from the nearest station, which is the BMKG Meteorological Station, in Kabu North Luwu.

Artificial Neural Network (ANN)

ANN is a concept in Engineering adopted to solve problems of Artificial Intelligence, by building a system with links that stimulate the human nervous system. It is based on a collection of connected units or nodes called artificial neurons. Neurons consist of inputs (dendrites), and one output (synapses through axons). The primary process of the human nervous system of the brain is the fundamental unit of information processing [14].

This research proposed an ANN model that used supervised learning with seven input parameters such as monthly rainfall t-3, t-2, t-1, and supporting parameters, namely: average exposure time, average humidity, average temperature, average wind speed, and finally T, as the target of rainfall in the month of occurrence. In defining the input layers: t-3 is the monthly rainfall three months before it occurs; t-2 is the monthly rainfall two months before it occurs; and t-1 is the monthly rainfall a month before it occurs. **FIGURE-1a** and **1b** show the differences about them, **FIGURE-1a** shows the human nervous system and **FIGURE-1b** shows the Artificial Neural Network.

Mislan (2015) only used t-2 and t-1 parameters without using climatological parameters. This study tried to complement previous researches [15]. ANN Structure and Architecture show in **FIGURE 2**.



(a) (b)

FIGURE 1. Neurons in the Human Brain (a) and ANN (b) [16]

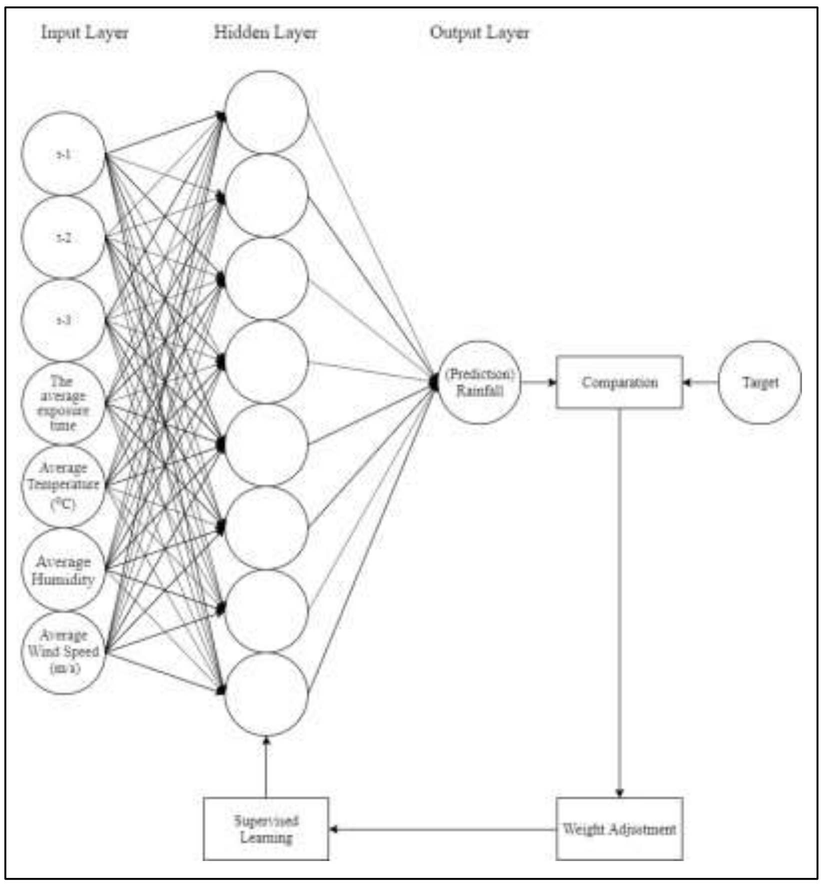


FIGURE 2. ANN Structure and Architecture

RESULT AND DISCUSSION

This chapter divided by three sub chapter, those are Determine Training Data and Test Data, Determine Hidden Neuron and Epoch, and Validation.

Determine Training Data and Test Data

The data used included climatological parameters such as: average exposure time, average humidity, average temperature, and average wind speed from BMKG Andi Jemma Meteorological Station in North Luwu Regency. The results were presented in a tabular form as shown in **TABLE-1** which comprises of training dataset for average temperature, humidity, exposure time and wind speed. Conversely, **TABLE-2** shows the parameters for the test dataset and their corresponding values. This research was implemented with Python 3.7 programming language, Anaconda 4.7.10 platform, and IDE Spyder 3.3.6. In addition, it was analysed using Tensorflow 1.14.0 and Hard 2.2.4 libraries which were computed with CPU.

TABLE 1. Training Dataset

No	Month - Year	Input							Target
		t-3	t-2	t-1	Average Temperature (°C)	Average Humidity (%)	Average Exposure Time (jam)	Average Wind Speed (m/s)	T
1	1-2008	294,9	122,7	116,3	26,90	81,48	5,10	1,71	136,1
2	2-2008	122,7	116,3	136,1	27,10	79,31	4,38	1,76	142,8
3	3-2008	116,3	136,1	142,8	26,76	82,71	4,81	1,23	225,2
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
86	2-2015	135	262,1	193	26,446	85,214	5,889	1,000	377

TABLE 2. Test Dataset

No	Month - Year	t-3	t-2	t-1	Average Temperature (°C)	Average Humidity (%)	Average Exposure Time (jam)	Average Wind Speed (m/s)	T
87	3-2015	262,1	193	377	26,976	83,931	5,675	0,968	392
88	4-2015	193	377	392	27,038	82,448	7,374	1,000	820,8
89	5-2015	377	392	820,8	27,121	81,483	6,596	1,033	202,2
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
108	12-2016	138,2	185,4	271,4	27,71	80,03	5,68	0,90	320,2

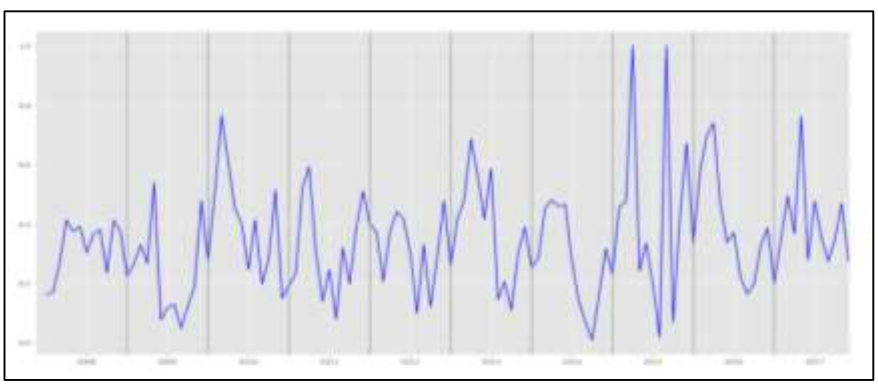


FIGURE 3. Graphic Actual Rainfall from 2008 to 2017

Determine Hidden Neuron and Epoch

Based on its ability, the ANN model was used to identify complex non-linear relationships between input variables to output variables, without the need to understand the nature of the physical process. It consists of a total of eight neurons comprising of seven input and one output layer. A hidden layer was selected and the number of neurons was experimentally determined by trial and error. It used the Rectified Linear function as the activation function and this was because it reduces the likelihood of the vanishing of the gradient in deep architecture. The Adadelta optimizer and different numbers of epochs was used to determine the architecture that works optimally. The Mean Square Error was calculated for evaluating the accuracy, by comparing the actual value with the value generated by the model.

The prediction model was selected by repeating the ANN model generation process based on the average accuracy. The epoch used varied from 250 to 4000. The first step was to determine the number of neurons in the hidden layer. **TABLE-3** shows the minimum average MSE value as 0.073460, with a total of eight neurons using 1000 epoch by default.

TABLE 3. Determine the Number of Neurons in the Hidden Layer

No	Hidden Neuron	Mean Square Error		Average
		Training	Test	
1	7	0,06286	0,0963	0,079580
2	**8**	**0,05937**	**0,08755**	**0,073460**
3	9	0,06440	0,10808	0,086240
4	10	0,06356	0,09140	0,077480
5	11	0,06515	0,13186	0,098505
6	12	0,07338	0,08572	0,079550
7	13	0,06851	0,10769	0,088100
8	14	0,07371	0,09716	0,085435

Furthermore, the next process was to ascertain the best epoch value by adopting the trial and error method. This was done using a different number of epochs, ranging from 250 to 5000. The result of the trial was shown in **TABLE-4**.

TABLE 4. Determine the Number of Epoch

No	Epoch	Mean Square Error		average
		Training	Test	
1	250	0,05909	0,08915	0,074120
2	**500**	**0,05865**	**0,08725**	**0,072950**
3	750	0,05930	0,08849	0,073895
4	1000	0,05937	0,08755	0,073460
5	2000	0,06115	0,09524	0,078195
6	3000	0,06250	0,09555	0,079025
7	4000	0,06644	0,12937	0,097905

TABLE-5 shows the minimum average value calculated from the training and test data as 0.072959, with epoch 500, which will be the benchmark for determining the number of neurons in the hidden layer.

TABLE 5. Determine the Number of Neurons in the Hidden Layer

No	Hidden neuron	Mean Square Error		Average
		Training	Test	
1	7	0,06164	0,12005	0,090845
2	**8**	**0,05865**	**0,08725**	**0,072950**
3	9	0,06308	0,09921	0,081145
4	10	0,06207	0,09172	0,076895
5	11	0,06159	0,11093	0,086260
6	12	0,06797	0,09063	0,079300
7	13	0,06373	0,10072	0,082225
8	14	0,06871	0,09265	0,080680

The simple linear regression was adopted to model the relationship between the variables (actual and predicted data), using the equation:

$$Y = a + bX \qquad \text{i)}$$

Where Y is the Response (Dependent), and X is the Predictor/Cause Factor (Independent) Variable. The value of 'a' is a constant while that of 'b' is a regression coefficient (slope), which is also the magnitude of the Response generated by the Predictor. They can be calculated using the mathematical formula below:

$$a = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} \qquad \text{ii)}$$

$$b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \qquad \text{iii)}$$

The value of 'a' was calculated to be 36.0718 while that of 'b' was 1.0882. Using the linear regression model to analyze the relationship between the actual and predicted data, the coefficient of determination was 0.48425 as shown in **FIGURE-4**.

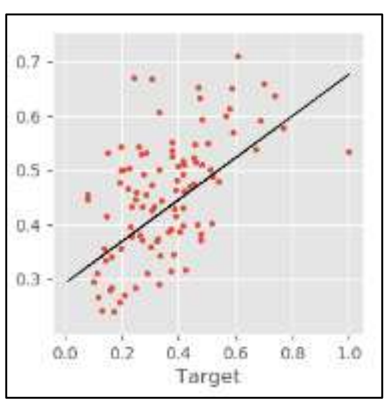


FIGURE 4. Graphic Actual Rainfall from 2008 to 2017

The results from training and validation stages agree with each other; therefore, it can be concluded that the ANN model was suitable for predicting rainfall.

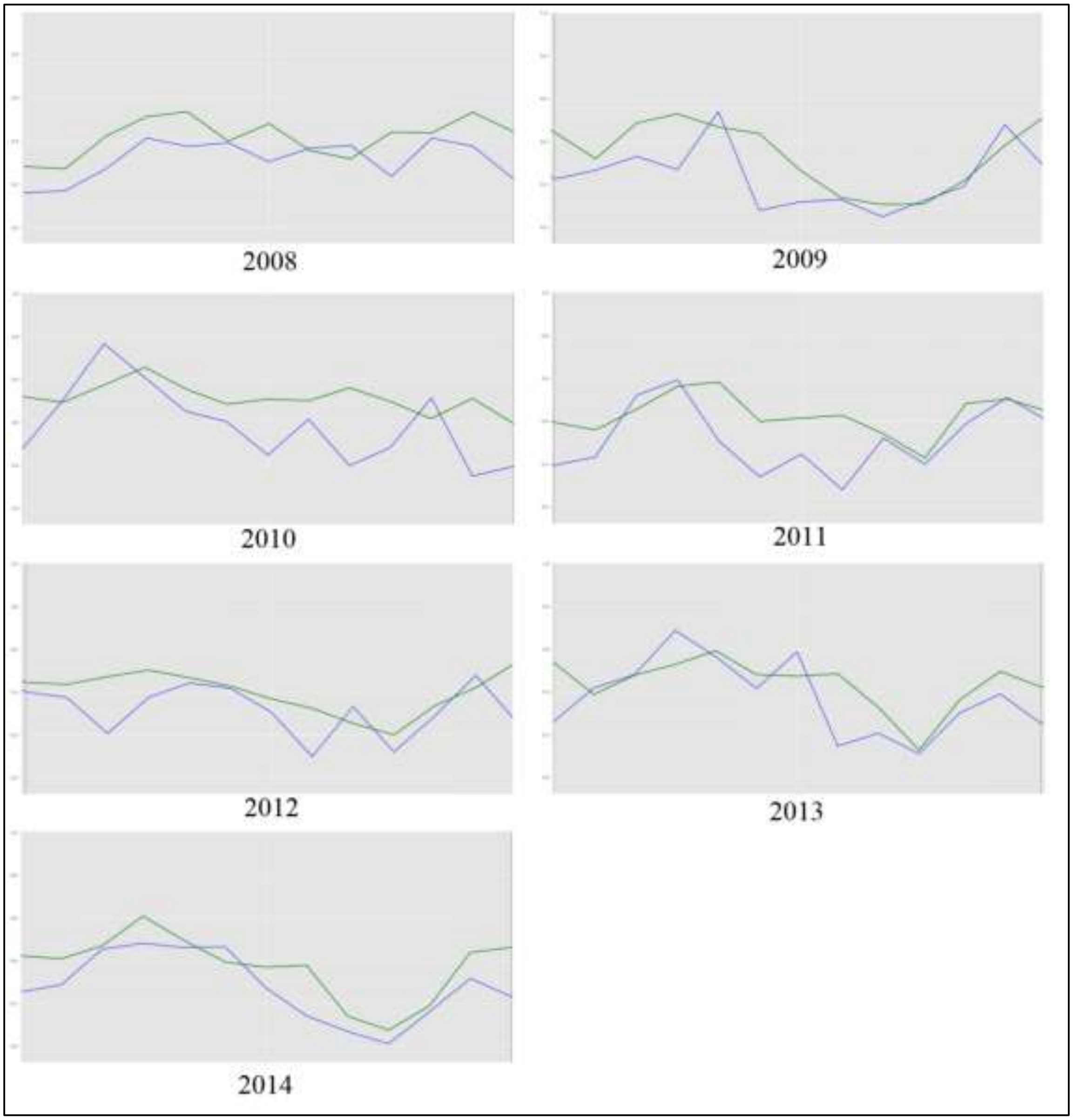


FIGURE 5. Training Plot

Caption

— : Actual Data

— : Training

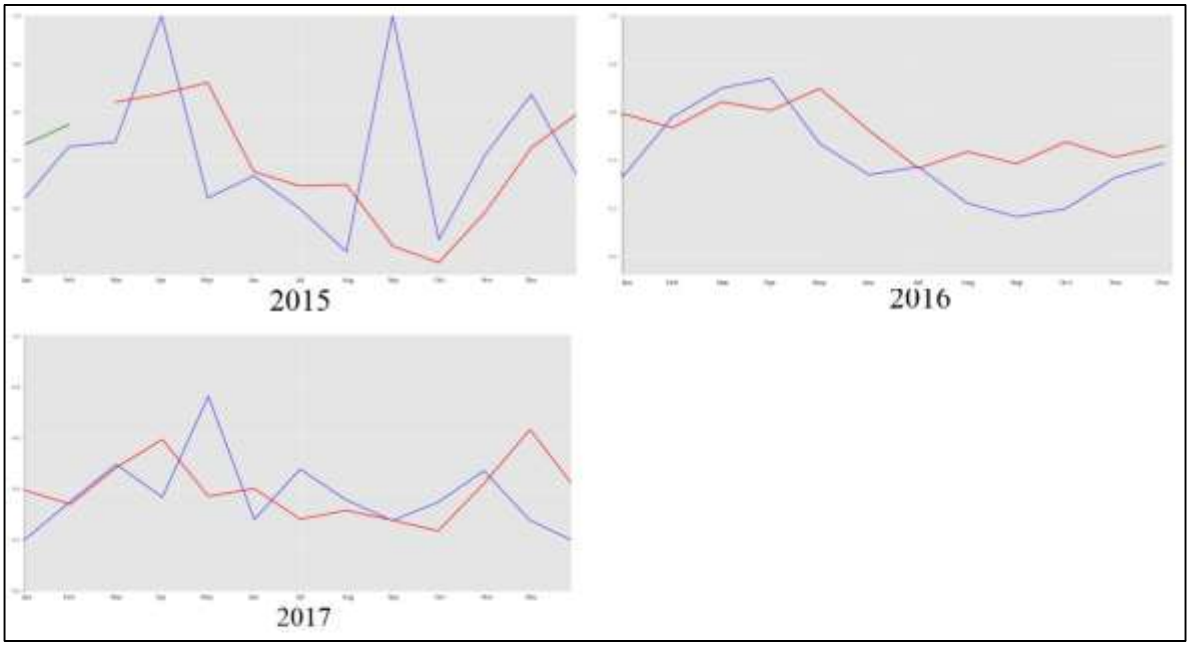


FIGURE 6. Prediction Plot

Caption
: Actual Data
: Result of Training (Rainfall Prediction)

Validation

In this section, there was a comparison between actual and predicted data obtained from 2018 and 2019 ANN model as shown in **TABLE-6**.

TABLE 6. Result Data from the ANN Model

No	Epoch	Mean Square Error		average
		Training	Test	
1	250	0,05909	0,08915	0,074120
2	**500**	**0,05865**	**0,08725**	**0,072950**
3	750	0,05930	0,08849	0,073895
4	1000	0,05937	0,08755	0,073460
5	2000	0,06115	0,09524	0,078195
6	3000	0,06250	0,09555	0,079025
7	4000	0,06644	0,12937	0,097905

FIGURE-7 shows a graphical representation of the actual and predicted data for 2018. By comparison, they both show a form of similarity in pattern which indicates the accuracy of the prediction model used.

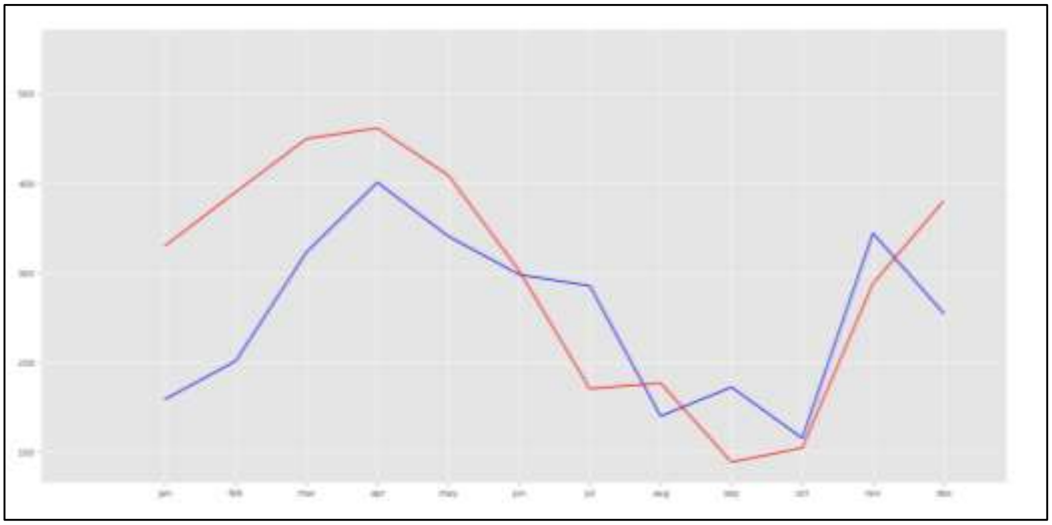


FIGURE 7. 2018 Prediction Chart

Caption

: Actual Data

: Result of Training (Rainfall Prediction)

The ANN model validation in 2018 resulted in a coefficient of determination, having a value of 0.5786.

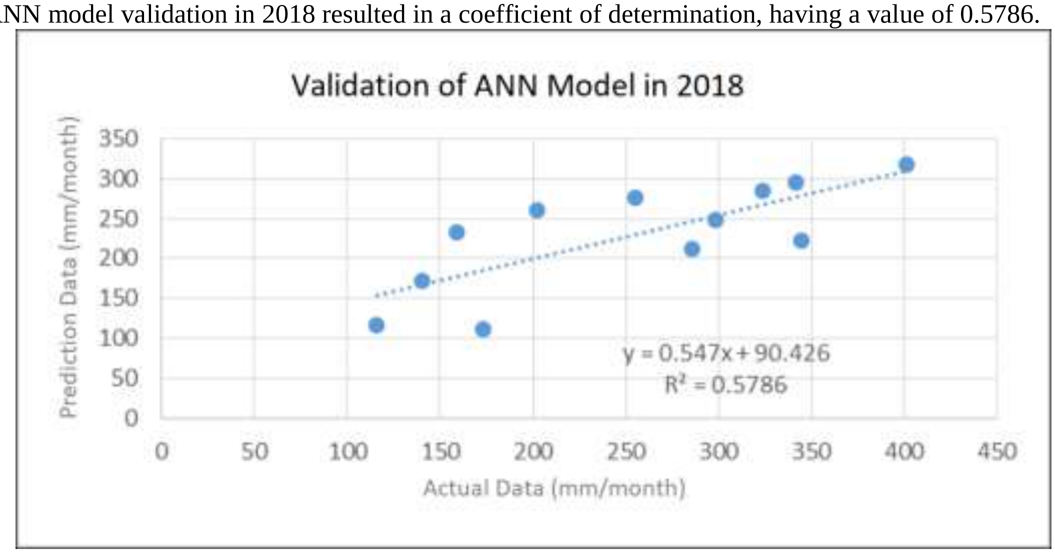


FIGURE 8. 2018 ANN Model Validation

FIGURE-9 shows the 2019 prediction chart, with the actual data ending in September. It also shows the point where the prediction data could not meet up to the high value of rainfall.

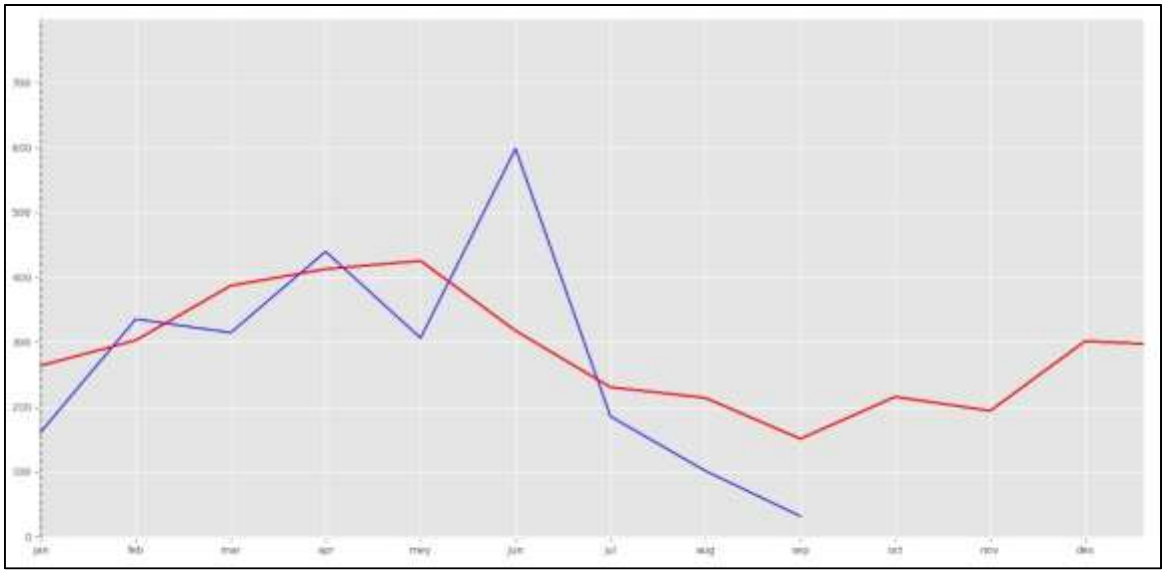


FIGURE 9. 2019 Prediction Chart

Caption

: Actual Data

: Result of Training (Rainfall Prediction)

In general, the data for 2019 had a smaller coefficient of determination than that of 2018.

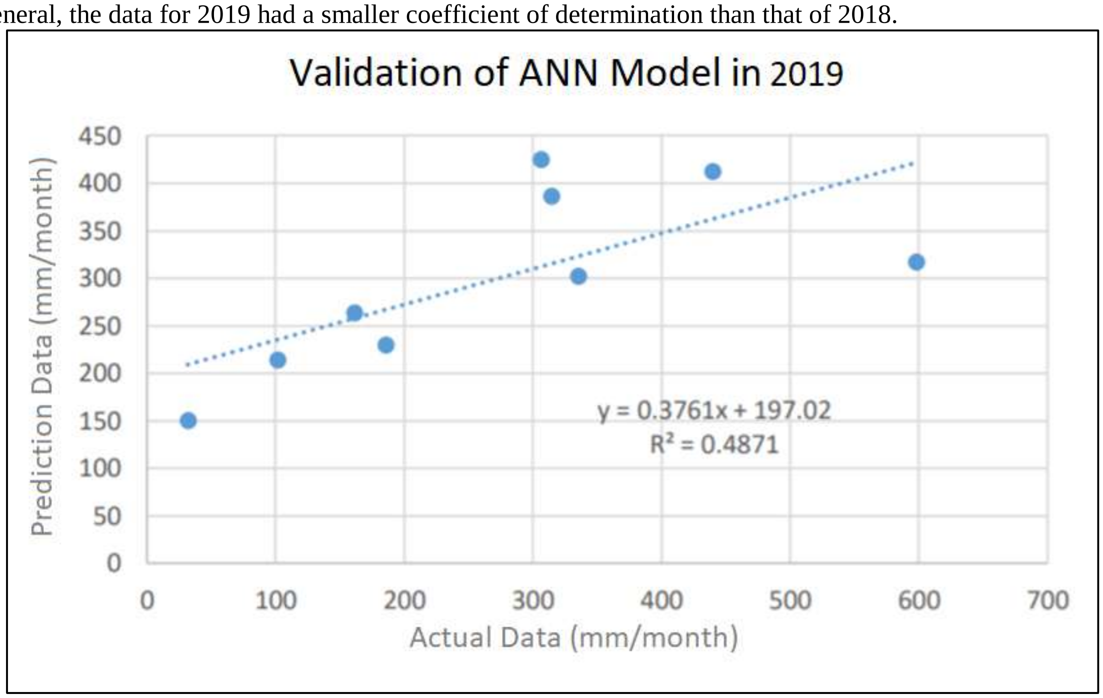


FIGURE 10. 2019 ANN Model Validation

CONCLUSIONS

In this research, an ANN algorithm was used to model and predict rainfall in PT. Vale Indonesia. After testing the architectures with different epochs, the best MSE value obtained was 0.05865 for training and 0.08725 for testing, with 7-8-1 architecture and train the model as many as 500 epochs. The results of this study shows that ANN models can be used as an algorithm that provides a good predictive accuracy.

ACKNOWLEDGMENTS

The authors express gratitude to the Department of PT. Vale Indonesia and Lembaga Penelitian dan Pengabdian Kepada Masyarakat (LPPM) UPN "Veteran" Yogyakarta for supporting and providing data needed to finish this study.

REFERENCES

[1] A. Rachmawati, "Rainfall Prediction in Pontianak City Using Weather Parameters as Predictors on Monthly, Ten Daily, and Daily Scales", 2015. [Online]. Available: http://jurnal.untan.ac.id/index.php/jpositron/article/view/11755. [Accessed: 07- Aug- 2020].

[2] I. Wahyuni, W. Mahmudy and A. Iriany, "Rainfall prediction in Tengger region Indonesia using Tsukamoto fuzzy inference system", 2016 1st International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2016. Available: https://www.researchgate.net/publication/308693809_Rainfall_Prediction_in_Tengger_Region_Indonesia_using_Tsukamoto_Fuzzy_Inference_System. [Accessed 7 August 2020].

[3] V. Nourani, "An Emotional ANN (EANN) approach to modeling rainfall-runoff process", Journal of Hydrology, vol. 544, pp. 267-277, 2017. Available: https://www.sciencedirect.com/science/article/abs/pii/S0022169416307399. [Accessed 7 August 2020].

[4] Q. Tan et al., "An adaptive middle and long-term runoff forecast model using EEMD-ANN hybrid approach", Journal of Hydrology, vol. 567, pp. 767-780, 2018. Available: https://www.sciencedirect.com/science/article/abs/pii/S0022169418300155. [Accessed 7 August 2020].

[5] I. Hadihardaja and S. Suktino, "Rainfall-Runoff Modeling Using Artificial Neural Network (ANN) with Backpropagation Method", Jurnal Teknik Sipil, vol. 12, pp. 249-258, 2005. Available: http://journals.itb.ac.id/index.php/jts/article/view/2681. [Accessed 7 August 2020].

[6] Y. Xiang, L. Gou, L. He, S. Xia and W. Wang, "A SVR–ANN combined model based on ensemble EMD for rainfall prediction", Applied Soft Computing, vol. 73, pp. 874-883, 2018. Available: https://www.sciencedirect.com/science/article/abs/pii/S1568494618305350. [Accessed 7 August 2020].

[7] S. Dharma, A. Putera and P. D. H. Ardana, "Artificial Nervous Network for Modeling Rainfall-Runoff in Watershed Areas in Bali Island", Bumi Lestari Journal of Environment, vol. 11, no. 1, pp. 9-22, 2011. Available: https://ojs.unud.ac.id/index.php/blje/article/view/80. [Accessed 7 August 2020].

[8] D. Susilokarti, S. Arif, S. Susanto and L. Sutiarso, "PREDICTION STUDY OF RAINFALL PREDICTION OF FAST FOURIER TRANSFORMATION (FFT), AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA) AND ARTIFICIAL NEURAL NETWORK (ANN)", Jurnal Agritech, vol. 35, no. 02, p. 241, 2015. Available: https://www.researchgate.net/publication/311001856_STUDI_KOMPARASI_PREDIKSI_CURAH_HUJAN_METODE_FAST_FOURIER_TRANSFORMATION_FFT_AUTOREGRESSIVE_INTEGRATED_MOVING_AVERAGE_ARIMA_DAN_ARTIFICIAL_NEURAL_NETWORK_ANN. [Accessed 7 August 2020].

[9] I. Famesa and F. Nhita, "RAINFALL PREDICTION USING HYBRID NEURAL NETWORK AND EVOLUTIONARY PROGRAMMING ALGORITHM", e-Proceeding of Engineering, vol. 2, no. 2, pp. 6872-6879, 2015. Available: https://openlibrary.telkomuniversity.ac.id/pustaka/files/102504/jurnal_eproc/prediksi-curah-hujan-menggunakan-algoritma-hybrid-neural-network-dan-evolutionary-programming.pdf. [Accessed 7 August 2020].

[10] R. Bekti, P. Susanti, and Suktino, "WEATHER PREDICTION WITH AUTOREGRESSIVE INTEGRATED MOVING AVERAGENEURAL NETWORK METHOD, AND ADAPTIVE SPLINES THRESHOLD AUTOREGRESSION IN JUANDA STATION SURABAYA", 2010. [Online]. DOI: http://dx.doi.org/10.30536/j.jsd.2010.v8.a1532. [Accessed: 07- Aug- 2020].

[11] V. Winarti, M. Jumarang and Apriansyah, "Forecasting Rain Events in Pontianak City with the JST-Fuzzy Logic Method", PRISMA FISIKA, vol., no. 2, pp. 117 - 123, 2018. Available: http://jurnal.untan.ac.id/index.php/jpfu/article/download/26522/75676577245. [Accessed 7 August 2020].

[12] E. J. Gumbel, “The Return Period of Flood Flows.” The Annals of Mathematical Statistics, vol. 12, no. 2, pp. 163–190, 1941, JSTOR, Available: www.jstor.org/stable/2235766. [Accessed 7 Aug. 2020.]

[13] Jayadianti, H., Cahyadi, T., Amri, N. and Pitayandanu, M., 2020. ARTIFICIAL NEURAL NETWORK COMPARATION METHOD ON RAIN PREDICTION - LITERATURE REVIEW. Jurnal Tekno Insentif, 14(2), pp.48-53. [Accessed 10 September 2020.]

[14] H. Kukreja, N. Bharath., C. S. Siddesh, and S. Kuldeep, "AN INTRODUCTION TO ARTIFICIAL NEURAL NETWORK", International Journal of Advance Research and Innovative Ideas in Education, vol. 1, no. 5, pp. 27-30, 2016. Available: http://ijariie.com/AdminUploadPdf/AN_INTRODUCTION_TO_ARTIFICIAL_NEURAL_NETWORK_1399.pdf. [Accessed 7 August 2020].

[15] Mislan, Haviluddin, S. Hardwinarto, Sumaryono and M. Aipassa, "Rainfall Monthly Prediction Based on Artificial Neural Network: A Case Study in Tenggarong Station, East Kalimantan - Indonesia", Procedia Computer Science, vol. 59, pp. 142-151, 2015. DOI: 10.1016/j.procs.2015.07.528 [Accessed 8 August 2020].

[16] J. Lin, "Artificial neural network related to biological neuron network: a review", Advanced Studies in Medical Sciences, vol. 5, pp. 55-62, 2017. Available: http://www.m-hikari.com/asms/asms2017/asms1-2017/p/linASMS1-2017.pdf. [Accessed 7 August 2020].

[17] H. Badrzadeh, R. Sarukkalige and A. Jayawardena, "Hourly runoff forecasting for flood risk management: Application of various computational intelligence models", Journal of Hydrology, vol. 529, pp. 1633-1643, 2015. Available: 10.1016/j.jhydrol.2015.07.057 [Accessed 7 August 2020].

[18] A. Edwin and O. Martins, "Stochastic Characteristics and Modelling of Monthly Rainfall Time Series of Ilorin, Nigeria", 2014. DOI: http://dx.doi.org/10.4236/ojmh.2014.43006

[19] R. S. Gautama, “Storm Rainfall Analysis: An Important Factor in Designing Mine Dewatering Facilities in Tropical Region.”, 6th International Mine Water Association Congress, vol. 2, pp. 235-242, 1997. Available: https://www.imwa.info/docs/imwa_1997/IMWA1997_Gautama_235.pdf

[20] S. Selvalingam and M. Miura, "STOCHASTIC MODELING OF MONTHLY AND DAILY RAINFALL SEQUENCES", Journal of the American Water Resources Association, vol. 14, no. 5, pp. 1105-1120, 1978. Available: https://www.researchgate.net/publication/230476864_STOCHASTIC_MODELING_OF_MONTHLY_AND_DAILY_RAINFALL_SEQUENCES1. [Accessed 7 August 2020].

[21] N. Ünal, H. Aksoy and T. Akar, "Annual and monthly rainfall data generation schemes", Stochastic Environmental Research and Risk Assessment, vol. 18, no. 4, pp. 245–257, 2004. Available: https://link.springer.com/article/10.1007/s00477-004-0186-4. [Accessed 7 August 2020].