Rainfall prediction with backpropagation method

E G Wahyuni¹*, L M F Fauzan¹, F Abriyani¹, N F Muchlis¹ and M Ulfa¹

¹Department of Informatics, Faculty of Industrial Technology, Universitas Islam Indonesia

* Corresponding author: elyza@uii.ac.id

Abstract. Rainfall is an important factor in many fields, such as aviation and agriculture. Although it has been assisted by technology but the accuracy can not reach 100% and there is still the possibility of error. Though current rainfall prediction information is needed in various fields, such as agriculture and aviation fields. In the field of agriculture, to obtain abundant and quality yields, farmers are very dependent on weather conditions, especially rainfall. Rainfall is one of the factors that affect the safety of aircraft. To overcome the problems above, then it's required a system that can accurately predict rainfall. In predicting rainfall, artificial neural network modeling is applied in this research. The method used in modeling this artificial neural network is backpropagation method. Backpropagation methods can result in better performance in repetitive exercises. This means that the weight of the ANN interconnection can approach the weight it should be. Another advantage of this method is the ability in the learning process adaptively and multilayer owned on this method there is a process of weight changes so as to minimize error (fault tolerance). Therefore, this method can guarantee good system resilience and consistently work well. The network is designed using 4 input variables, namely air temperature, air humidity, wind speed, and sunshine duration and 3 output variables ie low rainfall, medium rainfall, and high rainfall. Based on the research that has been done, the network can be used properly, as evidenced by the results of the prediction of the system precipitation is the same as the results of manual calculations.

1. Introduction

Rainfall prediction is the use of science and technology to predict the state of the earth's atmosphere in the future for a particular place. Currently, rainfall prediction is done using computer-assisted modeling. Although it has been assisted by technology but the accuracy can not reach 100% and there is still the possibility of error. Though current rainfall prediction information is needed in various fields, such as agriculture and aviation fields.

In the field of agriculture, to obtain abundant and quality yields, farmers are very dependent on weather conditions, especially rainfall. Indonesia as archipelago country which located in equator area is prone to climate change [1]. Several climate aspect which can be changed are rainfall pattern, sea level, and temperature. If rainfall conditions can be predicted beforehand, then farmers can decide the time of planting and right harvest time. In addition, farmers can also anticipate the occurrence of things that are not desirable. Unpredictable rainfall changes can disrupt cultivation periods leading to crop failure.

In aviation, passenger safety is a top priority. Rainfall is one of the factors that affect the safety of aircraft. Accuracy of rainfall prediction is needed to anticipate every possibility that will happen.
Accurate rainfall information can provide guidance on the selection of plane routes in order to fly safely and get to the destination at the right time. In addition, in case of bad weather it can be overcome even anticipated because of rainfall information that has been known.

Several studies have been conducted, such as: Implementation of Fuzzy Inference System on Rainfall Prediction in North Surabaya, Improvement of Abdulrahman Saleh Airport’s Weather Forecast Method with Neural Network Backpropagation Algorithm, and Implementation of Artificial Neural Network Backpropagation to Predict Weather (Case Study: Bengkulu City).

Navianti et al. [2] in the journal Application of Fuzzy Inference System on Rainfall Prediction in North Surabaya conducted a study aimed at predicting rainfall in North Surabaya by applying basic logic and fuzzy logic rules by applying Fuzzy Inference System method. This study uses six input variables that affect the occurrence of rain that is in the form of air temperature, relative humidity, wind speed, air pressure, total cloud layer, and the duration of solar irradiation. While the output variable in the form of rain prediction. The results of this study obtained the accuracy of rainfall prediction of 77.68% of eleven experiments and it can be concluded that the results of rainfall prediction have characteristics similar to the actual data of rainfall in North Surabaya. The advantages of this research is the use of Mamdani method that is more intuitive, covers a wide field, and according to the process of human information input so that the results obtained quite accurate, while the lack of use of the data used must be precise.

Yuniar et al. [3] Conducting research that aims to improve the method of weather forecasting at Abdulrahman Saleh Airport with neural network backpropagation algorithm. The variable input used in this study consists of three inputs namely air temperature, air humidity, air pressure and two output variables ie wind speed and rainfall. The results show the forecast has been close to the target value. The model has successfully performed the forecast process well with the MSE value of wind speed forecast of 0.0086 and the MSE forecasted precipitation value of 0.004846. The advantage of this research is to predict the weather for every day. While the drawback is to use only one year data earlier as input variables. The 2008 forecast uses 2007 input data, 2009 forecasts using input data for 2008, and so on.

Research conducted by Haryanto et al [4] aims to predict daily weather in Bengkulu City by implementing artificial neural network backpropagation method using temperature, humidity and rainfall data from 2008 to 2013 as input variables. While the output variables used in the study are daily weather categories such as very light rain, light rain, moderate rain, heavy rain, and very heavy rain .. The tests conducted in this study are: a). Maximum iteration of 100, 500, 1000, 1500, 10000, and 15000. b). Hidden layers: 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 c). Maximum error of 0.01 and 0.001. In the three tests that have been done the best results obtained by using the maximum iteration of 15000, 7 hidden layer, and maximum error 0.001. The advantages of this study is to compare the daily weather on the observation of Meteorology Climatology and Geophysics (BMKG) Bengkulu with daily prediction of weather with artificial neural networks. While the shortcomings of the research is only using three variables as input, and the used of data is still only as much as 5 years.

Research conducted by Nhita and Adiwijaya [5] A Rainfall Forecasting using Fuzzy System Based on Genetic Algorithm, Based on the result of experiment, it is concluded that the combination of GA and Fuzzy System prediction model have accuracy in several different population sizes and crossover probability more than 90%. This obtained accuracy is highly depending on data which been used as input and classification which been established due to the vast missing value in the data.

To overcome the problems above, then it’s required a system that can accurately predict rainfall. In predicting rainfall, artificial neural network modeling is applied in this research. The method used in modeling this artificial neural network is backpropagation method. Backpropagation methods can result in better performance in repetitive exercises. This means that the weight of the ANN
interconnection can approach the weight it should be. Another advantage of this method is the ability in the learning process adaptively and multilayer owned on this method there is a process of weight changes so as to minimize error (fault tolerance). Therefore, this method can guarantee good system resilience and consistently work well.

2. Methods

2.1. Backpropagation
Back propagation using gradient descent technique suffers from the scaling problem. It works well on simple training problems. However, as the problem complexity increases, the performance of back propagation falls off rapidly because of the fact that gradient search techniques tend to get trapped at local minima, thus making for very slow training [6]. If the number of hidden neurons is increased, the number of independent variable of the error function also increases and the computing time also increases rapidly [7].

The Backpropagation Training Algorithm is A supervised learning algorithm and usually used by perceptron with many layers for Changing the weights associated with neurons-neurons in the hidden layer [8]. Back propagation [9] requires that the activation function used by the artificial neurons (or "nodes") be differentiable. According to Fausett [10] before committing Training process, there are several network parameters which must be determined first, namely:

a. The level of training (learning rate) is symbolized with Parameters, must be given and have the positive value is less than 1. the higher the value, The faster the network's ability to learn. But this is not good, because error Resulting unevenly.
b. Fault tolerance (error tolerance), the smaller error, then the network will have a weight value which is more accurate, but will prolong the time training.
c. Maximum number of training process conducted (Maximum epoch,) is usually of great value and given to prevent iterating without end.
d. Threshold value or threshold value, denoted by θ.

2.2. System Analysis
Data in this research were not taken directly through location observation, but were taken from the official Meteorology Agency Japan website with the research object is climate in Japan. Research data is data of climate conditions for 26 years, from January 1991 to December 2016.

2.3. Recapitulation Data
Data in this research is secondary data. The data obtained from the official website of Meteorology Agency Japan recap with the Microsoft Excel program. The data recapped, analyzed, and evaluated to serve as input, namely temperature, relative humidity, sunshine duration, and wind speed. Data recapped then will be analyzed, as for the stages of the process is as follows:

a. Divide rainfall into three parts, there are low rainfall, medium rainfall, and high rainfall.
b. Determine the range for every part of rainfall, 0-200 mm for low rainfall, 201-400 mm for medium rainfall, and > 400 mm for high rainfall

3. Data Normalization
After data analysis, the data will be normalization to minimize data redundancy on the official Meteorology Agency Japan site to work optimally. The result of normalization of data is presented in Figure 1.
4. Result and Discussion

Prediction of rainfall using artificial neural network with backpropogation method run in Java. The value of learning rate 0.01, architecture (4-5-3) which are 4 inputs, 5 hidden layers, and 3 outputs with maximum of iterations 1500 and error tolerance 0.01. The program will stop if the error tolerance <0.01 or iteration (epoch) has reached its maximum value of 1500. The process taken to build rainfall prediction system with backpropogation method is training stage, testing stage, and prediction stage.

4.1. Training Stage

Training stage is a learning process to the pattern of data that will be recognized. This process is performed using training data, which stops if MSE < specified error (0.01) or a specified epoch (1500). The best result obtained is the MSE value of 0.34165862738135266 with learning rate 0.01 and stop at epoch maximum (1500) (Figure 2).
4.2. Testing Stage

Testing stage is used to validate the data that has been done in the training process by entering new data that has never been trained before to find out the value of error generated. The data used in the testing stage is random data in 2017 (Figure 3).

| Temp  | Wind Speed | R. Humidity | L. Of Radiation | Rainfall | Y1         | Y2         | Y3         | Prediction Results | Actual Predicted |
|-------|------------|-------------|-----------------|----------|------------|------------|------------|-------------------|------------------|
| 6.16  | 3.20       | 47.73       | 192.90          | 69.21    | 0.9760636769 | 0.03056924 | 0.009485317 | LOW               | LOW              |
| 6.71  | 3.42       | 49.04       | 171.62          | 61.97    | 0.95853236  | 0.04950892 | 0.007645096 | LOW               | LOW              |
| 5.77  | 3.52       | 54.15       | 171.37          | 112.58   | 0.950412971 | 0.09430797 | 0.006411695 | LOW               | LOW              |
| 14.85 | 3.51       | 59.06       | 176.25          | 130.48   | 0.826948892 | 0.10611942 | 0.002096701 | LOW               | LOW              |
| 19.24 | 3.35       | 64.23       | 177.82          | 141.75   | 0.65450991  | 0.330143816| 0.002008183  | LOW               | LOW              |

Figure 3. Testing Result

4.3. Prediction Stage

Predictions stage using feed forward method. The data used as input is data from 1991 - 2016 to predict the year 2017 (Figure 4 and Figure 5).

5. Conclusion

Based on the research that has been done with learning rate 0.01, the architecture (4-5-3) which are 4 inputs, 5 hidden layers, and 3 outputs with a maximum of iterations 1500 and a tolerance error 0.01 obtained the best result stops at maximum iteration conditions (1500) with the value of MSE 0.34165862738135266. The network is designed using 4 input variables, namely air temperature, air humidity, wind speed, and sunshine duration and 3 output variables ie low rainfall, medium rainfall, and high rainfall. The network can be used properly, as evidenced by the results of the prediction of the system precipitation is the same as the results of manual calculations in Microsoft Excel.

References

[1] Agricultural Research and Development Agency 2011 General guidelines for adaptation to climate change in the agricultural sector (Indonesia: Ministry of Agriculture)
[2] Navianti, D R, Usadha I G N R and Widjajati F A 2012 Jurnal Sains dan Seni ITS 1 A23
[3] Yuniar R J, Rahadi D and Setyawati O 2013 Journal EECCIS 7 65
[4] Haryanto S A F, Ernawati and Puspitaningrum D 2015 J. Rekursif 3 82
[5] Nhita F and Adiwijaya 2013 International Conference of Information and Communication Technology (ICoICT) 111
[6] Xiaofeng L 2001 J. Forecast. 20 69
[7] Huawang S and Yong D 2009 Int. Sym. Inform. Process 263
[8] Kusumadewi 2004 Membangun Jaringan Syaraf Tiruan (Menggunakan MATLAB & EXCEL LINK) (Yogyakarta: Graha Ilmu)

[9] Rojas R 2013 Neural networks: a systematic introduction. (Berlin: Springer Science & Business Media)

[10] Fausett L 1994 Fundamentals of Neural Networks Architectures, Algorithms, and Applications (London: Prentice Hall, Inc.)