PHOTON
Journal of Natural Sciences and Technology

p-ISSN: 2087-393X | e-ISSN: 2579-5953
http://ejurnal.umri.ac.id/index.php/photon

Artificial Neural Network based Rainfall Prediction using Back Propagation Technique in Pekanbaru city

Nurmala[1], Yulia Fitri[1]*, Sanya Gautami[2]

[1]Department of Physics, Faculty of Mathematics, Natural Sciences and Health, Universitas Muhammadiyah Riau, Indonesia
[2]BMKG Sultan Syarif Kasim II, Riau, Indonesia

Received 24 January 2024 | Accepted 20 March 2024 | Published 31 May 2024
DOI: https://doi.org/10.37859/jp.v14i2.6679

Keywords:

Rainfall; Prediction; Artificial Neural Network (ANN)

Abstract. The intensity of rainfall is a significant weather component that has a profound effect on natural disasters, particularly floods and landslides, particularly in Indonesia. Precise meteorological forecasts and comprehensive climatic data, including accurate rainfall projections, will effectively reduce the hazards associated with severe weather events. Prior studies have demonstrated the efficacy of the Backpropagation Neural Network (BPNN) approach in accurately forecasting rainfall. The objective of this work is to forecast the daily precipitation in Pekanbaru City using Neural Networks with the Backpropagation technique. The neural network model was constructed using supervised multilayer learning, initially with one hidden layer and subsequently expanded to two hidden layers, utilizing daily data spanning three years (2017-2019). The rainfall forecasting model was constructed by many iterative training and testing procedures. Forecasts of rainfall were categorized into six groups: cloudy, light rain, moderate rain, heavy rain, very heavy rain, and extreme rain. The forecast outcomes were shown using a MATLAB graphical user interface (GUI). While the prediction accuracy of 61% falls short of the national verification threshold of 75%, this work establishes a fundamental framework for the application of neural networks in weather forecasting. The outcomes can be enhanced by using more relevant data and using more precise training procedures to achieve more precise predictions.

*Corresponding author.
E-mail address: yuliafitri@umri.ac.id

1. Introduction

One of the key weather factors influencing climate type is rainfall. Indonesia, located along the equator, experiences relatively high annual rainfall ranging from 2,000 to 3,000 mm (Indrabayu *et al.*, 2011). Rainfall intensity is a major natural factor contributing to disasters such as floods and landslides. Excessive rainfall, when not supported by adequate environmental capacity, causes water to exceed the soil's absorption ability. As a result, surface runoff erodes the topsoil and carries materials to lower areas (Kompasiana, 2017).

The Meteorology, Climatology, and Geophysics Agency (BMKG), as the national authority in meteorology, climatology, air quality, and geophysics, has an essential role in providing information to relevant agencies and the public regarding climate changes, monitoring local and regional weather conditions, and issuing early warnings and rainfall forecasts for specific regions (BMKG, 2021).

Research on rainfall prediction is crucial to assist BMKG forecasters in predicting weather conditions and reducing the impact of extreme rainfall events such as floods and landslides. Moreover, rainfall significantly affects several vital sectors, including agriculture, transportation, and disaster management. Over the past five years, studies on rainfall prediction have demonstrated that the Backpropagation Neural Network (BPNN) method performs better than other techniques. Handayani & Ardi (2015) applied a Backpropagation Artificial Neural Network (ANN) using air temperature, humidity, and air pressure data, achieving 96% accuracy with an error rate of 0.001. Ritha (2016) compared two algorithms—Levenberg-Marquardt and Backpropagation—for rainfall prediction using rainfall, temperature, humidity, wind speed, and air pressure data. The results showed that the Backpropagation algorithm yielded a smaller error value (0.07876).

Sari (2019) conducted rainfall prediction research for Pekanbaru City using a multilayer Backpropagation ANN with three hidden layers and varying neuron numbers (10, 100, and 200). The study utilized ten years of data (2009–2018), with three input variables (monthly temperature, humidity, and wind speed) and one output variable (monthly rainfall). The model achieved a prediction accuracy of 71%, though the training process required a considerable amount of time (11 hours 33 minutes 5 seconds). The prolonged processing time was likely due to initial parameter selection, such as the number of hidden layers and epochs.

According to Fausett (1994) in Fundamentals of Neural Networks, a single hidden layer is generally sufficient to produce accurate pattern recognition. However, additional hidden layers may be used if the desired accuracy has not been achieved. Saiful (2012) later confirmed that increasing the number of neurons in hidden layers prolongs the training time of an ANN model.

Based on the findings of previous studies, this research aims to develop a rainfall prediction model for Pekanbaru City using the Backpropagation Neural Network (BPNN) method to obtain the most effective model configuration. The study focuses on optimizing the number of hidden layers and neurons to improve prediction performance while reducing training time.

The input data in this research consist of four daily meteorological variables: temperature, humidity, wind direction, and wind speed, while the output variable is rainfall. The selection of these variables was based on Sinurat's (2016) study on the influence of wind direction on rainfall in Kototabang, which utilized Equatorial Atmosphere Radar (EAR) and Optical Rain Gauge (ORG) data, concluding that wind direction is one of the main factors influencing rainfall.

As BMKG Pekanbaru Class I is responsible for daily rainfall forecasting, this research supports its operational activities by developing a daily rainfall prediction model using three years of data (2017–2019). The prediction results are displayed through a MATLAB-based Graphical User Interface (GUI) to facilitate easy access and interpretation for weather forecasters.

2. The Methods

This study employs a quantitative research method, which aims to predict daily rainfall based on secondary data of temperature, humidity, wind direction, wind speed, and rainfall in Pekanbaru City using an ANN with the Backpropagation algorithm.

The data used in this study were secondary data obtained from BMKG Sultan Syarif Kasim II Pekanbaru Station, consisting of daily records of temperature, humidity, wind direction, wind speed, and rainfall from January to December during the period 2017–2019. Each parameter consisted of 1,095 data points, resulting in a total of 4,380 input data (from four variables) and 1,095 output data (rainfall).

The BPNN architecture was designed to recognize and predict rainfall levels for upcoming years. The ANN model used in this research adopts a multilayer network structure, comprising four input variables—temperature, humidity, wind direction, and wind speed—and one output variable, which is rainfall.

3. Result and Discussion

3.1. One-hidden layer test

The training process of the rainfall prediction ANN model using one hidden layer was conducted with several neuron size variations of 25, 50, 75, 100, 125, 150, and 200 neurons. Each neuron variation in the single hidden layer architecture was run three times to obtain the optimal training weights and identify the best-performing rainfall prediction model.

An example of the ANN architecture for rainfall prediction using one hidden layer with 100 neurons is presented in Table 1. The training process in this study was performed using 10,000 epochs with the Trainscg (Scaled Conjugate Gradient) training function. The training results using a single hidden layer achieved a prediction accuracy of 45.05% with 50 neurons. Although this model produced a relatively low accuracy and a higher error rate, it required shorter computational time, averaging 5 seconds per run.

Previous research by Andriani *et al.* (2015) also reported that using a single hidden layer can significantly reduce computation time during model training. In his experiment, a single hidden layer produced an average running time of 10 seconds and a prediction accuracy of 42%, confirming that this configuration offers faster processing despite lower accuracy. Since the single hidden layer model in this study did not yield satisfactory prediction results, the number of hidden layers was increased to two in the subsequent training phase to improve model performance.

Table 1. Training results using one hidden layer.

Neurons	Accuracy (%)	Error (%)
25	29.24	73.61
50	45.05	54.95
75	36.41	63.59
100	30.6	69.4
125	30.6	69.4
150	30.94	69.06
175	36.97	60.03
200	26.39	73.61

3.2. Two-hidden-layer test

The subsequent training process was carried out using the same approach as the previous stage, employing a multilayer neural network with the BPNN method. In this stage, training was conducted using a pattern recognition network consisting of two hidden layers with varying numbers of neurons. The training was performed for 100,000 epochs using the Trainscg algorithm.

Each combination of neuron numbers in the two hidden layers was trained three times to evaluate the consistency and accuracy of rainfall prediction results. The best training performance was obtained with 150 neurons in the first hidden layer and 150 neurons in the second hidden layer. The optimal training accuracy was achieved at iteration 17,947, out of the maximum 100,000 epochs.

The training process required 18 minutes and 38 seconds to produce the final rainfall prediction model. This finding indicates that increasing the number of hidden layers significantly affects the computational time, as a larger network structure requires more time for training.

The performance of the ANN rainfall prediction model was further evaluated through regression plots generated during the training process. The regression plot illustrates the relationship between the network output (predicted rainfall) and the target values (observed rainfall from BMKG). As shown in Figure 1, the output data represented by the “o” symbol are closely aligned with the dotted line representing the target data, indicating a strong correlation. The regression coefficient (R) obtained from the training process was 0.95086, equivalent to a 95.086% correlation, demonstrating a strong linear relationship between predicted and actual rainfall values.

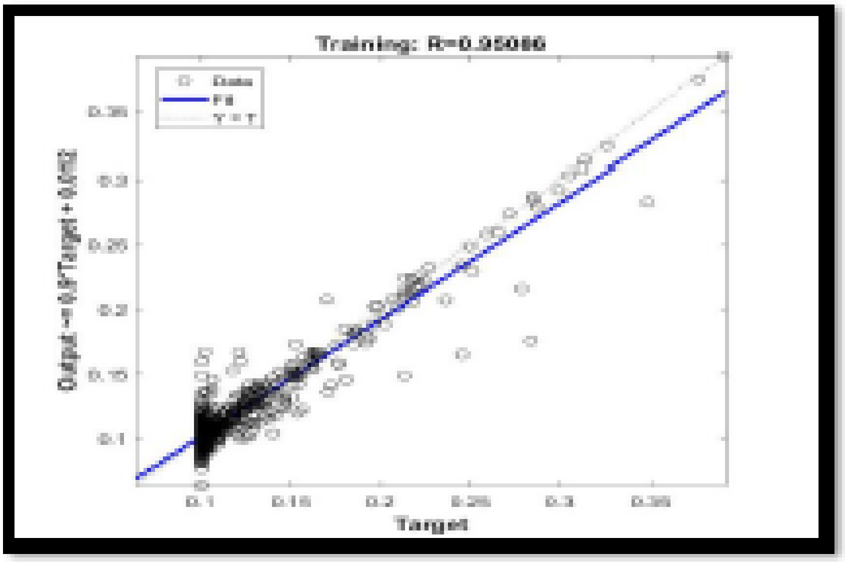

Figure 1. Training regression model with two-hidden layers.

The training results for the two-hidden-layer model shows that the highest prediction accuracy, 64%, was achieved with 150 neurons in both hidden layers. Based on these findings, the two-hidden-layer model performed better than the single-hidden-layer model, and thus, this configuration was selected for the testing phase of the study.

The testing phase was conducted to evaluate whether the trained model could accurately recognize the relationship between input and target data. A total of 20% of the dataset was allocated for testing, corresponding to 216 data points per variable. With four input variables, the total number of input data for testing was 864, while the output target consisted of 216 rainfall data points.

The input and target data were organized in Microsoft Excel and then imported into MATLAB for the testing process. The best-trained model (with two hidden layers of 150 neurons each) was used for testing to ensure consistency with the training configuration.

The output classification from the testing phase was evaluated based on predefined rainfall threshold categories, as follows:

Output = 0 → Cloudy
Output ≥ 0.5 → Light rain
Output ≥ 20 → Moderate rain
Output ≥ 50 → Heavy rain
Output ≥ 100 → Very heavy rain
Output ≥ 150 → Extreme rain

The model tested each data point (daily record) according to these classification thresholds.

The testing results showed a prediction accuracy of 60%. The predicted rainfall values were automatically categorized into their corresponding rainfall intensity classes. In comparison, Sari (2019) conducted monthly rainfall prediction with only two categories—rain or no rain—and achieved a 70% accuracy.

Despite the slightly lower numerical accuracy, the present study's model is more complex and detailed, as it predicts daily rainfall across six classification levels (cloudy, light, moderate, heavy, very heavy, and extreme). Therefore, the prediction accuracy of 60% can be considered fairly satisfactory, particularly when evaluated against the BMKG's national standard for rainfall forecast accuracy, which recognizes results above 60% as acceptable for operational use.

4. Conclusion

Based on the results of this study, the following conclusions can be drawn: The rainfall prediction training using one hidden layer achieved a prediction accuracy of 45.05% with 50 neurons. The best-performing BPNN structure was obtained using two hidden layers, each consisting of 150 neurons, with the training converging at 17,947 epochs out of a maximum of 100,000 iterations. The rainfall prediction model achieved a training accuracy of 63% and a testing accuracy of 61%, successfully identifying 132 rainy days out of a total of 216 testing data points. Additionally, the number of hidden layers significantly affects computation time—the greater the number of hidden layers, the longer the processing time required for model training.

References

Andriani, M., Indrabayu., dan Aren., S. A. 2015. Prediksi Pemakaian Obat di Instalasi Farmasi Rumah Sakit Pendidikan Dengan Menggunakan Metode Jaringan Syaraf Tiruan. Dielektrika. Vol. 2. No. 1.

Fauset, Laurene. 1994. Fundamentals of Neural Networks : Architectures, algorithms, and applications. United States. Prentice-Hall, Inc.

Handayani, L dan Adri, M. 2015 Penerapan JST (Backpropagation) untuk Prediksi curah hujan (Studi kasus: Kota Pekanbaru). Seminar Nasional Teknologi Informasi Komunikasi dan Industri.

Indrabayu, H. N., Pallu, M. S., dan Achmad, A. 2011. Prediksi Curah Hujan Wilayah Makassar Menggunakan Metode Wavelet-Neural Network. Jurnal Ilmiah Elektikal Enjiniring UNHAS. Vol. 9. No. 2. Hal 50-59.

Ritha, N. 2016. prediksi curah hujan menggunakan Algoritma levenberg Marquardt dan Backpropagation. Tanjung Pinang. Universitas Maritim Raja Ali Haji. Jurnal SUSTAINABLE. Vol. 5. No. 2.

Sari, W. N. 2019. Model Prediksi Curah Hujan Bulanan Untuk Pengelolaan Budidaya Kelapa Sawit Menggunakan Jaringan Saraf Tiruan (JST). Skripsi. Universitas Muhammadiyah Riau.

Saiful, A. 2012. Sistem Deteksi Dini Hama Wereng Batang Coklat Menggunakan Jaringan Syaraf Tiruan Backpropagation. Semarang. Universitas Negeri Semarang. Journal of Mathematics Vol. 1. No. 2.

Sinurat, N. 2016. Analis Arah Angin Terhadap Curah Hujan Menggunakan Equatorial Atmoshphare Radar (EAR) dan Optocal Rain Gauge (ORG) di Atas Koto Tabang Sumatera Barat. Pekanbaru. Universitas Riau.

Kompasiana, 2017. htpps://www.kompasiana.com/amp/arifmhmmd/2017/Reseach. Dampak Curah Hujan Research. Diakses Pada 23 Desember 2020.