BAREKENG: Journal of Mathematics and Its Applications
March 2023 Volume 17 Issue 1 Page 0449–0456
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss1pp0449-0456

APPLICATION OF THE BACKPROPAGATION METHOD TO PREDICT RAINFALL IN NORTH SUMATRA PROVINCE

Rinjani Cyra Nabila [1*], Arnita[2], Amanda Fitria[3], Nita Suryani[4]

[1,2,3] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, Medan State University*
Jl. Williem Iskandar Pasar V Medan Estate, Medan, 20221, Indonesia

*Corresponding author's e-mail: * **rinjanicyranabila@gmail.com***

Article History:

Received: 12th November 2022
Revised: 1st February 2023
Accepted: 12th February 2023

Keywords:

Implementation;
Backpropagation;
Prediction;
Rainfall.

ABSTRACT

Natural disasters are to blame for the high level of community loss. This is due to the community's lack of information about potential disasters around them. As a result, public understanding of disaster response is extremely low. As a result, weather information is critical for the smooth operation of human activities and activities, such as determining the amount of rainfall. The goal of this research is to identify the best model for predicting rainfall in North Sumatra Province and to forecast rainfall trends for the coming year. The rainfall time series data used in this study were collected from six stations in North Sumatra Province over the last ten years, including the Sibolga Meteorological Station, Aek Godang Meteorological Station, and Silangit Meteorological Station. Backpropagation is used in this study. Backpropagation is one of the methods used in artificial neural networks, which are usually divided into three layers: an input layer, a hidden layer, and an output layer connected by weights. During the training stage, the learning rate, iteration, and the number of nodes in the hidden layer were all tested. Following the training process, the best model will be used for testing. The best model was obtained using rainfall data from North Sumatra Province, with an optimal iteration of 1000 iterations, an optimal learning rate of 0.1 in the learning rate trial, and the best number of hidden 5 nodes. During the testing, the MSE values were 0.047 and 0.022, respectively, and the MSE squared value was 0.0022 and 0.00049.

How to cite this article:

R. C. Nabila, Arnita, A. Fitria and N. Suryani., "APPLICATION OF THE BACKPROPAGATION METHOD TO PREDICT RAINFALL IN NORTH SUMATRA PROVINCE," *BAREKENG: J. Math. & App.*, vol. 17, iss. 1, pp. 0449-0456, March 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

North Sumatra Province is one of the provinces on Sumatra Island's northern tip, with coordinates of 1° - 4° N and 98° - 100° E [1]. North Sumatra's geographical location is advantageous because it is located near the equator, is traversed by the Bukit Barisan Mountains, and is surrounded by the Malacca Strait and the Indian Ocean. As a result, the climate conditions for rainfall in North Sumatra exhibit global climate-influenced characteristics, such as the Indian Ocean Dipole (IOD) phenomenon [2]. Furthermore, regional-scale climatic factors such as monsoons, tropical disturbances, and regional convergence influence regional climate conditions. Cloud formation and rain have an impact on natural conditions, as well as the apparent motion of the sun [1]. Natural disasters are to blame for the high level of community loss. This is due to the community's lack of information about potential disasters around them. As a result, public understanding of disaster response is extremely low [3]. As a result, weather information is critical for the smooth operation of human activities and activities, such as determining the amount of rainfall. A detour is unavoidable due to changing weather conditions. After several days of rain, these deviations are visible. Rainfall is an intriguing climatic factor to investigate because it does not follow the same pattern across Indonesia [4]. Changes in rainfall patterns in North Sumatra may result in a longer rainy season, which may cause flooding in several locations. Flooding, particularly in North Sumatra, harms many people because it can halt the economic activity, disrupt transportation, damage urban infrastructure, and have other negative consequences[5]. Rainfall must be calculated in order to collect data, which is then shared with the public [6]. This weather forecast needs to be more accurate in order to facilitate human activities that are highly dependent on the weather [7].

An artificial neural network (ANN) is a computer system whose architecture and operation are inspired by knowledge of biological neurons in the brain. It is one of the artificial representations of the human brain that is constantly attempting to stimulate the human brain's learning processes. Artificial neural networks can also predict what will happen in the future based on patterns of past events. This is referred to as forecasting or prognosis [8]. Forecasting is a method of predicting what will happen in the future based on past data. Forecasting algorithms are generally used to estimate or predict an event or a specific event before it occurs [9]. This can be accomplished by using systematic modeling to feed data from the past into the future [10].

The backpropagation method is the most effective in predicting. Backpropagation is a supervised learning algorithm used in neural networks. Backpropagation is the difference between the network's ability to recognize training patterns and the network's ability to correctly respond to input patterns that are similar (but not identical) to training patterns. Backpropagation is the best training algorithm among the eleven, according to [11], with BPNN accuracy reaching 98.72% and 97.93%, respectively. The benefit of this method is that the forecaster's experience and knowledge can be formulated, and the forecasting rules can be changed very quickly [12].

We can see how the best backpropagation modeling is used to predict rainfall in North Sumatra Province and see the trend of rainfall in the coming year from this study. This can assist the government in developing policies and forecasting future disasters. In this case, the title "Implementation of the Backpropagation Method to Predict Rainfall in North Sumatra Province" can be used. The purpose of this study is to forecast rainfall in the coming year using time series data from previous years' rainfall (2012-2021). Predictions are made using time series analysis, which can provide information about trends, cycles, or fluctuations around the long-term average value [13].

2. RESEARCH METHODS

Backpropagation is used in this study to forecast rainfall in North Sumatra Province. Rainfall data were obtained from an official online data website of the BMKG and the Center for Meteorology, Climatology, and Geophysics Region 1 Medan. The data includes six stations in North Sumatra Province, including the Silangit Meteorological Station, Aek Godang Station, Sibolga Station, Binaka Station, Deli Serdang Station, and the Center for Meteorology, Climatology, and Geophysics Region 1 Medan. The data used is time series rainfall data from January 2012 to December 2021.

2.1 Data Preprocessing

At this point, the obtained data is preprocessed by looking at the data that is missing a value. When data is missing, the imputation method is used to find the average value in the missing value data. Furthermore, data normalization is performed after the missing value data is satisfactory. The goal of data normalization is to match the network output to the activation function.

2.2 Modeling Using the Backpropagation Method

Figure 1 shows the architecture used in this study.

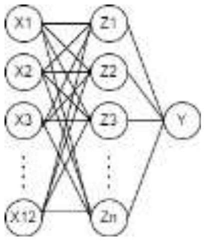


Figure 1. Backpropagation Network Architecture

Figure 1, The backpropagation network architecture is made up of three layers that are connected by weights: the input layer, the hidden layer, and the output layer. The training process determines the best modeling by determining the backpropagation parameters. Learning rate trials, the number of hidden layer nodes, and iterations will be tested during the training process. These parameters were tested with a learning rate ranging from 0.1 to 0.9 **[14]**, a total of 3.4.5 layers of hidden nodes, and 100, 500, and 1000 iterations **[15]**.

The process of predicting rainfall is divided into several stages, namely:

a) Data Training and Data Testing Distribution
At this point, the dataset is divided into 80% training data and 20% testing data **[16]**. This study's training data included rainfall data from January 2012 to December 2018, which were analyzed for each variable. Testing data was collected between January 2019 and December 2021. This was done using data from each station in North Sumatra Province.

b) Pattern formation for the backpropagation method
The pattern formation process in this case is accomplished by adjusting the weight value. Weights are initialized with fairly small random or random values. When the resulting error reaches the target error, the weight adjustment in pattern recognition is stopped. The error is computed following the forward propagation stage until the resulting error is equal to the target error.

c) Testing
Following the completion of the pattern formation process for each variable, the data processing results are subjected to a trial phase. At this stage, it is carried out using the best model obtained during the training stage. Tests are performed to determine the accuracy of the system method used to predict data, and see the results of predictions of rainfall in the coming year. At this point, the prediction results will be denormalized so that the data obtained is real.
The following formula calculates the denormalization process:

$$x = \frac{(max\ value - min\ value)(x' - 0{,}1)}{0{,}8} + min\ value \quad (1)$$

d) Evaluation
At this point, the test results are being forecasted for the following year. This stage also considers the smallest error value, or MSE.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(t_k - t_k)^2 \quad (2)$$

3. RESULTS AND DISCUSSION

Rainfall data from the Sibolga Meteorological Station, Aek Godang Meteorological Station, Silangit Meteorological Station, Meteorological Center, Climatology and Geophysics Region I Medan, Deli Serdang Climatology Station, and Binaka Meteorological Station are used to predict rainfall in this study. **Table 1** contains the normalization data for rainfall that will be used in the training and testing processes.

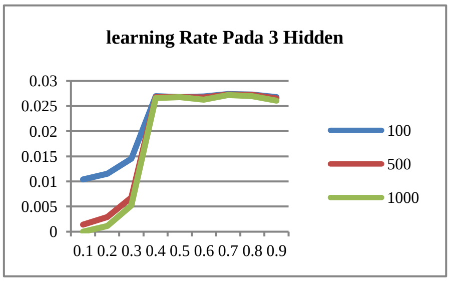
Table 1. Data on Rainfall Normalization for the Province of North Sumatra

Year	Month										
	Jan	**Feb**	**Mar**	**Apr**	**May**	**Jun**	**Jul**	**Aug**	**Sep**	**Oct**	**Nov**
2012	0.08582	0.19628	0.25822	0.33971	0.27506	0.12457	0.26645	0.33696	0.32220	0.45230	0.72902
2013	0.14474	0.29237	0.06599	0.19725	0.19149	0.17345	0.04034	0.35601	0.30694	0.53939	0.43835
2014	0.19033	0.17963	0.16529	0.20355	0.25976	0.06035	0.07488	0.30992	0.33086	0.27043	0.45465
2015	0.30246	0.10981	0.19164	0.28040	0.25374	0.03899	0.29237	0.19168	0.25609	0.45917	0.40354
2016	0.11526	0.31015	0.32151	0.22074	0.39395	0.24269	0.26714	0.22360	0.49403	0.40184	0.57319
2017	0.43348	0.20880	0.42216	0.28275	0.31668	0.10989	0.06677	0.39191	0.67219	0.35400	0.57826
2018	0.21448	0.25706	0.38236	0.32371	0.33329	0.15961	0.42220	0.24330	0.37854	0.70236	0.61693
2019	0.35176	0.18743	0.30227	0.24238	0.27873	0.37224	0.13844	0.21174	0.23426	0.70170	0.43970
2020	0.48352	0.11723	0.20826	0.39403	0.29164	0.37726	0.36378	0.17194	0.34369	0.37154	0.46775
2021	0.36988	0.00000	0.41671	0.31038	0.22603	0.22398	0.11255	0.52525	0.31664	0.27232	0.51957

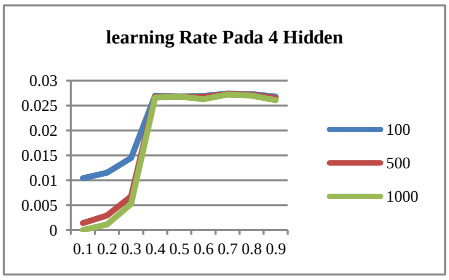
Data source: *(Data from the BMKG and the Center for Climatology, Meteorology, and Geophysics are available online for Region 1 Medan.)*

3.1 Backpropagation Modeling Determination

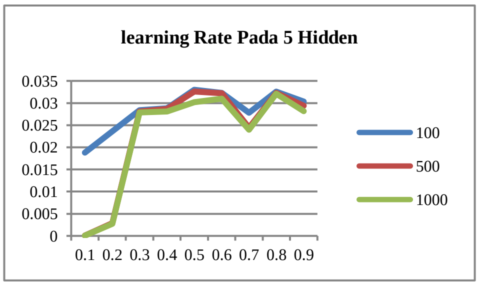
Formation of artificial neural networks through network training with artificial neural network architectures designed by testing hidden layers 3, 4, and 5 **[15]**. This training process is carried out in an experiment to see the effect of the learning rate value with a range (0.1 - 0.9) for each hidden layer **[15]**. It will conduct iteration trials to find the optimal number of iterations. Iterations of 100, 500, and 1000 are used for each hidden layer experiment in each rainfall data at each station. **Figure 2** depicts the effect of the learning rate on rainfall data from January 2012 to December 2019.



(a)



(b)

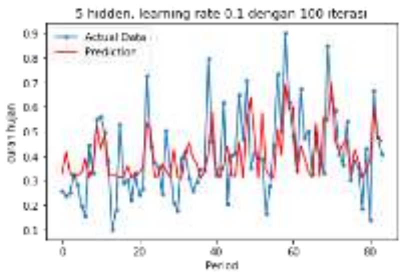


(c)

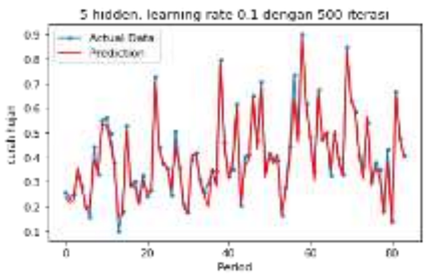
Figure 2. The Influence of The Learning Rate Value On (a) 3 Hidden Layers, (b). 4 Hidden Layers, And (c). 5 Hidden Layers

Figure 2, Depicts the relationship between the learning rate value and the number of nodes (3, 4, and 5 hidden layers). The graph shows how, with experiments on three hidden layers, a learning rate with a range of 0.2-0.9 obtained a more stable MSE, with no significant decrease or increase. The MSE value, however, tends to decrease at a learning rate of 0.1. Learning rates with a range of 0.3-0.9 obtained more stable MSE values, with no significant decrease or increase in the learning rate, but values with a range of 0.1- 0.2 MSE tend to be smaller. However, at a learning rate of 0.1 and an MSE value of 0.00006 in the hidden 5 layers, the smallest MSE value occurs.

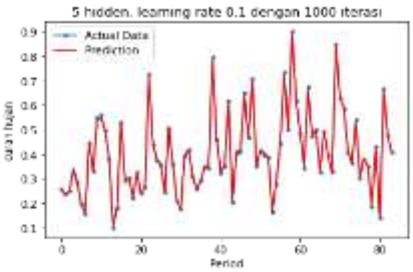
The best MSE value is obtained with a learning rate of 0.1 after testing the learning rate value on the number of hidden nodes determined. The optimal learning rate is 0.1 with 5 hidden nodes. A comparison of training results between 100, 500, and 1000 iterations can be seen after understanding the effect of the learning rate.



(a)



(b)



(c)

Figure 3. Rainfall Data Training for North Sumatra Province in 5 Hidden Iterations (a) 100, (b) 500, (c) 1000

Figure 3, Depicts the results of the training by comparing 100, 500, and 1000 iterations. The graph looks very good after the 1000th iteration. The shape of the actual line (blue) and the prediction line (red) are close together, indicating that the training results can meet the actual value. In 1000 iterations, the MSE is 0.0000038. According to the results of the experiments, the more iterations specified, the smaller the resulting error value.

3.2 Process of Testing

The best model obtained during the training process is used for testing, which is an architectural model with 12 inputs, 5 hidden layers, and 1 output layer (12-5-1). The data used in testing was never used in the training process. The test data is derived from a combination of rainfall data from six stations, with the average data calculated from January 2020 to December 2021. This testing process will yield a predicted value in 2022. Following the prediction of rainfall in 2022, the next step will be to test the prediction of rainfall in 2023 at each station using data from 2021 to 2022.

The MSE value for the test results based on rainfall data for 2020-2021 is 0.047, and the MSE squared value is 0.022. An MSE value of 0.0022 is obtained from the 2021-2022 data for predictions for 2023, with an MSE square of 0.00049.

Table 2 shows the results of the rainfall prediction in North Sumatra Province after the testing process.

Table 2. Prediction of Rainfall in North Sumatra Province

Month	Year	
	2022	2023
January	0,410327	0,454163
February	0,381673	0,392483
March	0,407148	0,555477
April	0,281296	0,36858
May	0,413352	0,267101
June	0,597875	0,271948

Month	Year	
	2022	2023
July	0,449739	0,412122
August	0,391506	0,534789
September	0,401561	0,476313
October	0,808859	0,69221
November	0,366155	0,561381
December	0,470088	0,641057

3.3 Data Denormalization

The next step is to denormalize the predictive data obtained from the testing process for each station. Denormalization attempts to restore yk's true value. The denormalization procedure is described in **Equation (1)**.

$$x = \frac{(max\ value - min\ value)(x\prime - 0{,}1)}{0{,}8} + min\ value \qquad (3)$$

Table 3 shows the effects of yk denormalization on rainfall prediction results for the province of North Sumatra.

Table 3. Results of Denormalized Rainfall Prediction in North Sumatra

Month	Year	
	2022	2023
January	230.5	249.5
February	218.2	222.8
March	229.2	293.2
April	174.9	212.5
May	231.8	168.8
June	311.4	170.9
July	247.5	231.3
August	222.4	284.2
September	226.8	259.0
October	402.5	352.1
November	211.5	295.7
December	256.3	330.1

Table 3, Shows the trend of rainfall in North Sumatra in 2022, with the highest rainfall occurring in October at 402.5 mm and the lowest rainfall occurring in April at 174.9 mm. In 2023, the highest rainfall was 352.1 mm in October, and there was no rainfall in the lowest category, but there was moderate rainfall of 168.8 mm in May. **Figure 4** shows the graph.

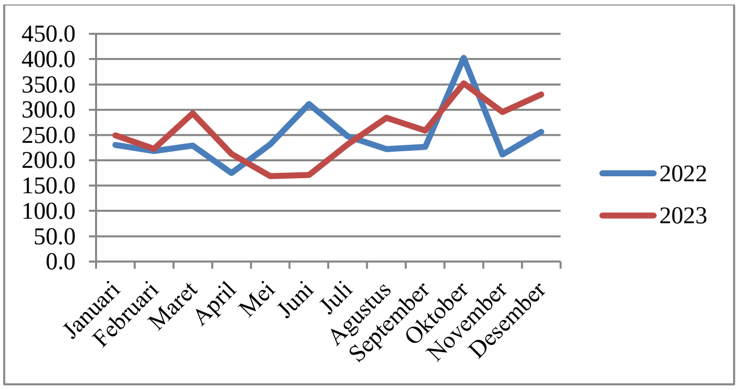


Figure 4. Rainfall Forecast for North Sumatra in 2022-2023

3.4 Evaluation

The training and testing process was completed successfully. Experiments in the architectural model training process with 3, 4, and 5 hidden layers at each station can be useful in predicting the artificial neural network. The 5 hidden layer experiments produce good training results. This is comparable to research by **[15]**, which attempted the same number of hidden. This demonstrates that having too few hidden layer neurons is bad because it can slow down the learning process, so the more hidden layers there are, the better the learning process and the lower the MSE value. Experiments were also conducted to test the effect of the learning rate value on each hidden layer. Experiments were carried out with values ranging from 0.1 to 0.9, as well as studies **[15]** and **[14]**.

The test results show that the model can predict rainfall data even if it is not exactly the same. However, the model can react when the data fluctuates. Using rainfall data from 2020 to 2021, the MSE value is 0.047, and the MSE squared value is 0.022. In the 2021-2022 data for predictions for 2023, an MSE value of 0.0022 is obtained with an MSE square of 0.00049.

4. CONCLUSIONS

The following conclusions can be drawn from the Backpropagation artificial neural network design results in predicting rainfall data in North Sumatra Province:

a) Based on the experimental results, the optimal learning rate value is 0.1, with a range of 0.1-0.9. The optimal number of hidden layers (5 hidden) yields the number of tested hidden layer nodes. This demonstrates that having too few hidden layer neurons is bad because it can slow down the learning process, which means that the more hidden layers there are, the better the learning process and the lower the MSE value. Experiment results show that the larger the epoch specified, the smaller the error value generated.
b) Rainfall prediction results were obtained at six stations in North Sumatra Province: Sibolga Meteorological Station, Aek Godang Meteorological Station, Silangit Meteorological Station, Meteorology, Climatology, and Geophysics Center Region I Medan, Deli Serdang Climatology Station, and Binaka Meteorological Station. According to the prediction results of 6 stations, the province of North Sumatra had the highest rainfall in 2022 occurring in October at 349.4 mm, while the rainfall in the low category does not exist, but the smallest rainfall in the medium category occurs in March of 162,5 mm. In 2023, the highest rainfall will be 371.6 mm in November, while the smallest rainfall for the low category will be 171.4 mm in May for the medium category.

REFERENCES

[1] B. Prasetyo, H. Irwandi, and N. Pusparini, "Karakteristik Curah Hujan Berdasarkan Ragam Topografi Di Sumatera Utara," *J. Sains Teknol. Modif. Cuaca*, vol. 19, no. 1, p. 11, 2018, doi: 10.29122/jstmc.v19i1.2787.
[2] I. Gustari, "Pantai Barat Sumatera Bagian Utara," *J. Meteorol. Dan Geofis.*, pp. 29–38, 2007.
[3] R. Rahmad, S. Suib, and A. Nurman, "Aplikasi SIG Untuk Pemetaan Tingkat Ancaman Longsor Di Kecamatan Sibolangit, Kabupaten Deli Serdang, Sumatera Utara," *Maj. Geogr. Indones.*, vol. 32, no. 1, p. 1, 2018, doi: 10.22146/mgi.31882.
[4] Z. Julisman and Erlin, "Prediksi Tingkat Curah Hujan di Kota Pekanbaru menggunakan Logika Fuzzy Mamdani," *J. SATIN - Sains dan Teknol. Inf.*, vol. 3, no. 1, pp. 65–72, 2014.
[5] Dynes Rizky Navianti, I Gusti Ngurah Ray, Farida Agustini W, "Penerapan Fuzzy Inference System Pada Prediksi Curah Hujan di Surabaya Utara," *J. Sains dan Seni ITS*, vol. I, no. 1, p. I, 2012.
[6] A. Roy and A. M. Hendriyawan, "Implementasi Metode Naive Bayes Classifier Untuk Perkitaan Cuaca," p. 12, 2019, [Online]. Available: http://eprints.uty.ac.id/2916/.
[7] R. J. Yuniar, D. R. S., and O. Setyawati, "Perbaikan Metode Prakiraan Cuaca Bandara Abdulrahman Saleh Dengan Algoritma Neural Network Backpropagation," *J. EECCIS*, vol. 7, no. 1, p. pp.65-70, 2013.
[8] P. N. L. Raja and S. Sinurat, "Implementasi Metode Backpropagation Untuk Memprediksi Penjualan Furniture Pada Pt . Indah Jaya Medan, " *Majalah Ilmiah INTI*, vol. 13, pp. 146–151, 2018.
[9] D. Ruswanti, "Pengukuran Performa Support Vector Machine Dan Neural Netwok Dalam Meramalkan Tingkat Curah Hujan," *Gaung Informatika*, vol. 13, no. mm, pp. 66–75, 2020.
[10] W. R. A. Situmorang and M. Jannah, "Implementasi Jaringan Syaraf Tiruan Memprediksi Hasil Panen Padi Pada Desa Pagar Jati Dengan Metode Backpropagation," *Jurnal Ilmu Komputer dan Sistem Informasi*, vol. 3, no. 3, pp. 167–175, 2021.
[11] P. Antwi *et al.*, "Estimation of biogas and methane yields in an UASB treating potato starch processing wastewater with backpropagation artificial neural network," *Bioresour. Technol.*, vol. 228, pp. 106–115, 2017, doi: 10.1016/j.biortech.2016.12.045.
[12] R. Novita, A. Putri, P. Studi, and S. Informasi, "Analisis Algoritma Backpropagation Neural Network dalam Permalan

Jumlah Benih Ikan," *Jurnal Ilmiah Rekayasa dan Manajemen Sistem Informasi*, vol. 7, no. 2, pp. 201–207, 2021.

[13] D. Susilokarti, S. S. Arif, S. Susanto, and L. Sutiarso, "Studi Komparasi Prediksi Curah Hujan Metode Fast Fourier Transformation (Fft), Autoregressive Integrated Moving Average (Arima) Dan Artificial Neural Network (Ann)," *J. Agritech*, vol. 35, no. 02, p. 241, 2015, doi: 10.22146/agritech.9412.

[14] D. Jauhari, Al. Hinawan, and D. Candra, " *Jurnal Teknologi Informasi dan Ilmu Komputer*, Volume 3 | Nomor | 2016 | Halaman -," vol. 3, pp. 3–6, 2016.

[15] B. B. Wiko Putra, M. A. Albar, and B. Irmawati, "Penerapan Algoritma Jaringan Saraf Tiruan Metode Backpropagation untuk Memprediksi Jumlah Nilai Ekspor di Provinsi NTB," *J. Teknol. Informasi, Komputer, dan Apl. (JTIKA)*, vol. 1, no. 2, 2019, doi: 10.29303/jtika.v1i2.50.

[16] I. Loelianto, M. S. S. Thayf, and H. Angriani, "Implementasi Teori Naive Bayes Dalam Klasifikasi Calon Mahasiswa Baru Stmik Kharisma Makassar," *SINTECH (Science Inf. Technol. J.*, vol. 3, no. 2, pp. 110–117, 2020, doi: 10.31598/sintechjournal.v3i2.651.