

ANALYSIS OF THE BACKPROPAGATION ALGORITHM IN PREDICTING WATER VOLUME OF PDAM TIRTAULI PEMATANG SIANTAR CITY

Saputra Ramadani^{1*}, Anjar Wanto¹, M.Safii¹

¹Sistem Informasi, STIKOM Tunas Bangsa

*email: *saputraramadani982@gmail.com*

Abstract: Increasing living standards cause an increase in the need for drinking water. However, current water supply estimates are still not optimal, with water production sometimes being more or less than requirements. To estimate the amount of water, an appropriate method is needed. The method used in this research is the back propagation algorithm artificial neural network method. When developing forecasts, past data is necessary to produce accurate results. This research aims to develop a predictive model that can estimate the volume of water that will be used by PDAM Tirtauli in the future. It is hoped that this predictive model can help PDAMs in planning more efficient water supply management and can reduce the potential for water supply shortages in the future. This research uses water distribution data for the 2015-2022 period. Training data starts in 2015-2021, testing data starts in 2016-2022. In this research, results were obtained using the Matlab R2011a application. In this research, the 5 architectures used are architecture 6-53-1, 6-58-1, 6-61-1, 6-81-1, 6-87-1. Based on these five architectures, the best architecture was obtained, namely architecture 6-87-1 with a root mean square error test value of 0.00010031 and an accuracy of 92%. The results achieved in 2023 are the total water volume of PDAM Tirtauli Pematangsiantar of 189,610,426.

Keywords: backpropagation; distribution; PDAM; prediction; water

Abstrak: Meningkatnya taraf hidup menyebabkan meningkatnya kebutuhan akan air minum. Namun, perkiraan pasokan air saat ini masih belum optimal, dengan produksi air kadang-kadang lebih atau kurang dari kebutuhan. Untuk memperkirakan jumlah air diperlukan suatu metode yang sesuai. Metode yang digunakan dalam penelitian ini adalah metode jaringan syaraf tiruan algoritma back propagation. Saat mengembangkan perkiraan, data masa lalu diperlukan untuk menghasilkan hasil yang akurat. Penelitian ini bertujuan untuk mengembangkan model prediktif yang dapat memperkirakan volume air yang akan digunakan oleh PDAM Tirtauli di masa mendatang. Model prediktif ini diharapkan dapat membantu PDAM dalam perencanaan pengelolaan pasokan air yang lebih efisien dan dapat mengurangi potensi kekurangan pasokan air pada masa yang akan datang. Penelitian ini menggunakan data sebaran air periode 2015-2022. Data pelatihan dimulai pada tahun 2015-2021, data pengujian dimulai pada tahun 2016-2022. Pada penelitian ini diperoleh hasil dengan menggunakan aplikasi Matlab R2011a. Pada penelitian ini 5 arsitektur yang digunakan adalah arsitektur 6-53-1, 6-58-1, 6-61-1, 6-81-1, 6-87-1. Berdasarkan kelima arsitektur tersebut diperoleh arsitektur terbaik yaitu arsitektur 6-87-1 dengan nilai uji root mean square error sebesar 0,00010031 dan mendapatkan akurasi sebesar 92%. Hasil yang dicapai pada tahun 2023 adalah total volume air PDAM Tirtauli Pematangsiantar sebesar 189.610.426.

Kata Kunci: air; backpropagation; distribusi; PDAM; prediksi

INTRODUCTION

Water is the main source of supply and is very important for daily human

needs and everyone has the right to use clean water [1]. Drinking water is a basic need that cannot be separated from human life. Its use is not only for household

purposes but also for other installations. Along with population growth, development progress and increasing living standards, the need for drinking water continues to increase [2]. Therefore, the community must be consistent in the quality of business management and drinking water services. is a company operating in the field of drinking water supply. One of the objectives of establishing PDAM is to meet the community's drinking water needs, including provision, development of infrastructure and services, as well as distribution of drinking water [3]. PDAM as one of the BUMDs can make a full contribution as a community service and can contribute to local original income (PAD). This makes the service quality of PDAM Tirtauli Pematangsiantar City and the clean water management company very important for the community. However, current water supply estimates are still not optimal, with water production sometimes being more or less than requirements. High water consumption causes the need for drinking water to continue to increase, while the supply of drinking water continues to decrease every year, along with the large amount of empty green land being used for housing or buildings [4]. This is of course a matter of wasting water because PDAM or consumers lack water. If this situation continues it will be a big problem for PDAM and the community [5].

Predictions are needed to calculate PDAM Tirtauli's monthly water distribution. When making predictions, you need to have the right way to complete them so that the predictions produce accurate results. Artificial neural networks are a great method that can be used to predict the problems described above. In artificial neural networks there are several methods, one of which is the method

used in this field research practice report, namely the backpropagation algorithm. The backpropagation algorithm is a popular and widespread mathematical tool used to predict and estimate time, as well as determine the results of non-linear functions [6]. Backpropagation is an algorithm that uses the error value in the output to change the inverse weight and uses a differential activation function for the forward step [7]. This algorithm is often used to solve complex problems. Indeed, this algorithm is trained using learning methods. By using this technique, it is hoped that a system can be created that can predict the distribution of PDAM water in PDAM Tirtauli, Pematangsiantar City and can help PDAM Tirtauli, Pematangsiantar City to calculate the amount of PDAM water distribution [8].

Some of the previous research that served as a guide for writing this article included research on predicting clean water needs in Malang district PDAMs. This research used training data and test data from 2013 to 2018 in the period January to December. This research resulted in a research accuracy rate of 89.72% [9]. In research predicting the installation of the number of clean water supply facilities in PDAM Pematang Siantar, in this research if the data was checked the accuracy was 89% [10].

METHOD

This chapter discusses the systematic procedures or methods used by researchers to find the truth of a phenomenon through logical considerations and supported by factual data as real evidence (objective, not personal hypothesis). Following are the steps that will be taken:

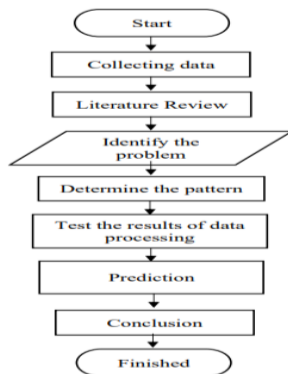


Image 1. Research Framework

The first step is to collect data. In this research, data was obtained from PDAM Tirtauli Pematangsiantar. Then the literature review aims to complete the background knowledge and theories used in this research. Apart from that, literature study is also the first step in the research process carried out in this article. The literature study used comes from journals. Then Identifying the Problem. This stage is carried out after obtaining a collection of data that will be processed at the data transformation stage according to the bot that has been determined. After that Determine the Pattern. At this stage a backpropagation model will be generated which will determine the pattern. Testing Data Processing Results. After completing the model determination, testing will be carried out using the Matlab R2011a application. After finding a model that passes the Matlab application test, predictions will be carried out to compare the model results with the highest and most accurate results. Final Evaluation.

Has a weight that will strengthen or weaken the signal, to determine the output, each neuron uses an activation function that is applied to the amount of input received [11]. ANN is determined by 3 things, namely [12] : (1) Pattern of connections between neurons (called network architecture), (2) Method for determining link weights (called train-

ing/learning method)., (3) Activation function, which is the function used to determine the output of a neuron.

Backpropagation Algorithm

Backpropagation is an algorithm that uses the error value in the output to change the inverse weight and uses a differential activation function for the forward step [7].

This algorithm uses the output error to change the value of the weights in a backward direction. The Backpropagation method involves three layers, namely: the input layer, where data is introduced to the network; hidden layer, where data is processed; and the output layer, where the results of the given input are produced [13].

Prediction

Prediction is the process of estimating future needs. Predictions are also estimates of future needs to meet demand for goods and services that take into account the criteria of price, quantity, time, place and quality. [7]. In predictions you don't have to give a definite answer about events that will happen in the future, but rather try to find accurate answers that might happen later [14].

Matlab

Matlab (Matrix Laboratory) is a high-level, closed, case-sensitive programming language in a digital computing environment developed by MathWorks [15].

RESULTS AND DISCUSSION

This research carried out several preparatory steps, including collecting data and distributing the data to the Matlab application, then looking for prediction values.

Table 1. Training Data Before Normalization

Data	2015	2016	2017	2018	2019	2020	2021
January	166,589	154,236	159,037	167,842	161,745	164,991	174,991
February	158,476	154,033	160,778	145,447	176,698	162,876	182,879
March	165,897	145,732	155,896	136,745	186,576	163,773	153,773
April	177,835	175,387	166,818	189,425	186,754	189,994	179,995
May	1,778,932	161,897	164,092	142,092	166,186	184,891	186,792
June	174,183	155,692	164,282	165,476	174,265	184,265	176,765
July	158,76	159,976	163,235	157,264	152,585	163,440	173,450
August	173,82	186,482	171,631	164,595	158,473	162,401	169,878
September	1,648,736	178,945	166,948	158,914	177,154	163,000	182,580
October	166,873	173,145	165,062	173,425	163,198	174,742	178,967
November	173,045	177,956	164,262	158,925	166,654	174,891	189,692
December	170,185	178,634	164,446	162,478	156,580	161,075	173,165

Data processing

The data used is data for 2015-2021, this data will be used as training data, test data and will be implemented in the Matlab application. As seen in Tables 1 and 2, the table displays the training data and test data before the data is normalized using the sigmoid function.

activation function is 1, so that the original data in integer form is converted into a decimal number with a range of [0,1]. [0.9] using the formula presented in equation 1 [16].

$$X_n = 0.8 * \frac{(x-a)}{(b-a)} + 0.1 \quad (1)$$

Information :

X_n = Normalized data results

X = Original data

a = The smallest value of data

b = The greatest value of data

Data Normalization

Data normalization is carried out using the backpropagation algorithm process using a binary sigmoid activation function with a range of $0 < x < 1$. The maximum value of the binary sigmoid

Table 2. Test Data Before Normalization

Data	2016	2017	2018	2019	2020	2021	2022
January	154,236	159,037	167,842	161,745	164,991	174,991	186,791
February	154,033	160,778	145,447	176,698	162,876	182,879	190,834
March	145,732	155,896	136,745	186,576	163,773	153,773	167,788
April	175,387	166,818	189,425	186,754	189,994	179,995	186,594
May	161,897	164,092	142,092	166,186	184,891	186,792	178,997
June	155,692	164,282	165,476	174,265	184,265	176,765	190,723
July	159,976	163,235	157,264	152,585	163,440	173,450	183,413
August	186,482	171,631	164,595	158,473	162,401	169,878	178,787
September	178,945	166,948	158,914	177,154	163,000	182,580	173,010
October	173,145	165,062	173,425	163,198	174,742	178,967	189,178
November	177,956	164,262	158,925	166,654	174,891	189,692	175,427
December	178,634	164,446	162,478	156,580	161,075	173,165	187,684

Where 0.8 is for determination,
Xn is for data normalization. From Table

3 we can see training data from 2015-
2021.

Table 3. Training Data After Normalization

Data	2015	2016	2017	2018	2019	2020	2021
January	0.1145	0.1085	0.11086	0.115149	0.11218	0.11376	0.11863
February	0.1105	0.1084	0.11171	0.10424	0.11946	0.11273	0.12247
March	0.1142	0.1043	0.10933	0.1	0.12428	0.11317	0.1083
April	0.1200	0.1188	0.11465	0.125663	0.12436	0.12594	0.12107
May	0.9	0.1122	0.11332	0.102605	0.11434	0.12345	0.12438
June	0.1182	0.1092	0.11342	0.113997	0.11828	0.12315	0.1195
July	0.1107	0.1113	0.11291	0.109996	0.10772	0.113	0.11788
August	0.1180	0.1242	0.117	0.113567	0.11059	0.1125	0.11614
September	0.8365	0.1205	0.11471	0.1108	0.11969	0.11279	0.12233
October	0.1146	0.1177	0.1138	0.117869	0.11289	0.11851	0.12057
November	0.1176	0.1200	0.1134	0.110805	0.11457	0.11858	0.12579
December	0.1162	0.1204	0.11349	0.112536	0.10966	0.11185	0.11774

Input Assignment

Table 5. Input Data

Variable	Data
X1	2015
X2	2016
X3	2017
X4	2018
X5	2019
X6	2020
X7	2021
X8	2022

In this research, before implementing the Matlab application, the input data will be determined first. We can see in the table below that the data entered

for this search contains 8 variables. Initialize and assign values to variables that will be used during calculations. Input variables are variables that are used as input during the calculation process.

Output Determination

The backpropagation algorithm requires output to achieve the desired result. The results are necessary to determine accurate future results. In this study, the results were determined by the learning rate ≤ 0.02 correct and ≥ 0.02 incorrect.

Table 4. Test Data After Normalization

Data	2016	2017	2018	2019	2020	2021	2022
January	0.3587	0.42971	0.55993	0.46976	0.517773	0.6657	0.84021
February	0.35569	0.45545	0.22871	0.69092	0.486493	0.7823	0.9
March	0.23293	0.38325	0.1	0.83702	0.499751	0.3518	0.55913
April	0.67153	0.54479	0.87916	0.83965	0.887567	0.7397	0.83728
May	0.47201	0.50448	0.17909	0.53545	0.812094	0.8402	0.72493
June	0.38023	0.50728	0.52495	0.65494	0.802843	0.6919	0.89836
July	0.44359	0.49180	0.40349	0.33428	0.494829	0.6429	0.79023
August	0.83562	0.61598	0.51191	0.42137	0.479457	0.59	0.72181
September	0.72415	0.54671	0.42789	0.69766	0.488328	0.7779	0.63637
October	0.63837	0.51882	0.64251	0.49125	0.661993	0.7245	0.87551
November	0.70952	0.50698	0.42805	0.54237	0.664194	0.8831	0.67212
December	0.71955	0.50970	0.4806	0.39337	0.459858	0.6387	0.85341

Architectural Design Analysis and Backpropagation Testing

The next step is to carry out testing using the Matlab R2011a application. From the experimental results, the best architecture was obtained, namely architecture 6-87-1 with a root mean square error value of 0.00010030 and an accuracy of 92%.

Table 6. Architectural training data 6-87-1

Data Name	Output	Error	SSE	Accuracy
January	0.1292	-0.0106	0.00011236	1
February	0.1267	-0.0042	0.00001764	1
March	0.0972	0.0111	0.00012321	1
April	0.1209	0.0002	0.00000004	1
May	0.1240	0.0004	0.00000016	1
June	0.1235	-0.0040	0.00001600	1
July	0.0999	0.0180	0.00032400	1
August	0.1087	0.0074	0.00005476	1
September	0.1218	0.0005	0.00000025	1
October	0.1037	0.0168	0.00028224	1
November	0.1337	-0.0080	0.00006400	1
December	0.1321	-0.0143	0.00020449	1
JLH SSE			0.00119915	
MSE			0.00009993	100%

Table 7. Architecture test data 6-87-1

Data Name	Output	Error	SSE	Accuracy
Januari	0.8663	-0.0261	0.00068121	1
Februari	0.8913	0.0087	0.00007569	1
Maret	0.5595	-0.0004	0.00000016	1
April	0.8399	-0.0026	0.00000676	1
Mei	0.7278	-0.0029	0.00000841	1
Juni	0.8948	0.0036	0.00001296	1
Juli	0.7756	0.0147	0.00021609	1
Agustus	0.7278	-0.0059	0.00003481	1
September	0.6341	0.0023	0.00000529	1
Oktober	0.8638	0.0118	0.00013924	1
November	0.6747	-0.0025	0.00000625	0
Desember	0.8575	-0.0041	0.00001681	1
JLH SSE			0.00120368	92%

Architectural Recapitulation

It can be seen in Figure 2 that carrying out architectural training 6-87-1

takes 02 minutes 06 seconds. As can be seen in Table 8, there are 5 architectures implemented in this research and present the lowest to highest levels of accuracy. The next step after testing data on the Matlab application is to evaluate the best architecture.

The evaluation process is carried out to determine which architecture is suitable for use in the prediction process. The table below shows the best architectures presented in the best architecture with a root mean square error value of 0.00010030. The 6-87-1 architecture is the best architecture because it provides the highest level of testing accuracy among other architectures and produces the smallest mean square error value among other architectures.

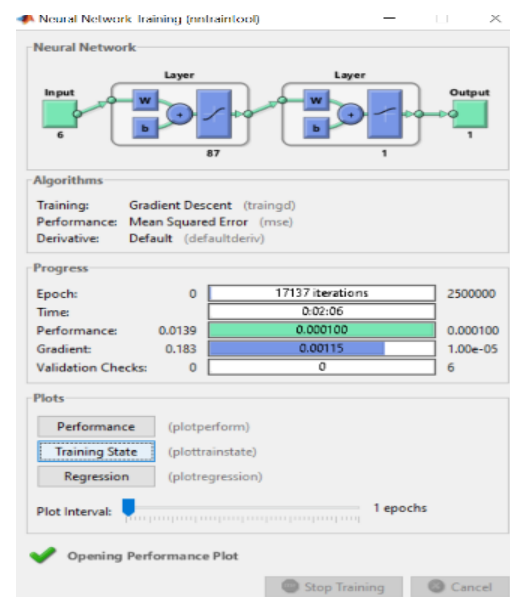


Image 2. Arsitektur 6-87-1

Table 8. Architectural Recapitulation

Arsitektur	Epoch	Time	Mse Pelatihan	Akurasi	Mse Pengujian	Akurasi
6_53_1	65295	08 M 02 D	0.00009987	92%	0.00039233	75%
6_58_1	34606	03 M 40 D	0.00010014	100%	0.00018148	83%
6_61_1	15620	01 M 42 D	0.00009970	92%	0.00057930	67%
6_81_1	23211	02 M 45 D	0.00010017	92%	0.00015124	83%
6_87_1	17137	02 M 06 D	0.00009993	100%	0.00010031	92%

Tabel 9. Water Distribution Predictions for 2023

No	Nama Data	Data Real	Target	Target Prediksi	Prediksi
1	Januari	186,791,923	0.84021	0.8663	189,863,938
2	Februari	190,834,775	0.9	0.8913	190,584,143
3	Maret	167,788,199	0.55913	0.5595	181,025,576
4	April	186,594,165	0.83728	0.8399	189,103,401
5	Mei	178,997,645	0.72493	0.7278	185,874,000
6	Juni	190,723,791	0.89836	0.8948	190,684,972
7	Juli	183,413,121	0.79023	0.7756	187,251,032
8	Agustus	178,787,012	0.72181	0.7278	185,874,000
9	September	173,010,153	0.63637	0.6341	183,174,669
10	Oktober	189,178,692	0.87551	0.8638	189,791,917
11	November	175,427,668	0.67212	0.6747	184,344,283
12	Desember	187,684,521	0.85341	0.8575	189,610,426

In this research, after getting the best architecture, the next step is a prediction process based on the best architecture value obtained. It can be seen from the table below that the actual data is taken from the test target data before normalization, the target data is taken from the test target data after normalization, and the predicted target data is taken from the best architectural output values obtained from the Matlab application.

CONCLUSION

It was concluded that the back propagation method can be used to predict the water quantity of PDAM Tirtauli Pematangsiantar city with valid data. Based on the model obtained, the 6-87-1 architecture is the model with the highest accuracy, namely 92%. From the research results it can be concluded that backpropagation can be used as a prediction method that helps predict things very easily.

This research only focuses on 1 PDAM in Pematangsiantar City. Therefore, it is recommended to conduct further research to collect a larger sample, for example by collecting PDAM samples throughout the province of North Sumatra.

BIBLIOGRAPHY

- [1] M. Adi and P. Hutabarat, "Penerapan Algoritma Backpropagation Dalam Memprediksi Jumlah Penduduk di Kecamatan Pematang Bandar Berdasarkan Nagori / Kelurahan," *J. Inf. Syst. Res.*, vol. 1, no. 2, pp. 63–69, 2020.
- [2] N. Afrida, S. Ramadani, I. Ambarita, and H. D. A. N. Pembahasan, "Prediksi Pemakaian Air Menggunakan Metode Backpropagation," *J. BIMASATI*, vol. 1, no. 3, pp. 105–110, 2022.
- [3] D. Agustina, M. Hafiyusholeh, A. Fanani, and D. Prasetijo, "Prediksi Distribusi Air Perusahaan Daerah Air Minum (PDAM) Tirta Dharma Kota Pasuruan Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation," *J. Ilm. Sist. Informasi, Teknol. Inf. dan Sist. Komput. P-ISSN*, vol. 18, no. 1, pp. 8–16, 2023.
- [4] Ekojono, Y. Yunshasnawa, and D. Mardhika, "Implementasi Metode Backpropagation pada Prediksi Pemakaian Air Perbulan," pp. 137–142, 2019.
- [5] B. Fachri, A. P. Windarto, and I. Parinduri, "Penerapan Backpropa

- gation dan Analisis Sensitivitas pada Prediksi Indikator Terpenting Perusahaan Listrik,” *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 2, p. 202, 2019, doi: 10.26418/jp.v5i2.31650.
- [6] Y. Franciska, H. Hayadi, And A. Setiawan, “Jaringan Syaraf Tiruan Untuk Memprediksi Penyakit Campak Menggunakan Algoritma Backpropagation,” *J. Ict Apl. Syst.*, Vol. 1, Pp. 43–46, 2022.
- [7] M. P. Sianipar and H. S. Tambunan, “Implementasi Jaringan Syaraf Tiruan Backpropagation Untuk Memprediksi Jumlah Pemasangan Instalasi Air Pada PDAM Tirtauli Pematangsiantar,” *J. Terap. Inform. Nusantara*, vol. 1, no. 9, pp. 489–498, 2021.
- [8] F. Rahmadani and A. M. H. Pardede, “Jaringan Syaraf Tiruan Prediksi Jumlah Pengiriman barang Menggunakan Metode Backpropagation (Studi Kasus : Kantor Pos Binjai),” Vol. 5, No. 1, 2021.
- [9] I. I. Ridho, A. Agung, G. Bagus, and A. P. Windarto, “Optimasi Fungsi Pembelajaran Jaringan Saraf Tiruan dalam Meningkatkan Akurasi pada Prediksi Ekspor Kopi Menurut Negara Tujuan Utama,” *J. BITS*, vol. 4, no. 4, 2023, doi: 10.47065/bits.v4i4.3240.
- [10] A. Wanto, “Analisis Prediksi Indeks Harga Konsumen Berdasarkan Kelompok Kesehatan Dengan Menggunakan Metode Backpropagation,” vol. 2, pp. 37–44, 2019.
- [11] R. S. Rosaamalia, Tiara. Santoso, “Penerapan algoritma backpropagation dan optimasi conjugate gradient untuk klasifikasi hasil tes laboratorium,” *J. GAUSSIAN*, vol. 11, pp. 506–511, 2023, doi: 10.14710/j.gauss.11.4.506-511.
- [12] I. I. Ridho, C. F. Ramadhani, and A. P. Windarto, “Penerapan Artificial Neural Network dengan Metode Backpropagation Dalam Memprediksi Harga Saham (Kasus : PT . Bank BCA , Tbk),” *J. Ris. Sist. Inf. Dan Tek. Inform.*, vol. 8, pp. 295–303, 2023.
- [13] A. Ikhwan, R. A. Putri, and M. Badri, “Implementasi Algoritma Backpropagation Neural Networks Untuk Memprediksi Hasil Kinerja Dosen,” vol. 4, no. 2, pp. 410–417, 2023, doi: 10.47065/josh.v4i2.2685.
- [14] P. I. Sijabat, G. W. Nurcahyo, and A. Sindar, “Algoritma Backpropagation Prediksi Harga Komoditi terhadap Karakteristik Konsumen Produk Kopi Lokal Nasional,” *J. Teknol. Inf. dan Digit. Zo.*, vol. x, no. x, pp. 96–107, 2020.
- [15] M. A. Al Ghifari, A. Panji Sasmito, and D. Rudhistiar, “Aplikasi Pendeteksian Kematangan Tomat Menggunakan Thresholding,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 1, pp. 294–300, 2022, doi: 10.36040/jati.v6i1.4606.
- [16] I. Permana and S. Febi, “Pengaruh Normalisasi Data Terhadap Performa Hasil Klasifikasi Algoritma Backpropagation,” *Indones. J. Inform. Res. Softw. Eng.*, vol. 2, no. 1, pp. 67–72, 2022.