

BACKPROPAGATION ALGORITHM IN PREDICTING THE AMOUNT OF WEST SIANTAR POPULATION GROWTH

Rika Setiana^{1*}, Yuegilion Pranayama Purba¹, M.Safii¹

¹ Information Systems, STIKOM Tunas Bangsa Pematang Siantar

email: rikasetiana957@gmail.com

Abstract: Indonesia is included in the category of developing countries and is a country with a relatively large population. According to data from the Pematang Siantar City Population and Civil Registration Service, West Siantar District is one of the sub-districts in Pematang Siantar City whose population continues to increase and this will lead to an increase in poverty and unemployment rates. To overcome the above problems, a method is needed to analyze the population growth of West Siantar, one of which is using the Backpropagation method. This research will use training data starting from 2017-2021 and test data from 2018-2022. The results carried out using MATLAB R2011a software show the best architecture 4-19-1 with an accuracy of 100 with an MSE number of 0.00010031375 and an epoch value of 124777. Based on the research carried out, the population of West Siantar in the next year is 86067 people. It is concluded that backpropagation can be used as a method that makes it easier to search for predictions and the level of accuracy obtained depends on the architecture used.

Keywords: Birth; Growth; Matlab; Predictions; Resident

Abstrak: Indonesia termasuk dalam kategori negara berkembang dan merupakan negara dengan jumlah penduduk yang relatif besar. Menurut data Dinas Kependudukan dan Catatan Sipil Kota Pematang Siantar, Kecamatan Siantar Barat merupakan salah satu kecamatan di Kota Pematang Siantar yang jumlah penduduknya terus meningkat dan hal ini akan berdampak pada peningkatan angka kemiskinan dan pengangguran. Untuk mengatasi permasalahan diatas diperlukan suatu metode untuk menganalisis pertumbuhan penduduk Siantar Barat, salah satunya adalah dengan menggunakan metode Backpropagation. Penelitian ini akan menggunakan data latih mulai tahun 2017-2021 dan data uji 2018-2022. Hasil yang dilakukan dengan menggunakan software MATLAB R2011a menunjukkan arsitektur terbaik 4-19-1 dengan akurasi 100 dengan angka MSE 0.00010031375 dan nilai epoch 124777. Berdasarkan penelitian yang dilakukan, populasi Siantar Barat di tahun berikutnya sebanyak 86067 orang. Disimpulkan bahwa backpropagation dapat digunakan sebagai metode yang memudahkan pencarian prediksi dan tingkat akurasi yang diperoleh tergantung pada arsitektur yang digunakan.

Keywords: Kelahiran; Matlab; Pertumbuhan; Penduduk; Prediksi

INTRODUCTION

The change in a population over time, which can be calculated as the change in the number of individuals within a population using "per unit of time" as the measurement, is referred to as population growth [1]. The country with the highest population density, ranked 4th in the world, is Indonesia [2]. One of the countries with the largest population in the world, approximately two hundred million people, and ranking fourth after the United States in the list of countries with the largest populations globally is Indonesia [3]. During this period, the population of Indonesia has been increasing from year to year, while the land area and cities have started to shrink. Along with the increasing population, the land area will gradually decrease due to the growing population density.

The population growth continues to increase in all countries, including Indonesia. Therefore, the government must undertake various efforts to control population growth in order to achieve balanced population growth and enhance the quality of the population [4]. Population growth in a region is crucial as it can influence the progress and well-being of that area [5]. If the population growth rate is too high, various issues such as high unemployment and poverty rates may arise in the city of Pematang Siantar. If the population growth rate is too high, various issues such as high unemployment and poverty rates may arise in the city of Pematang Siantar. On the contrary, if we can anticipate and allocate sufficient resources, we can minimize the negative impacts mentioned above. According to data obtained from the Population and Civil Registration Service of Pematang Siantar City, West Siantar District is one of the sub-districts

in Pematang Siantar City which has a very large population, such as in 2017 it had a population of 81,519 people, in 2018 83,755 people, in 2019 80,253 people, in 2020 79,744 people, in 2021 79,454 people, in 2022 82,657 people, it can be seen that in 2022 West Siantar sub-district will have an increase again, therefore a prediction of population growth is needed in the next year. Increasing population growth has economic, social, political and defense implications. As time goes by, population levels increase. This could be a problem like a population explosion. The city of Pematang Siantar is classified as a developing city and is a city with a relatively large population.

Therefore, to overcome the above problems, it is necessary to realize that as an alternative choice, it is necessary to apply and analyze appropriate prediction models so that local governments can predict population growth so that local governments can make the right decisions [6]. To overcome the above problems, a method is needed to obtain the best architecture for analyzing the population growth of West Siantar. Another way to estimate population growth is to use artificial neural networks with backpropagation. Based on previous research, there are several studies that predict the rate of population growth. In previous research, in predicting the development of the population of Ciamis city, it was explained that the backpropagation algorithm can predict the rate of population growth by producing a correlation value of 0.84796 [7]. Research on the prediction of the population of Pematang Bandar produced good accuracy results with an accuracy rate of 92.3% [8]. Based on previous research, artificial neural networks can solve complex problems related to

recognition, prediction and pattern recognition.

Artificial neural network (ANN) technology has developed rapidly in the field of prediction. ANN can make predictions based on past event data and related factors. The backpropagation neural network technique is often used to solve prediction problems. This technology can solve complex problems and refers to the process of identification, prediction and pattern recognition. In addition, backpropagation is fast, simple, and easy to program without any optimization parameters other than the number of flexible inputs, since no prior knowledge of the network is required and no special features of the function under study are required. In finding the best model, it is necessary to implement the basic data into the Matlab application to ensure accurate results once completed. This research will use Matlab training data and test data for data processing. Training data starts from 2017-2021 and test data starts from 2018-2022. In this research, predictions will be made of the population of West Siantar District, Pematang Siantar City, where the results of the prediction of the population of West Siantar will hopefully be able to help the Pematang Siantar government in formulating appropriate policies to prevent a population explosion so as not to cause negative impacts such as increasing poverty rates. and unemployment rates.

METHOD

In this research, population data was collected from the Pematang Siantar Civil Registration Population Service, West Siantar Regency. This research uses a training data set from 2017 to 2021, a

test data set from 2018 to 2022, and makes predictions in 2023. After receiving data from the Civil Registration Service, problem identification and decision making are carried out, as well as output and input to be able to test the Matlab application to obtain the optimal architecture. Predictions are then made based on previously obtained data based on the best architecture from Matlab. In selecting the best architecture, this research uses a fairly small error value, namely ≤ 0.002 to get the correct value or 1.

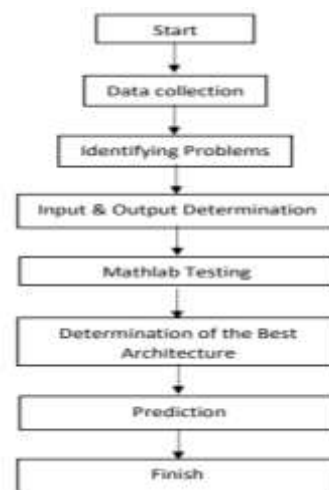


Image 1. Research Framework

Resident

Every person who has resided in a geographical area of the Unitary State of the Republic of Indonesia for more than 6 months and/or who has resided for less than 6 months but intends to remain is called a resident. The number of people living in an area at a certain time is the result of demographic processes such as birth rates, death rates and migration [9].

Artificial intelligence

Artificial intelligence is a general term that refers to technology that makes machines “smarter” [10]. In the Back-propagation Algorithm, data pre-processing is carried out using data nor-

malization techniques using the sigmoid activation function, changing the data at intervals of 0.1 to 0.9 [11].

The backpropagation artificial neural network (BPNN) technique was first introduced by Paul Werbos in 1974, reformulated by David Parker in 1982, and popularized by Rumelhart and McClelland in 1986 [12]. It can be seen in Image 2 that the backpropagation algorithm has several artificial neural network layers. The network used in this research has three layers, namely 7 input layers, 12 hidden layers, and 1 output layer.

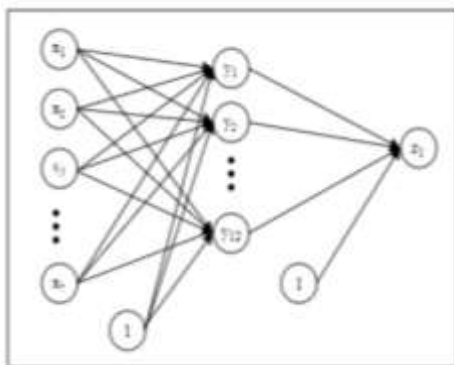


Image 2. Artificial Neural Network Architecture

Carrying out training on the backpropagation method includes 3 stages [13]:

1. forward propagation

This stage includes initializing the initial weights, calculating the number of values in the input layer and hidden layer, as well as calculating the activation value in the input layer and hidden layer.

2. Backpropagation

The backpropagation stage is the difference between the network output and the desired target, according to the error that occurred.

3. Change in Weight

Matlab

This software provides convenience and simplicity in solving problems related to vectors and matrices. Find the

inverse matrix and solve the linear equation. When processing data to obtain output values, use the Matlab application with the following formula [14]:

```
% Creating Multi Layer Neural Network (2,3,4,5(Free))
net = newff(minmax(p),[hidden layer,1],{'tansig','logsig'},'traingd');
% Generating weights and bias
net.IW{1,1}
net.LW{2,1}
net.b{1}
net.b{2}
% Fletcher-Reeves default parameter value (trainrp)
net.trainParam.epochs = 2500000;
net.trainParam.show = 1000;
net.trainParam.showCommandLine = false;
net.trainParam.showWindow = true;
net.trainParam.goal = 0.0001;
net.trainParam.time = inf;
net.trainParam.min_grad = 1e-05;
net.trainParam.max_fail = 6;
net.trainParam.searchFcn = 'srchcha'
% Performing Testing
net = train(net,p,t)
% View results when performance is found
[a,Pf,Af,e,perf] = sim(net,p,[],[],t)
%perform simulations using test data based on training results
[a,Pf,Af,e,perf] = sim(net,p1,[],[],t1)
```

RESULTS AND DISCUSSION

Before entering the Matlab application to select the best architecture for the overall analysis, a suitable identification process is required. Identifying problems in research including the root of the problem and developing a detailed problem formulation including the impact of the problem. Without accurately defining the problem, research

results can easily be wrong. Problem identification also requires a thorough understanding of the nature of the problem in order to produce the right solution to resolve the problem. After identifying the problem, the next step will be determining the input and determining the output.

Input Assignment

Determination of data is taken from the number of years included in the general population data of the Civil Registration Service of the Pematang Siantar Civil and Population Registration Service. Data uses data from 2018 to 2022. This data will guide decision making when making predictions using the backpropagation algorithm. Determining the input in this research is divided into training data and test data. Explains that the input data contained in this search has 6 variables with variables X1 to X6. Each variable has its own criteria.

Table 1. List of Input Predictions for the Population of West Siantar

No	Variables	Kriteria's name
1	X1	2017
2	X2	2018
3	X3	2019
4	X4	2020
5	X5	2021
6	X6	2022

The backpropagation algorithm requires an optimal architecture when carrying out prediction processing, paying attention to minimal error values and paying attention to the highest accuracy. This study uses a minimum error value of 0.02 for correct values (1) and 0.02 or higher for incorrect values (0). The smaller the minimum error obtained, the better the research.

Data processing

In the initial stage before input-

ting data into the Matlab application, we need to normalize the data using the Excel sigmoid function (which does not reach 0 or 1). This report uses the sigmoid function to separate the normalized data into two parts: training data and testing data. Pada penelitian ini data latih dimulai pada tahun 2017 hingga 2021 dan data uji dimulai pada tahun 2018 hingga 2022. Before separating training data and test data, the data is normalized first. Therefore, data normalization can be expressed with the following equation [15]:

$$X^1 = \frac{0.8(X-a)}{b-a} + 0.1 \tag{1}$$

Information :

x' : Data after normalization

: Data to be normalized

: The smallest data

b : The biggest data

Table 2. Normalization of Training Data

2017	2018	2019	2020	2021
0.4252	0.4392	0.4260	0.4237	0.4226
0.8578	0.9000	0.8761	0.8779	0.8868
0.1895	0.1890	0.1721	0.1697	0.1652
0.1166	0.1160	0.1060	0.1015	0.1000
0.5012	0.5181	0.4995	0.4990	0.4955
0.3780	0.3810	0.3526	0.3454	0.3420
0.1500	0.1552	0.1426	0.1405	0.1385
0.2393	0.2458	0.2334	0.2310	0.2269

Table 2 shows data that shows normalized training data using the sigmoid function. In ANN, data normalization is used to change data values (training data and test data) into a range of values that can be processed in ANN. ANN is only able to handle values between -1 and 1. Normalization at this stage is carried out to obtain data between 0 and 1 because the activation function used in this research is the sigmoid binary activation function which has values ranging from 0 to 1 Data

Normalization Method used in this research is the min-max method. The min-max data normalization method is a data normalization method by carrying out linear transformations on the original data. Table 3 shows the experimental data normalized to zero using the sigmoid function.

Best Architecture

In this research, the 4-19-1 architecture provided the best level of accuracy when testing data in Matlab. From Table 4 it can be seen that the 4-19-1 architecture produces training data with an accuracy level of 100% and an SSE Image of 0.00010014. From Table 5 it can be seen that architecture 4-19-1 is the best architecture with testing accuracy of more than 75%, namely 100%, mean square error of 0.00114545.

Tabel 3. Normalization of Testing Data

2018	2019	2020	2021	2022
0.4392	0.4260	0.4237	0.4226	0.4265
0.9000	0.8761	0.8779	0.8868	0.8883
0.1890	0.1721	0.1697	0.1652	0.2253
0.1160	0.1060	0.1015	0.1000	0.1301
0.5181	0.4995	0.4990	0.4955	0.5183
0.3810	0.3526	0.3454	0.3420	0.3456
0.1552	0.1426	0.1405	0.1385	0.1396
0.2458	0.2334	0.2310	0.2269	0.2279

Table 4. Architecture Training Data 4-19-1

No	Village name	Output	Error	SSE	Results
1	Kel. Banjar	0.4073	0.0153	0.00023409	1
2	Kel. Bantan	0.8871	-0.0003	0.00000009	1
3	Kel. Dwikora	0.1602	0.0050	0.00002500	1
4	Kel. Proklamasi	0.1150	-0.0150	0.00022500	1
5	Kel. Simarito	0.4998	-0.0043	0.00001849	1
6	Kel. Sipinggolpinggol	0.3573	-0.0153	0.00023409	1
7	Kel. Teladan	0.1354	0.0031	0.00000961	1
8	Kel. Timbang Galung	0.2194	0.0074	0.00005476	1
Number of SSE				0.00080113	100%
MSE				0.00010014	

Table 5. Best Architectural Testing Data

No	Village name	Output	Error	SSE	Results
1	Kel. Banjar	0.4271	-0.0006	0.00000036	1
2	Kel. Banten	0.8886	-0.0003	0.00000009	1
3	Kel. Dwikora	0.2074	0.0180	0.00032400	1
4	Kel. Proklamasi	0.1227	0.0074	0.00005476	1
5	Kel. Simarito	0.5186	-0.0004	0.00000016	1
6	Kel. Sipinggolpinggol	0.3445	0.0011	0.00000121	1
7	Kel. Teladan	0.1588	-0.0192	0.00036864	1
8	Kel. Timbang Galung	0.2352	-0.0073	0.00005329	1
Number of SSE				0.00080251	100%
MSE				0.00010031	

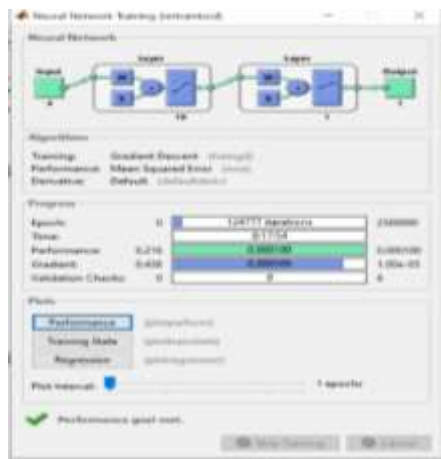


Image 3. Architectural Test Results 4-19-1

Seen in Image 3, the 4-19-1 architecture takes 17 minutes 54 seconds to carry out the train data process.

Architectural Recapitulation

The overall results for each architecture can be seen in the table below. It can be seen in the image below that the 4-19-1 architecture produces 100% accu-

racy. In this research, to achieve the best results, the author used 5 architectures which were tested using the Matlab R2011a application. When applying the architectural model, the results are different again and it can be seen in the previous table that the best architecture is obtained in the 4-19-1 model with a mean square error for training of 0.00010014125 and a root mean square error for testing of 0.00010031375 in 17 minutes 54 seconds at iteration 124777 epoch. To determine the average number of errors, you first need the total number of SSE, where the number of SSE is taken from the error results in Matlab and raised to the power so that the total number of SSE is divided by all the data, then you get the mean square error result. The table below shows the tested architecture. Table 6 shows that training accuracy can be better than testing accuracy.

Table 6. Architectural Recapitulation

Model	Epoch	Time	MSE Training	Accuracy	MSE Testing	Accuracy
4_7_1	286471	49 M 10 D	0.0001001	100%	0.0008845	75%
4_33_1	28686	02 M 6 D	0.0001001	100%	0.0006818	75%
4_21_1	31015	03 M 51 D	0.0000999	100%	0.0004251	75%
4_19_1	124777	17 M 54 D	0.0001001	100%	0.0001003	100%
4_68_1	32787	04 M 1 D	0.0001003	88%	0.0008843	88%

Prediction Results

Table 7. Prediction Results of Annual Population Number

No	Village name	Real Data	Target	Target Prediction	Prediction
1	Kel. Banjar	11977	0.426474	0.4271	12329
2	Kel. Bantan	23890	0.888331	0.8886	23611
3	Kel. Dwikora	6789	0.22534	0.2074	6958
4	Kel. Proklamasi	4332	0.130085	0.1227	4887
5	Kel. Simarito	14345	0.51828	0.5186	14566
6	Kel. Sipinggolpinggol	9890	0.345563	0.3445	10309
7	Kel. Teladan	4578	0.139622	0.1588	5770
8	Kel. Timbang Galung	6856	0.227938	0.2352	7637
Amount					86067

In research, when carrying out the prediction process, the maximum and minimum values are determined in the original data obtained from the target data before normalization. After getting the maximum value and minimum value, enter the target value obtained from the test data after normalization, then determine the target value that is expected to be obtained according to the best architectural and architectural data test results. The final step, make a prediction using real data, subtract 0.1 and multiply it by the maximum result value, then divide by the value 0.8, then add the minimum value from the real data. It can be seen in the table below 7 of the prediction results made based on the best architecture that has been obtained, table 7 shows that the population of West Siantar in the coming years will reach a population of 86,067 people, it can be concluded that the population of West Siantar in the next year will be based on The prediction that has been made is that the population will increase by 3401 people.

CONCLUSION

This research shows that the backpropagation algorithm can determine the optimal architecture for analyzing the

population in West Siantar in the future and obtain results for the population of West Siantar in the coming years, namely 86,067 people. In this research, looking at the level of accuracy, the 4-19-1 architecture is the best architecture because it achieves the highest accuracy, namely 100% accuracy and an MSE number of 0.00010031375 and an epoch value of 124777. 4-19 Architecture -1 also has the smallest error so this architecture accepts smallest error value. The smaller the minimum error value, the better the search. However, when viewed from a speed perspective, the 4-19-1 architecture is an architecture that takes quite a long time to test. In terms of speed of implementation of the training process, the 4-33-1 architecture is the fastest architecture, but in terms of accuracy the 4-33-1 architecture is very good with a low accuracy of 75%.

BIBLIOGRAPHY

[1] Purwadi, P. Sari, And N. Safitri, "Penerapan Data Mining Untuk Mengestimasi Laju Pertumbuhan Penduduk Menggunakan Metode Regresi Linier Berganda Pada Bps Deli Serdang," *Sains Dan Komput.*, Vol. 18, No. 1, Pp. 55–

- 61, 2019.
- [2] M. Ardiansyah, A. Alim, N. Fadilah, And N. Dhi, "Analisis Data Science Pada Struktur Data Kepadatan Penduduk Kota Tegal," *Jupti*, Vol. 1, No. 3, Pp. 35–41, 2022.
- [3] N. R. Rahmawati, "Prediksi Laju Pertumbuhan Penduduk Dengan Program Matlab Menggunakan Metode Jaringan Syaraf Tiruan," *J. Publ. Tek. Inform.*, Vol. 2, No. 1, Pp. 1–5, 2023.
- [4] A. K. Dewi, M. T. Furqon, And Wihandika, "Prediksi Laju Pertumbuhan Penduduk Menggunakan Metode Support Vector Regression," *Pengemb. Teknol. Inf. Dan Ilmu Komput.*, Vol. 4, No. 1, Pp. 421–427, 2020.
- [5] P. Kurniawan, H. Rossa, A. Permana, W. A. Ramadan, And B. W. Aji, "Prediksi Jumlah Penduduk Jakarta Selatan Menggunakan Metode Regresi Linear Berganda Multiple Linear Regression Method," *J. Sist. Dan Teknol. Inf.*, Vol. 10, No. 4, Pp. 518–523, 2022, Doi: 10.26418/Justin.V10i4.48331.
- [6] A. Agung And A. Putri, "Penerapan Data Mining Untuk Mengestimasi Laju Data Mining Usage To Estimate Civil Growth In Denpasar," *Jbase*, Vol. 6, No. 1, Pp. 37–44, 2023.
- [7] R. Armanda, Kurnia, "Prediksi Pertumbuhan Penduduk Kecamatan Cimaragas Kabupaten Ciamis Dengan Metode Artificial Neural Network," *J. Algoritm.*, Vol. 3, No. 2, Pp. 170–178, 2023.
- [8] 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.
- [9] H. Aulawi, R. Kurniawati, And V. V. Pratama, "Analisa Keputusan Pemilihan Jasa Ekspedisi Dengan Metode Ahp Dan Borda," No. 1, Pp. 23–29.
- [10] R. Roosdianto, A. Sari, And A. Satriansyah, "Rancang Bangun Aplikasi Sistem Informasi Absensi Karyawan," *J. Inti Nusa Mandiri*, Vol. 15, No. 2, Pp. 135–142, 2021.
- [11] A. F. Suahati, A. A. Nurrahman, And O. Rukmana, "Penggunaan Jaringan Syaraf Tiruan – Backpropagation Dalam Memprediksi Jumlah Mahasiswa Baru Predicting Number Of New Student Using Artificial Neural Network - Backpropagation," *J. Media Tek. Dan Sist. Ind.*, Vol. 6, No. 1, Pp. 21–29, 2022, Doi: 10.35194/Jmtsi.V6i1.1589.
- [12] J. Rinaldi, H. Haviluddin, And S. Pakpahan, "Algoritma Backpropagation Neural Network Dalam Memprediksi Harga Komoditi Tanaman Karet," *J. Ilm.*, Vol. 12, No. 1, Pp. 32–38, 2020.
- [13] B. Yanto, R. Hutagaol, And R. Rahman, "Analisis Optimasi Algoritma Backpropagation Momentum Dalam Memprediksi Jenis Tingkat Kejahatan Di Kecamatan Tambusai Utara Budi," Vol. 1, Pp. 47–60, 2022.
- [14] V. V. Utari, A. Wanto, I. Gunawan, And Z. M. Nasution, "Prediksi Hasil Produksi Kelapa Sawit Ptpn Iv Bahjambi Menggunakan Algoritma Backpropagation," *J. Comput. Syst. Informatics (Josyc)*, Vol. 2, No. 3, Pp. 271–279, 2021.

- [15] F. Khairati And H. Putra,
“Prediksi Kuantitas Penggunaan
Obat Pada Layanan Kesehatan
Menggunakan Algoritma
Backpropagation Neural
Network,” *J. Sistim Inf. Dan
Teknol.*, Vol. 4, Pp. 128–135,
2022, Doi:
10.37034/Jsifotek.V4i3.158.