International Journal of Information System & Technology
Vol. 3, No. 1, (2019), pp.11-17

Increasing Prediction Accuracy with the Backpropagation Algorithm (Case Study: Pematangsiantar City Rainfall)

Yogi Prayoga[1], Dedy Hartama[2], Jalaluddin[3], Sumarno[4], Zulaini Masuro Nasution[5]

[1,2,3,4,5]STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra, Indonesia

rayogayogi922@gmail.com

Abstract

The more advanced science and technology from various disciplines, currently rainfall can be predicted by carrying out various empirical approaches, one of which is by using Artificial Neural Networks (ANN). This study aims to apply ANN with backpropogation algorithm in predicting rainfall. The research data used is BPS data of the transfer city. The results of the study state that of the 6 models (4-5-1, 4-10-1, 4-25-1, 4-5-10-1, 4-5-25-1 and 4-5-50-1) architecture that was trained and tested using Matlab 6.1 application software, the results showed that the 4-5-25-1 architectural model was the best model for making predictions with 75% truth accuracy, Training MSE 0.001004582, Testing MSE 0.021882712 and Epoch 59,076 . It is expected that research can provide input to the government, especially BMKG Pematangsiantar city in predicting Rainfall based on computer science so as to improve the quality of services in the fields of Meteorology, Climatology, Air Quality and Geophysics in accordance with applicable laws and regulations.

Keywords: *Prediction, Backpropogation, ANN, Rainfall, Pematangsiantar*.

1. Introduction

Rainfall is water vapor that condenses and falls from the atmosphere to the earth in all its forms both liquid (rain) and solid (snow) in a series of hydrological cycles. The occurrence of high intensity rainfall can cause various kinds of hazards such as floods, landslides and winds tight which can later cause harm to living things around him. Pematangsiantar City is a city located near the equator on the line 2°53'20" - 3° 01'00" North Latitude and 99°1'00" - 99°6'35" East Longitude, and is in the middle of the region Simalungun Regency. Pematangsiantar City is classified into a tropical and flat area, with a temperate climate with an average maximum temperature of 30.3ºC and an average rainfall of 229 mm - 341 mm.

Pematangsiantar city is a city located above sea level with a tropical and flat area, temperate and has normal rainfall. The fact that occurs from satellite monitoring of the Badan Meteorologi, Klimatologi dan Geofisika (BMKG) that in 2018 Pematangsiantar city has a high intensity of rainfall which causes a great danger to the special life in Pematangsiantar city. The Badan Meteorologi, Klimatologi dan Geofisika (BMKG) is an official government agency that provides information services to the public regarding weather and rainfall. Therefore many say that the institution is superior in predicting rainfall. The institute provides information to the public regarding the latest rainfall forecasting on a daily, weekly, monthly time scale. But, all prediction models given by these institutions are statistical models [1].

With the advancement of science and technology from various scientific disciplines, currently rainfall can be predicted by taking a variety of empirical approaches, one of which is by using Artificial Neural Networks (ANN). ANN is one technique that can be used for prediction cases by learning to form a reference model based on training data, which is then followed by pattern matching [2], [3].

ISSN: 2580-7250 IJISTECH

One of the advantages of ANN is the ability to classify data that has not been given during the previous study [4] [5]. There are several methods in ANN one of which is the Backpropogation method. This method can reduce the error rate on a large scale because of the hidden layer. This hidden layer also reduces the error rate that exceeds the ability of the single layer network [2], [6]. Many studies related to backpropogation are able to solve prediction problems with high predictive accuracy values. One of them is a study conducted by [7] about the prediction of consumer price index based on health groups. The results of the study concluded that the backpropogation algorithm can be used in the prediction of the consumer price index based on health groups using 8 architectural models, namely: 12-5-1 with an accuracy rate of 58%, 12-26-1 = 58%, 12-29-1 = 75%, 12-35-1 = 50%, 12-40-1 = 42%, 12-60-1 = 67%, 12-70-1 = 92% and 12-75-1 = 50%. From the test model obtained the best model 12-70-1 with an accuracy rate of 92%, MSE 0.3659742 with an error rate used of 0.001 - 0.05. Based on this explanation, it is expected that research can provide input to the government, especially Pematangsiantar BMKG in predicting rainfall based on computer science so as to improve the quality of services in the fields of Meteorology, Climatology, Air Quality and Geophysics in accordance with applicable laws and regulations.

2. Research Methodology

2.1. Artificial intelligence

The intelligence referred to here refers to machines that are able to think, weigh the actions to be taken and are able to take decisions like humans do. [8].

2.2. Artificial Neural Networks

Artificial neural networks are created as a generalization of a mathematical model of human understanding based on the assumption of processing information called neurons where signals flow between neurons through a connecting connection, each connecting connection has a corresponding weight [9].

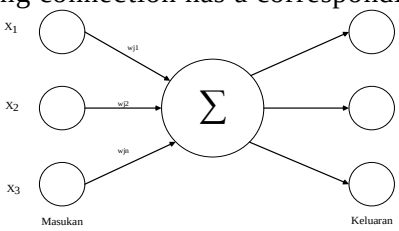

Figure 1. Neuron Model

2.3. Backpropagation Method

Each pattern of input and output information that is given into the artificial neural network is processed in neurons where the neurons are collected in layers called neuron layers. The layers making up the artificial neural network can be divided into 3, including: input layer, hidden layer dan output layer [10-15].

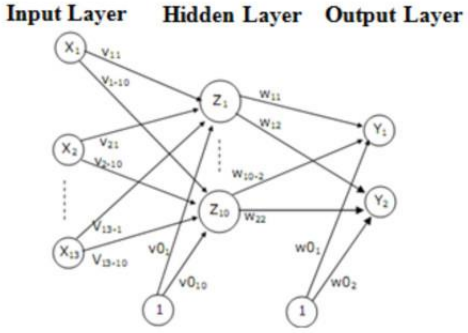

Figure 2. Backpropagation Network Architecture

Research methodology is the stage of conducting research in collecting data or information used in finding solutions to problems as shown in the following flowchat.

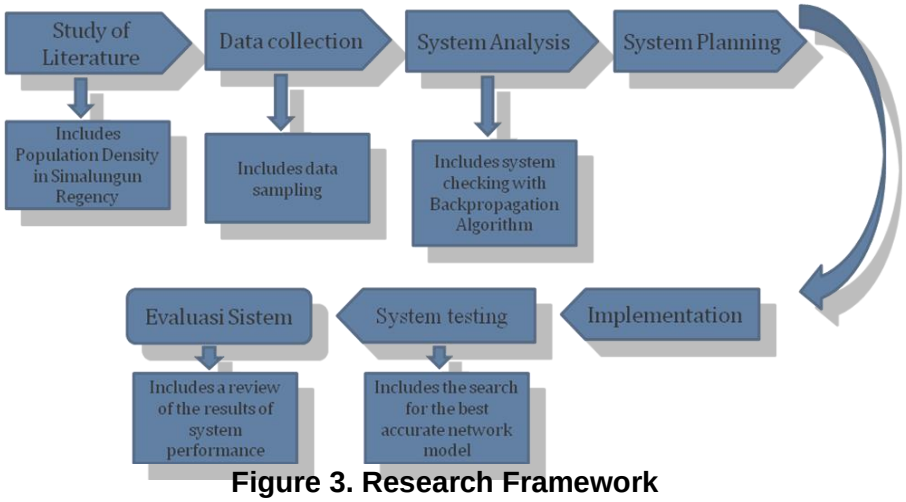

Figure 3. Research Framework

2.4. Data source

The source of research data was obtained from the Badan Pusat Statistic of Pematangsiantar city by url: https://siantarkota.bps.go.id/. The data used in this study is rainfall data during the period 2006-2015.

Table 1. Rainfall data (mm) Pematangsiantar city (2006-2015)

Month	Rainfall data (mm)								
	2006	2007	2008	2009	2010	2012	2013	2014	2015
January	198	164	119	410	144	90	480	57	148
February	277	70	163	55	56	237	367	119	56
March	74	116	355	404	197	274	208	115	139
April	378	347	253	309	139	341	386	309	211
May	811	271	183	353	96	221	246	347	339
June	271	259	201	41	372	93	115	132	153
July	117	197	344	126	355	291	133	159	84
August	151	461	544	148	408	143	235	386	204
September	167	373	574	478	372	340	221	235	236
October	406	253	413	342	243	204	427	401	211
November	210	287	319	254	477	285	392	194	403
Desember	237	193	205	202	370	230	560	266	-

3. Results and Discussion

3.1. Input and Target

The data used in this study are rainfall data for the period 2006-2015 obtained from the Badan Pusat Statistic of Pematangsiantar city by url: https://siantarkota.bps.go.id/. The data is then processed using the backpropogation method. So that the data can be recognized by artificial neural networks, the data must be represented in numerical form between 0-1, this is because the network uses the activation function of binary sigmoid (logsig) which has a range of values from 0-1.

3.2. Input Variable

Variables are needed as input. In this case the data was obtained from the Badan Pusat Statistic (BPS) of Pematangsiantar city with the subject of rainfall data (2006-2015). In this case the variable is determined by looking at the data dependence on

the research conducted. The data is divided into 2 parts including: Training data (2006-2010) and testing data (2011-2015). This data has a variable rainfall data input every month (January-December) in the last 10 years.

3.3. Target Variable

The target variable used in rainfall prediction in the city of Pematangsiantar is rainfall data.

3.4. Output Variable

The expected outcome at this stage is to form the best architectural model for predicting rainfall in Pematangsiantar city by using the backpropogation algorithm. The test results are as follows:

a) The output of this prediction is the best architectural pattern in predicting rainfall by looking at the minimum error.
b) Categorizing training outputs and testing the minimum error rate of the target as shown in the following table:

Table 2. Category Data

No	Information	*Error* Minimum
1	True	0.1 between 0.001and -(0.1 between 0.001)
2	False	> 0.1 and > (-0.1)

3.5. Data processing

Data processing is done with the help of Matlab 6.1 software applications. The data used is the 2006-2015 rainfall data. The data is divided into two parts including: Training data (2006-2010) and testing data (2011-2015) as shown below:

a) Training Data
   Input (X): X1 to X4 (rainfall data / year)
   Output (Y): Y (rainfall data)
b) Testing Data
   Input (X): X1 to X4 (rainfall data / year)
   Output (Y): Y (rainfall data)

After the data is divided into 2 parts, the data is then changed to 0-1 because the activation function used is binary sigmoid (logsig). This function has input and output values in the range 0-1. Following are the results of training data conversion and rainfall testing as shown in the following table:

Table 3. Training data (Conversion)

Month	Rainfall data (mm)				
	2006 (X_1)	2007 (X_2)	2008 (X_3)	2009 (X_4)	2010 (Y)
January	0,263	0,228	0,181	0,483	0,207
February	0,345	0,130	0,227	0,115	0,116
March	0,134	0,178	0,426	0,477	0,262
April	0,450	0,418	0,320	0,378	0,202
May	0,900	0,339	0,248	0,424	0,157
June	0,339	0,326	0,266	0,100	0,444
July	0,179	0,262	0,415	0,188	0,426
August	0,214	0,536	0,623	0,211	0,481
September	0,231	0,445	0,654	0,554	0,444
October	0,479	0,320	0,486	0,413	0,310
November	0,276	0,356	0,389	0,321	0,553
Desember	0,304	0,258	0,270	0,267	0,442

Table 4. Testing data (Conversion)

Month	Rainfall data (mm)				
	2011 (X_1)	2012 (X_2)	2013 (X_3)	2014 (X_4)	2015 (Y)
January	0,266	0,151	0,556	0,117	0,211
February	0,219	0,304	0,439	0,181	0,116
March	0,275	0,342	0,274	0,177	0,202
April	0,366	0,412	0,458	0,378	0,277
May	0,388	0,287	0,313	0,418	0,410
June	0,246	0,154	0,177	0,195	0,216
July	0,266	0,360	0,196	0,223	0,145
August	0,367	0,206	0,302	0,458	0,269
September	0,403	0,411	0,287	0,302	0,303
October	0,392	0,269	0,501	0,474	0,277
November	0,383	0,354	0,465	0,259	0,476
Desember	0,363	0,296	0,639	0,334	0,457

The process of converting training and testing data is carried out using the formula x1 = (0.8 (xa) / ba) + 0.1 where the value x1 is the value of the conversion results, x is the value to be converted, b is the maximum value of the entire data used and a is the value the minimum from the overall data used.

3.6. Discussion

Selection of the best architecture from 6 models (4-5-1, 4-10-1, 4-25-1, 4-5-10-1, 4-5-25-1 and 4-5-50-1) architecture who were trained and tested using Matlab 6.1 application software have different results in terms of epoch, accuracy, MSE training and MSE testing. The best model is used to predict rainfall in the city of Pematangsiantar. The assessment of the best architectural model is seen from several aspects such as epoch, minimum error and accuracy of truth as shown in the following table:

Table 5. Recapitulation of Architectural Models

Parameter	4-5-1	4-10-1	4-25-1	4-5-10-1	4-5-25-1	4-5-50-1
MSE Training	0,000998425	0,001007804	0,001000727	0,00100222	0,001004582	0,001003217
MSE Testing	0,037632091	0,032822047	0,040237025	0,02601923	0,021882712	0,028372541
Epoch	2.099.018	54.223	9.194	151.965	59.076	37.545
Akurasi	50%	58%	25%	50%	75%	58%

Based on the table, the selection of the best architectural model is 4-5-25-1 with an accuracy level of 75%, MSE Training 0.001004582, MSE Testing 0.021882712 and Epoch 59.076. Following are the complete results of the Matlab 6.1 application analysis.

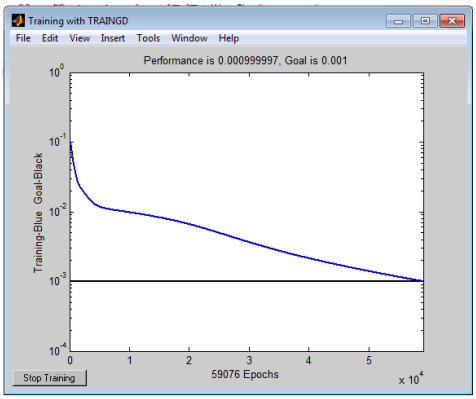

Figure 4. Architecture 4-5-25-1 achieving Goal

Following are the results of 4-5-25-1 architecture testing.

Table 5. Training & Testing Data and Error Results 4-5-25-1

Data Training				
No	**Target**	**ANN 4-5-25-1**		
		Output	**Error**	**SSE**
1	0,207	0,2317	-0,0247	0,0006094486
2	0,116	0,1764	-0,0608	0,0036985353
3	0,262	0,2775	-0,0154	0,0002378405
4	0,202	0,2428	-0,0410	0,0016795094
5	0,157	0,1002	0,0569	0,0032424890
6	0,444	0,4389	0,0050	0,0000249611
7	0,426	0,4272	-0,0010	0,0000009336
8	0,481	0,4802	0,0011	0,0000012071
9	0,444	0,4477	-0,0038	0,0000144696
10	0,310	0,2884	0,0215	0,0004609665
11	0,553	0,5433	0,0097	0,0000938382
12	0,442	0,3972	0,0446	0,0019907821
			Total	0,0120549811
			MSE	0,0010045818

Data Testing					
No	**Target**	**ANN 4-5-25-1**			
		Output	**Error**	**SSE**	**Prediction**
1	0,211	0,3177	-0,1065	0,0113488899	True
2	0,116	0,4429	-0,3273	0,1071354918	False
3	0,202	0,5210	-0,3192	0,1018770331	False
4	0,277	0,3110	-0,0344	0,0011817522	True
5	0,410	0,2428	0,1668	0,0278257061	False
6	0,216	0,1917	0,0247	0,0006082950	True
7	0,145	0,1912	-0,0465	0,0021645454	True
8	0,269	0,2364	0,0330	0,0010857453	True
9	0,303	0,2191	0,0835	0,0069718162	True
10	0,277	0,2841	-0,0075	0,0000558999	True
11	0,476	0,5153	-0,0392	0,0015363346	True
12	0,457	0,4291	0,0283	0,0008010370	True
			Total	0,2625925465	**75**
			MSE	0,0218827122	

4. Conclusion

Based on the results of the research conducted, conclusions can be drawn including:

a) Artificial neural network with backpropogation method can be applied to predict rainfall in Pematangsiantar city by using data from theBadan Pusat Statistic of Pematangsiantar city with the subject of rainfall data (2006-2015).
b) By using 6 models (4-5-1, 4-10-1, 4-25-1, 4-5-10-1, 4-5-25-1 and 4-5-50-1) architectures that are trained and tested using the Matlab 6.1 application software, the results obtained that the 4-5-25-1 architectural model is the best model for making predictions with 75% accuracy, MSE Training 0.001004582, MSE Testing 0.021882712 and Epoch 59.076.
c) From a series of model trials, adding hidden is not a solution in increasing prediction accuracy. This is evidenced by the results obtained by architectural models 12-25-1 have the least accuracy compared to architectural models 12-5-1 and 12-10-1.

References

[1] A. P. Windarto, "Penerapan Datamining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering Method," *Techno.Com*, vol. 16, no. 4, pp. 348–357, 2017.
[2] A. P. Windarto, "Implementation of Data Mining on Rice Imports by Major Country of Origin Using Algorithm Using K-Means Clustering Method," *Int. J. Artif. Intell. Res.*, vol. 1, no. 2, pp. 26–33, 2017.
[3] A. P. Windarto, "Penerapan Data Mining Pada Ekspor Buah-Buahan Menurut Negara Tujuan Menggunakan K-Means Clustering," *Techno.COM*, vol. 16, no. 4, pp. 348–357, 2017.
[4] D. Nabila Batubara, "Analisis Metode K-MEANS Pada Pengelompokan Keberadaan Area Resapan Air Menurut Provinsi," *Semin. Nas. Sains Teknol. Inf.*, pp. 345 – 349, Jul. 2019.
[5] C. Astria, "Metode K-Means Pada Pengelompokan Wilayah Pendistribusian Listrik," *Semin. Nas. Sains Teknol. Inf.*, pp. 306 – 312, Jul. 2019.
[6] T. Budiharjo, Soemartono, T., Windarto, A.P., Herawan, "Predicting tuition fee payment problem using backpropagation neural network model," *Int. J. Adv. Sci. Technol.*, 2018.
[7] T. Budiharjo, Soemartono, T., Windarto, A.P., Herawan, "Predicting school participation in indonesia using back-propagation algorithm model," *Int. J. Control Autom.*, 2018.
[8] A. P. Windarto, M. R. Lubis, and Solikhun, "Implementasi JST Pada Prediksi

Total Laba Rugi Komprehensif Bank Umum Konvensional Dengan Backpropagation," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 5, no. 4, pp. 411–418, 2018.

[9] A. P. Windarto, M. R. Lubis, and Solikhun, "Model Arsitektur Neural Network Dengan Backpropogation Pada Prediksi Total Laba Rugi Komprehensif Bank Umum Konvensional," *Kumpul. J. Ilmu Komput.*, vol. 5, no. 2, pp. 147–158, 2018.

[10] A. P. Windarto, L. S. Dewi, and D. Hartama, "Implementation of Artificial Intelligence in Predicting the Value of Indonesian Oil and Gas Exports With BP Algorithm," *Int. J. Recent Trends Eng. Res.*, vol. 3, no. 10, pp. 1–12, 2017.

[11] Sumijan, A. P. Windarto, A. Muhammad, and Budiharjo, "Implementation of Neural Networks in Predicting the Understanding Level of Students Subject," *Int. J. Softw. Eng. Its Appl.*, vol. 10, no. 10, pp. 189–204, 2016.

[12] S. Putra Siregar and A. Wanto, "Analysis Accuracy of Artificial Neural Network Using Backpropagation Algorithm In Predicting Process (Forecasting)," *Int. J. Inf. Syst. Technol.*, vol. 1, no. 1, pp. 34–42, 2017.

[13] Solikhun, A. P. Windarto, Handrizal, and M.Fauzan, "Jaringan Saraf Tiruan Dalam Memprediksi Sukuk Negara Ritel Berdasarkan Kelompok Profesi Dengan Backpropogation Dalam Mendorong Laju Pertumbuhan Ekonomi," *Kumpul. J. Ilmu Komput.*, vol. 4, no. 2, pp. 184–197, 2017.

[14] B. Febriadi, Z. Zamzami, Y. Yunefri, and A. Wanto, "Bipolar function in backpropagation algorithm in predicting Indonesia's coal exports by major destination countries," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 420, no. 12089, pp. 1–9, 2018.

[15] B. Fachri, A. P. Windarto, and I. Parinduri, "Penerapan Backpropagation dan Analisis Sensitivitas pada Prediksi Indikator Terpenting Perusahaan Listrik," *J. Edukasi dan Penelit. Inform.*, vol. 5, no. 2, pp. 202–208, 2019.

Authors

1st Author
Yogi Prayoga
STIKOM Tunas Bangsa, Pematangsiantar, North Sumatra