International Journal of Control and Automation
Vol. 11, No. 11 (2018), pp.57-68
http//dx.doi.org/10.14257/ijca.2018.11.11.05

Predicting School Participation in Indonesia using Back-Propagation Algorithm Model

Budiharjo[1], Triyuni Soemartono[1], Agus Perdana Windarto[2] and Tutut Herawan[3,4,5]

[1]Universitas Prof Dr Moestopo (Beragama), Jakarta, Indonesia
[2]STIKOM Tunas Bangsa, Pematangsiantar, Indonesia
[3]Universitas Negeri Yogyakarta, Indonesia
[4]Universitas Teknologi Yogyakarta, Indonesia
[5]Politeknik Negeri Malang, Indonesia
budiharjo@dsn.moestopo.ac.id, triyuni.soemartono@yahoo.com

Abstract

School participation rate has been known as one of the indicators of successful development of education services in Indonesia. The higher the school participation rate, the region is considered the most successful in providing access education services. This paper presents a study on predicting school participation in Indonesia. It is based on back-propagation algorithm model of artificial neural network. The data used is data from the National Statistics Agency of Republic Indonesia. These data are school participation rate of 2003 to 2017 that consists of 7-12 years old, 13-15 years old, 16-18 years old and 19-24 years old. The input variables used for the training are the school participation rate of 2004-2009 and as the target used for training school participation rate of 2010. Meanwhile the input variables used for the test are school participation rate on 2011-2016 and as a target used for testing school participation rate of 2017. The architectural model of training and testing are 4 architectures i.e. 6-2-1, 6-8-1, 6-8-2-1 and 6-2-8-1. The output generated is the best pattern of the ANN architecture. The best architectural model is 6-2-1 with epoch 1929, MSE 0.0051 and accuracy rate is 75%. The results show that, the prediction of school participation rate of 2018 for age 7-18 is above 88% and for age 19-24 is less than 14%.

***Keywords**: School Participation Rate; ANN; Backpropagation; Prediction*

1. Introduction

The Artificial Neural Networks (ANN) are frequently used to model complicated nonlinear processes. In theory, Multi-Layer Perceptron (MLP) neural networks can be used for non-linear functions and therefore the Backpropagation (BP) algorithms are common for training MLP neural networks. However, the BP algorithms have drawbacks (*e.g.*, depending on the choice of the initial weights; number of hidden neurons; having very slow convergence rapids, being sensitive to noises in the training data sets; and having poor generalization for complicated nonlinear functions [1]. School Participation Rate is known as one of the indicators of successful development of education services in a region of either Province, Regency or City in Indonesia. The higher the value of School Participation Rate (SPR), then the area is considered successful in providing access to education services [2].

The high School Participation Rate indicates greater opportunities for access to education in general. In the age group where the opportunity occurs can be seen from the amount of School Participation Rate in each age group [2]. School Participation Rate is a measure of the absorptive capacity of educational institutions towards school-age

Received (August 23, 2018), Review Result (October 15, 2018), Accepted (October 22, 2018)

ISSN: 2005-4297 IJCA

population. School Participation Rate is a basic indicator used to assess access of the population to educational facilities, especially for school-aged residents in a region. The higher the School Participation Rate, the greater the number of people who have the opportunity to receive education. From the above description the government must know the full School Participation Rate from all provinces in Indonesia in order to take the right policy in increasing the next School Participation Rate next year.

In order to find out the School Participation Rate, it is necessary to make an in-depth study on the prediction of School Participation Rate. Therefore, this work presents prediction method of School Participation Rate based on artificial neural network with Back-propagation method, where this work conducts training and testing to get the best prediction model.

The rest of this paper is organized as follow. Section 2 presents theoretical background on artificial intelligence. Section 3 present the proposed method for prediction. Section 4 presents obtainde results and following by discussion. Finally, Section 5 concludes this work.

2. Rudimentary

2.1. Artificial Intelligence

Artificial Intelligence (AI) is a field of study based on the premise that intelligent thought can be regarded as a form of computation - one that can be formalized and ultimately mechanized. To achieve this, however, two major issues need to be addressed. The first issue is knowledge representation, and the second is knowledge manipulation [2]. The main aim of Artificial Intelligence (AI) is to study how to build artificial systems that perform tasks normally performed by human beings. This concept was introduced in 1956 in the Darthmounth conference. From that moment on a lot of effort has been made and many goals have been achieved but unfortunately many failures as well. Today, the AI is a very important discipline and it includes a number of well-recognized and mature areas including Expert Systems [3,4,5], Fuzzy Logic [6,7,8,9,10], Genetic Algorithms [11,12,13], Language Processing, Logic Programming, Planning and Scheduling, Neural Networks and Robotics [13]. The general problem of simulating intelligence has been simplified to specific sub-problems which have certain characteristics or capabilities that an intelligent system should exhibit. The following characteristics have received the most attention:

a. Deduction, reasoning, problem solving (embodied agents, neural networks, statistical approaches to AI);
b. Knowledge representation (ontologies);
c. Planning (multi-agent planning and cooperation);
d. Learning (machine learning);
e. Natural Language Processing (information retrieval – text mining, machine translation);
f. Motion and Manipulation (navigation, localization, mapping, motion planning);
g. Perception (speech recognition, facial, recognition, object recognition);
h. Social Intelligence (empathy simulation);
i. Creativity (artificial intuition, artificial imagination); and
j. General Intelligence (Strong AI).

Classical AI approaches focus on individual human behavior, knowledge representation and inference methods. Distributed Artificial Intelligence (DAI), on the other hand, focuses on social behavior, *i.e.*, cooperation, interaction and knowledge-sharing among different units (agents). The process of finding a solution in distributed resolution problems relies on sharing knowledge about the problem and cooperation among agents. It was from these concepts that the idea of intelligent multi-agent technology emerged. An agent is an autonomous cognitive entity which understands its environment, *i.e.*, it can work by itself and it has an internal decision-making system that acts globally around other agents. In multi-agent systems, a group of mobile autonomous agents cooperate in a coordinated and intelligent manner in order to solve a specific problem or classes of problems [14].

2.2. Artificial Neural Networks

Artificial Neural Network (ANN) is a computational model, which is based on Biological Neural Network. Artificial Neural Network is often called as Neural Network (NN) can be depicted in Figure 1 as follow.

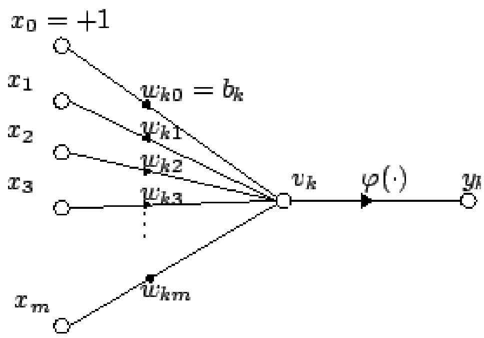

Figure 1. ANN

From Figure 1 above, to build artificial neural network, artificial neurons, also called as nodes, are interconnected [15,16]. The architecture of NN is very important for performing a particular computation. Some neurons are arranged to take inputs from outside environment. These neurons are not connected with each other, so the arrangement of these neurons is in a layer, called as Input layer. All the neurons of input layer are producing some output, which is the input to next layer. The architecture of NN can be of single layer or multilayer. In a single layer Neural Network, only one input layer and one output layer is there, while in multilayer neural network, there can be one or more hidden layer. An artificial neuron is an abstraction of biological neurons and the basic unit in an ANN [17,18]. The Artificial Neuron receives one or more inputs and sums them to produce an output. Usually the sums of each node are weighted, and the sum is passed through a function known as an activation or transfer function. The objective here is to develop a data classification algorithm that will be used as a general-purpose classifier. To classify any database first, it is required to train the model. The proposed training algorithm used here is a Hybrid BP-GA [19,20]. After successful training, user can give unlabeled data to be classified.

The synapses or connecting links: that provide weights, w_j, to the input values, x_j for j = 1,...,m. An additional function that sums the weighted input values to compute the input to the activation function as follow:

$$V = W0 + \sum_{j=1}^{m} W\, jXj \tag{1}$$

where, w0 is called the bias, is a numerical value associated with the neuron. It is convenient to think of the bias as the weight for an input x0 whose value is always equal to one, so that;

$$V = \sum_{j=1}^{m} W\, jXj \tag{2}$$

An activation function g: that maps *v* to *g*(*v*) the output value of the neuron. This function is a monotone function. The logistic (also called the sigmoid) function *g*(*v*) = (ev/(1+ev)) as the activation function works best. The practical value of the logistic function arises from the fact that it is almost linear in the range where g is between 0.1 and 0.9 but has a squashing effect on very small or very large values [21].

2.3. Architecture of Back-propagation

The back-propagation learning algorithm (BPLA)has become famous learning algorithms among ANNs. In the learning process, to reduce the in accuracy of ANNs, BPLAs use the gradient-decent search method to adjust the connection weights. The structure of a back-propagation ANN is shown in Figure 2.

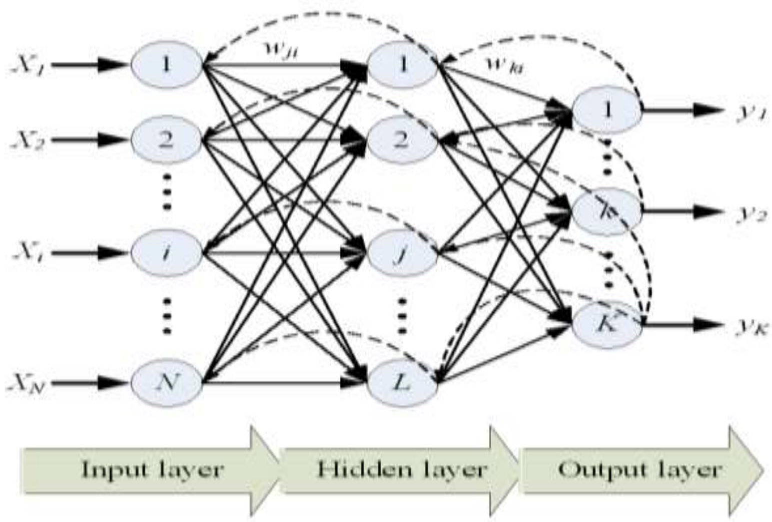

Figure 2. Back-propagation ANN

The output of each neuron is the aggregation of the numbers of neurons of the previous level multiplied by its corresponding weights. The input values are converted into output signals with the calculations of activation functions. Back-propagation ANNs have been widely and successfully applied in diverse applications, such as pattern recognition, location selection and performance evaluations [22]. There are several algorithms that can be used to create an artificial neural network, but the Back propagation was chosen because it is probably the easiest to implement, while preserving efficiency of the network. Backward Propagation Artificial Neural Network (ANN) use more than one input layers (usually 3). Each of these layers must be either of the following:

a. Input Layer – This layer holds the input for the network
b. Output Layer – This layer holds the output data, usually an identifier for the input.
c. Hidden Layer – This layer comes between the input layer and the output layer. They serve as a propagation point for sending data from the previous layer to the next layer [23].

2.4. Back-propagation Neural Network

Back-propagation is the most widely used algorithm for supervised learning with multi-layered feed-forward networks. The basic idea of the Back-propagation learning algorithm [24]. *Phases in Back-propagation Technique* algorithm can be divided into two phases: propagation and weight update.

Phase 1: Propagation
Each propagation involves the following steps:

1. Forward propagation of a training pattern's input is given through the neural network in order to generate the propagation's output activations.
2. Back propagation of the output activations propagation through the neural network using the training pattern's target in order to generate the deltas of all output and hidden neurons.

Phase 2: Weight Update
For each weight-synapse:

1. Multiply its input activation and output delta to get the gradient of the weight.
2. Bring the weight in the direction of the gradient by adding a ratio of it from the weight.

This ratio impacts on the speed and quality of learning; it is called the learning rate. The sign of the gradient of a weight designates where the error is increasing; this is why the weight must be updated in the opposite direction. The phases 1 and 2 are repeated until the performance of the network is satisfactory[25].

2.5. Evaluating the Performance of the Models

The main measures used for evaluating the performance of machine learning techniques for predicting the software effort are as follows[26]:

a. Sum Squared Error (SSE)

The sum squared error is defined as.

$$\sum_{i=1}^{n} (\text{Pi-Ai})^2 \tag{3}$$

where

Pi = Estimated value for data point i;
Ai = Actual value for the data point i;
n = Total number of data points.

b. Mean Squared Error (MSE)

The mean squared error is defined as.

$$\frac{1}{n}\sum_{i=1}^{n} (\text{Pi-Ai})^2 \tag{4}$$

where

Pi = Estimated value for data point i;
Ai =Actual value for the data point i;
n = Total number of data points.

3. Result and Discussion

3.1. System Planning

School Participation Rate data will be processed by Artificial Neural Network with back-propagation method. In order for data to be recognized by Artificial Neural Networks, the data should be represented in numerical form between 0 and 1, the both variables and their contents which are input data of School Participation Rate as pattern and output recognition which is the prediction of School Participation Rate obtained from architectural model best when determining the best pattern. This is because the network uses the binary sigmoid activation function (logsig) that ranges from 0 to 1. The values used are obtained based on the categories of each variable as well as to make it easier to remember in defining it.

The previous year's School Participation Variables (2004-2009) and (2011-2016) are the criteria to be used in predicting the future School Participation Rate which consists of 7-12 years old, 13- 15 years old, 16-18 years old year and 19-24 years old using Artificial Neural Network. Variables determined by looking at the dependence of data on research conducted. The criteria used are based on National Statistics Agency data from URL website: www.bps.go.id. The list of variables in predicting School Participation Rate is shown in Tables 1 and 2 as follow:

Table 1. School Input Prediction Training Prediction List

No	Variable	The year of School Participation Rate
1	X1	2004
2	X2	2005
3	X3	2006
4	X4	2007
5	X5	2008
6	X6	2009

Table 2. Predicted Testing List of School Participation Rate

No	Variable	The year of School Participation Rate
1	X1	2011
2	X2	2012
3	X3	2013
4	X4	2014
5	X5	2015
6	X6	2016

The input data is obtained from the National Statistics Agency website. School Participation Rate data from 2004 to 2017 are grouped by age. The sample data that used is the School Participation Rate data for 2004, from 7-12, 13-15 years, 16-18 years and 19-24 years old consisting of 6 training input data and 6 test data and 1 training target data and 1 test target data. This data will be transformed to a data between 0 and 1 before the training and testing using Artificial Neural Network Back-propagation method with the formula:

$$x' = \frac{0.8(x-a)}{b-a} + 0.1 \tag{4}$$

The target data is School Participation Rate of 2010 for training and School Participation Rate for 2017 for testing.

3.2. Data Processing

Data processing is running in Matlab software version 6.1. Sample data of School Participation Rate is grouped by age. This data will be used in training data and testing data. Sample data that has been processed and transformed is presented in Tables 3-6 as follows.

Table 3. Sample Prediction Training Data of School Participation Rate

No	Nama	Variabel						Target
		X1	X2	X3	X4	X5	X6	
1	Age 7-12	96,77	97,14	97,39	97,64	97,88	97,95	98,02
2	Age 13-15	83,49	84,02	84,08	84,65	84,89	85,47	86,24
3	Age 16-18	53,48	53,86	53,92	55,49	55,5	55,16	56,01
4	Age 19-24	12,07	12,23	11,38	13,08	13,29	12,72	13,77

Table 4. Transformed sample training data

No	Name	Variable						Target
		X1	X2	X3	X4	X5	X6	
1	Age 7-12	0.8885	0.8919	0.8942	0.8965	0.8987	0.8994	0.9000
2	Age 13-15	0.7658	0.7707	0.7713	0.7765	0.7788	0.7841	0.7912
3	Age 16-18	0.4887	0.4922	0.4928	0.5073	0.5074	0.5042	0.5121
4	Age 19-24	0.1064	0.1078	0.1000	0.1157	0.1176	0.1124	0.1221

Table 5. Sample prediction training data of School Participation Rate

No	Name	Variable						Target
		X1	X2	X3	X4	X5	X6	
1	Age 7- 12	97.62	98.02	98.42	98.92	99.09	99.09	99.14
2	Age 13-15	87.99	89.76	90.81	94.44	94.72	94.88	95.08
3	Age 16-18	57.95	61.49	63.84	70.31	70.61	70.83	71.42
4	Age 19-24	14.82	16.05	20.14	22.82	22.95	23.93	24.77

Table 6. Transformed sample training data

No	Name	Variable						Target
		X1	X2	X3	X4	X5	X6	
1	Age 7- 12	0.8856	0.8894	0.8932	0.8979	0.8995	0.8995	0.9000
2	Age 13-15	0.7942	0.8110	0.8210	0.8554	0.8581	0.8596	0.8615
3	Age 16-18	0.5092	0.5428	0.5651	0.6265	0.6293	0.6314	0.6370
4	Age 19-24	0.1000	0.1117	0.1505	0.1759	0.1771	0.1864	0.1944

3.3. Design of Artificial Neural Network Architecture

Network used to predict School Participation Rate with backpropagation with feedforward learning step. This network has several layers, they are the input layer (input), output layer (output) and some hidden layer (hidden). The hidden layer helps the network to recognize more input patterns compared to networks that do not have a hidden layer. The parameters in Back-propagation network formation use 6 input variables, 1 or more hidden layers and 1 output layer. The architectural model used to get the best architecture is 6-2-1, 6-8-1, 6-2-8-1 and 6-8-2-1. Neural Network to be built is back propagation algorithm with sigmoid activation function. The activation function in the

Artificial Neural Network is used to process the calculation of the actual value of the output on the hidden layer and calculate the actual value of the output on the output layer.

3.4. Defining Output

The expected result at this stage is the detection pattern determines the best value for predicting School Participating Rate. Test results are as follows:

a. To know the prediction of School Partitioning Rate is based on the results of the previous year's School Enrollment Rate. The output of these predictions is the best architectural pattern in predicting School Participation Rate for 7-12, 13-15, 16-18 and 19-24 years old by looking at minimum errors.
b. Categorization Output training and testing
   The category for output is determined by the minimum error rate of the target. Restrictions on these categories are listed in the following Table 7:

Table 7. Categorized Data

No	Description	Error Minimum
1	True	0.01 - 0.001
2	False	> 0.01

3.5. Artificial Neural Network Architecture Design

Design of artificial neural network architecture for training and test data, then used 6 input variables for training and 6 input variables for the test are grouped by age that are: (1) for 7-12 years old; (2) for 13-15 years old; (3) for 16-18 years old; and (4) for 19-24 years old. The following steps will be performed in the user back propagation algorithm with sigmoid activation function. Stages that must be done is as follows:

a. Initialization, is the stage where the variable values will be set or defined first, such as: the value of input data, weight, expected output value, learning rate and other data values.
b. Activation is the process of calculating the actual value of output on the hidden layer and calculate the actual output value of the output layer.
c. Weight Training, is the process of calculating the value of the gradient error on the output layer and calculating the value of the gradient error on the hidden layer
d. Iteration, is the final stage in the test, where if still the minimum error that is not expected to be found then back in the activation stage (activation).

3.5.1. Training and Testing Architecture 6-2-1

Here are the test results with 4 test data with a 6-2-1 test pattern. Test results data and Training can be seen in the Table 8 as follow:

Table 8. Training Results and Testing by Model 6-2-1

Training					**Testing**				
No	**Target**	**Output**	**Error**	**SSE**	**No**	**Target**	**Output**	**Error**	**SSE**
1	0.9000	0.8784	0.0216	0.0004665600	1	0.9000	0.8784	0.0216	0.0004665600
2	0.7912	0.8776	-0.0864	0.0074649600	2	0.8615	0.8785	-0.0170	0.0002890000
3	0.5121	0.3772	0.1349	0.0181980100	3	0.6370	0.7220	-0.0850	0.0072250000
4	0.1221	0.0813	0.0408	0.0016646400	4	0.1944	0.0829	0.1115	0.0124322500
Total				0.0277941700	Total				0.0204128100
MSE				0.0069485425	MSE				0.0051032025
					Accuracy of Truth **(%)**				**75 %**

3.5.2. Training and Testing Architecture 6-8-1

Here are the test results with 4 test data with 4-5-1 test pattern. Test results data and Training can be seen in the table as follows:

Table 9. Training Results and Testing by Model 6-8-1

Training					Testing				
No	**Target**	**Output**	**Error**	**SSE**	**No**	**Target**	**Output**	**Error**	**SSE**
1	0.9000	0.7746	0.1254	0.0157251600	1	0.9000	0.7766	0.1234	0.0152275600
2	0.7912	0.7938	-0.0026	0.0000067600	2	0.8615	0.8076	0.0539	0.0029052100
3	0.5121	0.6511	-0.1390	0.0193210000	3	0.6370	0.7345	-0.0975	0.0095062500
4	0.1221	0.2414	-0.1193	0.0142324900	4	0.1944	0.2830	-0.0886	0.0078499600
Total				0.0492854100	Total				0.0354889800
MSE				0.0123213525	MSE				0.0088722450
					Accuracy of Truth **(%)**				**75 %**

3.5.3. Training and Testing Architecture 6-2-8-1

Here are the test results with 4 test data with 6-2-8-1 test pattern. Test results data and Training can be seen in the table as follows:

Table 10. Training Results and Testing with Model 6-2-8-1

Training					Testing				
No	**Target**	**Output**	**Error**	**SSE**	**No**	**Target**	**Output**	**Error**	**SSE**
1	0.9000	0.7242	0.1758	0.0309056400	1	0.9000	0.7241	0.1759	0.0309408100
2	0.7912	0.8492	-0.0580	0.0033640000	2	0.8615	0.7567	0.1048	0.0109830400
3	0.5121	0.5448	-0.0327	0.0010692900	3	0.6370	0.6272	0.0098	0.0000960400
4	0.1221	0.1961	-0.0740	0.0054760000	4	0.1944	0.1731	0.0213	0.0004536900
Total				0.0408149300	Total				0.0424735800
MSE				0.0102037325	MSE				0.0106183950
					Accuracy of Truth **(%)**				**50%**

3.5.4. Training and Testing Architecture 6-8-2-1

Here are the test results with 4 test data with 6-8-2-1 test pattern. Test results data and Training can be seen in the table as follows:

Table 11. Training Results and Testing by Model 6-8-2-1

Training					**Testing**				
No	**Target**	**No**	**Target**	**No**	**Target**	**No**	**Target**	**No**	**Target**
1	0.9000	0.7631	0.1369	0.0187416100	1	0.9000	0.7634	0.1366	0.0186595600
2	0.7912	0.7830	0.0082	0.0000672400	2	0.8615	0.7800	0.0815	0.0066422500
3	0.5121	0.5029	0.0092	0.0000846400	3	0.6370	0.6156	0.0214	0.0004579600
4	0.1221	0.2594	-0.1373	0.0188512900	4	0.1944	0.2446	-0.0502	0.0025200400
Total				0.0377447800	Total				0.0282798100
MSE				0.0094361950	MSE				0.0070699525
					Accuracy of Truth **(%)**				**75%**

3.5.5. Selection of Best Architecture of Artificial Neural Networks

The result of Matlab 6.1 application software used for architectural model 6-2-1, architecture 6-8-1, architecture 6-2-8-1 and architecture 6-8-2-1 is obtained the best architectural pattern. From this pattern will be used to predict the School Participation Rate. Assessment of the best architectural model is seen from several aspects such as epoch, minimum error and accuracy of truth. The details results from this part can be presented in the following Table 12:

Table 12. Recapitulation of Architectural Model

Model	6-2-1	6-8-1	6-2-8-1	6-8-2-1
Epochs	1929	451	416	1054
MSE	0.0051032025	0.0088722450	0.0106183950	0.0070699525
Akurasi	75%	75%	50%	75%

From Table 11 above, it can be seen that the best architectural model that will be used to make predictions from a series of model trials is 6-2-1 with epoch 1929, MSE 0.0051032025 and accuracy rate is 75%.

3.5.6. Prediction of School Enrollment Rate by age category

The last stage is the process of predicting School Participation Rate by age category. This stage is done by using the best architects by entering the data of Participation Rate of previous year as input then we will get the next School Participation Rate. The formula used to predict total comprehensive income is 6-2-1 architectural model is:

$$x = ((x - 0.1)\ (x.\text{max}-\text{xmin}) / 0.8)) + x.\text{min} \tag{5}$$

where
x' = Normalization Data
$x.\text{max}$ = Maximum Original Data

x.min = Minimum Original Data

For more details please note the following Table 13.

Table 13. Predicted School Participation Figures with Model 6-2-1

No	Group	Year	Predicted School Participation Rate	Normalization (Y actual)	e	e2
1	Age 7-12	2018	96.86	0.8784	0.0216	0.0004665600
2	Age 13-15		96.94	0.8791	-0.0879	0.0077264100
3	Age 16-18		88.26	0.7968	-0.2456	0.0603193600
4	Age 19-24		13.15	0.0842	0.0379	0.0014364100
					MSE	0.0174871850
					Accuracy	75%

4. Conclusion

In this paper, we have presented a study on predicting school participation in Indonesia using back-propagation algorithm model. Based on the results and discussion above, the following conclusions can be drawn: (1) Adding a lot of hidden layers during training and testing is not a maximum result.; (2) After an experiment in the process of training and system testing conducted using Matlab 6.1 Application of Neural Network Model used is 6-2-1, model 6-8-1, model 6-2-8-1 and model 6 -8-2-1 can be obtained good results by looking at MSE the smallest test is 6-2-1; (3) With architectural model 6-2-1, it can predict School Participation Rate in Indonesia by showing 75% accuracy; (4) To improve the accuracy level, more complex training data is required.

Acknowledgement

This research is supported by Universitas Prof Dr Moestopo (Beragama) research grant.

References

[1] M. I. El Adawy, M. E. Aboul-Wafa, H. A. Keshk, and M. M. El Tayeb, "A SOFT-Back-propagation algorithm for training neural networks", In Proceedings of the Nineteenth National Radio Science Conference (NRSC'02), 2002, 397–404.

[2] Khalid, Mohd Sazali, et al. "Teaching and Learning Using Computers: How Should We Tread on Its' Changing Technology?." International Journal of Emerging Technologies in Learning 9.5 (2014).

[3] Chiroma, H., Abdulkareem, S., Abubakar, A.I. and Herawan, T., 2014. Kernel functions for the support vector machine: comparing performances on crude oil price data. In Recent Advances on Soft Computing and Data Mining (273-281). Springer International Publishing.

[4] Hakim, R.F., Sari, E.N. and Herawan, T., 2014. Soft Solution of Soft Set Theory for Recommendation in Decision Making. In Recent Advances on Soft Computing and Data Mining (313-324). Springer International Publishing.

[5] Lasisi, A., Ghazali, R. and Herawan, T., 2014. Comparative performance analysis of negative selection algorithm with immune and classification algorithms. In Recent Advances on Soft Computing and Data Mining (441-452). Springer International Publishing.

[6] Handaga, B., Herawan, T. and Deris, M.M., 2012. FSSC: An Algorithm for Classifying Numerical Data Using Fuzzy Soft Set Theory. International Journal of Fuzzy System Applications (IJFSA), 2(4), pp.29-46.

[7] Herawan, T., Abdullah, Z., Chiroma, H., Sari, E.N., Ghazali, R. and Nawi, N.M., 2014, September. Cauchy criterion for the Henstock-Kurzweil integrability of fuzzy number-valued functions. In Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on (1329-1333). IEEE.

[8] Ma, X., Qin, H., Sulaiman, N., Herawan, T. and Abawajy, J.H., 2014. The parameter reduction of the interval-valued fuzzy soft sets and its related algorithms. IEEE Transactions on Fuzzy Systems, 22(1), pp.57-71.
[9] Yanto, I.T.R., Ismail, M.A. and Herawan, T., 2016. A modified Fuzzy k-Partition based on indiscernibility relation for categorical data clustering. Engineering Applications of Artificial Intelligence, 53, pp.41-52.
[10] Abubakar, A.I., Zeki, A., Chiroma, H. and Herawan, T., 2014. Investigating Rendering Speed and Download Rate of Three-Dimension (3D) Mobile Map Intended for Navigation Aid Using Genetic Algorithm. In Recent Advances on Soft Computing and Data Mining (261-271). Springer International Publishing.
[11] Qin, H., Ma, X., Herawan, T. and Zain, J.M., 2012, May. An improved genetic clustering algorithm for categorical data. In Pacific-Asia Conference on Knowledge Discovery and Data Mining (100-111). Springer Berlin Heidelberg.
[12] Shah, H., Ghazali, R., Nawi, N.M., Deris, M.M. and Herawan, T., 2013. Global artificial bee colony-Levenberq-Marquardt (GABC-LM) algorithm for classification. International Journal of Applied Evolutionary Computation (IJAEC), 4(3), pp.58-74.
[13] Castaño, B., Moreno, Á., Carbajo, M. and de Pedro, J., 2008. Artificial Intelligence and Bluetooth Techniques in a Multi-user M-learning Domain. IJCSA, 5(1), pp.1-13.
[14] Dilek, S., Çakır, H. and Aydın, M., 2015. Applications of Artificial Intelligence Techniques to Combating Cyber Crimes: A Review. arXiv preprint arXiv:1502.03552.
[15] Abubakar, A.I., Khan, A., Nawi, N.M., Rehman, M.Z., Wah, T.Y., Chiroma, H. and Herawan, T., 2016. Studying the Effect of Training Levenberg Marquardt Neural Network by Using Hybrid Meta-Heuristic Algorithms. Journal of Computational and Theoretical Nanoscience, 13(1), pp.450-460.
[16] Chiroma, H., Abdul-kareem, S., Ibrahim, U., Ahmad, I.G., Garba, A., Abubakar, A., Hamza, M.F. and Herawan, T., 2015. Malaria severity classification through Jordan-Elman neural network based on features extracted from thick blood smear. Neural Network World, 25(5), p.565.
[17] Husaini, N.A., Ghazali, R., Nawi, N.M., Ismail, L.H., Deris, M.M. and Herawan, T., 2014. Pi-Sigma Neural Network For A One-Step-Ahead Temperature Forecasting. International Journal of Computational Intelligence and Applications, 13(04), p.1450023.
[18] Nawi, N.M., Rehman, M.Z., Aziz, M.A., Herawan, T. and Abawajy, J.H., 2014, November. Neural network training by hybrid accelerated cuckoo particle swarm optimization algorithm. In International Conference on Neural Information Processing (237-244). Springer International Publishing.
[19] Nawi, N.M., Rehman, M.Z., Aziz, M.A., Herawan, T. and Abawajy, J.H., 2014, November. An Accelerated Particle Swarm Optimization Based Levenberg Marquardt Back Propagation Algorithm. In International Conference on Neural Information Processing (245-253). Springer International Publishing.
[20] Chiroma, H., Abdul-Kareem, S., Muaz, S.A., Khan, A., Sari, E.N. and Herawan, T., 2014, October. Neural Network Intelligent Learning Algorithm for Inter-related Energy Products Applications. In International Conference in Swarm Intelligence (284-293). Springer International Publishing.
[21] Ganatra, A., Kosta, Y.P., Panchal, G. and Gajjar, C., 2011. Initial classification through back propagation in a neural network following optimization through GA to evaluate the fitness of an algorithm. International Journal of Computer Science and Information Technology, 3(1), pp.98-116.
[22] Che, Z.G., Chiang, T.A. and Che, Z.H., 2011. Feed-forward neural networks training: A comparison between genetic algorithm and back-propagation learning algorithm. International Journal of Innovative Computing, Information and Control, 7(10), pp.5839-5850.
[23] Choudhary, N.Y., Patil, M.R., Bhadade, U. and Chaudhari, B.M., 2013. Signature Recognition & Verification System Using Back Propagation Neural Network. International Jorunal of IT, Engineering and Applied Sciences Research (IJIEASR), 2(1), pp.1-8.
[24] D. E. Rumelhart and J. McClelland. Parallel Distributed Processing. 1986.
[25] Baboo, S.S. and Shereef, I.K., 2010. An efficient weather forecasting system using artificial neural network. International journal of environmental science and development, 1(4), p.321.
[26] 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. 189–204, 2016.

