Evaluation of Backpropagation Neural Network Models for Early Prediction of Student's Graduation in XYZ University

Ainul Yaqin[a,*], Arif Dwi Laksito[a], Siti Fatonah[a]

[a]Computer Science Faculty, University of AMIKOM Yogyakarta, Jl. Ringroad Utara, Yogyakarta, 55283, Indonesia
Corresponding author: [*]ainulyaqin@amikom.ac.id

Abstract— **The study period of the student in a tertiary institution is undoubtedly essential in implementing the objectives of the tertiary institution, particularly for the implementation of the study program, so that its outcomes will affect accreditation. Prediction of students' study period can be a reference for higher education institutions in making policies for the future. Based on XYZ University data, especially in the informatics study program, many students have the different generation and concentration therein. In the implementation of students in studying, several factors, including the value of the Grade Point Average (GPA), can affect the study period taken. Likewise, the institutions often do not understand the conditions or predictive value of students' study period on campus. The application of neural networks in predicting the students' study period at the XYZ University uses a network model with GPA values as input and 1 layer of hidden layers with 10, 50 and 100 neurons; learning rate values used are 0.01, 0.1 and 0.3 and 1 output target for the study period. Prediction results obtained the best results on the neuron network pattern 50 with 0.01 as a learning rate, which detail of MSE value, the training is 0,017516 and the testing is 0,047721, with an accuracy value of 77%.**

Keywords—**Prediction; study period; neural network; informatics.**

I. INTRODUCTION

There are three faculties at XYZ University: a faculty of computer science with 6 study programs, a faculty of economics & social with 6 study programs in it, and a faculty of science & technology with 3 study programs. XYZ University also has a postgraduate graduate program in Master of Informatics Engineering.

In Minister of Higher Education Decree No. 44 of 2015 concerning SN DIKTI Indonesia, the period of study of students in a tertiary institution is called the standard of study period based on references from various universities, as measured by the length of study. If the pre-college education is carried out for 12 years, then for a bachelor the length of study is four years", he explained," if the secondary education is 13 years, then bachelors must be a minimum of 3 years ".

Based on the Study Program Accreditation Instrument (SPAI) 4.0, the main point emphasizes the campus accreditation to the assessment instruments of graduates' results. Indeed, the study period is critical in implementing higher education management, especially for the implementation of the study programs. Then, the outcome can affect its accreditation. Accordingly, the students' study period's prediction can be a reference for higher education institutions in making policies for the future. Based on XYZ University data, especially in informatics study programs, many students have different generations and concentrations. In the implementation, students in studying, several factors can influence the study period taken. Beg also, the institution often do not understand the conditions or predictive period of student studies. Prediction of the study period in Haryati's research is based on GPA and Number of semester credit units (credits) with a load rule of at least 144 credits that are scheduled for eight semesters and can be reached in less than eight semesters at the maximum of 14 semesters[1]. Whereas in Riyanto's research, the study period's prediction is calculated based on nine attributes (sex, high school major, high school city, GPA value in semester 1 until 6) [2]. Nurhuda and Rosita's Research of study period uses several attributes, i.e., grade point, a cumulative semester of credits, economic status, and job status[3]. Whereas in this study, a prediction model for the study period was built using semester 1 to 4 GPA values based on the educational curriculum at XYZ University which emphasizes concentration selection in semester 4 lectures.

Haryati, et al. have used GPA values and the number of credits included in RapidMiner. The research used the C4.5 algorithm to know and analyze the level of timeliness of student studies [1]. Another study have used seven variable values in a competency examination in obstetrics and added to produce a knowledge Backpropagation algorithm to predict the value of the late obstetrics competency test results with an accuracy rate of 90% [4].

Asthana et al. concluded that in Backpropagation Neural Network, the hidden layer has a crucial role in the performance. It has been found that increasing the hidden layer also increases the accuracy performance slightly. On the contrary, the increase of the network complexity becomes the weakness of increasing the hidden layer [5].

Meanwhile, Chamsudin used three types of algorithms, namely Back Propagation, Quasi-Newton, and Lavemberg-Marquardt. This study forecasting results obtained by measuring the level I constellation forecasting by comparing the value of MAPE and MSE, of the three algorithms with the results Lavemberg-Marquardt algorithm generates the best MAPE value and MSE [6]. Priyanti used KNN algorithm to predict the amount of revenue with 83.62% accuracy value and compare it with the data from the UCI datasets, where the value of accuracy obtained at 79.18% [7]. Another study predicted students' academic performance in higher educational institutions using several attributes with a target value of CGPA at the end of semester 8. Their study talked about "Prediction of Student Academic Performance using Neural Network, Linear Regression and Support Vector Regression: A Case Study." The performance of the models was measured using the coefficient of correlation (R) and that of Root Means Square Error (RMSE) [8].

Umar predicts students' academic performance using Artificial Neural Networks [9]. The research used age, gender, residence, GPA semester 1 and number of subjects failed from the preceding semester as attributes and GPA in the second semester as a target. In this study, the neural network is used as a model; the confusion matrix is used as an evaluation method. Another research predicted graduation timeliness using the C4.5 algorithm with nine attributes and using SMOTE (Synthetic Minority Oversampling Technique) [10]. In comparison, Wibowo built an information system with the C4.5 algorithm to predict study time with application usability testing [11].

II. Material and Method

A. Neural Network

Neural Network is a biologically inspired computer program designed to simulate the way the human brain processes information. ANNs collect their knowledge by detecting data and relationships in data and learning (or being trained) through experience, not from programming [12].

1) Network Training: Neural Network Training has the purpose of seeking weight- weight contained in each layer. There are two types of training in artificial neural networks [13], which are Supervised Learning and Unsupervised Learning.

2) Neural Network Architecture: The division of neural network architecture can be seen from the framework and the interconnection scheme. The framework of artificial neural networks can be seen from the number of layers and the number of nodes in each layer. In Multilayer net network, there are one or more hidden layers [14]. The description of Multilayer net architecture can be seen in Figure 1 [15].
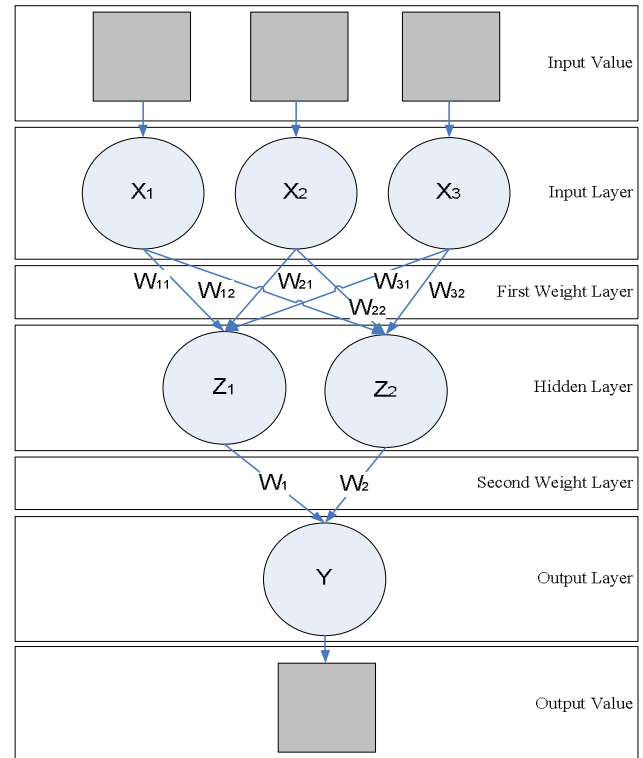

Fig. 1 Multilayer Net Architecture

3) Backpropagation: The algorithms are generally applied to the multilayer perceptron. Perceptron has an input, an output, and some lines that lie between the input and output. This middle layer, also known as hidden layers, can be one, two, three, and so on. The last layer output from the hidden layer is directly used as output from the neural network. [16]

A training algorithm for Backpropagation with one hidden layer (with a binary sigmoid activation function) as follows [17], Backpropagation with one layer can be seen in Fig 2.
Step 1: Initialize the weights set in small random values
Step 2: If the stop condition is false, work through steps 3-10.
Step 3: For each pair of training, work on steps 4-9.
- Feedforward
  Step 4: Each input unit (I = 1, ... n) receives input cues and is forwarded to hidden units.
  Step 5: Each hidden unit (Zj, j = 1, ... p) adds the weighted input cues.

$$\text{Z\_in}_{jk} = V_{oj} + \sum_{i=1}^{n} X_i V_{ij} \qquad (1)$$

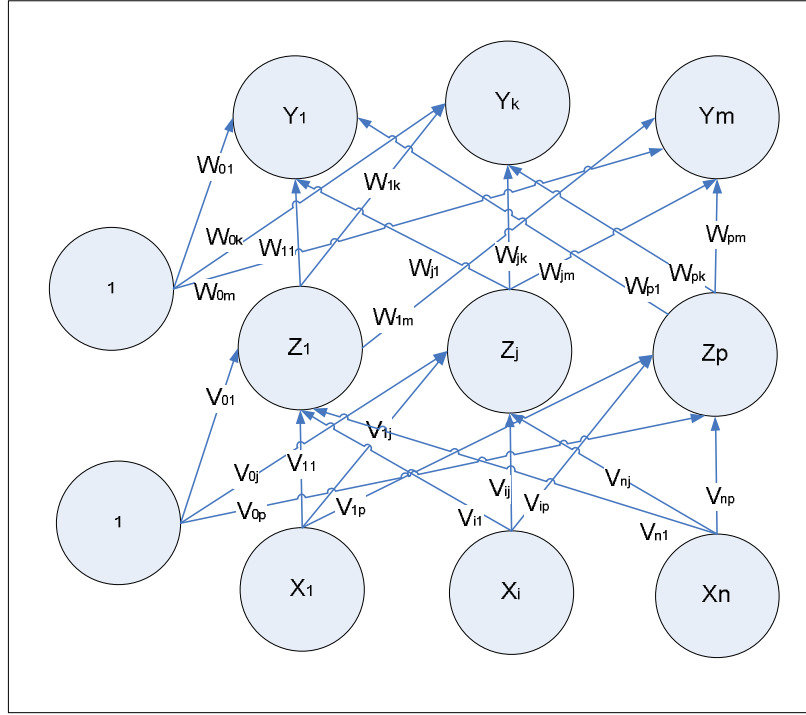By applying the activation calculation function:

$$Z_j = F(z - in_j) \qquad (2)$$

Fig. 2 Backpropagation with one hidden layer

Step 6: Each output unit (yk, k = 1, ... m) sums the weighted input cues.

$$y\_in_k = w_{ok} + \sum_{k=1}^{p} Z_j w_{jk} \qquad (3)$$

- Backpropagation
Step 7: Each output unit (yk, k = 1, ... m) receives a target pattern with the input training pattern. Calculate error value information:

$$\delta_i = (t_k - y_k)f'(y\_in_k) \qquad (4)$$

Calculate weight correction:

$$\Delta w_{jk} = a\delta_k x_j \qquad (5)$$

$$\Delta w_{ok} = a\delta k \qquad (6)$$

Step 8: Each hidden unit (zj, j = 1, ... p) sums the input delta (from the top layer units)

$$\delta\_in_j = \sum_{k=1}^{m} \delta_k w_{jk} \qquad (7)$$

Calculate the error value for the information:

$$\delta_j = \delta_{in_j} f(x\_in_j) \qquad (8)$$

Calculate the correction in the prerequisite's weights:

$$\Delta v_{ij} = a\delta_j x_i \qquad (9)$$

Update weights and prerequisites
Step 9: Each unit of output (yk, k = 1, ... m) update the weights and their prerequisites (j = 0, 1 ... ., p)

$$w_{jk}(baru) = w_{jk}(lama) + \Delta w_{jk} \qquad (10)$$

Each hidden unit (zj, j = 1, ... p) updates the weights and their prerequisites (j = 0, 1 ......, p)

$$v_{ij}(baru) = v_{ij}(lama) + \Delta V_{ij} \qquad (11)$$

Step 10: Stop the prerequisite test with:
x_1. xn: input
y_1.y_n: output
z_1.z_n: hidden layer value
v_{ij}: the weight between the input layer and the hidden layer
w_{jk}: the weight between the hidden layer and the output layer
δ: error information
α: speed or learning pace
μ: momentum

4) Mean Squared Error (MSE): According to William J. Stevenson and Sum Chee Chuong, they mentioned that there were three sizes of error used to sums up historical error. Those are Mean Absolute Deviation (MAD), the Mean Squared Error (MSE), and the Mean Absolute Percent Error (MAPE). MAD is the average absolute error; MSE is the quadratic error; and then, MAPE is the average percentage of absolute error. [18]. The Equation used to calculate MSE value can be seen in Equation (12).

$$MSE = \frac{\sum |aktual\ t - ramalan\ t|^2}{n} \qquad (12)$$

5) Data Preprocessing: Preprocessing Data is carried out to obtain more accurate analysis results using machine learning techniques or minimum data g. In some cases, preprocessing could make the value of the data smaller without changing the information it contains. There are some preprocessing of data carried out before applying a method, including the normalization or scaling is data change procedure so that it is on a specific scale [19]. This scale can be between (0,1), (-1,1) or any other scale you want. For example, we convert data into a scale or range of values between 0 to 1. In this case, the lower limit (BB) a is 0, and

the upper limit (BA) is 1. If the maximum value of each column is X_{max} and the minimum value is X_{min}, to change data to a new scale, for each data calculation can be done with Equation (13).

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} * (BA - BB) + BB \qquad (13)$$

B. Research Method

The research developed with methods consists of 3 phases, namely: data collection, data correlation analysis, NN prediction process algorithm, and testing. The flow of research is illustrated in Figure 3.
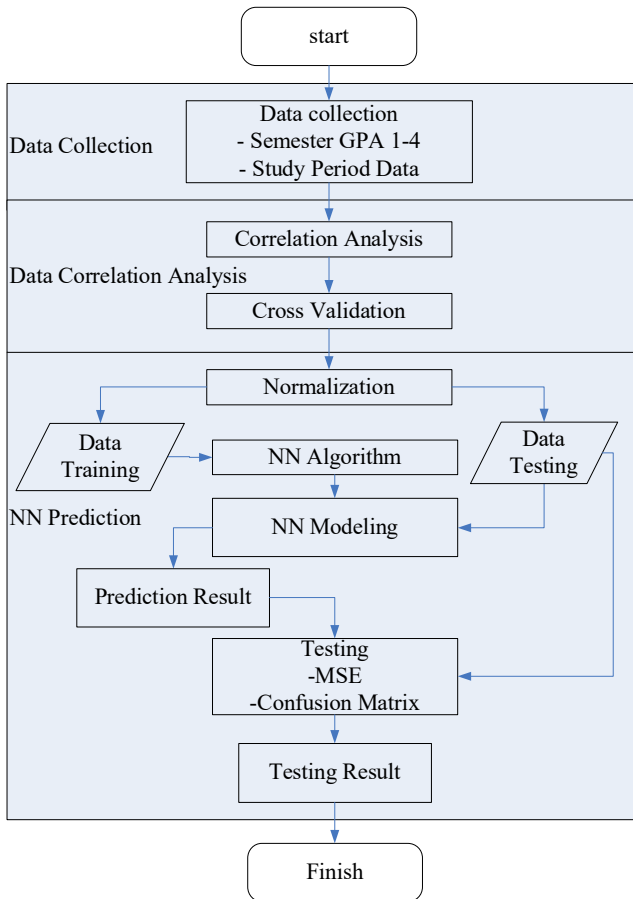

Fig. 3 Research method

C. Data Collecting

The data used in this study are students' GPA from 1st to 4th semester and data from the study period of students from 2013 to 2015, which will have a graduation period from 2017 to 2019 in the Informatics Study Program. The data is collected through the primary data obtained from interviews with the management of the informatics study program, in this case, data collected from the head of the study program or the study program secretary.

D. Data Correlation Analysis

The selection of data before the normalization process in the prediction model using the NN algorithm is carried out a correlation analysis used to measure the relationship between

1st to 4th semesters' GPA with the study period taken by students.

E. Data Regression Analysis

Data Regression Analysis is a process used to determine whether there is an influence on the value of GPA on the study period of students to complete college. The regression test that will be used is the F test, or testing the GPA semester 1 to 4 simultaneously on the Study Period.

F. Model Prediction using NN Algorithm

Prediction models using the NN algorithm are made to recognize net patterns. The model is determined after the training data, and test data are prepared. The characteristics and specifics used in the NN model architecture can be seen in Table 1.

TABLE I
NN MODEL CHARACTERISTIC AND SPECIFICATION

Characteristics	Specification
Layer	1 Hidden Layer
Neuron Input	4 Neurons (data input: GPA semester 1 to semester 4)
Output Neurons	1 Neuron (target study period prediction)
Hidden layer	Using three models using 10, 50 and 100 neurons
Activation Function	Binary Sigmoid
Maximum epoch	1000
Learning Rate	0.01, 0.1 and 0.3

G. Testing

Testing is done by using the MSE value test from the prediction data and then testing the prediction accuracy using the confusion matrix.

III. RESULTS AND DISCUSSION

A. Regression and Correlation Analysis

The use of correlation analysis aims to determine the degree of relationship between the first semester's GPA to the fourth semester and the study period. The correlation coefficient value of each GPA per semester with the study period has a negative relationship. It means that the lower GPA value the students achieve, the higher study period be responsible for the students, and conversely. The significance value is also indicated by a number below 0,05. So that it has a relationship between the GPA and the study period, the correlation test results can be seen in Table 2.

TABLE III
CORRELATION TEST RESULT

		GPA1	GPA2	GPA3	GPA4	SP
GPA1	PC	1	.558*	.538*	.573*	-.212*
	Sig.(2-tailed)		.000	.000	.000	.000
	N	619	619	619	619	619
GPA2	PC	.558*	1	.564*	.577*	-.240*
	Sig.(2-tailed)	.000		.000	.000	.000
	N	619	619	619	619	619

		GPA1	GPA2	GPA3	GPA4	SP
GPA3	PC	.536*	.564*	1	.787*	-.357*
	Sig.(2-tailed)	.000	.000		.000	.000
	N	619	619	619	619	619
GPA4	PC	.573*	.577*	.787*	1	-.309*
	Sig.(2-tailed)	.000	.000	.000		.000
	N	619	619	619	619	619
SP	PC	-.212*	-.240*	.357*	-.309*	1
	Sig.(2-tailed)	.000	.000	.000	.000	
	N	619	619	619	619	619

*. Correlation is significant at the 0.01 level (2-tailed)
SP: Study Period
PC: Pearson Correlation

B. Regression Analysis

The use of regression analysis aims to determine the effect of the GPA variable on the study period. In this article's analysis, the focus is on the test of influence simultaneously of GPA in the first semester to the fourth semester on the study period, which is, in Figure 3, indicated sig values >0,05 and Fcount 23,116> Ftable 2,372. In conclusion, there is a simultaneous influence of GPA on the study period. Regression analysis results can be seen in Table 3 below.

TABLE IIIII
REGRESSION ANALYSIS RESULT

	Model	Sum of Squares	df	Mean Square	F	Sig
1	Regression	3270.028	4	817.507	23.116	.000
	Residual	21714.599	614	35.366		
	Total	24984.627	618			

a. Predictors: (constant), GPA1, GPA2, GPA3, GPA4
b. Dependent Variable: Study Period

C. NN Model Design

Neural network model design is built with nine models where the input and output used are the same data. The following Table 4 illustrates the NN configuration used for the simulation process.

TABLE IVV
NN CONFIGURATION

Model	Hidden Layer Neuron	Activation	Max Epoch	Learning Rate	Training
1	10	Sigmoid Biner	1000	0.01	traingda
2	10	Sigmoid Biner	1000	0.1	traingda
3	10	Sigmoid Biner	1000	0.3	traingda
4	50	Sigmoid Biner	1000	0.01	traingda
5	50	Sigmoid Biner	1000	0.1	traingda
6	50	Sigmoid Biner	1000	0.3	traingda
7	10	Sigmoid Biner	1000	0.01	traingda
8	10	Sigmoid Biner	1000	0.1	traingda
9	10	Sigmoid Biner	1000	0.3	traingda

The NN model's number of input variables is 4, namely the semester 1 to 4 GPA with 1 target value is the student's study period. The architecture of the NN model is shown in Figure 5. Z1 represents the number of neurons... Zn, the value of n is based on the model in Figure 4.
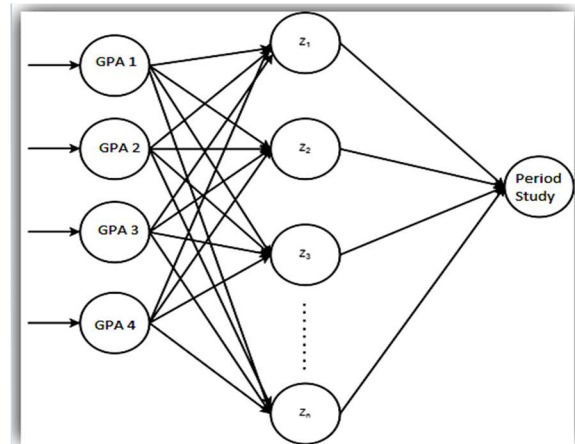

Fig. 4 NN Model Architecture

D. Data Preprocess

The data used in this study are GPA data and Study Period 2013-2015 with a total of 343 data, which will be divided into two parts, namely 300 training data and 43 data as testing data. Before data processing, the labeling for the study period above 48 will be labeled as 0 or late, and for graduates <= 48 months, it will be labeled as 1 or appropriate. The labeling is based on government regulations that the study period for the Bachelor's level is 4 years.

The activation function used on the network is a binary sigmoid function with a range of values from 0 to 1; therefore, data normalization is done using Equation 2.

In the preprocessing process also performed data normalization by scaling. Scaling is a procedure to change data so that it is on a specific scale [19]. This scale can be between (0,1), (1, -1), or other scale desired. If the maximum value for each column is X_{max} and the minimum value is X_{min} to change the data to a new scale, for each data can be obtained by Equation 1. The result of scaling data on preprocessing can be seen in Table 5.

TABLE V
SCALING DATA ON PREPROCESSING

NO	GPA1	GPA2	GPA3	GPA4	Study Period	Label
1	0.8214	0.9062	0.7142	0.7692	0.2162	1
2	0.7916	0.7812	0.7142	0.8974	0.2162	1
3	0.8214	0.7812	0.7499	0.7435	0.0270	1
4	0.7321	0.7499	0.7499	0.7692	0.2162	1
5	0.9702	0.9374	0.9285	0.9487	0.2972	1
6	0.9702	0.8749	0.8928	0.8461	0.1081	1
7	1	0.9687	0.9285	0.9487	0.1351	1
8	0.9702	0.8749	0.8214	0.8461	0.1891	1
9	0.7023	0.5624	0.3214	0.4615	0.7837	0
10	0.7619	0.7187	0.8214	0.8205	0.8378	0
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
342	0.8214	0.7812	0.8571	0.8974	0.3783	1
343	0.4940	0.4375	0.3571	0.7179	0.3783	1

E. NN Model Training Result

Measurement of the NN model's performance can be expressed by calculating the value of MSE, which is a measure of accuracy or the ability of the NN model to achieve

the intended target value, namely the value of the study period. For network, modeling is done using the R2017a MatLab tool. This study was conducted with nine patterns of data used for training by trying various values or quantities of learning rates 0.01, 0.1 and 0.3, then changing the number of neurons in the hidden layer using 10, 50, and 100 neurons. The training function used is Traingd. The network picture used starts from the network using 10 neurons as in Figure 5, while the network using 50 neurons in Figure 6 and the network with 100 neurons can be seen in Figure 7.
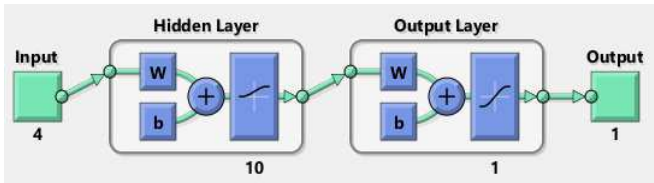

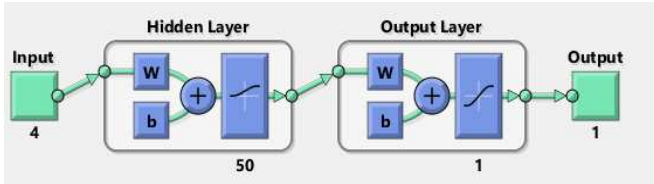Fig. 5 Network with 10 neurons

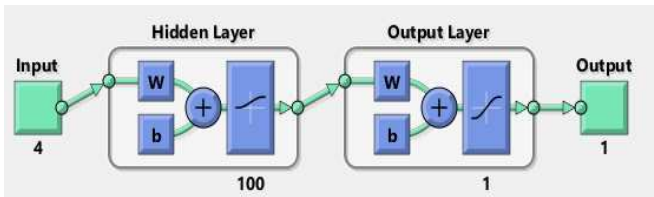Fig. 6 Network with 50 neurons

Fig. 7 Network with 100 neurons

The following are the results of training and testing on the NN models that were built following applied modeling can be seen in Table 6, Table 7, Table 8, and Table 9.

TABLE VI
MSE VALUE FROM TRAINING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0.019704	0.019211	0.020176
	50	0.017516	0.022593	0.01882
	100	0.019202	0.019169	0.019374

TABLE VII
MSE VALUE FROM TESTING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0,04950	0,05748	0,05139
	50	0,04772	0,05689	0,05267
	100	0,05234	0,05611	0,05902

TABLE VIII
RMSE VALUE FROM TRAINING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0,14037	0,13861	0,02018
	50	0,13235	0,15031	0,13719
	100	0,13857	0,13845	0,13919

TABLE IX
RMSE VALUE FROM TESTING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0,22249	0,23975	0,22669
	50	0,21845	0,23852	0,22950
	100	0,22878	0,23687	0,24295

F. Testing Phase

The network model's testing phase has been conducted training using data that is not included in the training process. The testing data used were 43 data, according to the appendix. Network output or output is the target data for the study period in the form of data normalization with Equation 13, and then the MSE value is calculated from the testing data. MSE value can be seen in Table 10 and RMSE in Table 11.

TABLE X
MSE VALUE FROM TESTING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0.049504	0.057481	0.051386
	50	0.047721	0.05689	0.052669
	100	0.052341	0.056109	0.059024

TABLE XI
RMSE VALUE FROM TESTING RESULT

		LR		
		0.001	**0.1**	**0.3**
Neuron	**10**	0,2224944565	0,2397513879	0,2266854612
	50	0,2184516343	0,2385166922	0,2294966998
	100	0,2287822923	0,2368740142	0,2429494077

To calculate the suitability of the results of testing the normalized value of the study period target data is returned in the form of normal data using inverse Equation 13. Study Period data from test results are labeled again with the number 1 if below is equal to 48 months and 0 if the study period is above 48, it is matched with real data from testing data to determine the suitability value. Graphically the real study period with network prediction with 10 neurons is shown in Figure 8. The network data graph with 50 neurons is seen in Figure 9 and the network data graph for 100 neurons seen in Figure10.
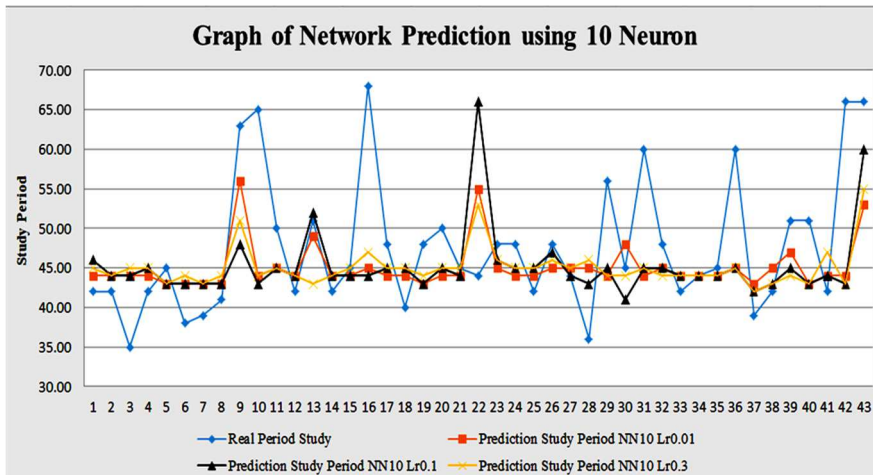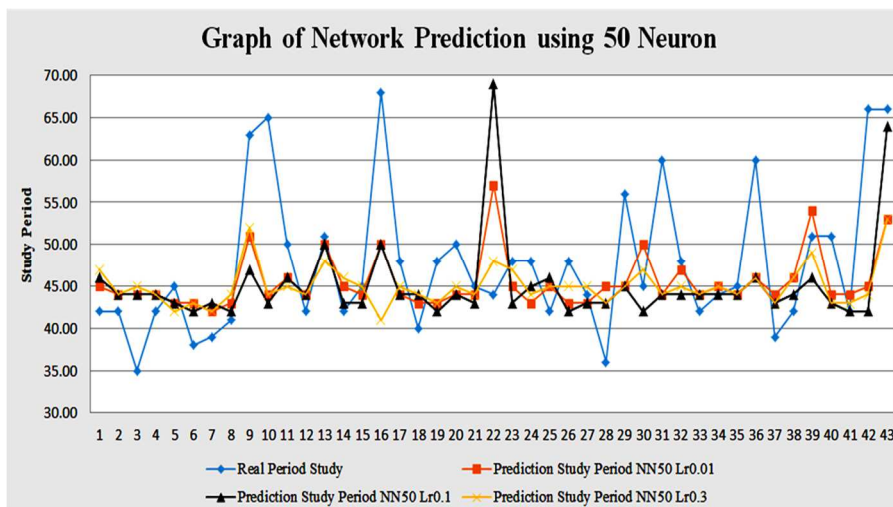
Fig. 8 Graph of network prediction using 10 neurons
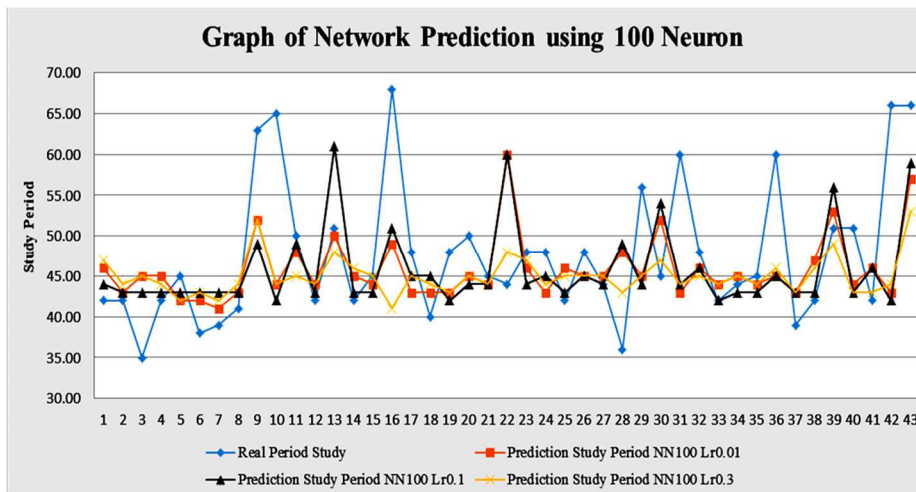

Fig. 9 Graph of network Prediction using 50 neurons

Fig. 10 Graph of network prediction using 100 neurons

H. Evaluation

In the process of evaluating this study, we used the evaluation matrix confusion. Confusion Matrix is a useful tool for analyzing how good a classifier is about tuple from different classes [20]. The evaluation results with the confusion matrix, based on Figure 8, 9, and 10, conformity testing is performed using a confusion matrix, the test results are displayed in Table 12 and graphically as follows in Figure 11.

616

TABLE XII
CONFUSION MATRIX TESTING

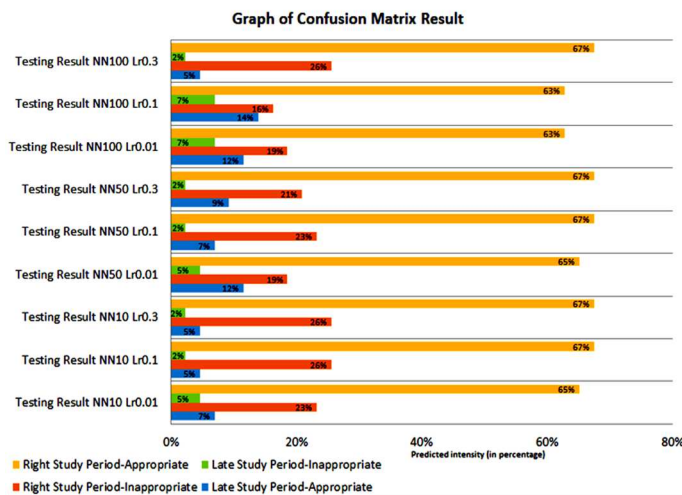| Network Model	Match	Not Match
Test Result nn10lr001	72%	28%
Test Result nn10lr01	72%	28%
Test Result nn10lr03	72%	28%
Test Result nn50lr001	77%	23%
Test Result nn50lr01	74%	26%
Test Result nn50lr03	77%	23%
Test Result nn100lr001	74%	26%
Test Result nn100lr01	77%	23%
Test Result nn100lr03	72%	28%

Fig. 11 Confusion matrix test result

IV. CONCLUSION

Based on the results, this study can be formulated as follows. MSE value represented the difference of the study period data between the real and the prediction. The value is both for training dan testing with the best result on the pattern of neuron network 50 with 0,01 as a learning rate. Based on the detail of MSE value, the training is 0,017516, and the testing is 0,047721. Then, the network pattern with 50 neurons with the training process using 0,01 as a learning rate produces a percentage of 77% compatibility. Generally, the study period's prediction results using neural networks have a good result with an average percentage of suitability above 70%.

Future research is still focused on the pattern of GPA in semester 1 to semester 4. Furthermore, it can be developed with different input values. Therefore, it is recommended to combine several algorithms for processing input data at the preprocessing stage, such as natural selection, genetic algorithm, or other algorithms that support neural networks to increase accuracy.

REFERENCES

[1] S. Haryati, A. Sudarsono, and E. (2015) Suryana, "Implementasi Data Mining untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5," *J. Media Infotama*, vol. 11, no. 2, pp. 130–138, 2015.

[2] V. Riyanto, A. Hamid and Ridwansyah, "Prediction of Student Graduation Time Using The Best Algorithm" *IJAIDM Indonesian Journal of Artificial Intelligence and Data Mining.*, vol. 2, no. 1, pp. 1–9, 2019.

[3] A. Nurhuda and D. Rosita, "Prediction Student Graduation on Time Using Artificial Neural Network on Data Mining Students STMIK Widya Cipta Dharma Samarinda," pp. 86–89, 2017.

[4] A. S. Kurniawansyah, "Implementasi Metode Artificial Hasil Ujian Kompetensi Kebidanan," *J. Pseudocode*, vol. V, no. 1, 2018.

[5] S. Asthana, D. Goyal, and A. Pandit, "Analysis on Multiple Hidden Layer Complexity of BPNN," *Int. J. Appl. Eng. Res.*, vol. 12, no. 14, pp. 4723–4728, 2017.

[6] C. Ahmad, "Implementasi Neural Network Untuk Memprediksi Jumlah Penderita Tuberculosis," *Emit. J. Tek. Elektro*, vol. 16, no. 01, pp. 43–50, 2015.

[7] E. Priyanti, "Implementasi Neural Network Pada Prediksi Pendapatan Rumah Tangga," *Swabumi*, vol. 6, no. 1, pp. 18–26, 2018.

[8] Efrem Yohannes Obsie and S. A. Adem, "Prediction of Student Academic Performance using Neural Network, Linear Regression and Support Vector Regression: A Case Study," *Int. J. Comput. Appl.*, vol. 180, no. 40, pp. 39–47, 2018.

[9] MA. Umar, "Student Academic Performance Prediction using Artificial Neural Networks: A Case Study," International Journal of Computer Applications., vol. 178, no. 48, pp. 24–29, 2019.

[10] Y. T. Samuel, J. J. Hutapea, and B. Jonathan, "Predicting the timeliness of student graduation using decision tree c4.5 algorithm in universitas advent Indonesia," *Proc. 2019 Int. Conf. Inf. Commun. Technol. Syst. ICTS 2019*, pp. 276–280, 2019.

[11] S. Wibowo, R. Andreswari, and M. A. Hasibuan, "Analysis and design of decision support system dashboard for predicting student graduation time," *Int. Conf. Electr. Eng. Comput. Sci. Informatics*, vol. 2018-Octob, pp. 684–689, 2018.

[12] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *J. Pharm. Biomed. Anal.*, vol. 22, no. 5, pp. 717–727, 2000.

[13] G. Florin, *Data Mining:Concepts, Models and Techniques*. Romania: Springer, 2011.

[14] D. Puspitaningrum, *Pengantar Jaringan Syaraf Tiruan*. Yogyakarta: Andi, 2006.

[15] A. Hasim, "Prakiraan Beban Listrik Kota Pontianak Dengan Jaringan Syaraf Tiruan (Artificial Neural Network)," *Cent. Libr. Bogor Agric. Univ.*, p. 1, 2008.

[16] M. T. Hagan, *Neural Network Design*. USA: PWS Publishing Co, 1996.

[17] F. L, *Fundamentals of Neural Networks Architectures, Algorithms and Applications*. London: Prantice-Hall, Inc, 1994.

[18] W. J. Stevenson and S. Chee Chuong, *Manajemen Operasi Perspektif Asia*. Jakarta: Salemba Empat, 2014.

[19] S. B, *Data Mining: Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu, 2007.

[20] J. Han, J, Kamber, M, & Pei, *Data Mining: Concept and Techniques, Second Edition*. Waltham: Morgan Kaufmann Publishers, 2006.