Journal of Engineering Science and Technology
Vol. 16, No. 3 (2021) 2028 - 2037
© School of Engineering, Taylor's University

# DATA MINING, NEURAL NETWORK ALGORITHM TO PREDICT STUDENT'S GRADE POINT AVERAGE: BACKPROPAGATION ALGORITHM

SELVIA LORENA BR GINTING*, M. ADITYA FATHUR RAHMAN

Department of Computer System, Universitas Komputer Indonesia,
Jl. Dipati Ukur No. 112 – 116, 40132, Bandung, Indonesia
*Corresponding Author: selvia.lorena@email.unikom.ac.id

**Abstract**

The Grade Point Average (GPA) for a tertiary institution is one of the elements used to measure the quality of teaching and learning. In the Department of Computer Systems, Universitas Komputer Indonesia, there are many students who have bachelor's degrees, but there are still many students who do not graduate on time with unsatisfactory GPA. By utilizing the pile of academic data from students who have graduated, modelling the data mining technique using the Artificial Neural Network (ANN) method with the backpropagation algorithm can be done. The purpose of this study is to test the ANN modelling with backpropagation algorithm in making a student's Grade Point Average (GPA) prediction information system. The student GPA prediction system is used as a supporting material for the academic advisor to provide evaluation guidance to students who are considered to have academic problems. The parameter attributes used in this study included Grade Point Semester (GPS) one to four as well as several course scores as data input and GPA as an output data. The test was carried out in four scenarios, where the best results of ANN were obtained in scenario-IV with the parameters of the number of input layer neurons 18, the number of hidden layer neurons 24, the number of output layer neurons 1, learning rate ($\alpha$) 0.15, 2000 iterations (epoch) and 591 data sets. The data is divided into 85% data set for training data as well as 15% data set for test data. The average accuracy obtained from this information system is 97.2%. It is expected that this study can be used as a reference for designing a data mining architecture model using the ANN backpropagation algorithm in creating values prediction information system.

Keywords: Artificial neural network, Backpropagation, Data mining, Grade point average.

## 1. Introduction

Student performance in a college is an assessment of the college quality [1]. One of the points that become the assessment of student achievement can be seen from the Grade Point Average (GPA) obtained by students [2]. GPA for higher education is one of the elements used to measure the quality of teaching and learning [2]. However, there are still many students who graduate not on time and their GPA are not satisfactory. However, there are still many students who do not graduate on time with unsatisfactory GPA. The factors that caused students' academic problems are usually students who have a small Grade Point Semester (GPS). It is because they sometimes faced a difficulty in certain subjects where these courses are related to other subjects in the next semester. Therefore, it caused a domino effect on GPS in certain semesters. In a university, piles of academic data for students who have graduated are usually left unattended and eventually become piles of data in the database [3]. Piles of academic data can be a useful information if it is managed and analysed using an algorithm that can explore the potential of existing information. Data mining techniques or knowledge discovery means finding or extracting knowledge patterns from a pile of data using algorithms [3-5].

Previous research stated that the results obtained when using backpropagation are able to predict with 97.07% accuracy [1]. Rivas also proposed that Artificial Neural Network (ANN) has the highest accuracy, namely 0.782 compared to other data mining methods (decision tree accuracy 0.705; random forest accuracy 0.755; extreme gradient boosting accuracy 0.765) [6]. Furthermore, according to Gedeon and Turner [7] the neural network can predict with an accuracy of 94%.

Based on previous research, the results obtained from the use of the data mining technique using ANN method with backpropagation algorithm have a good level of accuracy. It also can explore more new data patterns. Fault tolerance and error in the ANN method with the backpropagation algorithm can be considered as noise only. Based on related studies, the student GPA prediction system utilizes a pile of academic data from previous students of the Department of Computer Systems, Universitas Komputer Indonesia (UNIKOM). The system was built using the data mining technique using the ANN method with the backpropagation algorithm. The research attributes used to include several grades of certain subjects that did not change in the curriculum change in one to four semesters.

The purpose of this study is to test the ANN modelling with backpropagation algorithm in making a student's Grade Point Average (GPA) prediction information system. The student GPA prediction system is used as a supporting material for the academic advisor to provide evaluation guidance to students who are considered to have academic problems.

## 2. Research Method

There are many methods in data mining techniques, one of which is the Artificial Neural Network (ANN) that chosen in this study because of its potential and ability to predict accurately and tolerate errors as noise [8, 9]. ANN is an analysis modelling recognition of information pattern classification that is modelled like a human brain neural network (neurons) that can coordinate nerves effectively [8-11].

Backpropagation training algorithms are often used in studies that uses ANN method specifically in predicting a value [9]. The backpropagation training

algorithm is a supervised training algorithm. This training allows the weights and biases in the complex model to be updated with a smaller value [9]. The backpropagation algorithm has two phases, feedforward as the first phase which produces an output error. The second phase is backward which used to update the weights and biases values [9]. In general, the stages of the backpropagation training process are as follows [9]:

- Adjust initial weights and bias with small random values.
- Each input is received to the neuron which is then propagated to the next layer.
- In the hidden layer, each neuron is multiplied by the weight and the bias is added, as well as the neurons in the output layer whose final result is the output.
- Furthermore, the neuron output generated in the output layer compares with the desired output target and the output error results will be obtained. The equation is as follows:

$$\delta_m = (t_m - Y_m)f'(Y\_in_m) = (t_m - Y_m)Y_m(1 - Y_m) \quad (1)$$

- Update the weight and bias of the output layer times the output error.
- Multiply each weight and bias in the hidden layer by the output error, then the result is multiplied by the derivative of the activation function to get the output error in the hidden layer with the following equation:

$$\delta_n = \delta\_in_n f'(Z\_in_n) = \delta\_in_n Z_n(1 - Z_n) \quad (2)$$

- Update the hidden layer weight and bias multiplied by the error output in the hidden layer.
- Add up the weight and bias with the old for each layer.

$$W_{nm}(new) = W_{nm}(past) + \Delta W_{nm} \quad (3)$$

- These stages are carried out repeatedly until it meets the stop conditions iteration or minimum error.

Academic data of graduated students from the Department of Computer Systems Universitas Komputer Indonesia (UNIKOM) from class 2000 to 2015 were selected as data samples with the total of 591 respondents. It is divided into two parts, namely 85% for training data and 15% for testing data. The academic curriculum of the UNIKOM Computer Systems Department has changed four times from 2000 to 2020. Therefore, student academic data can be selected for data attributes to obtain the desired pattern of knowledge. The data attributes used in this study for input parameters use GPS one to four (X1-X4) and subject values (X5-X18), while the output parameters use GPA (Y).

The implementation of data mining techniques is depicted in the flow chart. The flow chart consists of two stages, namely the network training (See Fig. 1) and the network testing stage (See Fig. 2).

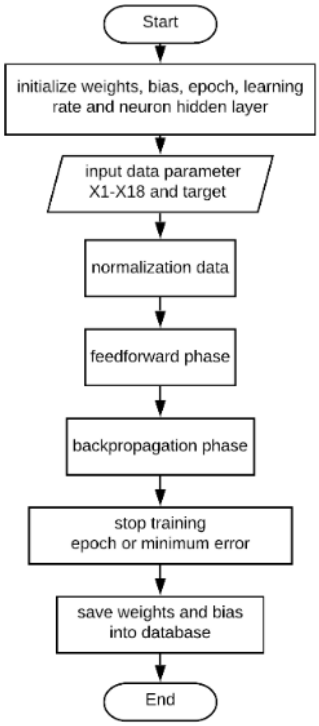


**Fig. 1. Training flow chart.**

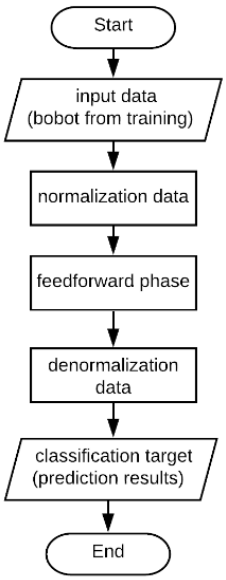


**Fig. 2. Testing flow chart.**

The stages of network training shown in Fig. 1 starts with initializing the parameters, including the number of neurons in the input, hidden, and output layer, as well as iteration (epoch), learning rate ($\alpha$), minimum error, initial weight, and initial bias. The training data set was going through data normalization to change the form of values ranging from zero to one. It was done to reduce complexity in learning or adjust weights and bias. The feedforward stage at the training stage is used to train the weight and bias patterns on the network and calculate the output error. Then, the backpropagation stage is used to update or correct weights and bias. The two phases are repeated until the stop conditions are met, namely iteration and minimum error. In the last stage, the results of network training weights and biases are stored in the database.

The network testing stage shown in Fig. 2 starts by entering the input data from 15% of the academic data dataset, then the data is normalized. The feedforward phase at the test stage is used to calculate the predicted output. The prediction output generated in the feedforward phase must denormalize the data in order to change shape the value back. The output classification is formed based on the results of the GPA prediction and reviewing the scores of several courses (parameters X5-X18), then four kinds of response outputs are obtained which are shown in Table 1.

**Table 1. Response output classification.**

| No | Condition | Response Output |
|---|---|---|
| **1.** | GPA $\geq$ 2.75 & All Courses ($X_5$-$X_{18}$) > 1 (D) | Students do not experience academic problems |
| **2.** | GPA $\geq$ 2.75 & Some Courses ($X_5$-$X_{18}$) $\leq$ 1 (D) | Students experience some academic problems and have to fix courses a, b, etc. |
| **3.** | GPA $\leq$ 2.75 & All Courses ($X_5$-$X_{18}$) > 1 (D) | Students experience academic problems and have below minimum scores are required to retake the courses |
| **4.** | GPA $\leq$ 2.75 & Some Courses ($X_5$-$X_{18}$) $\leq$ 1 (D) | Students experience academic problems and are required to improve the scores of courses a, b, c, d, e, etc. |

## 3. Results and Discussion

The GPA prediction information system for students using the artificial neural network method with the backpropagation algorithm was tested based on a predetermined architectural design. The test is carried out in four scenarios by comparing different Backpropagation architectural models (number of training data sets, neurons in hidden layers, and iterations). In this study, the scale of the backpropagation prediction output value was 0.00 - 4.00. For this reason, the learning rate ($\alpha$) is required to have a small value so that the output can be right on target. Therefore, the learning rate ($\alpha$) and the minimum error value can be equalized in testing.

### 3.1. Test scenario I

The test was carried out with 514 data sets for training data and 75 data sets for test data. The architectural parameters used include 18 neurons in the input layer, 24

neurons in the hidden layer, 1 neuron in the output layer, 0.15 learning rate ($\alpha$), 0.001 minimum error, and 200 iterations (epoch). The results of testing the scenario-I architectural model on the system obtained are shown in Fig. 3.

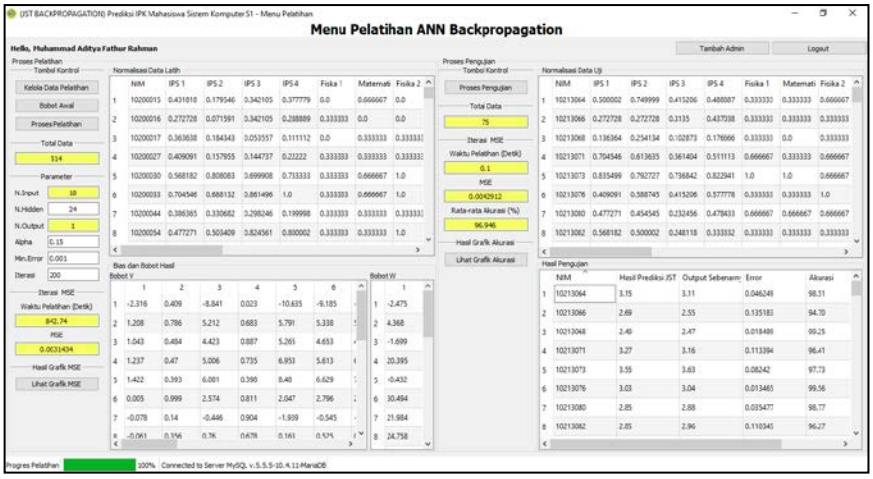


**Fig. 3. Results of training and test scenario-I.**

The results of the test resulted in the Mean Square Error (MSE) value of 0.0031434 which did not reach the minimum error limit. The resulting average accuracy value is 96.946% with the highest accuracy of 99.89% and the lowest accuracy of 89.82%. The test scenario-I is processed by iteration (epoch) of 200 which takes approximately 14 minutes.

## 3.2. Test scenario-II

The test was carried out with 442 data sets for training data and 147 for test data, where some of the training data were transferred to the test data. Architectural parameters used include 18 neurons in the input layer, 24 neurons in the hidden layer, 1 neuron in the output layer, 0.15 learning rate ($\alpha$), 0.001 minimum error, and 200 iterations (epoch). The results of testing the scenario-II architectural model on the system obtained are shown in Fig. 4.

The results of the scenario-II test resulted in the MSE value of 0.0030513 which did not reach the minimum error limit. The resulting average accuracy value is 96.361% with the highest accuracy of 100% and the lowest accuracy of 87.80%. The test scenario-II is processed by iteration (epoch) of 200 which takes approximately 11 minutes.

## 3.3. Test scenario-III

The test was carried out with 442 and 147 data sets for training and test data respectively, where some of the training data were transferred to the test data. Architectural parameters used include 18 neurons in the input layer, 36 neurons in the hidden layer, 1 neuron in the output layer, 0.15 learning rate, 0.001 minimum error, and 2000 iterations (epoch). The results of testing the scenario-III architectural model on the system obtained are shown in Fig. 5.

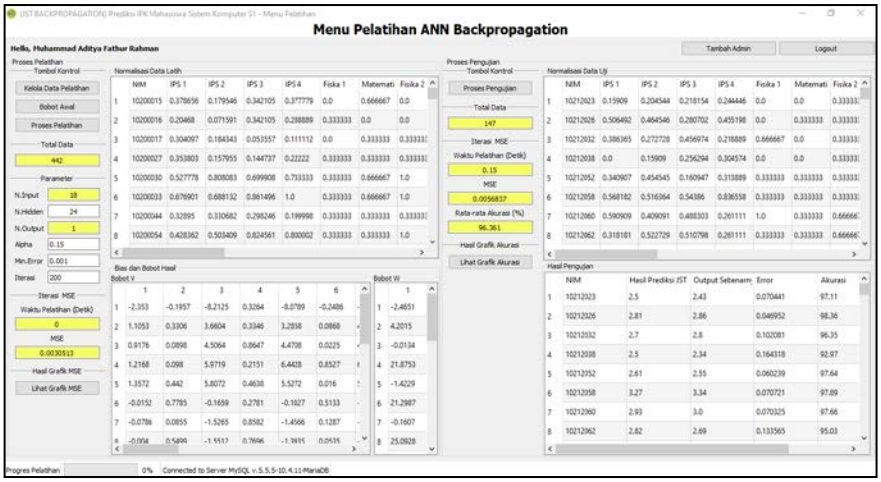


**Fig. 4. Results of training and test scenario-II.**

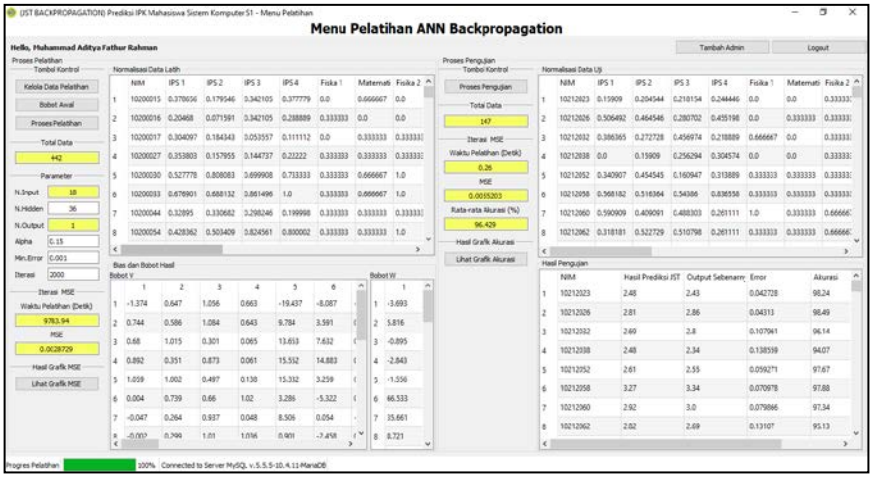


**Fig. 5. Results of training and test scenario-III.**

The results of the scenario-III test resulted in the MSE value of 0.0028729 which did not reach the minimum error limit. The resulting average accuracy value is 96.429% with the highest accuracy of 99.95% and the lowest accuracy of 88.20%. The scenario-III test is processed by iteration (epoch) of 2000 which takes approximately two hours and seven minutes.

## 3.4. Test scenario-IV

Test scenario-IV was carried out with 514 and 75 data sets for training and test data, respectively. Architectural parameters used include 18 neurons in the input layer, 24 neurons in the hidden layer, 1 neuron in the output layer, 0.15 learning rate ($\alpha$), 0.001 minimum error, and 2000 iterations (epoch). The results of testing the scenario-IV architectural model on the system obtained are shown in Fig. 6.

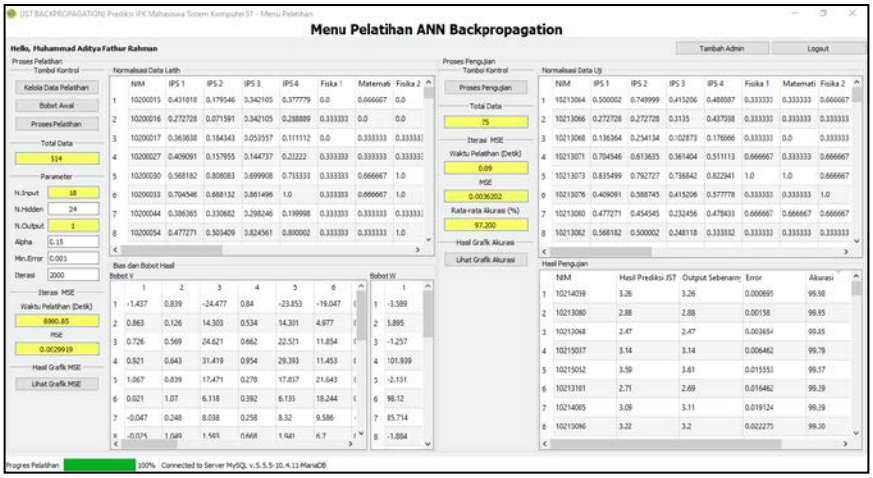


**Fig. 6. Results of training and test scenario-IV.**

The results of the scenario-IV test resulted in the MSE value of 0.0029919 which did not reach the minimum error limit. The resulting average accuracy value is 97.2% with the highest accuracy of 99.98% and the lowest accuracy of 90.7%. The scenario-IV test is processed by iteration (epoch) of 2000 which takes approximately two hours.

## 3.5. Results comparison

Based on the four test scenarios carried out, a comparison table of the backpropagation architectural model can be made from the four test scenarios. Table 2 presents a comparison of architectural models that include the amount of training data, the number of neurons in the hidden layer, learning rate ($\alpha$), and iteration (epoch).

**Table 2. Comparison of backpropagation architectural models.**

| Parameter | Scenario-I | Scenario-II | Scenario-III | Scenario-IV |
|---|---|---|---|---|
| **Training Data** | 514 | 442 | 442 | 514 |
| **N. Hidden Layer** | 24 | 24 | 36 | 24 |
| **Learning Rate** | 0.15 | 0.15 | 0.15 | 0.15 |
| **Iteration** | 200 | 200 | 2000 | 2000 |

Table 3 presents a comparison of the testing results the backpropagation architectural model from the four test scenarios that have been carried out. Parameters observed from the results of architectural model testing include MSE, average, highest, and lowest accuracy, as well as length of time in the network training process.

Based on Table 3, the backpropagation architectural model of the four test scenarios that the architectural model with the best performance has been carried out is scenario-IV testing. The average accuracy of the test is as high as 97.2% and MSE as small as 0.0029919. Thus, the scenario-IV testing architecture model is defined as the backpropagation architectural model of the GPA prediction information system for students.

**Table 3. Comparison of the results of the backpropagation architecture model.**

| Parameter | Scenario-I | Scenario-II | Scenario-III | Scenario-IV |
|---|---|---|---|---|
| **Training MSE** | 0.0031434 | 0.0030513 | 0.0028729 | 0.0029919 |
| **Average accuracy** | 96.946% | 96.361% | 96.429% | 97.2% |
| **Highest accuracy** | 99.89% | 100% | 99.95% | 99.98% |
| **Lowest accuracy** | 89.82% | 87.80% | 88.20% | 90.70 |
| **Training time** | ± 14 minutes | ± 11 minutes | ± 2 hours 7 minutes | ± 2 hours |

## 4. Conclusion

Testing the data mining architectural model using the Artificial Neural Network (ANN) with the backpropagation algorithm is tested with four scenarios. Where the best model performance results obtained an average accuracy of 97.2% and MSE 0.0029919 in the test scenario-IV. The architectural model includes 514 training data; 75 test data; 18 neurons in the input layer; 24 neurons in the hidden layer; 1 neuron in the output layer; learning rate ($\alpha$) 0.15; and iteration (epoch) 2000. The high average accuracy of backpropagation is influenced by the amount of training data and the number of neurons in each layer, which makes more patterns formed in the backpropagation. Furthermore, smaller learning rate ($\alpha$) with stabilization by a large number of training iterations makes the output value more accurate. The data mining technique of the ANN method Backpropagation algorithm is the right choice to predict values. As in this study, it predicts the student's Grade Point Average (GPA).

## References

1. Asogwa, O.C.; and Oladugba, A.V. (2015). Of students academic performance rates using Artificial Neural Networks (ANNs). *American Journal of Applied Mathematics and Statistic*, 3(4), 151-155.
2. Bacon, D.R.; and Bean, B. (2006). GPA in research studies: An invaluable but neglected opportunity. *Journal of Marketing Education*, 28(1), 35-42.
3. Tekin, A. (2014). Early prediction of students' grade point averages at graduation: A data mining approach. *Eurasian Journal of Educational Research*, 14(54), 207-226.
4. Ginting, S.L.B.; Ginting, Y.R.; Sutono.; and Rakhman, A. (2019). Data mining: The classification method to predict the types of motorcycle spare parts to be restocked. *Materials Science and Engineering*, 662(2), 1-7.
5. Ahmad, F.; Ismail, N.H.; and Aziz, A.A. (2015). The prediction of students' academic performance using classification data mining techniques. *Applied Mathematical Sciences*, 9(129), 6415-6426.
6. Rivas, A.; González-Briones, A.; Hernández, G.; Prieto, J.; and Chamoso, P. (2021). Artificial neural network analysis of the academic performance of students in virtual learning environments. *Neurocomputing*, 423, 713-720.
7. Gedeon, T.D.; and Turner, H.S. (1993). Explaining student grades predicted by a neural network. *Proceedings of the International Conference on Neural Networks.* Nagoya, Japan, 609-612.
8. Masoudi, S.; Sima, M.; and Tolouei-Rad, M. (2018). Comparative study of ann and anfis models for predicting temperature in machining. *Journal of Engineering Science and Technology*, 13(1), 211-225.

9. Habibagahi, G.; and Taherian, M. (2004). Prediction of collapse potential for compacted soils using artificial neural networks. *Scientia Iranica*, 11(1), 1-20.
10. Wang, T.; and Mitrovic, A. (2002). Using neural networks to predict student's performance. *Proceedings of the International Conference on Computers in Education*. Auckland, New Zealand, 969-973.
11. Sikder, M.F.; Uddin, M.J.; and Halder, S. (2016). Predicting students yearly performance using neural network: A case study of BSMRSTU. *Proceedings of the 5th International Conference on Informatics, Electronics and Vision*. Dhaka, Bangladesh, 524-529.