

Implementation of Neural Network Algorithms in predicting student graduation rates

Iin Fiqha^{1)*}, Gomal Juni Yandris²⁾, Fitri Aini Nasution³⁾

¹⁾²⁾³⁾Labuhanbatu University, North Sumatra, Indonesia

¹⁾ iinfiqha48@gmail.com, ²⁾ gomaljunianris@gmail.com, ³⁾ fitriaininasution689@gmail.com

Submitted: Jan 13, 2022 | **Accepted :** Feb 3, 2022 | **Published :** Feb 3, 2022

Abstract: Higher education institutions are required to be providers of quality education. One of the instruments used by the government to measure the quality of education providers is the number of graduates. The higher the graduation rate, the better the quality of education and this good quality will positively affect the accreditation value given by BAN-PT. Therefore, in this study, researchers provide input for research conducted at Bhayangkara University, Greater Jakarta to predict student graduation rates using the Neural Network algorithm. Neural Network is a method in machine learning developed from Multi Layer Perceptron (MLP) which is designed to process two-dimensional data. Neural Network is included in the type of Deep Neural Network because of the depth of the network level and is widely implemented in image data. Neural Network has two methods; namely classification using feedforward and learning stages using backpropagation. The way Neural Network works is similar to MLP but in Neural Network each neuron is represented in two dimensions, unlike MLP where each neuron is only one dimension. The prediction accuracy obtained is 98.27%. unlike MLP where each neuron is only one-dimensional. The prediction accuracy obtained is 98.27%. unlike MLP where each neuron is only one-dimensional. The prediction accuracy obtained is 98.27%.

Keywords: Pass Rate, Neural Network, Multi Layer Perceptron, Backpropagation, Prediction Accuracy

INTRODUCTION

Student graduation is one of the strategic goals of every higher education provider because it is one of the standards in the accreditation assessment instrument.

Bhayangkara University, Greater Jakarta, in this case the Management Study Program, Faculty of Economics, really needs a method that can facilitate the wishes of stakeholders at Bhayangkara University, Greater Jakarta, so that the graduation of management study program students can be predicted in advance.

Based on data obtained from the management study program, there are 185 students who have graduated with an average graduate of more than 4 years. The study period of more than 4 years in 2018 can of course be accelerated if stakeholders in this case officials in the management study program can predict student graduation early. Accelerating students to be able to graduate on time will certainly get a good rating on the accreditation of study programs by BAN-PT (Jefika et al., 2020).

Therefore, management study programs need to apply the right method to answer this problem by applying methods to predict student graduation. So that the problem of students who do not graduate on time can be reduced.

There are several previous studies conducted using the neural network algorithm method, namely:

1. Research conducted by Azahari and Yulindawati with the title Comparison of Naive Bayes Data Mining Algorithms and Neural Networks in Predicting Student Study Period. This study resulted in an accuracy of 72.58% (Ginting & Trinanda, 2014)
2. The research was conducted by Kusuma Dewi, Wing Wahyu Winarno, and M. Rudyanto Arief with the title Prediction of Student Graduation Using Artificial Neural Networks and Particle Swarm Optimization Methods. In this study, the prediction accuracy value of 87.31% was obtained. (Singh et al., 2020)

*name of corresponding author



3. The research was conducted by Abdul Rohman and Muhammad Rochcham with the title Comparison of Data Mining Classification Methods for Predicting Student Graduation. In this study, it was found that the predictive accuracy value of student graduation was 91.7% (Kusnadi & Jannah, 2021)

From the three studies above, testing was carried out by applying a learning rate that moved away from the number 0 and approached the number 1. Where we know the greater the value of the learning rate, the training process will run faster. However, if the value of the learning rate is relatively too large, in general the training process can exceed the optimal state when the minimum error value is reached. In other words, the learning rate affects the network accuracy of a system. The greater the value of the learning rate, the network accuracy will decrease, but conversely, if the learning rate is getting smaller, then the network accuracy will be greater or increase with the consequence that the training process will take longer (Witten et al., 2016)

This learning rate problem can be overcome by implementing a number of steps by reducing the number of non-important attributes in the research dataset. The importance of an attribute can be measured by its effect on accuracy and precision and the AUC value generated when the dataset is processed with Rapidminer (Wu et al., 2014)

The author in this case has applied the above method by reducing some unneeded attributes and absent values that do not move from 0 to 1.

LITERATURE REVIEW

Machine Learning

In the era of the popularity of science, machine learning can be said to be a branch of the artificial intelligence discipline that is widely used by business practitioners and academics. Ability to develop systems based on data, make many people use it as a method in research. There are many things that are necessary to learn in machine learning but basically there are 4 main things that we need to know, namely Supervised learning, Unsupervised Learning, Semi-supervised learning, and Reinforcement learning. (Fricles A Sianturi, Hasugian, Paska Marto, Simangunsong Agustina, 2019)

One technique to apply machine learning is supervised learning. As discussed earlier, machine learning without data will not work. Therefore, the first thing to prepare is data. The data will usually be divided into 2 groups, namely training data and testing data. The training data will be used to train the algorithm to find a suitable model, while the testing data will be used to test and determine the performance of the model obtained at the testing stage. (Witten et al., 2016)

From the model obtained, we can make predictions which are divided into two types, depending on the type of output. If the prediction results are separate, then it is called a classification process. For example, gender classification is seen from handwriting (output of male and female). Meanwhile, if the output is continuous, it is called a regression process. For example, the prediction of house price ranges in the city of Jakarta (output in the form of house prices) (Purba et al., 2019)

Data Mining

Data mining is a technology that combines traditional analytical methods with sophisticated algorithms to process large volumes of data. Data mining is a term used to find hidden knowledge in databases. Data mining is a semi-automated process that uses statistical, mathematical, intelligent, and machine learning techniques to extract and identify useful and useful information stored in large databases. (Meilina, 2015)

Some of the initial definitions of data mining focused on processing data in the automation process. The theory of data mining is widely used in the fields of marketing, sales, and customer service with the same goal of defining data mining as the process of exploring and analyzing large amounts of data automatically and semi-automatically so that meaningful patterns or rules are found. (Febriyani & Winanjaya, 2021)

The automated analysis performed by data mining exceeds that of the widely used traditional decision support systems. Data mining can answer business questions that have traditionally required a lot of time and high costs. Data mining explores databases to find hidden patterns, searches for information to predict what business people and education practitioners might forget because it lies beyond their expectations.

The rapid development in data processing technology has made it easier for humans to collect large amounts of data, thus generating a lot of data. Data mining is the process of automatically searching for information that is useful in large data stores. Other frequently used terms include knowledge discovery (mining) in databases (KDD) (Sulastri & Gufroni, 2017)

The terms data mining and Knowledge Discovery in Database (KDD) are often used interchangeably to describe the process of extracting hidden information in large databases. Actually the two terms have different concepts, but are related to each other. And one of the stages in the whole KDD process is data mining.

*name of corresponding author



Neural Network Algorithm

Neural network is an algorithmic model inspired by how neurons in the human brain work. Every neuron in the human brain is connected and information flows from each neuron (Believers & Riana, 2017)

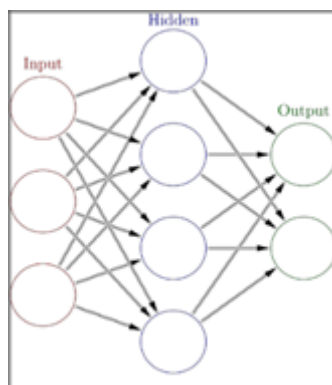


Fig. 1 Neural Network Model

METHOD

Research Stages

This research arranged through systematic stages with the aim that the research is directed. The stages of the research carried out are as follows.

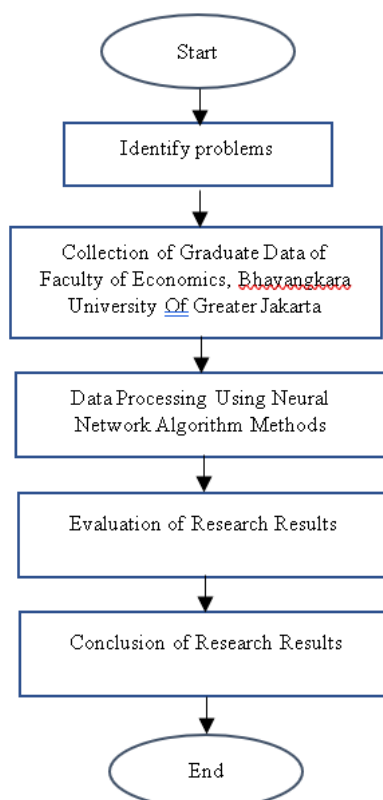


Fig. 2 Stages of Research

3.2. Research Instruments

Research instrument is a tool used in processing data into useful information. The data used in this study is data from graduates of the Management Study Program, Faculty of Economics, Bhayangkara University, Greater Jakarta, batch in 2014 and 2015. Furthermore, the data is processed using application Rapidminer Studio 9.3 to determine the accuracy of predicting student graduation rates.

*name of corresponding author



3.3. Data collection technique

The data collection method used in this research is to conduct a literature study on the academic data collected concerning student grades and credits at the Faculty of Economics, Bhayangkara University, Greater Jakarta.

1. Method of collecting data

This research done by using primary data. The primary data is in the form of student data from the 2014 and 2015 batches. From each of these data, 50% of the data is taken from testing. The data consists of several attributes, namely NIM, IP Semester 1 to 4, SKS Semester 1 to 4, Average GPA, and Total SKS.

2. Training, Learning, and Testing Data Collection Methods

Using Neural Network Algorithm This research uses a neural network as a training process from the results of data collection. In the training process, student graduation data will be processed first by changing the GPA value to number 1 or 0, Number 1 is defined as graduating on time and number 0 not graduating on time. From the results of the 0 and 1 classification, the data is then trained using a neural network.

The neural network algorithm training process requires three stages, namely feedward data input for training, backpropagation for error values and adjustment of the weight value of each node of each layer in the Neural Network. Starting with the input value feedward, each input unit i (x_i) receives an input signal which will then be transmitted to the hidden layer z_1, \dots, z_p . Then the j th hidden unit will calculate the signal value (z_j), which will be transmitted to the output layer, using the activation function f .

$$z_{in_j} = \theta_{1j} + \sum_{i=1}^n x_{ij} v_{ji} \quad (1)$$

And

$$z_j = f(z_j) \quad (2)$$

Where θ_{1j} = j th hidden unit bias. The bias value and initial weight can be taken randomly. Each k th output unit (Y_k).

$$Y_k = \theta_{2k} + \sum_j z_j w_{jk} \quad (3)$$

And

$$Y_k = f(Y_k) \quad (4)$$

Where $2k$ = k -th hidden bias unit. During the training process, each output unit compares the target value (T_m) for an input pattern in order to calculate the parameter values that will update (update) the weight values of each unit in each layer.

RESULT

Data Processing with Neural Network Method

The experiment was carried out using Rapidminer 9.3 tools and neural network algorithm methods. Parameters displayed are training cycle, learning rate, momentum, and hidden layer. 0.1 for learning rate and 0.1 for momentum. The experimental results can be seen in table 1 below.

Table 1
Trial Value Training Cycle

<i>Training Cycles</i>	<i>Learning Rate</i>	Momentum	Accuracy
200	0.1	0.1	98.19%
400	0.1	0.1	98.19%
600	0.1	0.1	98.27%
800	0.1	0.1	97.90 %

Based on the experimental results above, the largest training cycle value was selected, namely 600 with an accuracy rate of 98.27%.

This value is then used for experiments in determining the learning rate.

The value of the learning rate is obtained by entering a value from the range 0.1 to 1. While the value of 0.1 is used for the momentum value. The experimental results can be seen from table 2 below.

Table 2
Test Determination of Learning Rate

<i>Training Cycles</i>	<i>Learning Rate</i>	Momentum	Accuracy
600	0.1	0.1	98.27%

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

600	0.5	0.1	97.83%
600	0.7	0.1	97.83%
600	1	0.1	97.80 %

The learning rate value is selected based on the highest accuracy value generated. Based on these experiments, the learning rate with the highest accuracy is 0.1. Furthermore, this value will be used in determining the momentum value. The momentum value is determined by entering a value with a range of 0 to 0.9. The value of the training cycle and learning rate was chosen from the previous experiment, namely 600 and 0.1. The experimental results can be seen in table 3 below.

Table 3
Momentum Value Determination Experiment

<i>Training Cycles</i>	<i>Learning Rate</i>	Momentum	Accuracy
600	0.1	0	98.27%
600	0.1	0.3	97.83%
600	0.1	0.7	97.83%
600	0.1	0.9	97.80 %

Based on the experimental results above, the neural network parameters in this study used a value of 600 for the training cycle (epoch), 0.1 for the learning rate, and 0.1 for the momentum.

Meanwhile, to determine the number of hidden layers and hidden nodes, a trial was carried out using one and two hidden layers, while for hidden nodes using a value range of 1 to 24. The test results can be seen in table 4.

Table 4
One Hidden Layer Trial

Number of Hidden Nodes	Accuracy	Amount of Hidden	Accuracy
3	98.27	14	97.94%
7	97.93	17	97.94%
9	97.91	21	97.94%
13	97.94	24	97.94%

The best experimental results are with one hidden layer with the number of hidden nodes 3 with an accuracy of 98.27%. For the two hidden layers, 27 architectures were tested. The experimental results can be seen in table 5 below.

Table 5
Experiment with Two Hidden Layers

Number of Hidden Node Layer 1	Number of Layer 2 Hidden Nodes	Accuracy
1	1	97.94%
1	3	97.94%
1	5	97.94%
1	7	97.94%
1	9	97.94%
2	1	96.91%
2	3	97.94%
2	5	97.94%
2	7	97.94%
2	9	98.27%
3	1	96.91%
3	3	97.94%
3	5	97.94%
3	7	97.94%

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

3

9

97.94%

Based on the experimental results above for one hidden layer with the number of hidden nodes 3 does not produce a better accuracy than 98.27%. The number of input nodes in this study was determined by experimenting 2 times. The first one uses 6 input nodes (Gender, IP1, SKS2, IP2, SKS3, IP3). The second one uses 8 input nodes (Gender, IP1, SKS2, SKS3, IP3, SKS4, IP4). The two experiments above use a parameter value of 600 for the training cycle (epoch), 0.1 for the learning rate, hidden layer 1, hidden node 3, and 0.1 for momentum. For momentum, accuracy results are obtained as shown in table 6 below.

Table 6
Experiment Number of Input Nodes

Number of Input Nodes	Training Cycles (epoch)	Learning Rate	Momentum	Accuracy
4	600	0.1	0.1	96.90 %
6	600	0.1	0.1	97.96 %
8	600	0.1	0.1	98.27 %

Based on the experiment above, the number of input nodes 8 produces the greatest accuracy, namely 98.27, so in this study using 8 input nodes. The training process gets an AUC (Area Under Curve) value of 99.2% as seen from the ROC graphic image below.

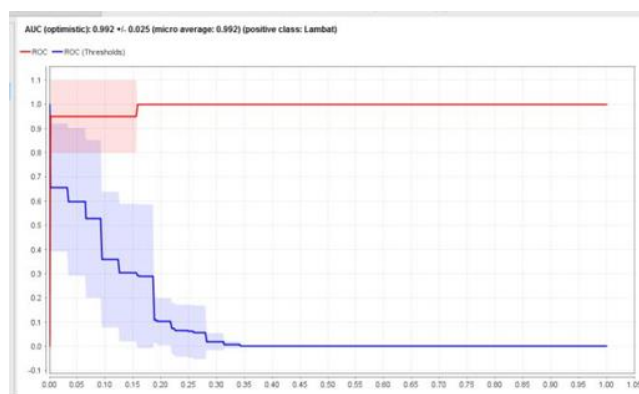


Fig. 3 AUC values in the ROC grafik chart

Testing the neural network algorithm method in determining the level of prediction accuracy, the confusionmatrix is obtained throughprocessing 344 data ie.

accuracy: 98.27% +/- 2.01% (micro average: 98.26%)

	true Cepal	true Lambat	class precision
pred Cepal	317	5	98.45%
pred Lambat	1	21	95.45%
class recall	99.69%	80.77%	

Fig. 4 Prediction accuracy results

$$Akurasi = \frac{(TN + TP)}{(TN + FN + TP + FP)}$$

$$Akurasi = \frac{(21 + 317)}{(21 + 1 + 317 + 5)} = 0,982558$$

Based on table 8 above, it is classified as follows:

True Positive (TP) = 21 records

True Negative (TN) = 317 records

False Negative (FN) = 1 record

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

False Positive (FP) = 4 records

The neural network architecture that is formed can be seen in Figure 5 below

Input Hide Output

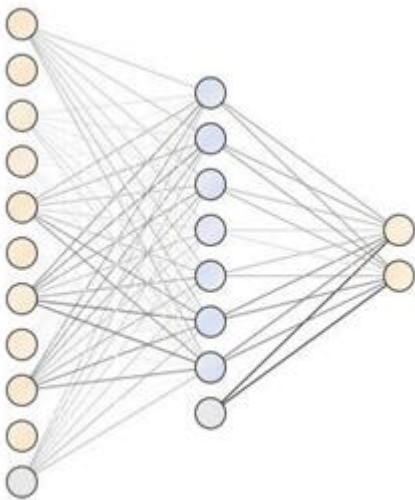


Fig. 5 Neural Network Architecture

Based on the results shown above, it is concluded that this study resulted in a prediction accuracy of student graduation of 98.27%. With such an accuracy value, it is certainly very appropriate to be used in assisting the Faculty of Economics, Bhayangkara University, Greater Jakarta in measuring student graduation rates.

DISCUSSIONS

From the results of research conducted at the Management Study Program, Bhayangkara University, Greater Jakarta, it can be concluded that:

1. The predicted results show results that illustrate that student graduation tends to be late because the dominant training data used has a late trend.
2. The number of inputs such as input nodes, training cycles, learning rate, momentum, number of hidden layers, and number of hidden nodes greatly affects the accuracy of the neural network algorithm.
3. From the results of the research conducted and the results of the application of the neural network algorithm, the most influential parameter is the student's IP (Achievement Index).
4. The results of the prediction accuracy of the application of the neural network algorithm to the prediction of graduation of management students at Bhayangkara University, Jakarta Raya University with a sample of the 2014 and 2015 batches were 98.27%.

With the results obtained from this study, the accuracy can be improved by reducing the number of

parameters and increase the number of input nodes and layers in the next study. The discussion and application of a comprehensive feature model is expected to increase the accuracy value which is getting closer to 100% in subsequent studies.

2. CONCLUSION

Based on the discussion from the previous chapter, it was concluded that the predicted results showed results that illustrated that students' graduations tend to be late because the dominant training data used has a late trend, in addition to the number of inputs such as input nodes, training cycles, learning levels, momentum, number of hidden layers, and the number of hidden nodes greatly affect the accuracy of neural network algorithms. The results of research conducted and the results of the application of neural network algorithms, the most influential parameter is the IP (Achievement Index) of students. With the calculation of predictions of the accuracy of the application of neural network algorithms to the graduation predictions of management students at Bhayangkara University, Universitas Jakarta Raya with batch samples in 2014 and 2015 was 98.27%. In this study also obtained accuracy that can be improved by reducing the number of parameters and increasing the number of input nodes

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

and layers in subsequent studies. Discussion and implementation of comprehensive feature models is expected to increase the accuracy value closer to 100% in future studies.

3. REFERENCES

- Febrivani, E., & Winanjaya, R. (2021). *Penerapan Data Mining Asosiasi Pada Persediaan Obat*. 3(3), 354–365.
- Fricles A Sianturi, Hasugian, Paska Marto, Simangunsong Agustina, N. B. (2019). Data Mining [Teori dan Aplikasi Weka. In -: *Vol. (Issue)*.
- Ginting, S. L. B., & Trinanda, R. P. (2014). Penggunaan Metode Naïve Bayes Classifier Pada Aplikasi Perpustakaan. *Expert Systems with Applications*.
- Jefika, M., Kosasi, H., Prayogi, G., & Dharma, A. (2020). *Prediksi Gelombang Corona Dengan Metode Neural Network*. 3(2), 102–107.
- Kusnadi, R., & Jannah, M. (2021). *Penerapan Jaringan Syaraf Tiruan Untuk Memprediksi Jumlah Bagasi Lion Air dan Batik Air di Bandar Udara Kualanamu Dengan Metode Backpropagation*. 3(3), 293–302.
- Meilina, P. (2015). Penerapan Data Mining dengan Metode Klasifikasi. *Jurnal Teknologi Universitas Muhammadiyah Jakarta*.
- Mukminin, A., & Riana, D. (2017). Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah. *Jurnal Informatika*.
- Purba, W., Siawin, W., & Hardih. (2019). Implementasi Data Mining Untuk Pengelompokan Dan Prediksi Karyawan Yang Berpotensi PHK Dengan Algoritma K-Means Clustering. *Jurnal Sistem Informasi Ilmu Komputer Prima (JUSIKOM PRIMA)*, 2(2), 85–90.
- Singh, H., Giawa, M. M., Zulfa, N., & Simbolon, Y. (2020). *Model Jaringan Syaraf Tiruan Dalam Pengenalan Penyakit Asam Lambung*. 2(1), 39–42.
- Sulastri, H., & Gufroni, A. I. (2017). Penerapan Data Mining Dalam Pengelompokan Penderita Thalassaemia. *Jurnal Teknologi Dan Sistem Informasi*, 03(02), 299–305. <https://doi.org/10.25077/TEKNOSI.v3i2.2017.299-305>
- Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. In *Data Mining: Practical Machine Learning Tools and Techniques*. <https://doi.org/10.1016/c2009-0-19715-5>
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2013.109>

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.