

Developing Study Program Choice by Using Neural Network

Arief Hermawan¹, Dian Yudhawati², Donny Avianto³, Fiandari Nurul Hidayah⁴
^{1,2,3,4}University of Technology Yogyakarta, Indonesia

ABSTRACT: In this research, a system that can predict a proper university study program for student has been developed. To build the system, a questionnaire was spread to 180 college students from 6 different study programs who have a good academic achievement. These questionnaires were intended to measure their academic competence. Then, answers of questionnaire were calculated to obtain the final score. Afterwards, half of obtained scores, were used for input data on neural network training process. The network training process has been done using back propagation algorithm. After training process was completed, it is known that resulted system can predict a proper study program for student with 100% accuracy rate when using training data and 67% accuracy rate when using untrained data.

Keywords: college student, dimensional reduction, neural network, stress resistance

I. INTRODUCTION

Nowadays, consistency between educational background and position that will be applied become a considered factor for someone to be hired or not. It aims to get prospective employees who are competent in their field. A survey by [1] which aims to determine what top entry employers want the most from potential employees was done in 2008. Based on those survey result, student's study program become the top of the list with total number of voters reaching 44%. This result shows that election of study program in university cannot be underestimated and can be a crucial event in student's life.

Another challenge that faced by students when choosing a study program is how to choose a proper university study program for them. Ideally, when a student has to choose his study program in university, he should choose a study program that fit with his potential and interest. This is because students usually give their maximum effort and full commitment when they were in the study program that fit with their personalities. On the other hand, according to [2], poor choice of study program would make them perform badly or decide to withdraw. Consequently, it is important for the students to choose a study program in university that suit with their personalities because the chosen programs are closely correlated to their career paths in the future [3].

However, many students are still don't know the proper study program for them when the time to choose is coming. This situation causes greater mistake possibility for student when choosing a study program. Moreover, there are many students that choose a study program based on their friends or parent suggestion without looking deeper on their own potential. If the suggestion is not aligned with student's potential, it often makes student confused and very hard to make a right decision. Therefore, a system that can predict a proper study program for students based on their potential were developed.

On this research, system was built using Artificial Neural Network (ANN) to give a university study program prediction. In order to give an accurate prediction, system must be trained until achieving some targets e.g. error tolerance and maximum epoch as defined before training started. Machine learning algorithm that used to build the system was Backpropagation Algorithm. To see a study program prediction result, student must answer some questions first. These questions are taken from test of potential academic which can measure the potential of a student in academic field. After then student has answered all of the questions, system will calculate test score and carry it on as an input data for input layer of neural network. This system is expected to be a tool for students who are still confused in determining their study program in university.

II. RELATED LITERATURES

Artificial Neural Network (ANN) is an information processing system that has been developed as a generalization of the mathematical model of human learning system [4]. Commonly, an ANN needs to be trained to make all the input data yielded an output as decided before [5]. Today, an ANN can be trained to solve problems that are difficult for conventional computers or even human beings [6]. Because its capabilities,

ANN has been used in various field such as recognition, identification, classification, control systems, and including prediction. An ANN prediction model can be trained to perform well in environments similar to where the training data are collected [7].

Prediction using an ANN not only commonly applied in finance and business field but also in educational field. Numerous research using ANN was conducted in various level of education. In a high school level, some researches have been conducted by [8] and [9]. On [8], a neural network based system was built to predict students' performance in their final exams. Data that be used as input for neural network are first grade high scores and answers of educational talent and motivation questions. A similar research also was done by [9]. However, in that research, the system was used for predicts students' performance in final exams only for a math course. To decided input data for neural network, a descriptive survey had been held on 192 students in third grade secondary school. From that survey, it has known some factors that influencing student's performance on math course final exams were level of parent's education, parent's occupation, family economic status, and number of family member. The research resulted a ANN model using 6 neurons in the input layer, and 3 neurons in the hidden layer and 1 neuron in the output layer that which can predict score for math course final exam with accuracy up to 95%.

Research about the use of ANN for predicting also conducted in higher level education such as university. In this level, ANN based prediction can be divided into 2 categories. The first category is prediction for students who has registered in university or usually called a college student. On this category, performance and achievement of college student were things which common to be predicted. Like the research that has been done by [10] and [11]. In those research, ANN was used to predict students' success in e-learning courses. System that was built can give 2 output groups i.e. group of students who successfully pass the course and group of students who failed to pass the course. This research very helpful for instructors in order to address specific needs from each group and change their teaching method accordingly. Experiment's result from [10] showed that students' grade prediction is possible to be done at early stage which in this case was the third week of the course. Furthermore, ANN in yielded system can predict students' grade accurately up to 92%. While experiment's result from [11] showed that the most powerful predictors of course outcome were the numbers of messages posted by the students and the contributions they made in team content creation projects. Accuracy rate which yielded in this research was higher than [10], reached 90%.

For the second category, prediction is subjected to high school graduate who had plan to continue in university or usually called prospective student. In this category, ANN is usually used to predict the most properly university study program for prospective student or whether a study program fits to prospective students. As researches has been done by [12] and [13]. Research by [12] was conducted at Faculty of Engineering and Information Technology in Gaza. In that research a Multilayer Perceptron ANN was developed and trained using data of five generations of graduates from the faculty. ANN that was built in this research needs 10 variables of input such as: high school score, type of high school, location of high school, student's gender, and some lessons' score. The ANN can return 4 type of prospective student's suitability level i.e. excellent, very good, good, and poor. Based on the experiment's result, it has known that ANN was able to predict suitability level of prospective student with accuracy rate reached 84.6%. Similar research was also done by [13] in University of Ibadan, Nigeria. On that research, ANN model, based on the Multilayer Perceptron Topology, was developed and trained using graduates' data from Department of Engineering's network input, was used 10 types of data which obtained from students' application / record cards like age of student, parents educational status, type and location of high school, etc., student's gender. Unlike in [12], output on this research only has 3 types i.e. good, average, and poor. After training and testing for the ANN, it has known that ANN can predict suitability of prospective student with accuracy rate up to 74%.

From earlier researches [12] and [13], it can be assumed that ANN prediction in second category is relatively more complex and more challenging. It can be seen from the number of the output which usually more than 2 types. Form researches above, it has also known that ANN with Multilayer Perceptron Topology can provide promising results. Therefore, in this current research, ANN will be used again for predicting on the second category. However, unlike the previous researches, now ANN will be given a prediction about the most properly study program for prospective student. The number of study program that can be options in prediction result comes from one faculty i.e. faculty of business and information technology. To make a prediction, ANN needs input data which is obtained from academic potential test.

III. RESEARCH METHOD

In this research, an ANN was built to predict the most properly study program for prospective students who want to continue their studies to university. This research was conducted at the University of Technology Yogyakarta, Indonesia. As a limitation, ANN that was built can only predict prospective students who want to registered in Faculty of Business and Technology Information. The faculty was chosen because its high interest of prospective students who want to register. Faculty of Business and Technology Information has 6 study

programs i.e. management, accounting, informatics engineering, and information system that belong to degree program and accounting and informatics management that belong to diploma program. Therefore, the six study programs are the possibility options for ANN output. The complete flow of this research can be seen in Fig.1 below.

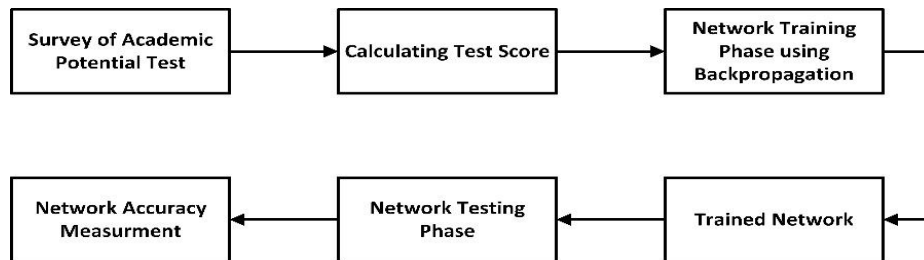


Figure 1. Process flow of current research

At beginning of the research, a survey was held using academic potential test. This survey was aimed to collect data about college students’ academic potential. Therefore, a questionnaire that contains academic potential test was distributed to 180 college students from 6 different study programs (each of 30). Prerequisite for college students who became a participant were at least in the 4th semester and have a GPA greater than or equal to 3.5 (scale of 4.0). Used assumption is if students meet the specified prerequisite, then they are considered to have chosen a study program that proper with their academic potential. The test itself consists of 60 multiple choice questions. These questions are divided into 3 categories i.e. verbal (question no. 1-30), numerical (question no. 31-40), and graphical pattern (question no. 41-60).

After all participant finished their test, the next step was calculating the test score. Because questions of the test are divided into 3 categories, the score calculation process also resulted 3 categories of score i.e. verbal score, numerical score, and graphic pattern score. These scores were then saved digitally in database and would become an input for artificial neural network.

In the next step, an artificial neural network was built and ready for training. The network has 3 input nodes which is can receive verbal score, numerical score, and graphic pattern score respectively. Furthermore, the network has a hidden layer which is contain 70 nodes and an output layer which contain 3 nodes. The nodes in output layer was used to encode all study program that possible to be resulted by network. A description of each code can be seen in Table 1.

Table 1. The encoding of study program

Study Program Code	Study Program Description
001	Accounting (Degree Program)
010	Management (Degree Program)
011	Informatics Engineering (Degree Program)
100	Information System (Degree Program)
101	Accounting (Diploma Program)
110	Informatics Management (Diploma Program)

After every code for all study program were determined, network training phase can be started. In the training phase, network used half of test score data that collected in earlier step. Network training has been done using backpropagation algorithm. As a result of this step, a network with best weight values in order to make a study program prediction was resulted.

To examined the performance of resulted network, two types of testing were conducted in network testing phase. The first network testing was conducted using the same data that used in network training phase. This test was aimed to verify the resulted network and making sure that the network was successfully trained with certain configuration of learning rate, maximum epoch, and error tolerance. On the other hand, the second network testing was conducted with purpose to explore resulted network capability in predicting a proper study program for prospective student. For that reason, another half of test score data were used in this test. Every network’s output in this step will be stored in database. In the last step, all of output from network testing will be matched to its original condition. This step was aimed to analyzed network capability and measured neural network accuracy rate in order to predicting study program.

IV. RESEARCH FINDING AND DISCUSSION

In the network training phase, 90 data which came from academic potential test results were used as input of artificial neural network. The data was then trained using backpropagation algorithm. During this phase, several configurations of the network training has been tested to found the best values for network weight. Screenshot of the system that was built in this phase can be seen in Fig.2 below.

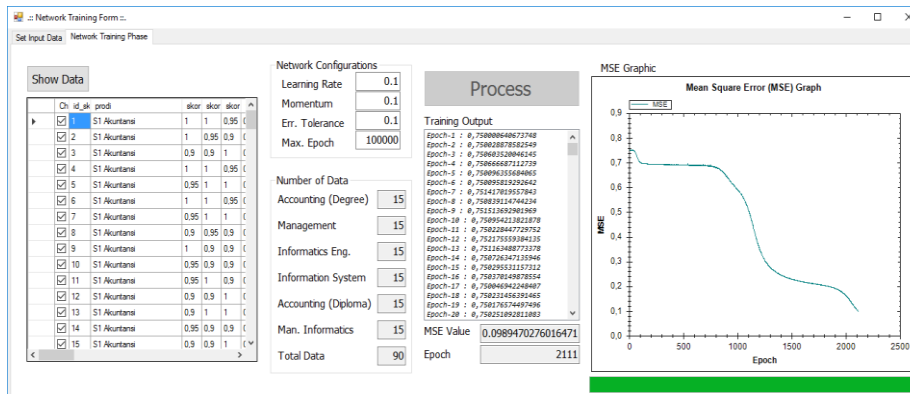


Figure 2. Screenshot for training phase

On the image above, graphical user interface that designed for training phase are shown. On the left side there is “Show Data” button which can show all test result data stored in database. In this part, user can manually choose data that want to be used in network training. Right on the next of this part, there are 2 parts i.e. “Network Configurations” and “Number of Data”. The network configurations part consists of learning rate, momentum, error tolerance, and maximum epoch that the value can be changed before network training is started. While on the number of data part, is shown summary of the amount of data that chosen as a training input. To start the network training process, user can click on the “Process” button in the middle. After network training process is completed, system will return 2 values i.e. Mean Squared Error (MSE) and epoch. Besides that, MSE that achieved during training process will also be shown at the right side after network training process is completed. Based on the experiment, it has known that best weight values for network were obtained with learning rate = 0.1; momentum = 0.1; and been trained up to 2111 epochs.

After trained network was obtained, the next step is testing it. In this step, 2 data set are used as an input of the network. The first data set contains 90 data of test result which has been used as input in training phase, so this set is named training set. Whereas, second data set contains another 90 data which never been used before and it is named testing set. Screenshot for network testing phase can be seen in the Fig. 3.

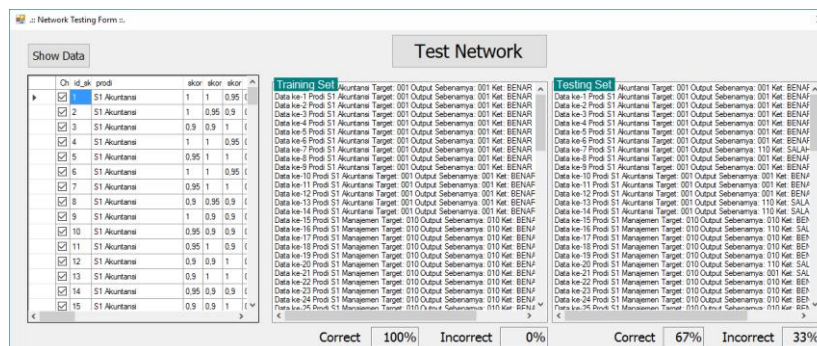


Figure 3. Screenshot for testing network

When “Test Network” button is clicked, trained network will be tested using training set and testing set. After testing process is completed, system will calculate the accuracy rate of system based on output from the network. The calculation result will be shown in textboxes in bottom side. Based on the experiment, artificial neural network that was built can predict all training set with 100% accuracy rate. However, when the network was tested using testing set, it only can predict with accuracy rate up to 67%. This may be caused by the existence of some survey participants which have same academic potential score result although they came from different study programs. This situation can make a confusing situation for the system during neural network training process and leads to decreasing of network accuracy rate.

In addition, system that was built also have a feature to start an academic potential test which is very useful for prospective students. Since academic potential test has 3 categories of question, the system also had 3 different appearances when the test was held. Sample for each appearance of the system can be seen in Fig. 4 through Fig. 6 below. After the prospective students had answered all questions of the test, system will automatically calculate scores for each category and shows the most appropriate study program for them. The capture of system that shows study program prediction as shown in Fig. 7. The user interface for this feature was made so easy to understand. Seven prospective student who has very basic computer competence can quickly adapt and understand how to use it.

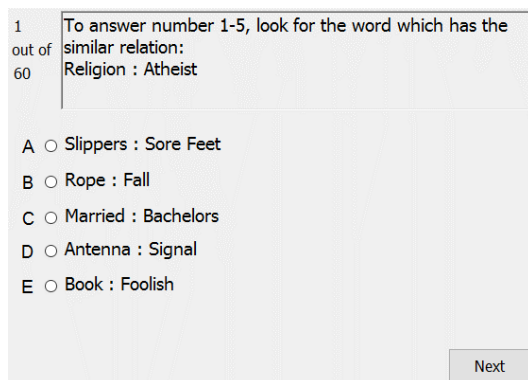


Figure 4. UI for verbal question

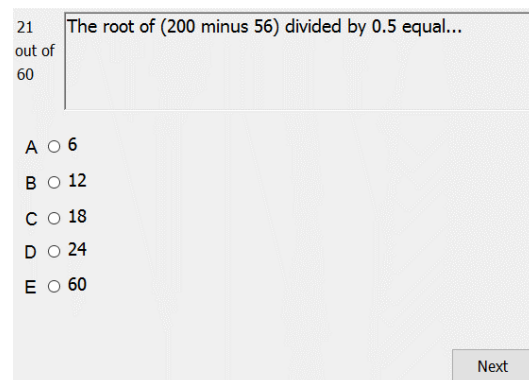


Figure 5. UI for numerical question

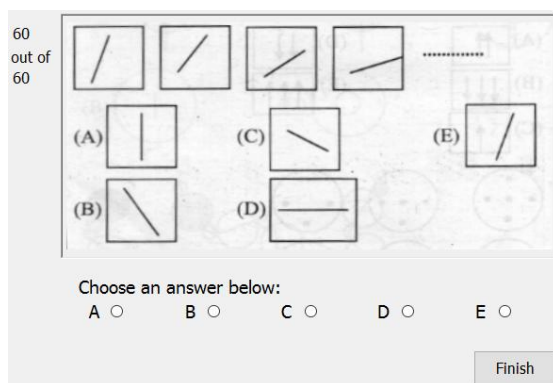


Figure 6. UI for graphic pattern question

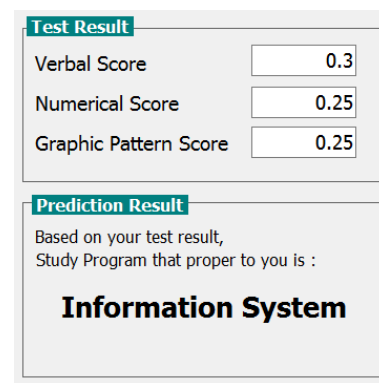


Figure 7. UI for prediction result

Figures above are screenshots that showing feature from resulted system. System can handle since the student start his academic potential test until prediction are given.

V. CONCLUSION AND FUTURE WORK

In this research, system that can predict a proper study program for prospective student has been built. The system can give 6 options of programs study as a prediction. Those study program are accounting, management, informatics engineering, information system, accounting (diploma program), and management informatics which came from 1 faculty. System that built used artificial neural network and backpropagation algorithm as a machine learning algorithm. The artificial neural network has 3 nodes in input layer, 4 nodes in hidden layer, and 3 nodes in output layer. In the input layer, every node will receive 3 types of score that resulted from academic potential test i.e. verbal score, numerical score, and pattern graphic score. The network has been trained with configuration of learning rate = 0.1 and configuration of momentum = 0.1. After 2111 epochs, the training was stop and yielded MSE = 0.0989. Trained network then was tested with 2 data set i.e. training set and testing set. The network can predict all data correctly when being tested using training set. When used a testing set, neural network that was built can predict correctly with accuracy rate up to 67%. System that was built also had a feature to held academic potential test through the system which is very helpful for prospective student.

In the future, some variables such as: student's gender, parent's occupation, and student's hobby can be added as an input of the network. The addition of variables is aiming to significantly improve neural network accuracy rate. Furthermore, the pattern of academic potential test answers can be used as input instead of test scores. Pattern of academic potential test answers is indicated have a high level of uniqueness so confusing situation during training phase can be avoided. It is also expected that answers from each study program have its own pattern. So the pattern of potential test answers can be used to predict study program with higher accuracy rate. Therefore, system that was built can very helpful for prospective students to avoid them from negative effect of poor study program choice.

REFERENCES

- [1] T. Hailey, The Factors Human Resource Managers Use to Determine Whether They Are Going to Hire a College Graduate and How They Communicate Their Expectations and Qualifications, *The University of Maryland McNair Scholars Undergraduate Research Journal*, 1(1), 2008, 143-154.
- [2] H. Christie, M. Munro, and T Fisher, Leaving University Early: Exploring The Differences between Continuing and Non-Continuing Students, *Studies in Higher Education*, 29(5), 2004,617-636.
- [3] N. Misran, S.N.S. Sahuri, N. Arsad, H. Husain, W.M.D.W. Zaki, and N.A. Aziz, Malaysian Matriculation Student's Factors in Choosing University and Undergraduate Program, *Asian Social Science*, 8(16), 2012, 222-231.
- [4] M. Patel, D. Mehta, P. Patterson, and R. Rawal, Applying Back Propagation Algorithm for Classification of Fragile Genome Sequence, *IOSR Journal of Computer Engineering*, 18(5), 2016, 1-10.
- [5] A. Hermawan, *Jaringan Saraf Tiruan* (Yogyakarta, ID: Andi Publisher, 2006).
- [6] A.A. Ajasa, L.A. Akinyemi, O.O. Shoewu, and A.A. Adenowo, Exploitation of Artificial Neural Networks Approach to Predict the Thermal Conductivity of Food Products in Nigeria, *American Journal of Engineering Research (AJER)*, 3(3), 2014, 22-29.
- [7] J.M. Mom, C.O. Mgbe, and G.A. Igwue, Application of Artificial Neural Network for Path Loss Prediction in Urban Macrocellular Environment, *American Journal of Engineering Research (AJER)*, 3(2), 2014, 270-275.
- [8] S.M. Mazinani, and S.A. Abolghasempur, Prediction of Success or Fail of Students on Different Educational Majors at the End of the High School with Artificial Neural Networks Methods, *International Journal of Innovation, Management and Technology*, 4(5), 2013, 461-465.
- [9] M. Sadat, A.T. Ashlaghi, and R. Radfar, Predict The Success of Students in Math's Course in Final Exams in Arak City with Neural Networks, *International Journal of Information, Security and System Management*, 4(1), 2015, 412-417.
- [10] I. Lykourantzou, I. Giannoukos, G. Mpardis, V. Nikolopoulos, and V. Loumos, Early and Dynamic Student Achievement Prediction in E-Learning Courses Using Neural Networks, *Journal of The American Society for Information Science and Technology*, 60(2), 2009, 372-380.
- [11] N.Z. Zacharis, Predicting Student Academic Performance in Blended Learning Using Artificial Neural Networks, *International Journal of Artificial Intelligence and Applications*, 7(5), 2016, 17-29.
- [12] S.A. Naser, I. Zaqout, M.A. Ghosh, R. Atallah, and E. Alajrami, Predicting Student Performance Using Artificial Neural Network: in the Faculty of Engineering and Information Technology, *International Journal of Hybrid Information Technology*, 8(2), 2015, 221-228.
- [13] V.O. Oladokun, A.T. Adebajo, and O.E. Charles-Owaba, Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course, *The Pacific Journal of Science and Technology*, 9(1), 2008, 72-79.