World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:8, 2014

Classifying Students for E-Learning in Information Technology Course Using ANN

S. Areerachakul, N. Ployong, S. Na Songkla

***Abstract*—**This research's objective is to select the model with most accurate value by using Neural Network Technique as a way to filter potential students who enroll in IT course by Electronic learning at Suan Suanadha Rajabhat University. It is designed to help students selecting the appropriate courses by themselves. The result showed that the most accurate model was 100 Folds Cross-validation which had 73.58% points of accuracy.

***Keywords*—**Artificial neural network, classification, students.

Open Science Index, Computer and Information Engineering Vol:8, No:8, 2014 publications.waset.org/9999134/pdf

I. Introduction

THE current curriculum in Suan Sunandha Rajabhat University includes Information Technology course (IT) into Global Education which all students in every faculty must enroll. Lecturers in IT are all assigned to teach this course. But the problems are the timing in classroom and shortage of numbers on teachers. The course lasts 3 hours per each period, so it is hard to set up enough classrooms to support all students.

From that challenge, a new way to filter students into IT course learning is created. An effective way to study this course can be done by self-study. Young generation love to perform self-study because they believe they have the potentials to do it alone. Moreover, students spend a lot of time using electronic devices after school to do assignments and research.

Therefore, the filter system will select students who will benefit most by performing self-study for this course. This model will also help with setting up classrooms' size and timing to support particular number of students. Moreover, it can help saving cost of expenses of the campus, such as overtime wages for teachers and officers. Not only saving for campus, students also spend less money on the course while receiving the same achievement.

Research team selected different factors to be the independent variables for creating the model with the IT final grades of students as the dependent variables. Weka Program was used as the model creator. Then the model was compared with other models created by Neural Network to find out the one with best accuracy ability. This final model will be used with new students for next semester. Neural Network was chosen in this research as it is the most accurate model and popular for data mining.

This paper is organized as follows: Section II describes the related work concerned this research. Section III demonstrated the methodologies used in the experiments. The experiments and results are shown in Section IV. Finally, Section V is the conclusion of the paper.

S.Areerachakul is with the Faculty of Science and Technology, Suan Sunandha Rajabhat University, Dusit, Bangkok 10300 Thailand (phone: 662-160-1226; e-mail: sirilak.ar@ssru.ac.th, sirilak.ar@gmail.com).

N.Ployong is with the Information Technology Center, Suan Sunandha Rajabhat University, Dusit, Bangkok 10300 Thailand (phone: 662-160-1231; e-mail: nat.pl@ssru.ac.th, taeployong@gmail.com).

S.Na Songkla is with the Faculty of Industrial Technology, Suan Sunandha Rajabhat University, Dusit, Bangkok 10300 Thailand (phone: 662-160-1231; e-mail: nat.pl@ssru.ac.th, taeployong@gmail.com).

II. Related Work

Data mining is very popular among commerce, marketing, financial and other types of business. As well as education field, data mining has been brought to use as data analyzing for search engine, such as students' record, GPA, etc. This kind of data wasn't used much in the past. Common techniques to be used for data mining in education are searching for association rule, data categorizing and data prediction. Predicting student graduation rate in institutes of higher education is of great value to the institution and an enormous potential utility for targeted intervention. During the past decade a number of researchers applied various methodologies in order to predict enrollment rates, persistence rates, and/or graduation rates [1]. Naciye Hardalaç et al. [2] demonstrate that machine learning can be used to classify students who had backgrounds in positive sciences (including engineering, science and math disciplines) vs. social sciences (including arts and humanities disciplines) by the help of musical hearing and perception using artificial neural networks. Dominic-Palmer Brown et al. [3] investigates the application of the snap drift neural network (SDNN) to the provision of guided student learning in formative assessments. SDNN is able to adapt rapidly by performing a combination of fast, convergent, minimal intersection learning (snap) and Learning Vector Quantization (drift) to capture both precise sub-features in the data and more general holistic features. Prediction of academic students, for improving student achievement and improve teaching quality of teachers is important. Xianmin Wei [4] adopts forward neural network model, using the Levenberg-Marquardt algorithm to calculate the optimal weights of model, to achieve the mapping of impact on the behavior of normal university students academic. Valquíria R. C. Martinho et al. [5] present school dropout is one of the most complex and crucial problems in the field of education. It permeates the several levels and teaching modalities and has generated social, economic, political, academic and financial damage to all involved in the educational process. Therefore, it becomes essential to develop efficient methods for prediction of the students at risk

Open Science Index, Computer and Information Engineering Vol:8, No:8, 2014 publications.waset.org/9999134/pdf

of dropping out, enabling the adoption of proactive actions to minimize the situation. Thus, this work aim to present the potentialities of an intelligent system developed for the prediction of the group of students at risk of dropping out in higher education classroom courses. The system was developed using a Fuzzy-ARTMAP Neural Network, one of the artificial intelligence techniques, which makes the continued learning of the system possible. This research was developed in the technology courses of the Federal Institute of Mato Grosso, based on the academic and socioeconomic records of the students. Xiujuan FAN et al. [6] was respective advantages of the fuzzy analysis and neural network with respect to evaluation are adopted herein to establish the fuzzy neural network model for comprehensive quality evaluation on college students.

III. Methodology

In this section, we demonstrated Multilayer perceptron neural network.

A. Multilayer Perceptron Network

The artificial neural network (ANN) or neural network in short, is inspired by simulating the function of a human brain. A neural network can be used to represent a nonlinear mapping between input and output vectors. Neural networks are among the popular signal-processing technologies. In 0engineering, neural networks serve two important functions: as pattern classifiers and as nonlinear adaptive filters [7], [8]. A general network consists of a layered architecture, an input layer, one or more hidden layers and an output layer [9]. The Multilayer perceptron (MLP) is an example of an artificial neural network that is used extensively to solve a number of different problems, including pattern recognition and interpolation [10], [11]. Each layer is composed of neurons, which are interconnected with each other by weights. In each neuron, a specific mathematical function called the activation function accepts input from previous layers and generates output for the next layer. In the experiment, the activation function used is the hyperbolic tangent sigmoid transfer function [12] which is defined as in (1):

$$f(n) = \frac{1 - e^{-2s}}{1 + e^{-2s}} \tag{1}$$

where $s_i = \sum_{i=1}^{n} w_i x_i$, in which w_i are weights and x_i are input values.

The MLP is trained using the Levenberg–Marquardt technique as this technique is more powerful than the conventional gradient descent techniques [10].

The Levenberg-Marquardt (LM) algorithm [13] is the most widely used optimization algorithm. It outperforms simple gradient descent and other conjugate gradient methods in a wide variety of problems. If a function $V(x)$ is to be minimized with respect to the parameter vector $\underline{x}$, then Newton's method would be:

$$\Delta \underline{x} = -[\nabla^2 v(\underline{x})]^{-1} \nabla v(\underline{x}) \tag{2}$$

where $\nabla^2 v(\underline{x})$ is the Hessian matrix and $\nabla v(\underline{x})$ is the gradient. If $v(\underline{x})$ reads:

$$v(\underline{x}) = \sum_{i=1}^{N} e_i^2(x) \tag{3}$$

then it can be shown that:

$$\nabla v(\underline{x}) = J^T(\underline{x}) \underline{e}(\underline{x}) \tag{4}$$

$$\nabla^2 v(\underline{x}) = J^T(\underline{x}) J(\underline{x}) + S(\underline{x}) \tag{5}$$

where $J(x)$ is the Jacobian matrix

$$J(x) = \begin{bmatrix} \frac{\partial_{e1}(\underline{x})}{\partial x_1} & \frac{\partial_{e1}(\underline{x})}{\partial x_2} & \cdots & \frac{\partial_{e1}(\underline{x})}{\partial x_N} \\ \frac{\partial_{e2}(\underline{x})}{\partial x_1} & \frac{\partial_{e2}(\underline{x})}{\partial x_2} & \cdots & \frac{\partial_{e2}(\underline{x})}{\partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial_{eN}(\underline{x})}{\partial x_1} & \frac{\partial_{eN}(\underline{x})}{\partial x_2} & \cdots & \frac{\partial_{eN}(\underline{x})}{\partial x_N} \end{bmatrix} \tag{6}$$

and

$$s(\underline{x}) = \sum_{i=1}^{N} e_i \nabla^2 e_i(\underline{x}) \tag{7}$$

For the Gauss-Newton method it is assumed that $s(\underline{x}) \approx 0$, and (2) becomes:

$$\Delta \underline{x} = \left[J^T(\underline{x}) J(\underline{x})\right]^{-1} J^T(\underline{x}) \underline{e}(\underline{x}) \tag{8}$$

The Lavenberg-Marquardt modification to the Gauss-Newton method is:

$$\Delta \underline{x} = \left[J^T(\underline{x}) J(\underline{x}) + \mu I\right]^{-1} J^T(\underline{x}) \underline{e}(\underline{x}) \tag{9}$$

The parameter μ is multiplied by some factor (β) whenever a step would result in an increased $V(\underline{x})$.

When a step reduces $V(\underline{x})$, μ is divided by β. When the scalar μ is very large the Levenberg-Marquardt algorithm approximates the steepest descent method. However, when μ is small, it is the same as the Gauss-Newton method. Since the Gauss-Newton method converges faster and more accurately towards an error minimum, the goal is to shift towards the Gauss-Newton method as quickly as possible. The value of μ is decreased after each step unless the change in error is positive; i.e. the error increases. For the neural network-mapping problem, the terms in the Jacobian matrix can be computed by a simple modification to the back-propagation algorithm [14].

Open Science Index, Computer and Information Engineering Vol:8, No:8, 2014 publications.waset.org/9999134/pdf

IV. Expreiment and Results

A. Data Selecting

To create the filter model for this research, raw data would be set to be ready for inputting into Weka program to make the model that covers all sample groups. There are 4,955 sets of information from IT course final result, in GE courses and general admission score of students in 2011. The data consisted of 11 factors 1) final scores of IT course 2) students' GPA from their secondary schools 3) students' faculties 4) students' scholarships status 5) students' diploma 6) father's salary 7) mother's salary 8) personal health condition 9) physical conditions 10) parents' marital status and 11) students' grade in IT.

There were also missing data in each factor, so we decided to deleted missing data in factor 1, 2 and 11 since data in these 3 factors can't be replaced. Eventually, there were 2,969 sets of data left. However, there were still missing data in parents' salary, so we had to replace the missing information and delete them before comparing the effect on accuracy.

Replacing missing data was done by inputting data in ordinal scale and mode value. The mode value of fathers' income "under 12,500 baht per month" was input into 762 sets of data. As well as mother's income, the same "under 12,500 baht per month" was input into 525 sets. After replacing missing data, there were 2,969 sets left, and they were kept separately for the next step. Next, missing data of father's and mother's salary were deleted. 1,682 sets were left. After preparing 2 sets of data, they were tested to find out whether they were related to the accuracy as the whole. Then the sets of data with replacing data and the one without were compared to find out the effect on accuracy in order to select the best one for model creating. The most accurate model was chosen as the classifying filter.

B. Result

The analysis of related factors by Neural Network technique using the 10 folds Cross-validation check showed the satisfying result that all factors were related and affected the accuracy value. The accuracy value would be decreased when some factors were cut out, so all factors should stay for the most accurate value. The result displays in Table I.

The comparison of replacing data and deleting data to find the effect on accuracy value by Neural Network had impacted on all models around 0.88%, which is in the acceptable level. The result displays in Table II.

Therefore, the set of data with replacing data were selected to be the main sets because they consisted of more information than the missing one. Result displays in Table III.

The result has shown that the most accurate model for analyzing with Neural Network is the 100 folds Cross-validation that has 73.58% of accuracy. As a result, it was chosen to create the filter model.

TABLE I
Result of Factors Analysis by Using Neural Network and 10 Folds Cross-validation Check

Factors	Accuracy of Neural Network (%)
All factors	**71.80**
General admission test	70.22
Secondary school grade	71.30
Faculties	68.10
Scholarship Status	71.47
Diploma	71.53
Father's income	71.75
Mother's income	71.02
Health condition	70.07
Physical condition	71.38
Parents' marital status	69.89

TABLE II
The Comparison of Replacing Data and Deleting Data to Find the Effect on Accuracy Value by Neural Network

Model	Replacing data accuracy (%)	Deleting data accuracy (%)	Discrepency
5 Folds	72.88	73.12	0.24
10 Folds	71.80	72.17	0.37
100 Folds	73.58	73.30	0.28
10 %	71.03	68.16	2.87
20 %	68.75	70.35	1.60
66 %	73.24	73.07	0.17
Separated data into 80/20	72.05	71.42	0.63

TABLE III
Data Analysis by Neural Network

Model	Accuracy (%)
5 Folds	72.88
10 Folds	71.80
100 Folds	73.58
10 %	71.03
20 %	68.75
66 %	73.24
Separated data into 80/20	72.05

V. Conclusion

100 Folds Cross-validation Model analyzing by Neural Network had the most accuracy value compared to other sets because all data were divided into 100 portions and used alternately between teaching and testing. That's why it was more accurate compared other sets.

But the highest accuracy value was only 73.58% because of missing some raw information from the start. However, this research has some suggestions in terms of data collecting. If there are any information factors which are automatically collected without filling in paper form, it can be more accurate. Manual information filing can be missed or incorrect, and it will affect the accuracy of the model.

For GPA sections, the data could be divided only into 3 sections. If it could be divided into more sections, it could also support the work of the model.

The comparison between replacing data and missing data showed only slight differences. Various sets of data for future research will be supporting researchers to get better result.

ACKNOWLEDGMENT

The authors would like to thank Suan Sunandha Rajabhat University for scholarship support. Thanks to the Information Technology Center at Suan Sunandha Rajabhat University, for the provided data.

REFERENCES

[1] Stamos T. Karamouzis, "Sensitivity Analysis of Neural Network Parameters for Identifying the Factors for College Student Success", *Journal of 2009 World Congress on Computer Science and Information Engineering*, USA, April 2009, pp. 671-675.

[2] Naciye Hardalaç, Nevhiz Ercan, Fırat Hardalaç and Salih Ergüt, "Classification of Educational Backgrounds of Students Using Musical Intelligence and Perception with the Help of Artificial Neural Networks", *Journal of 36th ASEE/IEEE Frontiers in Education Conference*, USA, October 2006, pp. 9-14.

[3] Dominic-Palmer Brown, Chrisina Draganova and Sin Wee Lee, "Snap-Drift Neural Network for Selecting Student Feedback", *Proceedings of International Joint Conference on Neural Networks*, USA, June 2009, pp. 391-398.

[4] Xianmin Wei, "Student Achievement Prediction Based on Artificial Neural Network", *Journal of 2011 International Conference on Internet Computing and Information Services*, Hong Kong, September 2011, pp. 485-487.

[5] Valquíria R. C. Martinho, Clodoaldo Nunes and Carlos Roberto Minussi, "An Intelligent System for Prediction of School Dropout Risk Group in Higher Education Classroom based on Artificial Neural Networks", *Journal of 2013 IEEE 25th International Conference on Tools with Artificial Intelligence*, USA, November 2013, pp. 159-166.

[6] Xiujuan FAN, Runping HAN and Guifang WANG, "Fuzzy Neural Network Model for Comprehensive Quality Evaluation on College Students", *Journal of 2010 International Conference on Intelligent Computation Technology and Automation*, Chaina, May 2010, pp. 375-378.

[7] D.Marquardt, "*An Algorithm for Least Squares Estimation of Non-Linear Parameter*", J. Soc. Ind. Appl.Math., pp. 1963.

[8] L.Fausett, "*Fundamentals of Neural Networks Architecture.Algorithms and Applications*", Pearson Prentice Hall, USA, 1994.

[9] M.J. Diamantopoulou, V.Z. Antonopoulos and D.M. Papamichail "The Use of a Neural Network Technique for the Prediction of Water Quality Parameters of Axios River in Northern Greece", *Journal 0f Operational Research*, Springer-Verlag, Jan 2005, pp. 115-125.

[10] D. Anguita, S.Ridella and F.Rivieccio, "K-folds Generalization Capability Assessment for Support Vector Classifiers", *Proceeding of International Joint Conference on Neural Network*, Canada, 2005, pp. 855-858.

[11] Li-hua Chen, and Xiao-yun Zhang, "Application of Artificial Neural Network to Classify Water Quality of the Yellow River", *Journal 0f Fuzzy Information and Engineering*, Springer-Verlag, Jan 2009, pp. 15-23.

[12] M.J. Diamantopoulou, V.Z. Antonopoulos and D.M. Papamichail, "The Use of a Neural Network Technique for the Prediction of Water Quality Parameters of Axios River in Northern Greece", *Journal 0f Operational Research*, Springer-Verlag, Jan 2005, pp. 115-125.

[13] S.H.Musavi and M.Golabi, "Application of Artificial Neural Networks in the River Water Quality Modeling: Karoon River,Iran", *Journal 0f Applied Sciences*, A sian Network for Scientific Information, 2008, pp. 2324-2328.

[14] Chi Zhou, Liang Gao and Chuanyong Peng, "Pattern Classification and Prediction of Water Quality by Neural Network with Particle Swarm Optimization", *Proceedings of the 6th World Congress on Intelligent Control and Automation*, China, June 2006, pp. 2864-2868.

Open Science Index, Computer and Information Engineering Vol:8, No:8, 2014 publications.waset.org/9999134/pdf