Indonesian Journal of Electrical Engineering and Computer Science
Vol. 5, No. 3, March 2017, pp. 684 ~ 690
DOI: 10.11591/ijeecs.v5.i3.pp684-690

Using Neuro-Fuzzy Technique to Classify and Predict Electrical Engineering Students' achievement Upon Graduation Based on Mathematics Competency

Usamah bin Mat*, Norlida Buniyamin
Engineering & Technical Education Research Group (EnTER), Faculty of Electrical Engineering,
Universiti Teknologi MARA, 40200 Shah Alam, Selangor, Malaysia
*Corresponding author, e-mail: usamahmat@gmail.com

Abstract

This paper discusses the findings of a case study that uses neuro-fuzzy tool to classify and predict Electrical engineering students graduation achievement based on mathematics competency. In this study, achievement upon graduation and mathematics grades were classified as the key performance index. It's based on longitudinal progress and cross validation model on two mathematics subjects, semesters' performance, and graduation achievement of electrical students. The outcomes indicated that there is a correlation between mathematics competency with electrical engineering performance, and it's interesting to note that weak and satisfactory students in mathematics are not able to achieve first class upon graduation, and yet there is small percentage of excellent and good students in mathematics couldn't graduate with high achievement. The findings conclude that the combination of statistical analysis and machine learning can help us to extract knowledge and enable university management to help low achievers at early stage. It's hoped that the findings can help faculty management to review mathematics curriculum with respect to increasing range of engineering field.

Keywords: *Educational Data Mining, Neuro-fuzzy classification, ANFIS, Mathematics Competency, Students' Achievement*

1. Introduction

Evaluation of an engineering student's learning achievement in any fundamental subject is a way to determine the performance level of the students in relation to learning outcomes. Mathematics as one of the fundamental subjects; electrical engineering students ought to have strong knowledge and skill in mathematics in order to have ability in problem solving and critical thinking. Fluency in mathematics is an essential weapon of a modern graduate engineer and required in large number of engineering courses [1].

Engineering is a profession directed towards the application and advancement of skills based on distinctive knowledge in mathematics, science and technology. Mathematics competency in engineering enables the ability to understand, judge, and use mathematics in variety of engineering context and situations [2], as a result Engineering institutions always strive to provide engineering students with mathematics curriculum that match the current demand of engineering profession.

Researchers are observing how mathematics is important in engineering education; investigating what topics should be taught, what skills students should acquire, how competency is evaluated, etc. In the past decades, a considerable amount of research exists about how engineering students perform in relation with mathematics ability. According to Uysal, insufficient skill in basic mathematics will cause problems for students who are majoring in engineering where the most important skills required of engineering students are problem solving and creative thinking [3]. Other researchers stressed that different learning outcomes of mathematics preparation will affect engineering students performance, and there is a strong need to improve mathematics curriculum [4, 5]. Nevertheless, most of the previous researches were based on statistics figures, surveys and experts opinions.

In addition, students profile has changed and there is now a wide range of engineering field due to increasing areas that requires engineering expertise. Though engineering education

Received November 30, 2016; Revised February 6, 2017; Accepted February 22, 2017

has managed to keep up, but it has become increasingly clear that the large data stored should be utilized to enable improvement of engineering education. Therefore, new approaches are needed to utilize large educational data stored and improve mathematics courses to be relevance to engineering changes. In fact, advance technology and machine learning existed several decades ago for business analysis, while recently education engineering endowed with Educational Data Mining (EDM) [6, 7].

EDM is a discipline that uses data from education setting such as universities and colleges to gain knowledge for better education planning [8]. Kumar used Decision Tree algorithm to predict the course and program outcome [9]. Murugananthan and Shiva Kumar used K-means unsupervised classifiers and processed via cubical structure universal index to propose an automated learning system in EDM [10]. Thakar et al presented a comprehensive survey on performance analysis and prediction in EDM for the last decade stated that "there are large number of factors that play significant role in prediction, suitable technique are required to measure, monitors, and infer these factors for prediction [11].

This paper presents a case study based on analysis of mathematics competencies using Adaptive Neuro-fuzzy inference System (ANFIS) to classify and predict Electrical Engineering students' graduation achievement. It is placed within EDM extracts hidden knowledge from existence data. This study is part of an ongoing research that aimed to create a tool to extract hidden information from data that will enable university management to predict final achievement and improve learning performance by providing early intervention to students that are predicted to achieve poor grades upon graduation.

2. Mathematics In Electrical Engineering, UiTM

The Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) designed the degree to be a four year programme, and provides students with an intensive mathematics education in the first year of learning to prepare students to acquire excellent skills in problem solving and critical thinking [12]. Mathematics_1 (MT1) is design to be taken in Semester_1 (Sem1) along with Electrical Engineering Circuits, Fundamental of Electronics, Signal and System, and communication theory; MT1 consists of trigonometric integrals, trigonometric substitution, and integration of rational function, differential equations, etc, whereas Mathematics_2 (MT2) is designed to be taken in Semester_3 (Sem3) along with Electrical Engineering Material, Signal and System_2 Digital System, and English. MT2 consists of partial derivative, multiple integral, vector analysis, and infinite series, etc.

3. Methodology

Longitudinal progress of mathematics tracked by cross validation performance of Graduation Grade is the methodology used and shown in Figure 1. The methodology focused on the following factors: performance of students in Mathematics, performance of Semesters' CGPA and graduation grade, mathematics correlation with graduation grade, and factors influence academic achievement.

3.1. Longitudinal Progress

A pre-processing and investigation analysis of classification performance of MT1, and parameters measured of MT2 and CGPA performance in Sem1, Sem3 and graduation Achievement (GA). Mathematics performance is classified into four classes to reflect GA set by UiTM grading system [12]. MT1 and MT2 marks were categorized into four classes as shown in Table 1. A Bachelor's (honors) Degree at UiTM is awarded and classified into four graduations class achievement as shown in Table 2.

Table 1. Classification label for mathematics marks

Performance group	Marks
Excellent	75 – 100
Good	65 – 74
Satisfactory	55 – 64
Weak	0 – 54

Table 2. Classification label upon graduation*

Class of Degree	Range of CGPA
First Class	3.50 – 4.00
Second Upper Class	3.00 – 3.49
Second Lower Class	2.20 – 2.99
Third Class	2.00 – 2.19

* Minimum Grade for graduation is 2.00

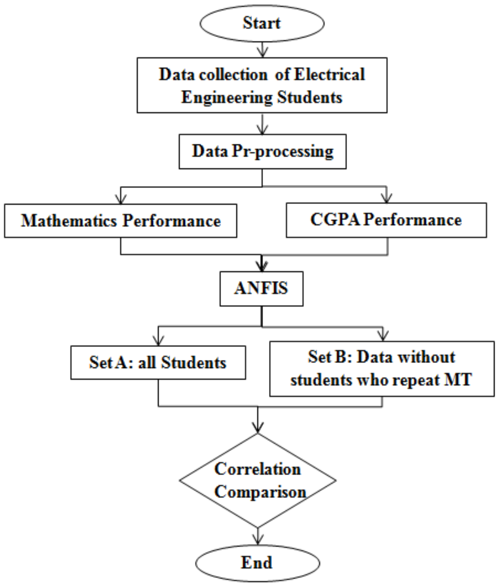

Figure 1. Methodology Flowchart

3.2. Cross Validation

There are many techniques exist for classification and prediction, and selecting a suitable method for a certain task is not easy as there is no generalized rule on selection. Basically, this research needs good support of visual analyses, identifying number of computational analyses, and functional usability, also the research required wide range of platform, open source, and extensive range of algorithm for better data preparation and friendly user interface. As a result, the most suitable technique for this study is ANFIS tool. The model used as multi input and single output system (MISO) [13, 14]. MT1 and MT2 are input parameters and graduation grade is output parameter as shown in Figure 2.

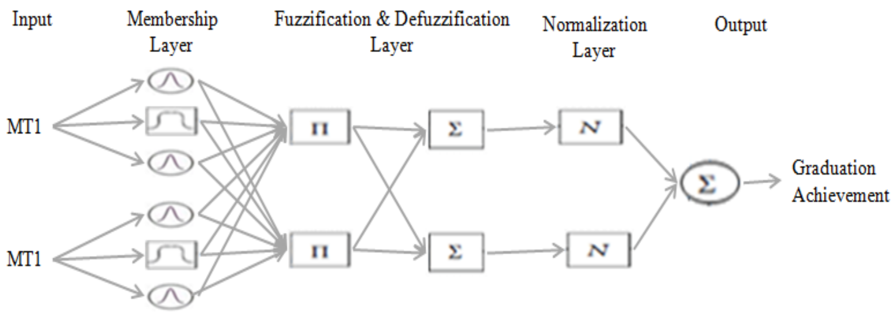

Figure 2. Input Output Diagram

4. Data Processing
4.1. Data Collection
Students' data is usually in the form of numeric or categorical. Numeric value such as test results, and grade point are discreet or can be in form of continuous such as spending time on solving test, or in library [6]. Numeric data was gathered from Students Information Management System (SIMS) at UiTM. The extracted data is the performance grades of electrical engineering student for three successive intakes. Around 12,202 records were collected from 391 electrical students, and for this paper 2346 recorded data of MT1, MT2, and Sem1, Sem3, and Graduation grade were used.

4.2. Data Splitting
Building a computational model with high prediction and generalization capability is main purpose of this forecasting model. In this case, a common splitting for supervised learning method is used called cross-validation. This method will help to build a model with high ability to generalize extracted knowledge and avoid overfitting [15]. Cross validation is to split the data into training and validation data. Training data is a critical set; where machine learns to provide prediction relationships. Whereas validation is to ensure the operation is robust, correct and useful data. Therefore, this research split data into 75% training data, and 25% validation data, and the selection of data is done by Simple Random Sampling (SRS).

5. Data Investigation
Longitudinal analysis of MT1 marks performance is shown in Table 3. A total of 391 students were grouped into four categories of marks. First group (75.00 to 100.00), second group (65.00 to 74.99), third group (55.00 to 64.99), and last group (0.00 to 54.99), and total students are 256, 64, 51, and 20 respectivly.

The third culumn shows MT2 progress distribution of each group. Almost half of MT1 first group performed differently in MT2 as seen there is only 55% who is able to keep their performance in Excellent level, while 45% fall into other lower performance. However, there are improvement of performance from lower group of MT1, as Excellent in MT2 can come from third anf forth groups. On the other hand, Weak students in MT2 come from all MT1 groups.

Last culmun shows GA performance in each MT1 group. In the first group there is 15.23% chieved First Class, while 56.64% achieved Second Upper class, 27.34% second Lower Class, and yet 0.39% graduated with CGPA lower than 2.20. In second group 3.13% only achieved First Class, and almost 50% achieved Second Upper Class. Moreover, none of students in third and fourth group could graduate with First Class award, and there are only 27.45% and 15% achieved Second Upper Class respectively.

Table 3. MT1 performance group

MT1 Performance Marks	Total number of students	Performance achievement in MT2		Graduation Achievement	
from 75.00 to 100.00	256 students*	55.08%	Excellent	15.23%	First Class
		28.13%	Good	56.64%	Second Upper Class
		10.55%	Satisfactory	27.34%	Second Lower Class
		6.25%	Weak	0.39%	Third Class
from 65.00 to 74.99	64 students*	39.06%	Excellent	3.13%	First Class
		32.81%	Good	48.44%	Second Upper Class
		20.31%	Satisfactory	42.19%	Second Lower Class
		7.81%	Weak	4.69%	Third Class
from 55.00 to 64.99	51 students	23.53%	Excellent	0.00%	First Class
		37.25%	Good	27.45%	Second Upper Class
		23.53%	Satisfactory	66.67%	Second Lower Class
		15.69%	Weak	5.88%	Third Class
from 0.00 to 54.99	20 students	15.00%	Excellent	0.00%	First Class
		20.00%	Good	15.00%	Second Upper Class
		50.00%	Satisfactory	80.00%	Second Lower Class
		15.00%	Weak	5.00%	Third Class

* There are some students who fail to graduate

Figure 3 shows average performance of MT1 in four categories based on CGPA of Sem1, Sem3, and GA. Performance of each semester is gradually increased from Weak to Excellent. However, looking into each group separately; there was a slight drop of average performance in Excellent group from Sem1 to Sem3 to GA; from 3.23 to 3.20 to 3.16 respectively. Moreover, there is a fluctuating performance in Weak Student; from Sem1 of 2.18 to increase in Sem3 of 2.72 and then dropped in GA of 2.68 average performance. Good and Satisfactory class experienced gradual increase from Sem1 to Sem3 and GA from 2.78, 2.92, to 2.95, and from 2.45, 2.77, to 2.80 respectively.

Figure 3. Average performance and achievement

6. ANFIS Finding

In this paper two types of data sets were tested. First data set (Set A) contains all students and test the model randomly and without filtering MT performance, whereas second data set (Set B) MT performance was filtered and excluded students who were not able pass MT from first attempt. Minimum training error was recorded with number of iteration, membership function was set up to four classes, and Gaussian membership was selected to identify mean and standard deviation. Backpropagation and least square technique were used in the model.

Table 4. ANFIS training and testing configuration

Training data		
Data	Set A	Set B
Epoch	1 to 61	1 to 49
RMSE for 1^{st} epoch	0.26	0.25
Membership	Gaussian	Gaussian
Type	hybrid	hybrid
Training data	75%	75%
Coefficient of correlation (R)	0.68	0.65
RMSE	0.25	0.24

Table 4 shows ANFIS results for training data. The results show promising model that able to train and produces coefficient of correlation (R) 0.6781 and 0.6471, and RMSE of 0.25 and 0.24, for set A, and B respectively. Therefore, the model is able to learn and understand performance of MT, and correlate it with GA.

Furthermore Table 5 shows validation data results. It can be seen that R is equal to 0.03 and RMSE is equal to 5.99 for Set A, but for R is equal to 0.46, and RMSE is equal to 0.36 for Set B. Therefore, set A experienced over fitting, and sensitivity of validation error. The model for Set A could not predict students' achievement due to irregular data came from unbalance performance of students who repeated MT.

Table 5. Configuration of validation data

Validation data		
Data	Set A	Set B
Coefficient of correlation	0.03	0.46
RMSE	5.99	0.36

As it can be seen also from Table 5; set A and set B iteration error started at similar magnitude of 0.26 and 0.25 respectively. Set B error dropped faster than first set A that implies the good fit model could understand and learns faster than overfitted data even though R is higher in training data, in addition the model shows that Mathematics contain critical features that affect GA.

Figure 4 shows scatter plot of actual and predicted Graduation Grade output for Set B. The good fit model is where validation error is low and slightly higher than training error.

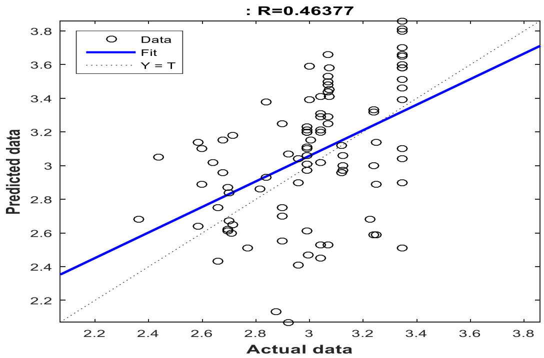

Figure 4. Regression between actual and predicted validation output

Table 6 shows confusion matrix that describe prediction performance of GA for set B where threshold is taken at 0.126, total of true prediction is 75.79%, while total false prediction is equal to 24.21%. True first class predication is 10.53%, and false predictions are 4.21%, and 1.05% for 2nd upper and 2nd Lower class respectively. In 2nd upper prediction; true prediction is the highest with 37.89%, and false prediction is 7.37%. at 2nd lower true prediction is 27.37% and false is 9.74%. and for Third class there is only false prediction with 2.11%.

Table 6. Prediction Accuracy

		Predicted			
		First	2ndUpper	2nd Lower	Third
Actual	First	10.53%	4.21%	1.05%	0.00%
	2ndUpper	0.00%	37.89%	7.37%	0.00%
	2nd Lower	0.00%	9.47%	27.37%	0.00%
	Third	0.00%	0.00%	2.11%	0.00%

7. Discussion

These findings indicate that ANFIS is able to classify and predict Students performance based on mathematics competences. The model experienced overfitting in set A, this sensitivity reflects the characteristic of students' performance as it's different from what the model learned and expected. Therefore, when we retraced the irregular data, we found out that some students were weak in mathematics, but graduated with better achievement. This can be explained positive intervention by faculty with regard to under prepared students.

Moreover, total of 81.84% of excellent and good students in MT1, and 59.84% graduated above 3.00; this can point out that students have good command in Mathematics, and prepared for electrical engineering study. However, the fluctuating CGPA performance from Sem1 to Sem3, and GA indicates that there are other factors may affecting their performance such as the core courses that were taken in the first or third semester.

On the other hand, there is a noticeable decreasing of achievement for excellent and good student in MT1, besides that a couple of students could not graduate as their CGPA is lower than 2.00. These confrontations imply that students may seem to be able to find correct solution to test and exam questions using familiar steps and procedure, yet they lack of deep conceptual understanding.

8. Conclusion

This paper evaluates ANFIS in extracting hidden knowledge from data available from students' management system. The combination of longitudinal progress and cross validation proved the ability of extracting features that affects students' performance based on mathematics competency. We investigated the pattern of mathematics performance and found that students with well conceptual understanding and skill of mathematics will be able to graduate with higher achievement. Also, it is suggested to classify GA to five classes which will break up large number of students in second upper and second lower class into second, third and forth class.

The findings prompted us to further investigate how mathematics affects core courses in engineering, and compare their sensitivity effect on students' achievement upon graduation.

References

[1] L Mustoe, D Lawson. Mathematics for the European engineer. A Curriculum for the Twenty-First Century. A Report by the SEFI Mathematics Working Group. Brussels: SEFI. 2002.

[2] E Łobos, J Macura. Mathematical competencies of engineering students. 2010.

[3] F Uysal. The Mathematics Education For The Engineering Students of 21st Century. *Online J. New Horizons Educ.* 2012; 2: 65-72.

[4] S Firouzian, Z Ismail, RA Rahman, YM Yusof. *Mathematical Learning of Engineering Undergraduates.* Procedia-Social and Behavioral Sciences. 2012; 56: 537-545.

[5] NA Ismail, ZM Nopiah, I Asshaari, H Othman, NM Tawil, A Zaharim. *Mathematical Performance of Engineering Students in Universiti Kebangsaan Malaysia (UKM).* Procedia-Social and Behavioral Sciences. 2012; 60: 206-211.

[6] U bin Mat, N Buniyamin, PM Arsad, R Kassim. *An overview of using academic analytics to predict and improve students' achievement: A proposed proactive intelligent intervention.* In Engineering Education (ICEED), 2013 IEEE 5th Conference on. 2013: 126-130.

[7] ZZ Manseur, A Ieta, R Manseur. *Work in progress—Mathematics preparation for a modern engineering program.* In Frontiers in Education Conference (FIE), 2010 IEEE. 2010: 1-2.

[8] I Saleh, Si Kim. A fuzzy system for evaluating students' learning achievement. *Expert Systems with Applications.* 2009; 36: 6236-6243.

[9] SA Kumar. Edifice an Educational Framework using Educational Data Mining and Visual Analytics. 2016.

[10] V Murugananthan, B ShivaKumar. An adaptive educational data mining technique for mining educational data models in elearning systems. Indian Journal of Science and Technology. 2016; 9.

[11] P Thakar. Performance Analysis and Prediction in Educational Data Mining: A Research Travelogue. 2015.

[12] FoE Engineering. Students' Handbook Faculty of Electrical Engineering. UiTM. 2011.

[13] N Buniyamin, U bin Mat, PM Arshad. *Educational data mining for prediction and classification of engineering students achievement.* In 2015 IEEE 7th International Conference on Engineering Education (ICEED). 2015: 49-53.

[14] UB Mat, N Buniyamin, PM Arshad. Educational Data Mining Classifier For Semester One Performance to Improve Engineering Students Achievement. *Middle-East Journal of Scientific Research.* 2016; 24: 338-346.

[15] B Abidin, RM Dom. *Prediction of preclinical academic performance using ANFIS model.* Int Proc Econ Dev Res. 2012; 41: 105-109.