



## PREDICTING STUDENT PERFORMANCE IN ENGINEERING EDUCATION USING AN ARTIFICIAL NEURAL NETWORK AT TSHWANE UNIVERSITY OF TECHNOLOGY

G.M. Kanakana<sup>1\*</sup> and A.O. Olanrewaju<sup>2</sup>

<sup>1</sup>Department of Industrial Engineering  
Tshwane University of Technology, South Africa  
[kanakanamg@tut.ac.za](mailto:kanakanamg@tut.ac.za)

<sup>2</sup>Department of Industrial Engineering  
Tshwane University of Technology, South Africa  
[olanrewajuoa@tut.ac.za](mailto:olanrewajuoa@tut.ac.za)

### ABSTRACT

Student achievement is a national and institutional phenomenon: studies have found that engineering programs are particularly vulnerable to underachievement, with a national throughput rate of 17% after 5 years of study [1]. The severity of the problem has led to the exploration of techniques which can be used to predict student performance after access to higher education.

In this respect Artificial Neural Network (ANN) and linear regression models were used to predict student performance during semester 1 of 2008. Data received from the Tshwane University of Technology was utilized for this study. The total Average Point Scores (APS) students obtained in grade 12 was employed as input variable. The results indicate a better agreement between ANN model prediction and observed values compared to those in the linear regression. This demonstrates that the ANN-based model developed may well be able to predict student performance in semester 1 with high accuracy.

---

\*Corresponding author

## 1. INTRODUCTION

The importance of higher education is well established; it has been reported by the Council of Higher Education that in countries which have managed to sustain high levels of economic growth with significant improvements in the standard of living, the majority of their population are those who have given priority to excellent education, training, higher education and training in particular as an agent of socio-economic change and development [1].

The core function of higher education is knowledge creation and high-level human capacity development; this is essential to address social problems in a context of economic globalization, to establish a productive niche and to avoid being uncompetitive compared to developed countries. Since South Africa is a developing country confronting challenges of skills development, poverty and inequalities, higher education must play a pivotal role in the development of the country. This was validated by the emphasis the government has placed on access to such education, in the white paper of 2007.

Scott[1] argued that there is a special need for higher education to produce an appropriate mix of good graduates, sufficient to meet the country's development requirements. A skills shortage of engineers, technologists and technicians was reported by Roodt [2]. Scott indicated that only 17% of engineering students graduate from a University of Technology, and only after 5 years in the program, which is supposed to take 3 years, while 14% are still registered, 59% have dropped out and 10% have changed qualifications.

The grade point average (GPA) or average point system (APS) is said to be the best indicator of student performance; however the study by Baars [3] indicated that the GPA or APS alone cannot accurately predict this. Student performance is the result of a very complex interaction between a range of student-related aspects such as ability, motivation, ambition, study skills, learning styles, personality traits, time spent on study related activities and external factors such as elements of the learning and social environment [ 4, 5, 6, 7, 3 ].

In South Africa, the average point system (APS) is used to determine access to higher learning institutions; therefore it is essential to determine the usefulness of APS scores to ensure that students are allowed access to the right program. In engineering in particular, minimum requirements for Mathematics, Science and English have been used by higher education institutions to determine access to engineering programs.

With some of the literature outlining that GPA or APS is of the essence, while other researchers argue that this statement is not correct, the study sought to validate whether APS alone can be used to accurately predict students' performance once they access higher education.

The study proposed the use of the multilayer perceptron (MLP) of an artificial neural network and regression analysis models to predict student performance using APS as inputs and their first year results as output. The results, indicated by the performance index, coefficient of correlation  $R^2$ , and mean absolute percentage error (MAPE), reported that the APS scores are not adequate to predict the students' average performance.

## 2. DATA AND METHODS

The study was conducted at Tshwane University of Technology, Faculty of Engineering and the Built Environment in Pretoria, South Africa. Its engineering curriculum consists of two

years of theoretical subject studies and one year of practical studies. The theoretical years are aimed at providing students with fundamental engineering concepts and specialized content related to their area of interest. The focus in the last year is placed on practical experience; students are expected to gain knowledge through solving real life problems in an industry related to the student's choice of program.

## 2.1 Data

Data used was obtained from 49 students who entered the department of Industrial Engineering, Tshwane University of Technology in 2008. Only the APS scores and pass rates of semester 1 were utilized for this study; the descriptive statistics of the data are reported in table 1 below.

	Inputs (APS Scores)			Output
	Maths	Science	English	Students' performance
Mean	72.2	60.4	65.1	60.2
Median	74.5	64.5	64.5	59.5
Standard Deviation	10	9.1	9.2	5.6
Minimum	44.5	44.5	54.5	48
Maximum	84.5	74.5	84.5	72.37

**Table 1: Descriptive statistics of the data used**

## 2.2 Method

Artificial neural network and regression analysis were used to analyze the data. These techniques are known for predicting future performance from current data.

### 2.2.1 Artificial neural networks

Artificial neural networks are computer models built to emulate the human brain, and are well known, massively parallel, computing models which have exhibited excellent behavior in solving problems. Neural networks are employed because they improve upon the error of the models developed by linear regression [8]. The goal of an artificial neural network is to map a set of input patterns onto a corresponding set of output patterns. The network accomplishes this mapping by learning from a series of past examples and defining the input and output sets for a given system [9].

The architecture of this network consists of three layers, namely the input, hidden and output layer, with each layer containing one or more neurons, in addition to bias neurons connected to the hidden and output layers. The computational procedure of the network is described below [10].

$$Y_j = f(\sum_i w_{ij} X_{ij}), \quad (1)$$

where  $Y_j$  is the output of node  $j$ ,  $f(\cdot)$  the transfer function,  $w_{ij}$  the connection weight between node  $j$  and node  $i$  in the lower layer and  $X_i$  the input signal from the node  $i$  in the lower layer. The back propagation is based on a steepest descent technique with a momentum weight (bias function) which calculates the weight change for a given neuron. It is expressed as follows [10, 11]: let  $\Delta w_{ij}^p(n)$  denote the synaptic weight connecting the

output of neuron  $i$  to the input of neuron  $j$  in the  $p$ th layer at iteration  $n$ . The adjustment  $\Delta w_{ij}^p(n)$  to  $w_{ij}^p(n)$  is given by

$$\Delta w_{ij}^p(n) = -\eta(n) \frac{\partial E(n)}{\partial w_{ij}^p}, \quad (2)$$

where  $\eta(n)$  is the learning rate parameter. By using the chain rule of differentiation, the weight of the network with the back propagation learning rule is updated using the following formulae:

$$\Delta w_{ij}^p(n) = \eta(n) \delta_j^p(n) X_i^{p-1}(n) m(n) \Delta w_{ij}^p(n-1), \quad (3)$$

$$\Delta w_{ij}^p(n+1) = w_{ij}^p(n) + \Delta w_{ij}^p(n), \quad (4)$$

where  $\delta_j^p(n)$  is the  $n$ th error signal at the  $j$ th neuron in the  $p$ th layer,  $X_i^{p-1}(n)$  is the output signal of neuron  $i$  at the layer below and  $m$  is the momentum factor.

For the neural network design, the number of hidden neurons was determined by comparing the performance of different cross-validated networks, with 1-18 hidden neurons, and choosing the number that produced the greatest network performance. This resulted in a network with three input neurons (mathematics, science and English APS scores), eight hidden neurons and a single output (students' average performance). In the analyses, network parameters of learning rate and momentum were set to 0.05 and 0.7, respectively. Variable learning rates with momentum (trainlm) as the network's training function, tansig and purelin as activation functions for hidden and output layers were utilized. The data used by the network must be scaled for the network to be effective. In theory the inputs to the network can be any value; however, scaling values to the same order of magnitude (generally in the range 0 to 1 or -1 to 1) enables the network to learn relationships quicker [12]. In this paper, we scaled the data to the range -1 to 1 so as to use a consistent scaling regime for inputs and outputs.

### 2.2.2 Regression analysis

Regression analysis is appealing because it provides a conceptually simple method for investigating functional relationships among variables [13]. The existence of the relationship is observed by the square of the correlation coefficient ( $R^2$ ). The power of a regression analysis lies in its ability to detect observed relationships in a set of data. When the value is nearer to 1.00, the regression relationship being assessed is more reliable [14]. The data was used to fit the following equation

$$y = a + bx_1 + cx_2 + dx_3 \quad (5)$$

Where  $a$ ,  $b$ ,  $c$ , and  $d$  are the regression coefficients,  $y$ , students' average performance, which is the response variable,  $x_1$ ,  $x_2$ ,  $x_3$  represent the APS scores for mathematics, physical science and English. The regression solves for unknown coefficients  $a$ ,  $b$ ,  $c$ , and  $d$  by minimizing the sum of the squares of the data from the model. A linear regression model to predict student performance is given below

$$y = 44.28 - 0.0419x_1 + 0.1484x_2 + 0.1541x_3 \quad (6)$$

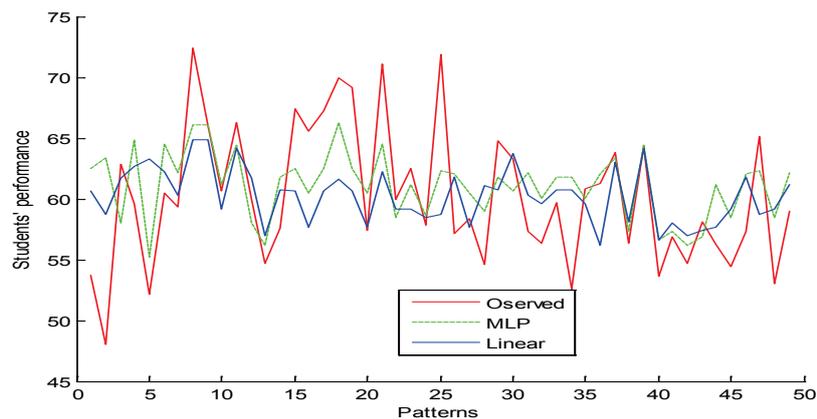
Mean absolute percentage error (MAPE), which is a measure of accuracy in a fitted series value in statistics, was made use of for a comparison of the prediction performances of the models. MAPE usually expresses accuracy as a percentage

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \left| \frac{A_i - P_i}{A_i} \right| \times 100 \quad (7)$$

where  $A_i$  is the actual value and  $P_i$  is the predicted one.

### 3. ANALYSIS AND RESULTS

In this work, visual inspection, mean absolute percentage error (MAPE), as well as the coefficient of correlations  $R^2$  were employed to check the performance, as well as the accuracy, of the models used for prediction. Figure 2 indicates that the ANN prediction seems to be closer to the observed data.



**Figure 2: Result of the network. Comparison between the target, ANN and regression**

MODEL	MAPE	$R^2$
Linear Regression	6.580	0.1577
Artificial Neural Network	6.226	0.3258

MAPE = mean absolute percentage error,  $R^2$  = correlation coefficient

**Table 2: Performance indices for models**

The  $R^2$  value for linear regression is 0.1577 with MAPE of 6.580. ANN had  $R^2$  of 0.3258 with MAPE of 6.226. Results of the two models used in this study are compared. In this case, the  $R^2$  and MAPE values obtained by ANN show more successful results as compared with the linear regression model.

Using both models to predict the students' performance, the table below (table 3) indicates predicted values of student performance which are not closer to actual student performance. Statistically, 3 standard deviation is used to determine the capability of the process. For these models, the standard deviations are less than the 3 standard deviation. For linear regression the standard deviation is 2.2 while for ANN, it is 2.8 which indicates poor prediction due to fewer factors.

	Measured Students' performance	Linear	ANN
Mean	60.2	60.3	61.1
Median	59.5	60.7	61.8
Standard Deviation	5.6	2.2	2.8
Minimum	48	56.2	55.2
Maximum	72.37	64.8	66.3

**Table 3: Predicted values**

#### 4. CONCLUSION

The study by Baars indicated that GPA or (APS) alone cannot accurately predict students' performance. Literature argues that other variables also contribute to such performance.

The results of this study indicate that ANN can predict student performance more accurately than regression analysis. The R value for ANN and regression models is less than 0.5 and is closer to 0. Therefore we can state with confidence that there is a weak relationship between input and output variables; in our case, as mentioned, inputs comprise APS scores while outputs consist of student performance.

These results of the  $R^2$  and MAPE of ANN indicate that ANN can be used successfully for the prediction of this performance, though the result of the coefficient of correlation is far from 1.00. Thus, we can conclude that APS scores by themselves are not adequate to determine the performance of students. This confirms literature studies which argue that APS alone cannot accurately predict students' performance after entering higher education. If this is the case, the next research questions will be, firstly: why higher education institutes utilize APS scores as the minimum requirement to access higher education? Secondly, what are the other factors which play pivotal roles in student performance after entering higher education?

#### 5. REFERENCES

- [1] Ian Scott, Yeld, N. & Hendry, J. 2007. Higher Education Monitor A Case for Improving Teaching and Learning in South Africa Higher Education. Pretoria: The Council on Higher Education
- [2] Roodt, J. & Du Toit, R. 2009. *Engineers in a developing country* Cape Town, HSRC Press.
- [3] Baars, G. J. A. 2009. *Factors Related to Student Achievement in Medical School* [Online]. Lemma, Utrecht. Available: <http://hdl.handle.net/1765/17103> [Accessed].
- [4] Benseman, J., Coxon, Eve, Anderson, H. & Anae, M. 2006. Retaining non-traditional students: lessons learnt from Pasifika students in New Zealand. *Higher Education Research & Development*, 25, 147-162.
- [5] Nariman, F. 2007. Retaining Students-and Their Hopes and Dreams. *ASEE Prism*, 16, 64.
- [6] Araque, F., Roldán, C. & Salguero, A. 2009. Factors influencing university drop out rates. *Computers & Education*, 53, 563-574.



- [7] **Brown, A. R., Carole, M. & Charles, W.** 2005. Influence of African American Engineering Student Perceptions of Campus Climate on Graduation Rates. *Journal of Engineering Education*, 94, 263.
- [8] **Clouster, D. L. & Jurs, P. C.** 1996. The simulation of  $(^{13}\text{C})$  nuclear magnetic resonance spectra of dibenzofurans using multiple linear regression analysis and neural networks. *Analytica Chimica Acta*, 321, 127-135.
- [9] **Amir Heydari, Shayesteh, K. & Kamalzadeh, L.** 2006. Prediction of Hydrate formation temperature for natural gas using artificial neural network. *Oil and Gas Business*.
- [10] **Hsu, C.-C. & Chen, C.-Y.** 2003. Regional load forecasting in Taiwan-applications of artificial neural networks. *Energy conversion and Management*, 44, 1941-1949.
- [11] **Huang, H., Hwang, R. & Hsieh, J.** 2002. A new artificial intelligent peak power load forecaster based on non-fixed neural networks. *Electric Power Energy Syst*, 24, 245-50.
- [12] **Hart, A.** 1992. Using neural networks for classification tasks - some experiments on datasets and practical advice. *J.Opl. Res. Soc.*, 43, 215 - 226.
- [13] **Chatterjee, S. & Hadi, A. S.** (eds.) 2006. *Regression Analysis by Example: A John Wiley and Sons inc.*
- [14] **Tomera, J. F. & Harakal, C.** 1997. Multiple linear regression analysis of blood pressure, hypertrophy, calcium and cadmium in hypertensive and nonhypertensive states. *Food and chemical Toxicology*, 35, 713-718.