

Full Length Research Paper

Predictive modelling and analysis of academic performance of secondary school students: Artificial Neural Network approach

Amoo M. Adewale, Alaba O. Bamidele and Usman O. Lateef

Department of Computer and Information Sciences, Tai Solarin University of Education, Ogun State Nigeria.

Received 7 June 2017; Accepted 1 July, 2017

The need for educational managers and policy makers to design and implement pedagogical and instructional interventions necessitates the development and validation of a predictive model for analyzing the academic performance of students before gaining admission into university. This study adopts feed-forward neural network to establish and analyze the complex nonlinear relationship that exist between cognitive and psychological variables that influence the academic performance of secondary school students. The sample space comprises of 120 students selected from four randomly selected secondary schools in Ibadan-North Local Government Area of Oyo State. Students' performances in five science subjects at 2015 West African Examination Council (cognitive factor) and psychological factors (age, gender, status of school and parent occupation) were used as inputs to the proposed model while the performance at Post-Unified Tertiary Matriculation Examination (Post-UTME) served as the target output. Simulated results from the study showed that artificial neural network (ANN) is efficient at clustering students into different categories according to their predicted level of performance. Study of this type will enable educational planners and curriculum developers provide better educational services as well as customize assistance to students when they gain admission into universities.

Key words: ANN, students' performance, university admission, predictive model, Ibadan.

INTRODUCTION

Students play vital role in the life of a university because the ranking of universities in Nigeria depends largely on the qualities of their graduates among other factors. However, the quality of students admitted into university affects the quality of research, teaching within the institution, and perhaps has an overall effect on the

expected level of services rendered to the nation by the universities.

According to Usman and Adenubi (2013), Kalejaye et al. (2015) and Oladokun et al. (2008) learners' academic performance should be monitored as early as possible in their academic career so as to foster national growth and

*Corresponding author. E-mail: amoomutiu@yahoo.com.

Author(s) agree that this article remain permanently open access under the terms of the [Creative Commons Attribution License 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

development as learners eventually become key players in the affairs of the country in all sectors of the economy.

Recently, research has shown that the quality of undergraduates in Nigeria Universities is low (Adeleke et al., 2013) due to inadequacies of the present admission system, and increasing gap between the number of students seeking admission and the total available admission space. This has posed a serious challenge on curriculum developers and other educational managers. They are now faced with the challenge of redesigning and implementing pedagogical and as well as instructional interventions that will enhance teaching and learning in tertiary institutions. To achieve the aforementioned goal, there is need to devise a means of assessing and analyzing students' academic performance early enough, and as well classify them into different categories so as to provide better educational services that is suitable for each group. It is therefore imperative to design and implement a predictive model for early academic performance prediction of students before gaining admission into the university. Many reasons have been advanced for learners' academic performance prediction. These include the following:

- (1) Prediction helps admission officers to distinguish between suitable and unsuitable candidates for admission, and identify candidates who would likely do well in school (Ayan and Garcia, 2013).
- (2) Results from the academic performance prediction can be used for classifying learners into groups, offer additional support, customize assistance and provide adequate learning resources (Usman and Adenubi, 2013)
- (3) Prediction results can be used to cluster learners. This clustering mechanism can be used by the teacher or instructor to select educational material suitable to the abilities of each group of students (Lykourantzou et al., 2009; Kalejaye et al., 2015)
- (4) Predicting students academic performance assist higher institutions of learning tackle academic failures and provide a basis for making the process of teaching and learning more effective, efficient and reliable (Kalejaye et al., 2015).
- (5) Modelling students' achievement is useful tool for both educators and students. This can help have better understanding of students' weakness and bring about improvement. It helps differentiate between the fast learners and slow learners (Adeleke et al., 2013; Nghe et al., 2007).
- (6) The result from academic performance prediction can be used to formulate policy that students who have no tendency of doing well in school be discovered at early stage of academic pursuit, thereby preventing continuous waste of human and material resources on such non-productive students (Abass et al., 2011) etc.

It can be inferred from the aforementioned reasons that learners' academic performance prediction helps improve

the instructional strategies and develop curriculum that will produce better learners and creative thinkers.

A number of variables have strong influence on the academic capabilities of learners, particularly, in their secondary school lives. Variables such as students' prior knowledge and prior performance contribute significantly the prediction accuracy of the model that predicts student academic performance. Psychological variables, such as gender, parent occupation/literacy level, and status of school (private/public) are controversial predictors for academic achievement.

Prior knowledge and performance variables used in this study include students O'Level grades in five (5) science subjects. These were extracted from the June 2015 West African Examination Council results released in November, 2015. These subjects include mathematics, English language, physics, chemistry and biology.

Identifying and choosing effective modelling approaches is vital in developing a predictive model. Various mathematical models, such as regression has been employed in constructing predictive models. This model has advantages and disadvantages. For instance, regression one of the commonly used approaches to constructing predictive models, is easy to understand and provides explicit mathematical equations.

However, regression should not be used to estimate complex relationships and is susceptible to outliers because the mean is included in regression formulas. Therefore, any predictive model designated to analyzing the students' academic performance must take the cognizance of the complex and nonlinearity relationships that exist among these variables.

Artificial neural networks (ANNs) have found wide applications in areas such as pattern recognition, image processing, optimizations, prediction, and control systems. ANN, which is a simplified model of the biological nervous system (brain), can handle this type of problem because of its function approximation capabilities. Early studies (Oladokun et al., 2008; Lykourantzou et al., 2009; Kalejaye et al., 2015) have applied ANN to estimate learners' achievement with reasonable level of precisions and Accuracies. In this research, an attempt was made to develop a predictive model for analyzing the academic performance of students early enough, before admitting such students into the university.

The main aim of this research is to model an ANN that can be used to predict a student's performance before admitting the student based on some given pre-admission data of the student, as this will enable educational managers customize assistance based on student's predicted level of performance. The specific objectives of this research, however, include:

- (1) To identify some suitable variables that affect student's academic performance.

(2) To transform these factors into forms suitable for an adaptive system coding in the field of Artificial Intelligence.

(3) To validate and evaluate the performance of ANN model.

The artificial neural networks: Brief introduction

ANN consists of a set of highly interconnected entities, called Processing Elements (PE) or unit. The structure and function of ANN is inspired by the biological central nervous system (brain). Each unit is designed to mimic its biological counterpart, the neuron or neural node. Each accepts a weighted set of inputs and responds with an output.

ANNs address problem that are often difficult for traditional computers to solve, such as speech and pattern recognition, weather forecasts, sales forecasts, scheduling of buses, power loading forecasts, electricity consumption prediction, student enrollment projection, clustering and early cancer detection. ANNs develop solution fast and have less reliance on domain experience.

ANNs develop solution to any problem under the consideration that one may not get thorough understanding of the subject matter (Usman and Alaba, 2014; Usman and Adenubi, 2013; Oladokun et al., 2008). The origin of the neural network can be traced to 1940s when two researchers, Warren McCulloch and Walter Pitts, tried to build a model to simulate how biological neurons work. Though the focus of this research was on the anatomy of the brain, it turns out that this model introduced a new approach for solving technical problem outside neurobiology. Several architectures for the ANN exist. These include feed-forward, feed-backward, single-layer, recurrent, radial basis function network, and self-organizing maps (Lykourantzou et al., 2009).

Among the neural network architectures, feed-forward is most commonly used. Feed-forward neural networks (FFNN) tend to be straight forward network that associate inputs with outputs. According to Haykin (1999), FFNN consists of one or more hidden layers of neurons. In this type of network, node connections, called *synapses*, do not form a directed cycle.

The goal of the FFNN training is to minimize the mean square error (MSE) between its desired and predicted results, by adjusting the network synaptic weights and other network parameters. More specifically, these network parameters are adjusted based on the back-propagation algorithm.

In this algorithm, information is passed forward from the input nodes, through the hidden layers, to the output nodes, and the error between the desired and the network response is calculated. Then, this error signal is propagated backwards to the input nodes. A popular

approach to optimize the performance of back-propagation is the *Levenberg-Marquardt* algorithm, which has been found to increase the speed of convergence and effectiveness of the network training.

Optimization of the algorithm is achieved with *Gradient Descent* algorithm. Several examples of Gradient Descent exist. These are TRAINLM, TRAINOSS, TRAINSCG, TRAINGDM and TRAINGDA etc. By this approach, input vectors are applied to the network and calculated gradients at each training sample are added to determine the change in synaptic weights and biases.

The FFNN parameters are estimated based only on the training dataset, and the performance of the network is evaluated by computing the MSE on the validation dataset. The detail backpropagation algorithm is presented in the next section.

Backpropagation algorithm

During ANN training phase, hidden layer allows ANN to develop its own complex internal representation capability. This internal representation allows the hierarchical network to learn any patterns that exist in the input-output vectors. This training process is called Backpropagation algorithms. Let us consider three-layer FFNN with input layer '*i*' nodes, hidden layer having '*h*', and output layer with '*o*' nodes. The basic Backpropagation algorithm structure is given as (Rajasekaran and Pai, 2012):

Initialize the weights

Repeat

For each training pattern

Train on that pattern

End

Until the error is acceptably low

It is important to normalize the inputs and outputs with respect to their minimum and maximum values because ANN works better if the input and outputs lie between 0-1.

METHODOLOGY

This study proposed the use of FFNN to develop, implement and analyze a predictive model for solving the problem of poor academic performance in tertiary institutions. Specifically, the study explored the possibility of FFNN to predict the performance of students before admitting such students into university. This will enable policy makers and educational managers to customize assistance in terms of curriculum and materials.

A total of 120 students' academic records in five (5) science subjects were extracted from the results of 2015 West African Examination Council. This data was obtained from four (4) randomly selected secondary schools in Ibadan North Local Government Area of Oyo state, including: Methodist Grammar

School, Bodija, Ibadan; Emmanuel College, Orita UI, Ibadan; Tafseer Model College, Ibadan; and Community Grammar School, Mokola, Ibadan. Grades obtained in science subject like Mathematics, English, Physics, Chemistry and Biology, and as well, psychological variables: Age, parent occupation and school status (private/public) were assigned numerical values, analyzed and harmonized into a manageable number of suitable patterns for coding within context of the ANN modeling.

Simulated post-UTME scores of the selected students were assumed to be target values. Implementation of the predictive model was done using '*nntool*' of MATLAB version 7.6.0 software. Table 1 summarized the numerical representation of possible O'Level grades and other variables used.

The input variables

The input variables are the dataset used as input to the predictive model as well as the target values. They can easily be obtained from students' register and academic records. As shown in Table 1, input variables include:

- (1) WAEC results in Mathematics, English language, Physics, Chemistry and Biology
- (2) Post-UTME grades
- (3) Student's Gender
- (4) Age (Age of student at the time of writing WAEC)
- (5) Parent occupation
- (6) School status (whether private or public)

Numerical descriptions of each variable were normalized to generate values between 0 to 1, which make them suitable for coding, and analysis within the context of ANN modelling.

The output variables

Output represents the performance of a student in post-UTME. The domain of the output variable has been grouped into five categories and assigned numerical values. As can be seen from Table 1, post-UTME scores in the range of 70 to 100% is Grade A and assigned 5 points on the numerical scale. Scores in the range of 60-69% is B and assigned 4 points, and so on.

Network topology

ANNs are useful only when the processing units (neural nodes) are organized in a suitable manner to accomplish a given pattern recognition task. The arrangement of the processing units, connections, and pattern input-output is referred to as topology. Choosing the topology of ANN is a challenging issue (Oladokun et al., 2008; Kalejaye et al., 2015).

Several options could be considered, each with inherent strengths and weaknesses. For example, some ANNs trade off speed for accuracy, while some are capable of handling static variables. Hence, in order to arrive at an appropriate topology, several options were considered. Due to the static nature of the selected variables and kind of data used, complex topology such as radial basis function (RBF) neural network, recurrent neural network (RNN) and self-organizing map (SOM) could not be used.

This study maintained feed-forward neural network (5-5-5-1). Our choice was borne out the fact that it is easy to use and can approximate any input-output map with reasonable accuracy. There are two hidden layers of five neural nodes each in the chosen topology. Selection of number of nodes in the hidden layer is a

difficult task. Having a small number of neural nodes in the hidden layers will slow down the processing capabilities.

On the other hand, large number of neural nodes in these layers will progressively slow down the training speed. In order to arrive at the best trained network, trainings were performed in a number of iterations. We first of all started with the small number of hidden nodes and increase the number progressively until a threshold level was reached at hidden nodes equals to ten.

Network learning method and validation process

Learning methods in ANN can broadly be classified into three basic types: supervised, unsupervised and reinforced. In a supervised method, every input pattern that is used to train the network is associated with an output pattern, which is the target pattern. A teacher is assumed to be present during the learning process, when a comparison is made between the network's computed output and the desired output, to determine the error.

In unsupervised method, the target output is not present and there is no teacher to present the desired patterns and hence, the ANN learn of its own by discovering and adapting to structural features in the input patterns. Though, a teacher is present in reinforced method, he does not present the target output but only indicates if the computed output is correct or not (Rajasekaran and Pai, 2012).

In this study, a supervised method is adopted. The usual practice here is to divide the data into three categories: the training set, verification set and testing set. The training set enables the system to observe relationship between input data and simulated outputs, so that it can establish a mapping function between the input and target output.

The proposed FFNN was trained with the number of runs set to three and training epoch set to terminate at 1000 iterations, thereby varying the parameters shown in Table 2. The objective of doing so is to determine which configurations of parameter produce the optimal results.

RESULTS AND DISCUSSION

After the training and cross-validation, the proposed FFNN was tested with test data and the results obtained. It was observed that the best training run was completed at epoch equals to 11 iterations, with performance of 0.03386 after 6 validation check. None of the training run exceed 1 minute.

Figure 1 shows the performance plot of the best training configuration. The best trained FFNN showed highly positive correlation coefficient (Table 3) when compared with other configurations used during the training session. The correlation coefficient was computed by comparing FFNN outputs with the target values. FFNN with the optimal configuration was then used to simulate the results shown in Figure 2.

On the secondary school basis, the ANN prediction accuracy for Methodist Grammar School, Bodija, Ibadan ranges from 66.7 to 80.0%; Emmanuel College, Orita-UI, Ibadan ranges from 73.3 to 86.7%; Tafseer Model College, Ibadan ranges from 86.7 to 93.3%; and Community Grammar School, Mokola, Ibadan ranges from 80.0 to 86.7%. The comparison is further

Table 1. Numerical representation of input variables.

S/N	Input variable	Domain value	Numerical value	
1.	WAEC results	Mathematics	A1	5
			B2-B3	4
			C4-C6	3
			D7-E8	2
			F9	1
		English	A1	5
			B2-B3	4
			C4-C6	3
			D7-E8	2
			F9	1
		Physics	A1	5
			B2-B3	4
			C4-C6	3
			D7-E8	2
			F9	1
		Chemistry	A1	5
			B2-B3	4
			C4-C6	3
D7-E8	2			
F9	1			
Biology	A1	5		
	B2-B3	4		
	C4-C6	3		
	D7-E8	2		
	F9	1		
2.	Post-UTME Scores	70-100	A	5
		60-69	B	4
		50-59	C	3
		45-49	D	2
		0-44	E	1
3.	Gender	Male	-	2
		Female	-	1
4.	Age	18 years –above	-	2
		Below 18 years	-	1
5.	Parent occupation	Civil servant	-	3
		Artisan	-	2
		Trading	-	1
6.	School status	Private	-	2
		Public	-	1

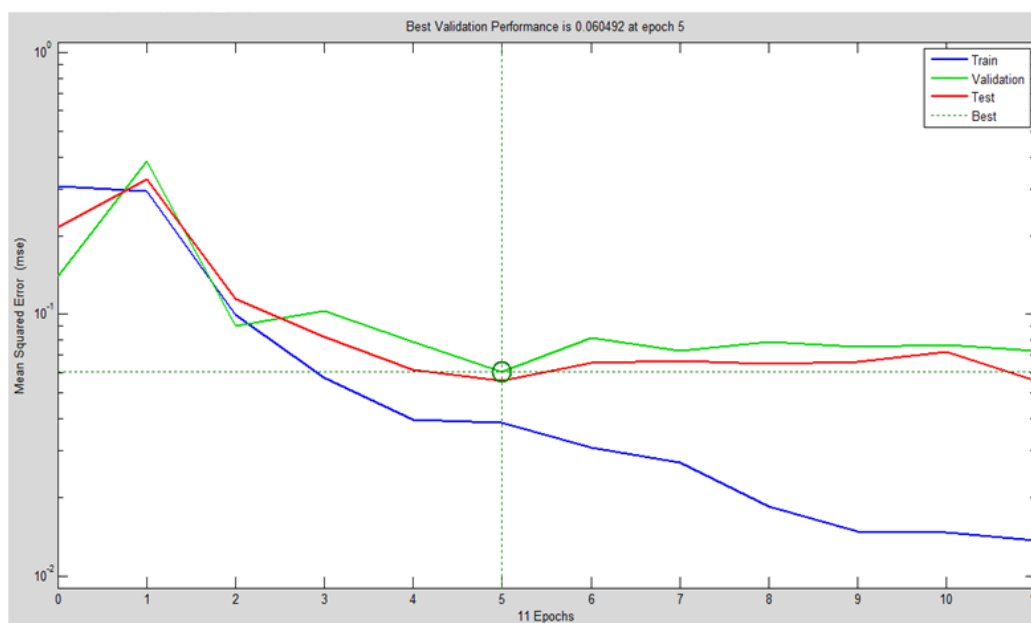
summarized in the confusion matrix as shown in Table 4.

Predicted results from the ANN are shown along the diagonal of the matrix while the desire values can be

obtained by summing up the values in the matrix. Data obtained from the source does not reflect distinction and good grades, hence, 0 in these cases. One gets 120

Table 2. ANN parameter adjustment.

Parameters	Values
Training algorithm	Levenberg-Marquardt
Training functions	TRAINLM, TRAINOSS, TRAINSCG
Performance function	Mean square error (MSE)
Adaption learning function	LEARNGA, LEARNGD, LEARNGDM
Number of layers	3
Transfer function	TANSIG, LOGSIG, PURELIN
Epoch	1000 iterations

**Figure 1.** Performance FFNN (TRAINLM, LEARNGD and TANSIG).**Table 3.** Correlation coefficient of different configuration of FFNNs.

FFNN configurations	Correlation coefficient
TRAINOSS, LEARNGD, LOGSIG	0.8663
TRAINLM, LEARNGD, TANSIG	0.9610
TRAINSGC, LEARNGDM, PURE	0.8496

either by summing along the diagonal of the matrix or all the values in the matrix. This represents the total number of data points used in the study. FFNN predict accurately 44 credit grades out of 50, 34 fair grades out of 38, and 30 fail grades out of 32. This gives an accuracy of 88% for credit grades, 89.5% for fair grades and 93.8% for fail. On the overall, ANN has been able to predict accurately 108 grades out of 120 grades representing 90%

accuracy. This result indicates that Artificial Neural Network's performance is fair going by similar results from the literature (Oladokun et al., 2008; Abass et al., 2011; Lykourantzou et al., 2009; Kalejaye et al., 2015; Wang and Mitrovic, 2002).

The implication of the study is that, if ANN is incorporated into the admission system, it will assist stakeholders in the education sector to identify the weak

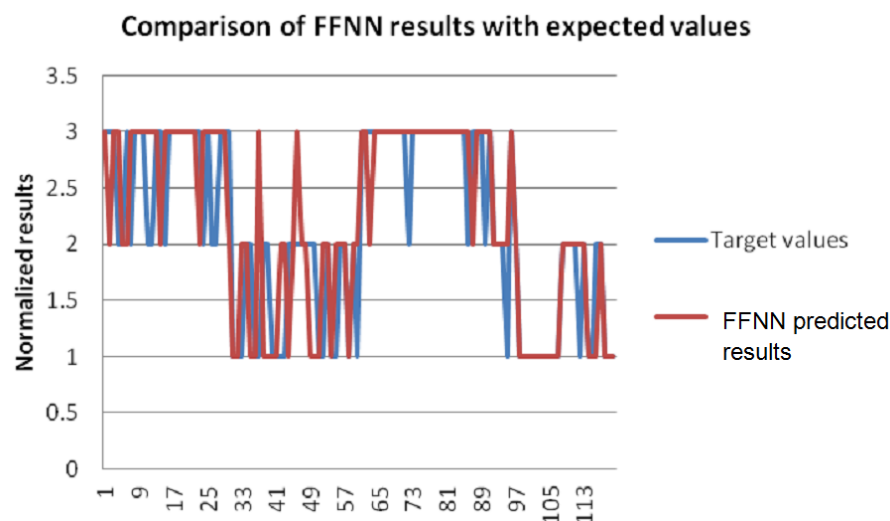


Figure 2. Comparison of results.

Table 4. Confusion matrix showing the results from the test of FFNN.

Predicted/ target output	Distinction	Good	Credit	Fair	Fail
Distinction	0	0	0	0	0
Good	0	0	0	0	0
Credit	0	0	56	2	1
Fair	0	0	-6	34	1
Fail	0	0	0	2	30

students before admitting them into the university. This will also enable them customize assistance in terms of curriculum, material and methods of teaching. Consequently, the effectiveness of a university admission system will be enhanced.

CONCLUSION

In this study, an attempt has been made to exhibit the potential of artificial neural network at predicting the secondary student post-UTME grade before admitting such student into the university system. Specifically, the model was developed using feed-forward neural network architecture and based on some selected input variables identified in the previous section. Clearly, ANN model achieved an accuracy of 90%, which shows the potential efficacy of ANN as a predictive model, a clustering instrument and a selection criterion for candidates seeking admission into a university. In spite of the high level prediction accuracy of ANN in nonlinear phenomena, however, the model does not easily allow the identification of how predictor variables are related to

one another in the explanation of the academic outcome. In other words, ANN model does not specify an explicit mathematical model for the relationship inputs and outputs, hence the need for further research in this regard.

CONFLICT OF INTERESTS

The author has not declared any conflicts of interest.

REFERENCES

- Abass O, Oyekanlu EA, Alaba OB, Longe OB (2011), Forecasting Student Academic Performance using Case-Base Reasoning". International conference on ICT for Africa. 23-26th march, 2010 Otta Nigeria, pp.105-112.
- Adeleke RA, Ruzaini AA, Hongwu Q (2013), Risk Status Prediction and Modelling of Students' Academic Achievement-A Fuzzy Logic Approach. Research Inventory: International Journal of Engineering Science, 3(11):07-14.
- Ayan MNR, Gracia MTC (2013), Prediction of University Students' Academic Achievement by Linear and Logistic Models. The Spanish Journal of Psychology, 11(1):275-288.
- Haykin S (1999), Neural Networks: A Comprehensive Foundation (2nd Edition). Macmillan College Publishing Company, New York.

- Kalejaye BA, Folorunso O, Usman OL (2015), Predicting Students' Grade Scores using Training Functions of Artificial Neural Network. *Journal of Natural Sciences Engineering and Technology*, 14(1):25-42.
- Lykourantzou I, Giannoukos I, Mpardis G, Nikolopoulos V, Loumos V (2009), Early and Dynamic Student Achievement Prediction in E-Learning Courses Using Neural Networks. *Journal of the Association for Information Science and Technology*, 60(2):372-380.
- McCulloch WS, Pitts WA (1943), A logical Calculus of Ideas Imminent in Nervous Activity. *Bull. Math. Biophys.* 5:115-133.
- Nghe NT, Janecek KP, Haddawy P (2007), A comparative Analysis of Techniques for Predicting Academic Performance. In, *Proceedings of 37th ASEE/IEEE Frontiers in Education Conference*, pp. 659-669.
- Oladokun VO, Adebajo AT, Charles-Owaba OE (2008). Predicting Students Academic Performance using Artificial Neural Network: A case study of an Engineering Course. *Pacific Journal of Science and Technology*, 9(1):72-79.
- Rajasekaran S, Vijayalakshmi Pai GA (2012), *Neural Networks, Fuzzy Logic, and Genetic Algorithms: Synthesis and Applications*. PHI Learning Private Limited, New Delhi-110001, ISBN: 978-81-203-2186-1.
- Usman OL, Adenubi AO (2013), Artificial Neural Network (ANN) model for Predicting Students' Academic Performance. *Journal of Science Education and Technology*, (JOSIT) 13(2):61-71.
- Usman OL, Alaba OB (2014), Predicting Electricity Consumption Using Radial Basis Function (RBF) Network. *International Journal of Computer Science and Artificial Intelligence*, 4(2):54-62.
- Wang T, Mitrovic A (2002). Using Neural Networks to Predict Students' Performance. Paper presented at the International Conference on Computers in Education, Auckland, New Zealand.