Predicting Students' Academic Performance using Artificial Neural Network: A Case Study of an Engineering Course.

V.O. Oladokun, Ph.D.[*], A.T. Adebanjo, B.Sc., and O.E. Charles-Owaba, Ph.D.

Department of Industrial and Production Engineering, University of Ibadan, Ibadan, Nigeria.

[*]E-mail: vo.oladokun@mail.ui.edu.ng

ABSTRACT

The observed poor quality of graduates of some Nigerian Universities in recent times has been partly traced to inadequacies of the National University Admission Examination System. In this study an Artificial Neural Network (ANN) model, for predicting the likely performance of a candidate being considered for admission into the university was developed and tested.

Various factors that may likely influence the performance of a student were identified. Such factors as ordinary level subjects' scores and subjects' combination, matriculation examination scores, age on admission, parental background, types and location of secondary school attended and gender, among others, were then used as input variables for the ANN model. A model based on the Multilayer Perceptron Topology was developed and trained using data spanning five generations of graduates from an Engineering Department of University of Ibadan, Nigeria's first University.

Test data evaluation shows that the ANN model is able to correctly predict the performance of more than 70% of prospective students.

(Keywords: university admissions, student performance, Artificial Neural Networks, ANN, tertiary education, predictive models)

INTRODUCTION

The main objective of the admission system is to determine candidates who would likely do well in the university. The quality of candidates admitted into any higher institution affects the level of research and training within the institution, and by extension, has an overall effect on the development of the country itself, as these candidates eventually become key players in the affairs of the country in all sectors of the economy.

Recently, however, there has been a noticeable slide in the quality of graduates of some Nigerian universities. The inadequacies of the present university admission system, among other factors, have been blamed for this decline. Due to the increasing gap between the numbers students seeking admission and the total available admission slots, there has been a corresponding increased pressure on the process. This pressure has lead to rampant cases of admission fraud and related problems.

In Nigeria, students are required to enter secondary school after spending a minimum of six years of Primary Education and passing a prescribed National Common Entrance Examination. A student then spends a minimum period of six years in Secondary School at the end of which he or she takes the General Certificate of Education Examination (GCE), also known as the Senior Secondary Certificate Examination (SSCE) or the Ordinary Level Exams. A maximum of nine and a minimum of seven subjects are registered for in the examination with Mathematics and English Language being compulsory. Nine possible grades are obtainable for each subject; these are A1, A2, A3 (distinctions grades) C4, C5, C6, (credit grades), P7, P8 (pass grades), and F9 (Failure).

Before a candidate can be admitted into any university, he/she is expected to pass, at credit level, some number of relevant subjects including Mathematics and English Language in the General Certificate Examinations (GCE) (JAMB, 2005). A second admission requirement is the Universities Matriculation Examination (UME), which was first conducted in 1978 by the National Admissions and Matriculation Board. The UME process involves the implementation of cut-off marks and certificate requirements. However it has been observed that desperate candidates are able to manipulate the system. It has become obvious that the present process is not adequate for selecting potentially good students. Hence there is the need to improve on the sophistication of the entire system in order to preserve the high integrity and quality for which Nigerian Universities were noted for in the seventies and eighties.

It should be noted that this feeling of uneasiness of stakeholders about the traditional admission system, which is not peculiar to Nigeria, has been an age long and global problem. Kenneth Mellamby (1956) observed that universities worldwide are not really satisfied by the methods used for selecting undergraduates. While admission processes in many developed countries has benefited from, and has been enhanced by, various advances in information science and technology, the Nigerian system has yet to take full advantage of these new tools and technology.

Hence this study takes an engineering approach to tackling the problem of admissions by seeking ways to make the process more effective and efficient. Specifically the study seeks to explore the possibility of using an Artificial Neural Network model to predict the performance of a student before admitting the student.

Intuitively one expects the performance of a student to be a function of some number of factors (parameters) relating to the background and intelligence of said student. It is however obvious that it will be quite difficult finding an analytical (or a mathematical) model that may acceptably model this performance/factors relationship. However one practical approach for predicting the performance of a student may be by 'extrapolating' from historical data of past students' background and their associated performances.

A practical approach to this type of problem is to apply conventional regression analysis in which historical data are best fitted to some function. The result is an equation in which each of the inputs x_j is multiplied by a weight w_j; the sum of all such products and a constant θ then gives an estimate of the output $y = \sum_j w_j x_j + \theta$.

The drawback here is the difficulty of selecting an appropriate function capable of capturing all forms of data relationships as well as automatically modifying output in case of additional information, because the performance of a candidate is influenced by a number of factors, and this influence/relationship is not likely going to be any simple known regression model.

An artificial neural network, which imitates the human brain in problem solving, is a more general approach that can handle this type of problem. Hence, our attempt to build an adaptive system such as the Artificial Neural Network to predict the performance of a candidate based on the effect of these factors.

STUDY OBJECTIVES

The objectives of this study are: 1) to determine some suitable factors that affect a students performance, 2) to transform these factors into forms suitable for an adaptive system coding, and 3) to model an Artificial neural network that can be used to predict a candidate's performance based some given pre admission data for a given student.

STUDENT PERFORMANCE: A BRIEF REVIEW

The literature is replete with various works bordering on university admission, student performance, and related problem. In 1954, the University of New Zealand Council for Educational Research investigated the relationship between academic standards of students on entrance and their first year university work. The study found that the median correlation found among the many sets of variables representing general school performance and general university performance was indicated by a *tau* coefficient of 0.36 for the first year students undertaking their studies on a full time basis (Maidment ,1968).

In 1975, Bakare summarized the factors and variables affecting students performance into the intellective and non-intellective factors, emphasizing that the intellectual abilities were the best measure (Bakare 1975). He categorized causes of poor academic performance into four major classes:

1) Causes resident in society
2) Causes resident in school
3) Causes resident in the family
4) Causes resident in the student.

Studies such as (Lage and Tregelia, 1996) and (Dynan, 1977) looked at a more general aspects of success while Anderson et al., 1994 studied the effect of factors such as gender, student age, and students' high school scores in mathematics, English, and economics, on the level of university attainment. According to their study, students who received better scores in high school also performed better in university. Another aspect discovered was that men had better grades than women and choose to drop from school less often.

Adedeji (2001) sought to find out a correlation between students matriculation exam (UME) scores and their academic performance in Nigerian universities, using the Faculty of Technology, University of Ibadan, Nigeria as a test case. He investigated the relationship between students' UME scores, first, second, and final year Grade Points (GP) with the use of a simple correlation and regression analysis. He concluded in his research that there exists a positive relationship between students admission scores and their undergraduate performance. However, recent trends after Adedeji's study indicates the unreliability of the UME scores.

THE ARTIFICIAL NEURAL NETWORKS: AN INTRODUCTION

Inspired by the structure of the brain, a neural network consists of a set of highly interconnected entities, called Processing Elements (PE) or units. Each unit is designed to mimic its biological counterpart, the neuron. Each accepts a weighted set of inputs and responds with an output. Neural networks address problems 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, early cancer detection, etc. (Adefowoju and Osofisan, 2004; Emuoyibofarhe, 2003; Principe, 1999; Principe et al., 2000; Oladokun et al., 2006; and Adepoju ,Ogunjuyigbe, and Alawode 2007).

A neural network is a more general method of regression analysis. Some of the advantages of the network over conventional regression include the following:

1) There is no need to specify a function to which the data are to be fitted. The function is an outcome of the process of creating a network.

2) The network is able to capture almost arbitrarily nonlinear relationships.

3) With Bayesian methods, it is possible to estimate the uncertainty of extrapolation.

The complexity and flexibility of the relationship that can be created is thus tremendous. Another desirable feature of network models is that they are readily updated as more historical data becomes available; that is, the models continue to learn and extend their knowledge base. Thus artificial neural network model are referred to as adaptive systems. This similarity to the human brain enables the neural network to simulate a wide range of functional forms which are either linear or non-linear. They also provide some insight into the way the human brain works. One of the most significant strengths of neural networks is their ability to learn from a limited set of examples (Principe et al., 2000; Anderson et al., 1994).

Generally a neural network consists of **n** layers of neurons of which two are input and output layers, respectively. The former is the first and the only layer which receives and transmits external signals while the latter is the last and the one that sends out the results of the computations. The n-2 inner ones are called hidden layers which extract, in relays, relevant features or patterns from received signals. Those features considered important are then directed to the output layer.

Sophisticated neural networks may have several hidden layers, feedback loops, and time-delay elements, which are designed to make the network as effective as possible in discriminating relevant features or patterns. The ability of an ANN to handle complex problems depends on the number of the hidden layers although recent studies suggest three hidden layers as being adequate for most complex problems, (Adefowoju and Osofisan, 2004).

There are feed-forward, back-propagation, and feedback types of networks depending on the manner of neuron connections. The first allows only neuron connections between two different layers. The second has not only feed-forward but also 'error feedback' connections from each of the neurons above it. The last shares the same features as the first, but with feedback connections, that permit more training or learning iterations before results can be generated.

ANN learning can be either supervised or unsupervised. In supervised learning, the network is first trained using a set of actual data referred to as the training set. The actual outputs for each input signal are made available to the network during the training. Processing of the input and result comparison is then done by the network to get errors which are then back propagated, causing the system to adjust the weights which control the network (Hertz, 1991; Antognetti and Milutinovic, 1991).

In unsupervised learning, only the inputs are provided, without any outputs: the results of the learning process cannot be determined. This training is considered complete when the neural network reaches a user defined performance level. Such networks internally monitor their performance by looking for regularities or trends in the input signals, and make adaptations according to the function of the network. This information is built into the network topology and learning rules, (Antognetti and Milutinovic 1991; Olmsted 1999).

Typically, the weights are frozen for the application even though some network types allow continual training at a much slower rate while in operation. This helps a network to adapt gradually to changing conditions. For this work, supervised training is used because it gives faster learning than the unsupervised training.

In supervised training, the data is divided into 3 categories: the training, verification, and testing sets. The Training Set allows the system to observe the type of relationships between input data and outputs. In the process, it develops a relationship between them. A heuristic states that the number of the training set data should be at least a factor of 10 times the number of network weights to adequately classify test data (Principe et al., 1999). About 60% of the total sample data was used for network training in this work.

The Verification Set is used to check the degree of learning of the network in order to determine if the network is converging correctly for adequate generalization ability. Ten percent of the total sample data was used in this study. The Test/Validation Set is used to evaluate the performance of the neural network. About 30% of the total sample data served as test data.

METHODOLOGY

Through extensive search of the literature and discussion with experts on student performance, a number of socio-economic, biological, environmental, academic, and other related factors that are considered to have influence on the performance of a university student were identified. These factors were carefully studied and harmonized into a manageable number suitable for computer coding within the context of the ANN modeling. These influencing factors were categorized as input variables. The output variables on the other hand represent some possible levels of performance of a candidate in terms of the present school grading system.

The Input Variables

The input variables selected are those which can easily be obtained from students' application/ record cards in the student's department. The input variables are:

1) UME score,
2) O/level results in Mathematics, English Language, Physics, and Chemistry,
3) Further mathematics,
4) Age of student at admission,
5) Time that has elapsed between graduating from secondary school and gaining university admission,
6) Parents educational status,
7) Zonal location of student's secondary school,
8) Type of secondary school attended (privately owned, State or federal government owned),
9) Location of university and place of residence, and
10) Student's Gender.

These factors were transformed into a format suitable for neural network analysis. The domain of the input variables used in this study shown in Table1.

Table 1: Input Data Transformation.

S/N	Input variable	Domain		
1	UME score*	Score	Normalized score	
2	O/level results	Math	A1–A2 A3–C4 C5–C6	1 2 3
		English	A1–A2 A3–C4 C5–C6	1 2 3
		Physics	A1–A2 A3–C4 C5–C6	1 2 3
		Chemistry	A1–A2 A3–C4 C5–C6	1 2 3
3	Further math	Present and passed Present, not passed Not present		1 2 3
4	Age at entry	Below 23 years 23 years – above		1 2
5	Time before admission	1 year 2 years 3 years – above		1 2 3
6	Educated parent(s)	Yes No		1 2
7	Zone of secondary school attended	South-west South-south East North Lagos		1 2 3 4 5
8	Type of Secondary school	Private State Federal		1 2 3
9	Location of school	Located in home state Outside home state		1 2
10	Gender	Male Female		1 2

** Since the general University Matriculation Examination performance may vary yearly normalizing is necessary. The* ***normalized score =(candidate score)/ (average score for the class).***

The Output Variable

The output variable represents the performance of a student on graduation. The output variable is based on the current grading system used by the university. However, for the scope of this project, the domain of the output variables represents some range of Cumulative Grade Point Averages (CGPA).

Table 2: Output Data Transformation.

S/N	Output Variable	Domain	
		Class	CGPA
1	GOOD	1st Class 2nd Class Upper	6.0 – 7.0 4.6 – 5.9
2	AVERAGE	2nd Class Lower	2.4 – 4.5
3	POOR	3rd Class Pass	1.8 – 2.3 1.0 – 1.7

The classification of output variable domain chosen above, that is 1st class and 2nd class upper as 'GOOD', 2nd class lower as 'AVERAGE', and 3rd class and pass as 'POOR', follows the practice of classifying candidates into these domains by most employing companies and postgraduate institutions, using the order stated above

Topology of the Network

After the data has been transformed and the method of training has been chosen, it is necessary to then determine the topology of the neural network. The network topology describes the arrangement of the neural network. Choosing the topology of the neural network is a difficult decision (Bose and Liang,1996; Emuoyibofarhe et al., 2003, and Oladokun et al., 2006). The network topologies available for are numerous; each with its inherent advantages and disadvantages. For example, some networks trade off speed for accuracy, while some are capable of handling static variables and not continuous ones. Hence, in order to arrive at an appropriate network topology, various topologies such as Multilayer Perceptron, recurrent network, and time-lagged recurrent network were considered. Due to the nature of our case study data, which is static and not sufficiently large to enable the use of complex topologies, the Multilayer Perceptron was selected.

Multilayer Perceptron

Multilayer Perceptrons (MLPs) are layered feed forward networks typically trained with static backpropagation.

These networks have found their way into countless applications requiring static pattern classification. Their main advantage is that they are easy to use, and that they can approximate any input/output map. The key disadvantages are that they train slowly and require lots of training data (typically three times more training samples than network weights) (Adefowoju and Osofisan, 2004).

The Network Layers and Processing Elements

The next step in building the neural network model is the determination of the number of processing elements and hidden layers in the network. Selection of the number of processing elements and hidden layers is a delicate one because having a small number of hidden layers in a neural network lowers the processing capability of the network. Similarly, a large number of hidden layers will progressively slow down the training time.

In determining the number of hidden layers to be used, there are two methods in the selection of network sizes: one can begin with a small network and then increase its size (i.e. Growing Method); the other method is to begin with a complex network and then reduce its size by removing not so important components (i.e. Pruning Method) (Hertz, 1991). The Growing Method was used in the building of the neural network model. Hence, the experimentation involves starting with no hidden layers and then gradually increasing them.

Trade-offs have to be made in determining the number of processing elements (PE). This is because, a large number of PE's can give the network a possibility of fitting very complex discriminate functions, and also involves a large number of weights. It has been shown that having too many weights can lead to poor generalization (Adefowoju and Osofisan, 2004). On the other hand, having too few PE's reduces the discriminating power of the network.

Since it is not possible to set the number of PE's analytically, the number of PE's is also varied in the study from 1 to 5 nodes, to arrive at the best performance network. The experiment is thus started with a small number of PE's, and observations made on the behavior of the learning curve.

If the final training error is of a small and acceptable value, then the network has the right number of PE's. However, if the final error is large, then one of two things has happened: either the learning curve has found itself in a local minimum or the network lacks enough capability to get the problem solved, so the number of PE's should be increased

The Data Set Grouping

In supervised training, the data is divided into 3 categories; the training set, verification set and the testing set. The training set enables the system to observe relationships between input data and resulting outputs, so that it can develop relationship between the input and the expected output.

A heuristic states that the number of the training set data should be at least a factor of 10 larger than the number of network weight to accurately classify test data with 90% accuracy (Adefowoju and Osofisan 2004). A total of 112 students records were used in the analysis. About 56% of the total data (i.e. 62 candidates \) were used as the training set, 30% (i.e. 34 candidates) as the testing set, and 14% (i.e. 16 candidates) used for cross validation.

Neural Network Topology

After the data classification, the neural network topology was built based on the Multilayer Perceptron with two hidden layers and five processing elements per layer.

Network Training and Validation Process

The network was trained with the number of runs set to three and the Epoch set to terminate at 1000. The training performance is then evaluated using the following performance measures:

The Mean Square Error (MSE):

$$MSE = \frac{\sum_{j=0}^{P} \sum_{i=0}^{N} (dij - Y_{ij})^2}{N\,P}$$

where:

p = number of output of processing element.
N= no of exemplars in the data set.
Yij=network output for exemplars i at processing element j,
dij=desired output for exemplars i at processing element j,

NETWORK TESTING

After the training and cross Validation, the network was tested with the Test data set and the following results were obtained. This involves given the input variable data to the network without the output variable results. The output from the network is then compared with the actual variable data. The comparison is summarized in the matrix bellow.

Table 3: Results from Testing.

Output / Desired	*Good*	*Average*	*Poor*
Good	9	3	1
Average	2	8	0
Poor	0	4	7

The network was able to predict accurately 9 out of 11 for the good data (which represents candidates with either a 1st Class or 2nd Class upper), 8 out of 15 of the Average data (which represents candidates with a 2nd Class lower) and 7 out of 8 of the Poor data (which represents candidates with a 3rd Class or Pass) used to test the Network's topology. This gives an accuracy of 82% for Good, 53% for Average and 88% for the Poor classification. This indicates an accuracy of about 74% for the Artificial Neural network's which is a fair performance going by similar results from the literature (Emuoyibofarhe et al., 2003; Adefowoju and Osofisan, 2004; and Oladokun et al., 2006).

CONCLUSION AND RECOMMENDATIONS

This study has shown the potential of the artificial neural network for enhancing the effectiveness of a university admission system. The model was developed based on some selected input variables from the pre admission data of five different sets of university graduates. It achieved an accuracy of over 74%, which shows the potential efficacy of Artificial Neural Network as a prediction tool and a selection criterion for candidates seeking admission into a university.

One limitation of this model stems from the fact that not all the relevant performance influencing factors are obtainable from the pre-admission record forms filled by the students. A model incorporating the use of results from a carefully designed oral interview administered to the students may likely be an improvement over the present model. Also the extension this research to non-engineering departments is recommended.

The current admissions system should be reviewed in order to improve the standard of candidates being admitted into the institution. A more adequate ANN may be very useful for such an exercise.

REFERENCES

1. Adedeji, O.B. 2001. "A Study Of The Relationship Between Students UME Results And Their Undergraduate Performance". Department Of Industrial And Production Engineering, University Of Ibadan. Unpublished Bsc Project.

2. Adefowoju, B.S. and Osofisan A.O. 2004. "Cocoa Production Forecasting Using Artificial Neural Networks". International Centre for Mathematics and Computer Science Nigeria. ICMCS117-136.

3. Adepoju, G.A., Ogunjuyigbe, S.O.A., and Alawode K.O. 2007. "Application of Neural Network to Load Forecasting in Nigerian Electrical Power System". *The Pacific Journal of Science and Technology*. 8(1):68-72 .

4. Anderson, G., Benjamin, D., and Fuss, M. 1994. "The Determinant Of Success In University, Introductory Economics Courses". *Journal Of Economic Education*. 25:99-120.

5. Anderson, G. and Mcneil, G. 2003. "Artificial Neural Networks Technology". Available At http://www.Dacs.Dtic.Mil/Techs/Neural/Neural_Toc.html\

6. Antognetti, P. and Milutinovic, V. 1999. *Neural Networks: Concepts, Applications, and Implementations*. Volumes I-IV. Prentice Hall: Englewood Cliffs, NJ.

7. Bakare, C.C. 1975. "Some Psychological Correlates of Academic Success and Failure". *African Journal of Educational Research*. 2.

8. Bose, N.K. and Liang, P. 1996. *Neural Networks Fundamental With Graphs Algorithms And Applications*. Mcgraw-Hill: New York, NY.

9. Dynan, K. and Rouse, C. 1997. "The Under Utilization Of Women In Economics: A Study Of Undergraduate Economics Students". *Journal Of Economic Education*. 28:50-68.

10. Emuoyibofarhe, O.J., Reju, A.S., and Bello, B.O. 2003. "Application Of ANN To Optimal Control Of Blending Process In A Continuous Stirred Tank Reactor(CSTR)". *Science Focus; An International Journal Of Biological And Physical Science*. 3:84-91.

11. Hertz, J.K. 1991. *Introduction to the Theory Of Neural Computation*. Addison –Wesley: New York, NY.

12. Joint Matriculation and Examination Board (JAMB). 2005. http://www.Jambng.Com/History/

13. Lage, M. and Tregelia, M. 1996. "The Impact Of Integrating Scholarship On Women Into Introductory Economics: Evidence From One Institution". *Journal Of Economic Education*. 27:26-36.

14. Mellamby, K. 1958. *The Birth Of Nigeria's University*. Methuen & Co.: London, UK.

15. Oladokun, V.O., Charles-Owaba, O.E. and Nwaozuru, C.S. 2006 "An Application of Artificial Neural Network to Maintenance Management". *Journal of Industrial Engineering International*. 2(3) South Tehran, 19-26.

16. Olmsted, D.D. 2003. *History and Principles of Neural Networks*. http://www.Neurocomputing.Org/History/History.html

17. Principe, J. 1999. *Neural Networks and Adaptive Systems*. John Wiley and Sons: New York, NY.

18. Principe, J.C., Eliano, N.R., and Lefeebvre, W.C. 2000. *Neural Networks And Adaptive Systems: Fundamental Through Simulation*. John Wiley and Sons: New York, NY.

ABOUT THE AUTHORS

V.O Oladokun, lectures in the Department of Industrial and Production Engineering, Faculty of Technology, University of Ibadan and currently serves as the Sub-Dean (Undergraduate) of the Faculty. He graduated with a B.Sc. (Hons) in Mechanical Engineering from the Obafemi Awolowo University (OAU) Ile-Ife. He has an M.Sc. and Ph.D. in Industrial Engineering from the University of Ibadan. He is a Member, Nigerian Society of Engineers; and a Member, Nigerian Institution of Engineering Management. He research interest includes enterprise model development and production systems optimization in a developing economy.

A.T. Adebanjo works as an Industrial Engineer with a multinational manufacturing company in Nigeria. He has a B.Sc. Degree in Industrial and Production Engineering from the University of Ibadan, Nigeria.

O.E. Charles-Owaba is an associate professor of Industrial and Systems Engineering at the University of Ibadan, Ibadan, Nigeria. In 1975, 1978 and 1981, respectively, he obtained his B.S., M.S., and Ph.D. degrees in Industrial Engineering at Texas Tech University, Lubbock, Texas, USA. He is a Member, Nigerian Institute of Industrial Engineers and the Nigerian Society of Engineers; and Fellow, Nigerian Institution of Engineering Management. His research interest includes the design of optimal management structures for business and public service organizations. He has over fifty articles in learned journals or peer reviewed conference proceedings and has authored five books including the popular civil service series for improving efficiency and effectiveness in government services

SUGGESTED CITATION

Oladokun, V.O., A.T. Adebanjo, and O.E. Charles-Owaba. 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.

Pacific Journal of Science and Technology