Accra Bespoke Multidisciplinary Innovations Conference (ABMIC)

& The Africa AI Stakeholders' Summit   14th December, 2021

Predicting Students' Academic Performance Using Artificial Neural Network: A Review of Intrinsic and Extraneous Factors

Aisida Akinwale Mayomi
Department of Information Technology
National Open University of Nigeria
Nigeria.

E-mail: aisidamayomi@gmail.com
Phone Number: +2347064794479

Predicting Students' Academic Performance Using Artificial Neural Network: A Review of Intrinsic and Extraneous Factors

Aisida Akinwale Mayomi

ABSTRACT

Students' academic performance is a measure of how the student has performed for a period in the academic parlance. It can be measured as semester-based, session-based or throughout the duration of the course of study. In recent time research, Artificial Neural Network (ANN) has been used globally, there are successful implementations in a wide range of classification, and prediction to be more efficient compared with other classifiers. Predictions are aimed at forecasting what might happen in the future by the means of estimating the likelihood of a certain events that has already occurred. We review literatures on the use of ANN for Predicting Students' Academic Performance Using Artificial Neural Network and also explore programming languages of choice for same. A number of intrinsic and extrinsic factors affecting students' performance were also identified

Keywords: Review, Data Mining, Approaches, Prediction, Factors, Students' Academic Performance, Artificial Neural Network

1. INTRODUCTION

Artificial Neural Network is a computational algorithm inspired by humans' central nervous systems, the structure and functions of biological neural networks. It works exactly like the way human brain works to processing of information. It includes a large number of connected processing units or nodes that work together to process information (Capizzi, *et al.,* 2015). ANN simulates the behaviour of biological systems composed of neurons. ANNs are computational models. It is capable of machine learning as well as pattern recognition and it can also apply for regression of continuous target attributes (Jumaat, 2014).

1.1 Artificial Neural Network (ANN) Defined
Jabin (2014) also defines Artificial Neural Network as an information processing technique that works like the way human brain processes information. And he emphasised that ANN includes a large number of connected processing units that work together to process information and also generate results that are meaningful from the processed information. ANN can be applied for both classification and prediction.
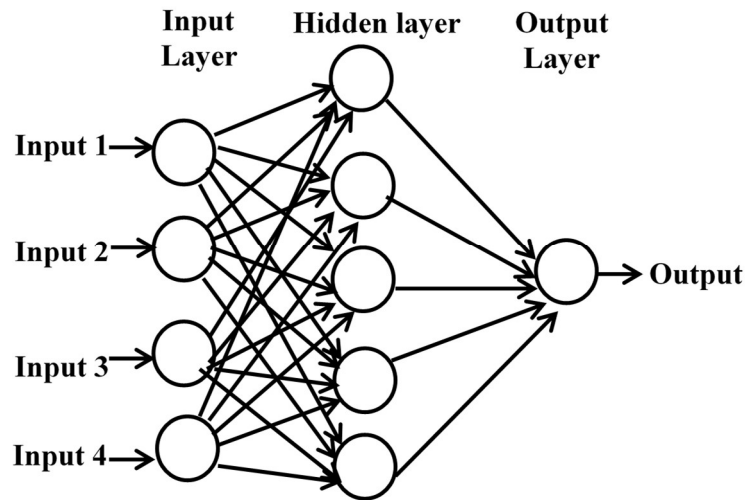
Figure 1: A Simple Artificial Neural Network with one Hidden Layer

Students' academic performance is a measure of how the student has performed for a period in the academic section (Leon, 2018). It can be measured on semester based, section based or throughout the student years of study, which can be gotten at the point of graduation. Several attributes such as Unit test, Assignment, Attendance, are attributes that contributes immensely to a university students' academic performance measure (Borkar and Rajeswari, 2014).

Students' academic performance known by a performance scale that varies from one school to another. Some school uses seven (7) point scales, some uses five (5) point scale while some four (4) point scale. This scale is known as Grade Point Aggregate (GPA) that can be gotten at the end of evaluation of a semester (half of an academic section) or Cumulative Grade Point Aggregate (CGPA) gotten at the end of evaluation of one academic section. In the area of study (National Open University), the performance is measured on a scale of five (5) Points and it is calculated as follows:

Each Course has a specific unit attached to it; these units vary from one (1) unit to six (6) units. A student has the liberty to register different subjects from their current level with a minimum of fourteen (14) total load units (TLU) and maximum of twenty four (24) total load units (TLU). Each subject result has a grade that is determined by the score a student has in a particular subject at the end of evaluation. From seventy five (75) to one hundred (100) is grade "A" an equivalence five (5) points, sixty (60) to sixty nine (69) is grade "B" an equivalence of four (4) points, fifty (50) to fifty nine (59) is grade "C" an equivalence of three (3) points, forty five (45) to forty nine (49) is grade "D" an equivalence of two (2) points, forty (40) to forty four (44) is grade "E" an equivalence of one (1) point and less than forty (40) is grade "F" an equivalence of zero (0) point. The Total Credit Point (TCP) is the sum of the products between the grade point and the course unit.

The GPA then is gotten by dividing the TCP by the TLU and it is for a particular semester.

$$GPA = \frac{TCP}{TLU}$$ (1.1)

The CGPA is the cumulative of Equation 2.1. It is the summation of the TCP a student has earned, divided by the summation of the TLU a student has offered for all the semesters or sections used in the school.

$$CGPA = \frac{\sum TCP}{\sum TLU}$$ (1.2)

A CGPA values from 4.5 to 5.0 is a first class grade, 3.5 to less than 4.5 is Second class upper grade, second class lower is CGPA value between 2.5 to less than 3.5, the third class equally known as pass has a value ranging from 1.5 to less than 2.5, while any CGPA value less than 1.5 is a fail.

1.2 Programming Languages for Implementing ANN
Python is the choice programming language for the implementation of the desktop application software. It is a High level programming language with formatting styles that makes it easy to read and to write; with several associated libraries and tools that performs different duties. Some of these are as follows:

Scikit-learn
Scikit-learn is also known as sklearn, it is a library in Python that provides simple and efficient tools for machine learning. sklearn implements several algorithms for classification, regression, clustering, dimensionality reduction, model selection and pre-processing.

TensorFlow
TensorFlow is an open source Python library for dataflow programming across a range of task. It is a symbolic math library and also used for machine learning application such as neural network.

Keras
Keras is an open source neural network library that is written in Python, capable of running on top of TensorFlow to provide fast experimentation with deep learning Neural Network. It focuses on being user friendly, modular and extensible.

NumPy
NumPy is the basic package for scientific computing in Python. It provides a powerful N-dimensional array object which is the fundamental data container for Scikit-learn. The array object is accompanied by several functions for manipulating it. It also provides some useful linear algebra, Fourier transform and random number facilities. NumPy is a fundamental library in Python as many other libraries rely on its services.

Pandas
Panda is a library in Python Programming language; it provides high performance, easy-to-use data structures and data analysis tools. It is also use to read data of several data types into a Python environment.

SciPy

SciPy is a Python library that provides various easy-to-use and efficient numerical routines such as numerical integration and optimization. It is a fundamental library for scientific computing.

PyQt5
PyQt5 is a set of Python bindings for the cross-platform Qt library. It is a plug-in that can be used to develop Graphical User Interface (GUI) applications in Python. It implements over 400 classes and over 6,000 functions. These include GUI widgets, database access, XML parser, Scalable Vector Graphics support and ActiveX controls. It has a module that connects the Python program to the GUI designer for actual functionality.

Integrated Development Environments (IDE)
The Integrated Development Environments (IDE) of choice for the implementation of the software isSpyder and Qt Designer.

Spyder
Spyder is an open source Python IDE designed for scientific programming. It integrates some common open-source libraries and also contains the interactive Python (IPython) console as it engage in syntax highlighting. Spyder also features multiple consoles and a GUI variable editor.

Qt Designer
Qt Designer is a toolkit for designing graphical user interfaces (GUI) using Qt Widgets. It allows building GUIs in a way that it can be reviewed and altered until the intended GUI is achieved. It can be easily integrated with Python code and its use of slots and signals makes it easy to assign behaviour to GUI elements.

2. RELATED WORKS ON STUDENT ACADEMIC PERFORMANCE PREDICTION

Prediction is given a specific instance of an event happening in the future instance. It is often time a generalised idea gotten from existing occurrence of that event in the time past, and often time it is correct while it could be wrong sometimes (Kabakchieva, 2012). Predicting students' academic performance is a field that researchers has shown a whole lot of interest to research on because it is very important to education environment, especially academic institutions worldwide, as the strength of the institution is dependent on the performance of their students and the application of their knowledge acquired (Gerritsen, 2018).

Paris *et al.* (2010) use data mining methods to classify students' data so as to predict their grades. The research is useful a lot to identify the weak students with low grades and do not understand their subjects easily, from the brilliant ones who has good grade and can assimilate very fast. It equally helps the administration in care of students' academic performance to take helpful measures at early stages to correct and produce better results in their area of study through organising remedial means for the weak students' category.

Sembiring (2012) also uses Kernel K-Means and Support Vector Machine to predict students' academic performance in his research conducted at University of Malaysia Pahang, using a dataset from student academic databases of the institution and he surveyed the essential motivation and social stand of the undergraduate students in the second semester and third session 2009/2010 academic section. The research understood the relationship between psychometric factors of student (the predictors) with their respective academic performance.

The model used for the prediction is Kernel K-means algorithm for clustering and Smooth Support Vector Machine for classification.

Mishra *et al.* (2014) equally researched on Mining Students' data for their academic performance prediction, they created a model using Waikato Environment for Knowledge Analysis (WEKA) and then analyses the data of 250 Students of Master of Computer Association (MCA) students' from Guru Gobind Singh Indraprastha (GGSIP) University using Random Tree and J48 algorithms. The performance of both the algorithm was satisfactory. Random Tree has higher accuracy of 94.418% while J48 has accuracy of 88.372%.

AlMalaise *et al.* (2014), also in their research to predict students' academic performance with the use of Multi Agent Data Mining. The prediction was done based on the students' data having very high prediction accuracy. And it provides help to the students with low performance through optimization rules. The implementation was evaluated by investigating the prediction accuracy of Adaboost.M1 and LogitBoost with C4.5 single classifier method. The results show that using SAMME Boosting algorithm/technique improves the prediction accuracy and outperformed C4.5

Machine learning was used to predict whether a student will be successful academically or not. The research focus on the comparison between two machine learning methods (Decision trees and naïve Bayes classification) in terms of the rate they improve the students' performance prediction, With the use of two different datasets, 353 and 265 students of different levels raw data, the result shows that naïve Bayes classification and decision trees has prediction accuracy of 98% and 78% respectively (Pojon, 2017).

Yehuala (2015) also applied the decision trees and Naïve Bayes algorithms to predict the likelihood of success or failure at DebreMarkos University (DMU). The dataset he used comprise of all 11,873 undergraduate students of the University. Using the factors such as; student result, gender, class population, number of courses given in a semester, and the student course of study as the major factors affecting the student performance. His highest prediction accuracy was 92.34% obtained with the use of decision tree algorithm using 10-fold cross validation.

3. FACTORS AFFECTING STUDENT GENERAL ACADEMIC PERFORMANCE – STUDENTS'-BASED

Several effort has been put in place to ensure that factors that are responsible for the performance of students' academic performance is known, so as to improve the student performance by putting the necessary resources together to help promote the positive influencing factors and discourage the factors contributing negatively (Syed and Raza, 2006).Some of these factors in categories and as follows:

3.1 Factors That Affect Academic Performance
Matthew *et al.*(2018) propose some factors that are common to students and are likely to have an influence on the student performance and these factors often times can be controlled by students. And they are:

Age
This is the number of years the student has lived from birth to present day, and it can be measured in Years. This can be acquired with the Equation 2.2.1.1

$$Age = Present\ Year - Year\ of\ Birth \qquad (3.1)$$

Some students gain admission into tertiary institution early while it takes some students a long time to gain admission into the university. This is likely to have effect on the academic performance of a student in the institution.

Gender
This tells the fact of a student being a Male or Female individual.

Students' Study Habit
This is the amount of student's effective study put in place for every of his/her finite set of courses (subjects) registered per section. This is relative to the frequency of reading, hours spent reading and the period students read. Some students find it difficult to read while for others it is convenient. This difference in students is likely to contribute to help those that study well to perform better.

Students' Fear and Perception about their course of study
This is the sensitivity student's has based on the knowledge and views they have about their department, when a student has a positive perception, it is likely to help the students in their respective subjects offered in department, and a negative perception may affect the performance of the student.

Students' Attendance in Classes
This is the measure of time students attends classes, this shows the level of effort and seriousness of students the student shows towards their studies. The numbers of classes attended and their punctuality to the classroom. A student who attends classes regularly is likely to have a good performance.

Students' Health
This is the measure of how healthy the student is, a student medical condition could be that the student fall sick often, get tired easily, on a compulsory medication, had an accident during assessment period, psychological ailment and other related health issues might influence the student academic performance.

Students' Communication Skill
This is the ability of the student to present their ideas in an understandable way, give response to questions in classes, and ask questions when he/she is not clear about a topic. Communication skills aid interactions between students and their lecturers. A research conducted to find out the factors that contributes to secondary school students' academic performance in Ilala district, focused on group discussion and emphasised that students with poor knowledge of vocabulary or terminology used in their various subject finds it difficult to make good and logical statements about the subjects and it causes a defect on the performance of the students greatly in the environment (Maganga, 2016).

Tutorials and Extra Classes
These are the extra effort students engage in to ensure they have a better understanding of all their respective courses (subjects) offered within the academic section. These include extra classes (tutorials), assistance from course mates, senior colleagues and lecturers. Tutorials

are gotten in different way, we have the online tutorial where the student register for courses online, watch tutorial video, download and study online course materials and interact in educational forums where questions could be asked and response will be given by experts in that field of study. There is also offline tutorials, where tutorials are given by physical tutors in a class room and interaction between students' and tutor are directly. And finally, there is mobile tutoring, instructions are gotten from mobile application and the student follows the instructions one after another. In this medium, tutors are absent.

Extracurricular Activities
This is the amount of other activities within and outside school that is allowed other than academic activities such as: sport/games activities, social activities and religious activities. Most of these activities takes time and reduced the time a student has for academic related works. The more activities a student engage in, the less time he/she has for academic activities which is likely to tell on the performance of the student.

4. FACTORS THAT CAN AFFECT STUDENTS ACADEMIC PERFORMANCE – LECTURERS'-BASED

Some factors are common to lecturers only and may or may not be controlled by the students. These are likely to influence the performance of students(Matthew et al., 2018). These factors include:

Lecturers Attitude
This is the measure of lecturer aggressiveness, ability to interact and encourage students to increase their importance to the likeness of his/her course (subject), interest to explain things to the simplest form and to the understanding of the students.

Lecturers Teaching Styles
This is the pattern of lecturer(s) teaching. In the likes of if he/she present a well-structured slide and visual aid to aid the lecture, explain lectures explicitly, dishes out voluminous course materials, carry the class along while doing the presentation (teaching) and a quality assurance that the lecturers understand the course they are teaching.

Lecturers Availability
This is the presence of a lecturer and how accessible he/she was when needed by their student(s), some lecturers are always available while some are not. And this can likely affect the performance of a student.

Lecturers Dedication to Work
This is the devotion of the lecturer(s) to the course they teach. It includes their readiness to explain things in its simplest form and extra effort put in place to ensure their student(s) good performance(s), their punctuality to classes and their interest to the course they tutor. Some lecturers do not come to classes regularly and some do not spend much time in classes and it is likely to have effect on the performance of students.

5. UNIVERSITY FACTORS THAT CAN AFFECT STUDENTS ACADEMIC PERFORMANCE

Some factors are also peculiar to the university structure and facilities and are likely to influence the academic performance of students. Singh *et al.* (2016) in their research to investigate the factors that influence the performance of students' academic performance of

200 management students from ten (10) different management institutes of Haryana state, they emphasised on the relationship between the institution of learning facilities and the students' academic performance using statistical techniques, and then concludes that the performance of the management students under evaluation can significantly be improved by the institution providing the students' with appropriate learning facilities required. The University factors considered are:

University Facilities
This measures the availability of various and learning facilities that aid learning such as: well-equipped practical laboratories, lecture hall/classroom (sufficient and conducive), provision of visual aid and sound systems. Some university does not possess necessary facilities their respective department used and this has made the university not to be conducive for learning and it is equally likely to affect student academic performance.

Students Class Population
This is the fraction of a lecturer to the students' population he/she teaches in classes per time. An overpopulated class will be stuffy and not conducive; this also could affect the way students assimilate in classes and could possibly affect academic performance of a student.

Lecture Time Duration
This measure how conducive and convenient the lecture schedule of the University was. Lecture time ranges from one hour to two hours. In the University under consideration (National Open University), lecture duration per week is dependent on the unit of the course, one unit course takes on hour of lecture in a week, two unit courses also takes two hours of lectures in a week.

Akessa and Dhufera (2015) in their research to evaluate the factors that influence students' academic performance in Rift Valley University (RVU), Jimma, Ethiopia, states that University facilities may lack some resources like: reference materials, well organized laboratory equipment's and computer laboratories, and lack of interest to some subject matter (specific field of study). This was the problem that has been observed at RVU Jimma campus, where their survey is carried out. And these also were common to the environment where this research is carried out National Open University.

Family Factors That Can Affect Students Academic Performance
Some family related factors that are likely to affect students' academic performance are

Family Income
This is how large the income of the parent/guardian of the student(s) was, in relation to the finance used to support the student(s) at school. The family that has large income are likely to support the students financially. According to Escarce (2003), he researched on factors that influence the educational opportunities available to adolescent students and on the chances of the student educational success. He adds that due to residential differences based on low income, students usually attend schools with lower funding levels and this have reduced achievement motivation and much higher risk of educational failure when he compared the results with their more affluent counterparts, low-income adolescents receive lower grades, earn lower scores on standardized test and are much more likely to drop out of school. Sentamu(2003) agrees with him as he argued that social class determines what school a child will attend and whether the child will pass an examination or not.

Family Stress
This is the amount of trouble students get as a result of the news from their respective home, pressure from family members and broken home (divorced parents). This might affect the mental being of the students and also can affect their academic performance.

Parent Level of Education
This is the degree of the students' education levels; it shows either the student(s) parent(s) are illiterate or literate. In some family they have a legacy that is strong in terms of education qualification, this may encourage a student to beat records that has been set in the family, and parents with university experience may have the privilege to be able to give precise advice to their children on school activities and could possibly help the performance of a student. Owens (1999) in her research that explores the beliefs about academic achievement, studied the relationship between students' parents or guardians educational status attainment to the academic achievement of the students and she established a fact that the educational status of parent or guardian does have a relationship with their children academic performance

Students' Proper Guidance
This is the level of support, advice and guidance given to students from their parents/guardians. Some parent takes their time to ensure they know the activities their children engaged in at school, their friends and show concerns by giving advice to them to motivate them. This is likely to contribute to student academic performance when they felt love and support from family members.

Kyoshaba (2009) investigated factors that affect students' academic performance of 357 undergraduate students in Uganda Christian University (UCU), he established that there is a positive relationship between parents' socio economic status and students' academic performance, from his analysis he concludes that it is very important to the academic performance of the student as parents provides financial and psychological support for their children through and this creates an environment that encourage skill development and learning instinct necessary for success at school.

5.1 Other Factors That Can Affect Students Academic Performance
Some other factors that contribute to Students' academic performance may include:

Internet Access
The availability of access to the internet in due time helps the study of a student and it get them informed about what is evolving and how fast the development is taking place.

Electricity
This measures the rate of power supply availability to help student(s) study. Electricity is needed within and outside the school environment for effective study process and completion of assignments. Some of these factors have been considered by several researchers to emphasis on the rate they contribute to the student academic performance and their respective contribution, to know if the contribution is positive or negative(Asif et al, 2017).

7. CONCLUDING REMARKS

In this work, we identified the application of Artificial Neural Networks for predicting students' academic performances. We highlighted several factors that must be taking into consideration when building a machine learning algorithm for the purpose of predicting performance. A number of programming languages that can be used to implement ANN were also highlighted.

BIBLIOGRAPHY

1. Akessa, G. M. (2015). Factors that Influences Students Academic Performance : A Case of Rift Valley University. Journal of Education and Practice, 6(22), 55–64.
2. Alexandridi, A., Stylios, C. D., & Manis, G. (2003). Neural Networks and Fuzzy Logic Approximation and Prediction for HRV Analysis, 312–317.
3. AlMalaise, A., Malibari, A., & Alkhozae, M. (2014). Students' Performance Prediction System Using Multi Agent Data Mining Technique. International Journal of Data Mining & Knowledge Management Process (IJDKP), 4(5), 1–20. https://doi.org/10.5121/ijdkp.2014.4501
4. Ani, R., Sasi, G., U, R. S., & Deepa, O. S. (2016). Decision Support system for diagnosis and prediction of Chronic Renal Failure using Random Subspace Classification, 1287–1292.
5. Asif, R., Merceron, A., Ali, S. A., & Haider, N. G. (2017). Analyzing undergraduate students' performance using educational data mining. Computers and Education, 113, 177–194. https://doi.org/10.1016/j.compedu.2017.05.007
6. Hackeling, G. (2014). Mastering Machine Learning with scikit-learn. Retrieved from http://books.google.com/books?id=fZQeBQAAQBAJ&pgis=1
7. Kabakchieva, D. (2012). Student performance prediction by using data mining classification algorithms. International Journal of Computer Science and Management Research, 1(4), 686–690.
8. King, C., Lee, I.-C., & Hong, Y. (2018). Reliability Analysis of Polymeric Materials, 1–13.
9. Maganga, J. H. (2016). FACTORS AFFECTING STUDENTS' ACADEMIC PERFORMANCE: A CASE STUDY OF PUPLIC SECONDARY SCHOOLS IN ILALA DISTRICT, DAR-ES-SALAAM, TANZANIA A DISSERTATION SUBMMITED IN PARTIAL FULFILMENT FOR THE REQUIREMENTS OF THE MASTER DEGREE IN EDUCATION ADMINISTARATION, PL.
10. Malik, S., Singh, S. P., & Singh, P. (2016). Research Paper Factors Affecting Academic Performance of Students. Indian Journal of Research, (April), 0–3.
11. Martha, K. (2009). Factors affecting academic performance of undergraduate students at Uganda Christian University. Educational Management, (December), 1–92.
12. Matthew, F. T., Adepoju, A. I., Ayodele, O., Olumide, O., Olatayo, O., Adebimpe, E., … Funmilola, E. (2018). Development of Mobile-Interfaced Machine Learning-Based Predictive Models for Improving Students' Performance in Programming Courses. International Journal of Advanced Computer Science and Applications, 9(5), 105–115. https://doi.org/10.14569/IJACSA.2018.090514
13. Nigel Mathers, N. F., & Hunn, A. (2010). Surveys and Questionnaires. Policy, 2–3.

14. Pedregosa, F., Weiss, R., & Brucher, M. (2011). Scikit-learn : Machine Learning in Python, 12(January), 2825–2830. https://doi.org/https://dl.acm.org/citation.cfm?id=2078195

15. Pojon, M. (2017). Using Machine Learning to Predict Student Performance. India, (June).

16. Razavi, S. V, Jumaat, M. Z., & Ei-Shafie, a H. (2011). Using feed-forward back propagation (FFBP) neural networks for compressive strength prediction of lightweight concrete made with different percentage of scoria instead of sand. International Journal of Physical Sciences, 6(6), 1325–1331. https://doi.org/10.5897/IJPS11.204

17. Tang, Z., & Fishwick, P. A. (1993). Feedforward Neural Nets as Models for Time Series Forecasting. ORSA Journal on Computing, 5(4), 374–385. https://doi.org/10.1287/ijoc.5.4.374

18. Venugopalan, S., Xu, H., Donahue, J., Rohrbach, M., Mooney, R., & Saenko, K. (2015). Translating Videos to Natural Language Using Deep Recurrent Neural Networks. Naacl Hlt, (June), 1494–1504. https://doi.org/10.3115/v1/N15-1173

19. Yilmaz, I., Erik, N. Y., & Kaynar, O. (2010). Different types of learning algorithms of artificial neural network (ANN) models for prediction of gross calorific value (GCV) of coals. Scientific Research and Essays, 5(16), 2242–2249.