IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 25, Issue 1, Ser. I (Jan. – Feb. 2023), PP 11-16
www.iosrjournals.org

Artificial Neural Network Analysis for Predicting The Performance Of Students In An Examination

Nurudeen Ajibola IROMINI[1], Leonard Iyase OGEWEY[2], Adedoyin Simeon ADEYEMI[3], Sarafa Rotimi OLATUNJI[4]

1,2,4. Department of Computer Engineering, Federal Polytechnic, Offa, Kwara State.
3. Department of Computer Science, Federal Polytechnic, Offa, Kwara State.

Abstract
The academic performance of students in education is one of the important criteria for determining the quality of students going into the labor market. There are several factors affecting the academic performance of students in higher institutions of learning and these factors have seriously led to an increase in failure rates of students in their examinations. Some of the factors affecting the performance of students are the wrong time of giving orientation to students, poor guidance, poor communication skills, attitude and mentality of students towards their studies, joining of bad gangs, and wrong use of technology, just to mention a few. Academic performance of HND I students of Computer Science, Federal Polytechnic, (FEDPOFFA) were predicted using an artificial neural network model in terms of their total scores in the examination. Artificial neural networks are forecasting methods that are based on simple mathematical models of the brain. They allow complex nonlinear relationships between the response variable and its predictors. In this paper, the performance of students in the class of 183 students was analyzed using an artificial neural network model. The model was designed using their practical scores, continuous assessment score, and their exam scores that were collected in the Computer Architecture course. 70% of the data were used for training, 15% were used for validation, and the remaining 15% was used for testing. The performance of the students and the regression analysis of these students were carried out using an artificial neural network. The network model was then applied to a certain set of students and it was able to accurately predict the students' total scores.
Keywords: *Artificial Neural Network, regression, performance, algorithm*

Date of Submission: 05-01-2023 Date of Acceptance: 19-01-2023

I. Introduction

Student performance describes the outcome of students in their examination and the extent to which the students as well as lecturers in an institution achieve their educational goals to achieve their desired performance. The performances of students play an important role in producing competent graduates who will become leaders in our country. Education is very important to every Nigerian citizen to achieve essential knowledge and to understand varieties of subjects to be applied in daily life. Education can be viewed as a process through which a society passes on knowledge, values, and skills from one generation to another. The way new skills, knowledge, and values can be acquired is regarded as learning.

This paper discussed various factors that amounted to poor performance of students in various polytechnics, especially Federal Polytechnic, Offa, Kwara State. Among the factors that are affecting the performance of students is their inability to use and adapt technology to the area of their education. Students prefer using their phones for social media activities instead of using them to search and explore course materials for their area of study. Because of this, much of their time is being spent on reading irrelevant things on their phones and little or no time is spent on reading their notes and textbooks let alone doing more findings or research on their phones or laptops. Artificial Neural Network (ANN) Model was also designed to determine the grades of students in their examination using the practical scores, the continuous assessment, and their scores in the examination.

Determining students' performance cannot be taken with levity hands as it is one of the yardsticks to identify good and quality students who will become the country's future leaders.

II. Related Works

Neural Network is another popular technique used in educational data mining. The advantage of the neural network is that it can detect all possible interactions between the predictor's variables (Shahiri et al 2015). Simulation approaches such as Artificial Neural Network (ANN) model are widely used by researchers to solve a variety of problems in science and engineering, particularly in some areas where the conventional modeling methods fail (Mohd-Noor et al, 2015). Artificial Neural Network (Chima and Duroha, 2019) is a more appropriate model for forecasting capital markets such as stock and currency, Government institutions, and financial institutions. Nowadays accounting and financial classification and prediction problems (Trifonof et al , 2021) are a high challenge and researchers use different methods to solve them. Many researchers have used ANN to analyze traditional classification and prediction problems in accounting and finance.

Examination practices, poor supervision, and impersonation, among others are some of the factors affecting the quality of assessing students' education in secondary schools (Bappah et al., 2021). According to Mustapha and Khan (2012), Class schedules, Class size, textbooks, homework, the environment of the class, the technology used in the class, exam systems, extra-curricular activities, family and work activities, finance, among others are some of the factors affecting students' academic performance.

Lidia, Ford, and Tokuro (2021) performed a systematic literature review on machine learning algorithms in determining the performance of students. Eleven selected research articles were selected from the databases of IEEE access and science direct. These articles were chosen not only according to the exclusion and inclusion criteria but also by using search results. The articles were carefully selected based on the year of publication, the type of article, and the abstract. The set of articles considered for review is from 2019 to 2021. One of the criteria used for the selection of the publications is exclusion criteria. The exclusion criteria are for the publications that are not available in full text, not in English as well as the publications that did not discuss predictions of students' performances and achievements based on machine learning algorithms. All the considered publications were grouped according to research questions. This research found that the prediction of the success of student learning is mostly done by classification machine learning algorithms. ANN, Naïve Bayes, Logistic regression, SVM, and Decision trees are mostly used in classification algorithms. In most of the research papers considered, students' success is predominantly classified into two or three categories which are pass/fail and pass/fail/excellent.

Okubo et al (2017) predicted the performance of students by Recurrent Neural Network (RNN) using multiple course data. In this paper, the learning log from 937 students who attended one of the six courses by two teachers was considered. The prediction performance and multiple regression analysis on these data were conducted. For the prediction performance analysis, the year 2015 courses data which are courses 1 to 5 were used for the training data, and the year 2016 courses data which are courses 5 and 6 were used as test data. When comparing the results of prediction performance by RNN and multiple regression analysis, the accuracy of prediction is the same after the fifth week. However, the accuracy of prediction by RNN before the fourth week is higher than that of multiple regression analysis.

The prediction of students' yearly performance using a neural network was carried out by sikder, Udin, and Halder (2016). Data from 120 students from the Department of Computer Science and Engineering of Bangabandhu Sheikh Mujibir Rahman Science and Technology University were collected via an online survey where fourteen factors were considered including not only academic data but also other personal information such as family education, living area, social media interaction, drug addiction, and others. MATLAB was used as a data mining tool to filter unnecessary data and most of the data were converted to numeric data as MATLAB only works with numeric data. These data were divided into two: the first thirteenth (13) factors were used for training while the remaining fourteenth factor was used as target data. In the study, Neural Network was used to predict students' yearly performance in the form of Cumulative Grade Point Average (CGPA). The results from ANN were compared with the original CGPA using the results from seven students. The highest accuracy percentage was 99.986 when the student's original result is 2.7 while the lowest accuracy percentage is 91.547 when the student's original result is 3.83.

III. Methodology

(i) **Data Collection:** The data used were collected from department of Computer Science, Federal Polytechnic Offa for students in Higher National Diploma I i.e. HND I. The data collected for these students were practical scores, continuous assessment scores, and exam scores. The data collected were scores of students that offered Computer Architecture course in the department.

The input variables to the ANN network are the practical scores, C.A. scores, and exam scores while the total scores were determined by the network which is the output variable. The network topology chosen for the networks was Multilayer Perceptron.

(ii) **Data Processing**: After the data collection, there is sometimes a need to filter some unnecessary data. Here, all the data needed are already in the required format. The data which are practical scores, continuous assessment scores, test scores, and the total scores are numerical values which are the format that Artificial Neural Network in MATLAB can work on. Since four factors are put into consideration, these factors were divided into two parts. The first part is the INPUT part which uses three factors from the four factors considered. The second part is called the target value. This is the essential part as it is the part that the neural network will work on to determine the pattern of the data and predict results based on the input parts. The inputs to the neural network model are practical scores, continuous assessment scores, and test scores while the output of the network is the total scores.

(iii) **Neural Network Model:** Neural Network has to be designed to make a prediction. The first step here was that we designed a neural network and this network was trained by the already preprocessed data. The supervised neural network was used and the network was trained using the Levenberg-Marquardt backpropagation algorithm.

Supervised learning is the machine learning system of inferring a function from labeled trained data. In a supervised learning system, two kinds of data are provided which are inputs and outputs. The system processes the inputs and compares its verdict outputs against the required outputs. Sometimes system finds some errors. These errors are back-propagated to the system. Then the system adjusts the weights which control the network. This process happened recursively as the weights are frequently tweaked.

The data which permits the training process is called the training set. In the training phase of the network, some dataset is preprocessed frequently as the connection weights are always experienced. Sometimes, the training process continues some days. This process is stopped only when the system reaches some statistically desired output or expected accuracy. However, some networks never learn because the input data cannot find the desired output. Networks cannot manipulate an adequate amount of data so part of the data can be held back as a testing phase. This network consists of many layers including multiple nodes in each layer that are capable of storing information.

IV. Implementation

In the implementation of our system, the preprocessed dataset was used in MATLAB. Then the neural network was designed with 5 neurons in the hidden layer as shown in Figure 1. We added the Levenberg-Marquardt backpropagation algorithm as our training algorithm while our training function trainlm function. The dataset was divided into three parts which are training, validation, and testing parts. Out of 195 datasets considered, 75% were used for training, 5% were used for validation and the remaining 20% were used for testing. Errors were found in the difference between the target results and the output results. Figure 2 shows the different error histograms with 20 bins.

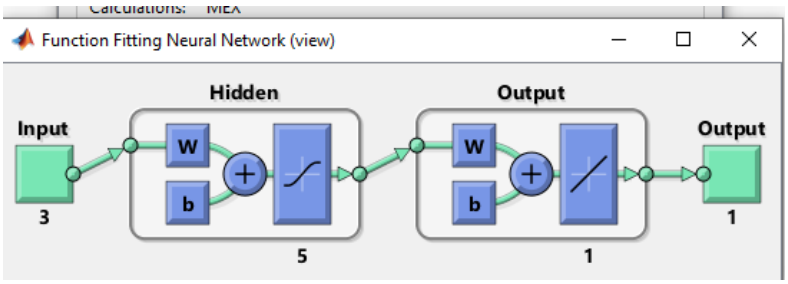

Figure 1: ANN Model

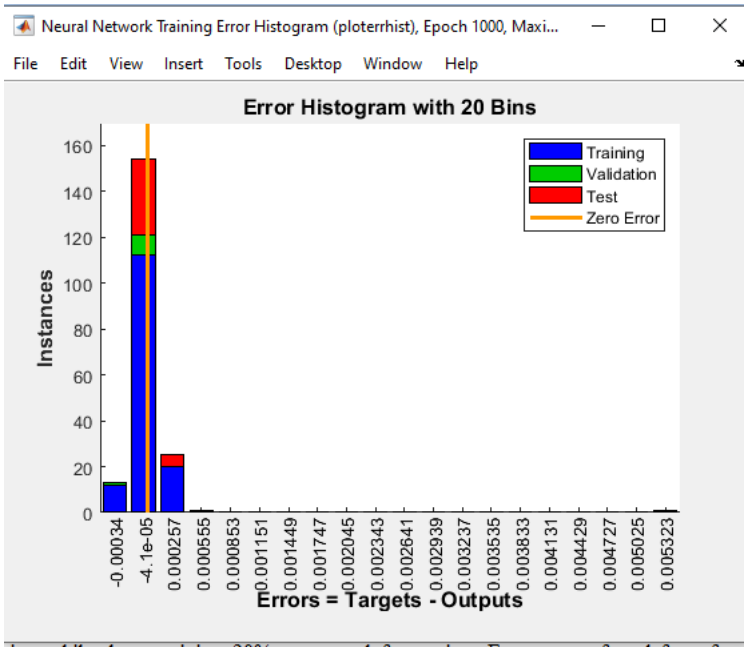

Figure 2: Errors Histogram

Each time the errors between the targets and the outputs were computed after 1000 iterations, all of the errors were within -0.00034 and 0.005323 and most of the errors were within -0.000041. This showed that the outputs are greater than the targets according to the errors formula as shown in equation 1.

$$errors = \texttt{targets} - \texttt{outputs} \quad (1)$$

V. Results

The predicted results were determined after running our network model. The result was best validated at epoch 1000 as shown in Figure 3 which is the performance of the neural network. It showed that the best validation performance is 2.206e-8 (0.00002206) at epoch 1000 which gives a performance accuracy of about 99.9%

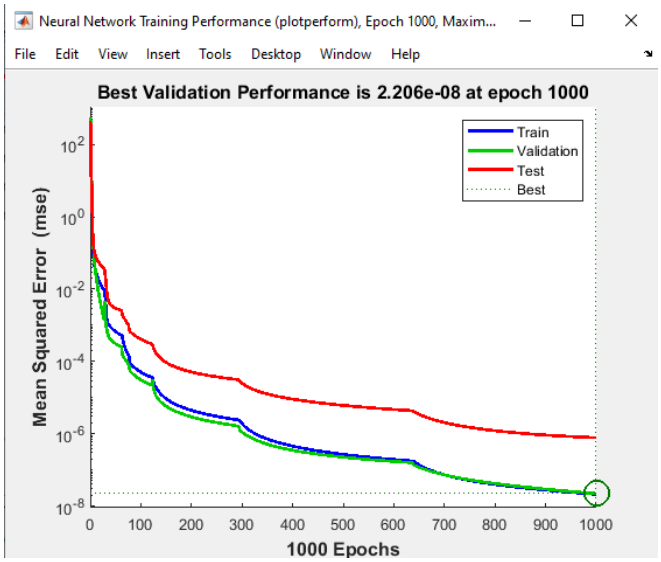

Figure 3: Neural Network Training Performance

For the dataset regression, three parts were involved which are training, testing, validation, and all phase regression analysis of the network as shown in Figure 4 where the dotted lines represented the target output which is the perfect result and the solid lines represent the best fit linear regression line between outputs and the targets. The R-value is an indication of the relationship between the outputs and the targets. The R-value is 1 in all the plots which indicated the exact linear relationship between the outputs and the targets.

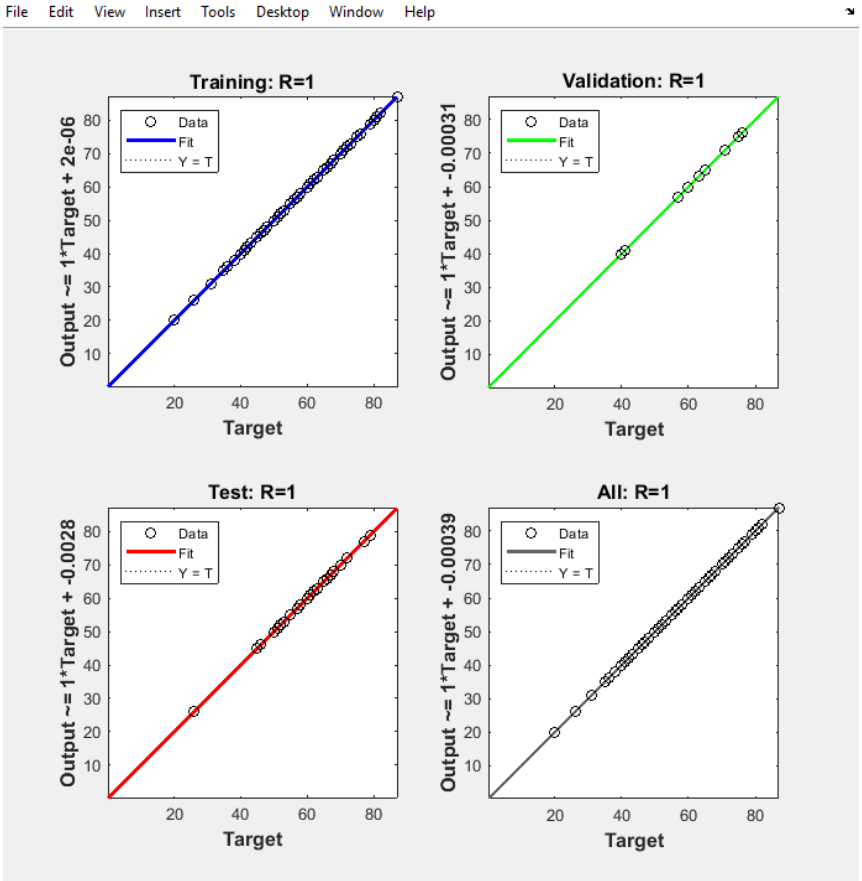

Figure 4: Regression Analysis of the Network

The network was tested to determine the total score of students in the Computer Architecture course having considered their practical scores, C.A. scores, and exam scores. Table 1 shows the students' original result and the predicted result by the Artificial Neural Network (ANN).

Table 1: Original and ANN results

Practical score	C.A.	Exam	Total	ANN Output
23	15	34	72	71.9995
21	10	29	60	59.9999
15	9	51	75	75.0012
11	14	22	47	46.9999

VI. Conclusion

The paper has treated the performance of students in the Computer Architecture course after considering their scores in practical, continuous assessment, and examination. All these are the factors that determine the total scores of students in some of the courses. Artificial Neural Network is regarded as one of the ways of determining the performance of students. This paper will help in the calculation of the total scores of students after giving the required inputs to carry out the calculation. In this paper, 195 datasets of students of Computer Science students were collected. 75% of this dataset was used for training, 5% was used for validation and 20% was used for testing while training the network. Our network was able to determine accurately the total scores of given students using the practical scores, C.A., and exam scores. This network can also be applicable to carry out any similar works. The Artificial Neural Network (ANN) is very good at making predictions.

References

[1]. Bappah A., Magaji M.A., and Usman B. (2021). Factors Affecting Students' Performance in Higher Institutions of Learning in Gombe State, Nigeria. International Journal of Research and Review, Vol., 8; Issue 3. www.ijrrjournal.com. E-ISSN: 2349-9788; P-ISSN: 2454-2237.

[2]. Chima A.N. and Duroha A.E. (2019). Artificial Neural Network Application in Prediction – A Review. African Journal of Computing & ICT, Vol. 12, No 4, pp 75 85.
[3]. Lidia S., Ford L.,and Tokuro M. (2021). Machine Learning Algorithm to Predict Students' Performance: A Systematic Literature Review.TEM Journal, Volume 10, Issue 4, Pages 1919-1927, DOI:10.18421/TEM104-56.
[4]. Mohd Noor C.W., Mamat R.,Najafi G.,WanNik W.B., and Fadhil M. (2015). Application of Artificial Network for Prediction of Marine Diesel Engine Performance. 3rd International Conference of Mechanical Engineering Research (ICMER 2015). IOP Conf. Series: Materials Science and Engineering 100 (2015) 012023; doi:10.1088/1757-899X/100/1/012023.
[5]. Mustapha and Khan (2012). Factors Affecting Student's Academic Performance. Global Journal of Management and Business Research. Volume 12, Issue 9 Version 1.0. Online ISSN: 2249-4588 & Print ISSN: 0975-5853.
[6]. Okubo F., Yamashita T., Shimada A and Konomi S. (2017). Student's Performance Prediction Using Data of Multiple Courses by Recurrent Neural Network. Proceeding of the 25th International Conference on Computers in Education, New Zealand: Asia-Pacific Society for Computers in Education, Pages 439 – 444.
[7]. Shahiri A.M., Husain W. and Rashid N.A. (2015). A Review on Predicting Student's performance using Data Mining Techniques. The Third Information Systems Conference, Prodedia Computer Science 72(2015) 414-422. www.sciencedirect.com.
[8]. Sikder F., Udin J., and Halder S. (2016). Predicting Students Yearly Performance Using Neural Network: A case Study of BSMRSTU. 5th International Conference on Informatics, Electronics, and Vision (ICIEV). 978-5090-1269-5/16/$31.00, IEEE. Pages 524-529.
[9]. Trifonof R., Yoshinof R., Pavlova G. and Georgi (2021). Artificial Neural Network Intelligentg Method for Prediction. Mathematical Methods and Computational Techniques in Science and Engineering, AIP Conf. Proc. 1872, 020021-1-020021-6; doi:10.1063/1.4996678.

Nurudeen Ajibola IROMINI, et. al. "Artificial Neural Network Analysis for Predicting The Performance Of Students In An Examination." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 25(1), 2023, pp. 11-16.