



## 1. INTRODUCTION

Universities play a remarkable role in the development of a country by producing skilled graduates for the country. The graduation rate is low as compared to the enrolment rate in higher institutions. Academic failure is the main reason for non-degree completion. Students' retention and high academic performance are significant for students, academic and administrative staff of universities [1, 5].

To achieve better learning outcomes, the choice of instructional interventions must take into account the diverse academic backgrounds and varied performance of students in relevant courses because each student will have a different reaction to them. Prediction of students' academic performance has been proposed using Structural Equation Modelling (SEM), and machine learning techniques against the conventional techniques include traditional statistical methods, discriminant analysis, and multiple linear regressions used on existing systems. Furthermore, these proposals include the use of Bayesian model, random forest classifier and ensemble learning.

The study seeks to explore the possibility of using an Artificial Neural Network model to predict the academic performance of a student. The results of these predictive models can help the instructor determine whether or not a pedagogical and instructional intervention is needed. 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 the said student. It has been suggested that the variables have different effects on different learning subjects. Identifying and choosing effective modeling approaches is also vital in developing predictive models.

The declining rate of students' performances in institutions of higher learning has posed a significant threat to the educational system around the globe. Several factors are contributory to this menace. Some of these factors may include a poor prior academic background, lack of interest or focus, distraction from campus activities, and improper selection during the admission process [2, 10].

According to [5, 11], learners' academic performance should be monitored as early as possible in their academic career to foster national growth and development as learners eventually become key players in the affairs of the country in all sectors of the economy. This research aims to use an Artificial Neural Network model to predict students' academic performance based on some identified attributes or factors which affect student academic performance. This aim is achievable through the following objectives:

- i. review existing systems to identify problems associated with them
- ii. identify appropriate perceptron algorithm for developing the predictive models
- iii. determine appropriate predictor variables / independent variables that can be used as the inputs of predictive models.

iv. implement the developed models using the data collected and predict the academic performance of students

v. evaluate the developed model against conventional traditional statistical methods techniques

## 2. RELATED WORKS

Neural networks have been used by analysts to predict student learning outcomes. [6, 9] used grades from a multiple-choice test as inputs to neural networks that predict students' final achievement. [3, 8] used the average point scores of grades 12 students as inputs to a multilayer perception neural network and predicted first-year college student achievement with high accuracy [7] found that cumulative grade point average (CGPA) is the most influential attribute because it determines future educational and career mobility. According to their findings, the neural network has the highest prediction accuracy by (98%) followed by the decision tree by (91%). Support vector machine and k-nearest neighbor had the same accuracy (83%), while naive Bayes gave lower prediction accuracy (76%).

### 2.1 Fundamentals of Artificial Neural Networks

An Artificial Neural Network (ANN) is a statically oriented modeling tool. It is a similitude of the biological nervous system. The basic processing element in the ANN is known as the neurons [4, 6]. The neuron is not the same as the neuron in the human body but in terms of functionality, it works in the same manner. Hence, the name is called artificial neurons. It has a normal range of output between (-1, +1). It could also be (0, 1) [6]. The neuron can be viewed as a processor that computes the sum of weighted inputs and then applies a non-linear transfer function to the computed sum. The transfer function could be Tang-sigmoid or logistic-sigmoid.

Artificial neural networks (ANNs) gained their popularity in many application areas such as pattern recognition, image processing, optimizations, prediction, and control systems, [8]. An Artificial Neural Network is specified by:

- Neuron model: the information processing unit of the neural network,
- Architecture: a set of neurons and links connecting neurons. Each link has a weight
- Learning algorithm: used for training the neural network by modifying the weights to model a particular learning task correctly on the training examples.

## 3. METHODOLOGY

The research explored the functionality of artificial neural networks using Feed Forward Neural Network, the hybridization of FFNN with Back Propagation Technique to predict the students' performance. Thus,

a tool that could automatically recognize timely students' performance and students with learning problems is very crucial for educators as shown in figure 1.

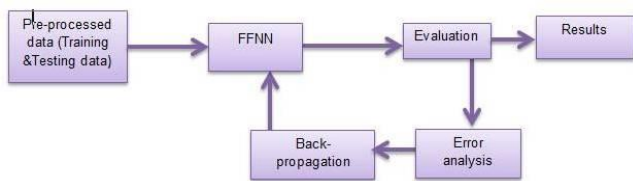


Fig. 1: System Conceptual Architecture

### 3.1.1 Perception Algorithm of FFNN (Student Performance Prediction)

A complete diagram depicting the Neuron of FFNN model is shown in figure 2 comprises:

1. For every input, multiply that input by its weight.
2. Sum all of the weighted inputs.
3. Compute the output of the perceptron based on that sum passed through an activation function (the sign of the sum).

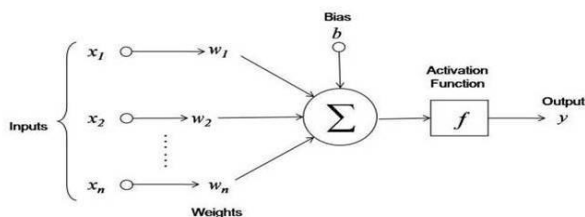


Fig. 2: Neuron of FFNN Model

$$y = f(a)$$

$$a = \sum_{i=1}^n (x_i w_i) + \theta \quad (1)$$

The sum of the inputs  $x_i$  multiplied by their respective weights  $w_i$ :

$$a(x, w) = \sum_{i=1}^n (x_i w_i) \quad (2)$$

This binary relationship of whether to exercise an option or not, can be computed by the sigmoid activation function:

$$Y_j(x, w) = 1 / 1 + e^{-a(x, w)} \quad (3)$$

But the target output of this network lies within the range of -1 and 1. Therefore, the hyperbolic tangent sigmoid function is used.

$$y_j(x, w) = \frac{e^{a(x, w)} - e^{-a(x, w)}}{e^{a(x, w)} + e^{-a(x, w)}} \quad (4)$$

The error function for the output of each neuron can be defined as:

$$E = \frac{1}{2} \sum_j^i \{t_j - y_j(x, w)\}^2 \quad (5)$$

$$E = \frac{1}{2} \sum_j^i (\text{target} - \text{predicted})^2$$

$$E = \frac{1}{2} \sum_j^i \left( \text{Target} - \frac{e^a - e^{-a}}{e^a + e^{-a}} \right)^2$$

Where  $a = (x_i w_{ij} + \text{bias})$

The weights can be adjusted using the method of gradient descent:

$$\Delta w_{ij} = -\eta \frac{\delta E}{\delta w_{ij}} \quad (6)$$

### 3.1.2 Data Collection

The researchers used a survey design with questionnaires to obtain data on attributes that potentially affect the students' academic performance as well as information on the demographic characteristics of university students, like gender, age, region, family system, parent marital status. Also, Feed Forward Neural Network enhances obtaining information on some quantitative variables related to student performance using a 5-point selective opinion composed of 1 (strongly agree), 2 (agree), 3 (neutral), 4 (disagree), 5 (strongly disagree). Items are designed to assess five dimensions associated with student academic performance; Demographic, Personality, Institution, Psychological, and IQ ability influences as shown in table 1.

Some students from different categories were examined using diverse numbers of questions. The results obtained showed students where: X represents the number of questions answered wrongly Y represents the number of questions answered correctly

Table 1: Coefficient Correlation

Variables	X	Y	XY	X <sup>2</sup>	Y <sup>2</sup>
Psychological domain	43	99	4257	1849	9801
Personality domain	19	66	1365	441	4225
Demographic domain	25	79	1975	625	6241
Institutional domain	42	75	3150	1764	5625
Students' Intellectual ability	57	87	4959	3249	7569
	<b>186</b>	<b>406</b>	<b>15706</b>	<b>7928</b>	<b>33461</b>

$$r = \frac{n(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[n\sum X^2 - (\sum X)^2][n\sum Y^2 - (\sum Y)^2]}} \quad (7)$$

The data collected was validated using equation (7) and found to have a coefficient correlation of 0.850. The basis for the input dataset is to train the neural network using the hybridization of FFNN with a backpropagation algorithm for efficient prediction and establish the level of association (performance) among metrics variables towards the aim of the study.

The following variables were used:

- Psychological domain: Self-efficacy, achievement, goal, interest

- Personality domain: Motivation, study time, online activities, attendance.
- Demographic domain: Age, gender, ethnicity, parent marital status.
- Institutional domain: Course program, learning environment, institutional support, course workload.
- Students' Intellectual ability Score

Each of the domains contributes to the performance measurement of the students and is made up of attributes that work individually and jointly for learners' success. The students' Intellectual ability Scores are measured based on intelligent tests answers from the questionnaire.

$$\text{Score} = \frac{\text{Questions answered correctly}}{\text{Total questions given}} * 100 \quad (8)$$

Table 2: The Performance Score

Variables	X	Y	$((Y/(X+Y))*100)$
Psychological domain	43	99	69.72
Personality domain	19	66	77.65
Demographic domain	25	79	75.96
Institutional domain	42	75	64.10
Students' Intellectual ability	57	87	60.42
	<b>186</b>	<b>406</b>	<b>68.58</b>

### 3.1.3 Classification

Classification maps data into a pre-defined group of classes. Base on the scope of this research, three classes were identified as 'pass', 'average', and 'fail'. The class label was defined as '1', '0' and '-1' for a pass, average, and fail.

### 3.1.4 Evaluation of Prediction Accuracy

The prediction accuracy of the model was examined by using error evaluation criteria, RMSE and R<sup>2</sup> formula:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (\text{predict value} - \text{target value})^2}{n}} \quad (9)$$

R<sup>2</sup> (Coefficient of determinant). The value of this coefficient is always between zero and one, and its proximity to one represents a better performance of the model.

$$R^2 = 1 - \frac{\sum_{i=1}^n (\text{target value} - \text{predict value})^2}{\sum_{i=1}^n (\text{target value}^2) - \frac{(\sum_{i=1}^n (\text{predict value}))^2}{n}} \quad (10)$$

$$\text{Percentage prediction accuracy} = \frac{\text{FFNN outputs}}{\text{Total predictions}} * 100\% \quad (11)$$

### 3.1.5 Design Analysis

Choosing the topology of the neural network is a difficult decision [12, 13], hence the reason for three (3) layers of the Neuron of FFNN model; input layer,

hidden layer, and output layer. The processed data were fed into the selected FFNN models (predicting students' performance). In this study, the input neurons were variables from respondents which is determined by  $X_i = \{X_1, X_2, X_3, \dots, X_n\}$  where  $i$ , is the number of variables (input neurons). The dataset was divided into three portions (70:15:15) as a required step taken when using MATLAB 2019a simulation tool, 70% is the training data set, 15% is the validation dataset and 15% is the testing dataset. The MLP network was trained using Levenberg Marquardt's backpropagation algorithm in the MATLAB R2019a tool. This algorithm (Levenberg Marquardt backpropagation algorithm) was chosen because it offers a numerical solution to the problem of minimizing a nonlinear function and it has a fast and stable convergence. Each of the inputs was assigned weights, added up with a bias, and passed through a transfer function which was then forwarded to the hidden layer. The tangent function was used between the input layer and the hidden layer, then the output was modified by a non-linear function before being outputted and the result was compared with the output of experimental data by propagating the error. This was done by adjusting the weights until the minimal error was attained in the network. Figure 3 shows the structure of the FFNN model.

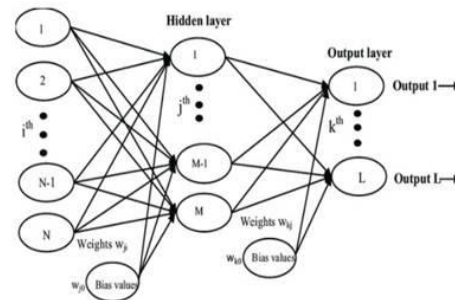


Fig. 3: Structure of the FFNN Model

Figure 4 shows the backpropagation network model of Feed Forward Neural Network FFNN where weight is adjusted to reduce the error. The process starts with a set of randomly generated weights, the results obtained are compared to the expected results, and the difference in results is used as input to the error function for learning. After obtaining a good and satisfactory set of weights, the trained model is ready to predict automatically any set of a given dataset.

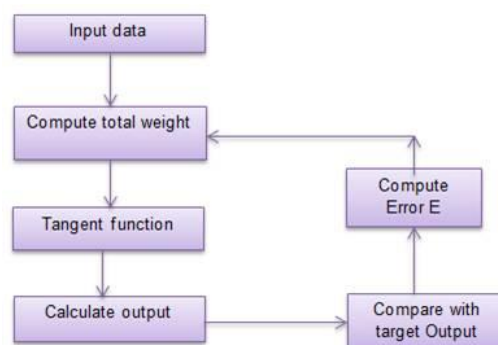


Fig. 4: Back Propagation Network Model

The results of the training phase of the simulation are presented in Figure 5 with the number of iterations, performance plot, and training regression plot inclusive.

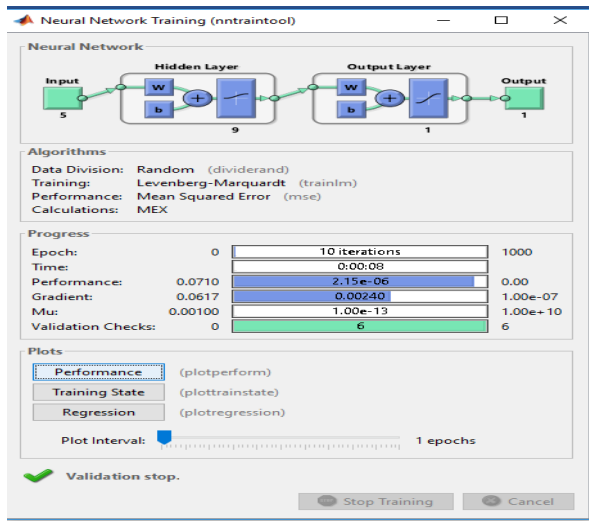


Fig. 5: FFNN Training Interface Summary

The training interface summary of 10 epoch iteration in eight seconds after six validation checks, indicating the performance and gradient values. The training, which was carried out on some iterations (10 epochs) does not exceed 8 seconds with a performance of 2.15e-06 and gradient of 0.00240 after 6 validation checks.

Figure.6 shows the performance plot of FFNN which was able to meet the objective of the training by minimizing the MSE efficiently by 0.00013398 at 4 epochs.

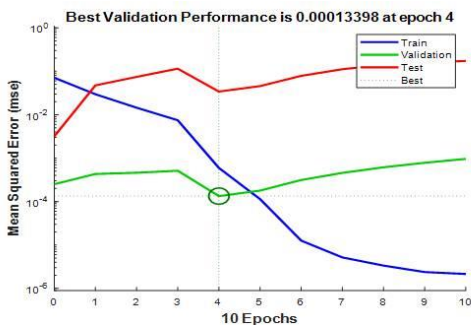


Fig. 6: FFNN Training Performance Plot

The FFNN training Regression plot as depicted in figure 7 shows the relationship between the variables that are made up of the inputs and most importantly, how the inputs and outputs are related. The dashed line in the plot indicates how the expected output (target output) relates to the trained outputs (neural outputs). The solid line indicated the linear regression between outputs and targets. The regression value, R, shown in Figure 4.2 for the training, testing, and validation had the value of 0.99487 that falls within the range  $-1 \leq R \leq 1$ .

This is because the network effectively learned the relationship between the inputs and the output. Consequently, the performance error was quite insignificant as shown in Figure 8. Similarly, the regression plot of Figure 7 was also a strong indication of the fact that the network had effectively learned the relationship between the input variables and the output.

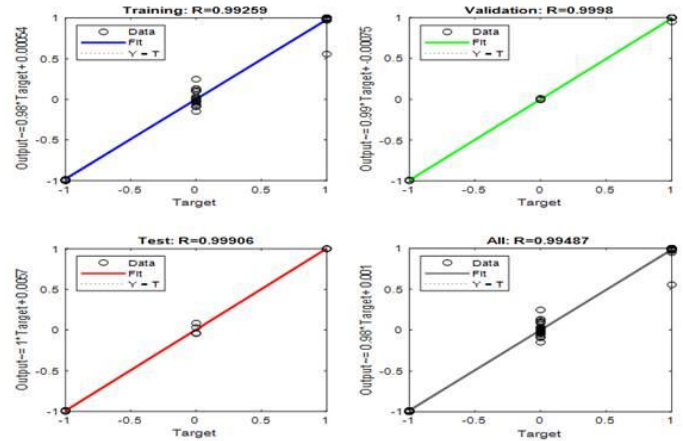


Fig. 7: FFNN Training Regression Plot

### 3.1.6 Testing the Network

The trained network is the dataset (inputs). The inputs were randomly selected from the given inputs to serve as a controlled prediction experiment. For the controlled prediction experiment, the value was selected and the expected outputs are shown in table 3.

Table 3: Sample of Network Outputs and Targets

Target output	Network output (FFNN)	Error output
1	0.99998	0.000014
1	1.0000	0.0000
0	0.032469	-3.25E-02
-1	-0.99963	-3.70E-04

Table 4 shows the evaluation results of the Feed Forward Neural Network model and Error function outputs used to test the Network's topology. The fact that the predicted value tallied with the expected value was enough to give the needed confidence that the network had learned the relationship between the inputs and the outputs.

Table 4: Evaluation Results

Type of Evaluation	Results
R <sup>2</sup>	0.999
RMSE	0.083
Percentage accuracy	99.4%
-1	-3.70E-04

The results from the research showed the efficiency of the model to predict student academic performance.

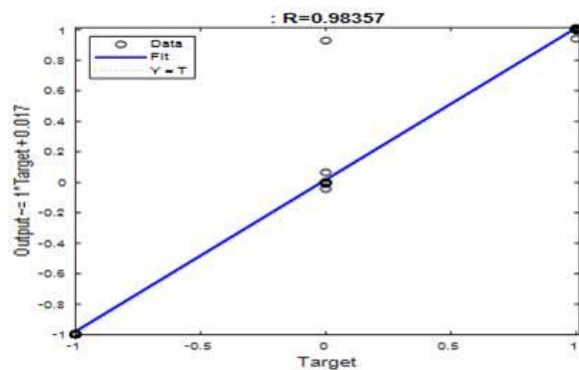


Fig. 8: Regression Plot of the Test FFNN

Regression R Values measure the correlation between outputs and targets. An R-value of 1 means a close relationship, 0 a random relationship.

Regression (R) =  $9.8357 \times 10^{-1}$  (0.98357) the correlation coefficient (R) between the target and the predicted values is 0.98357.

The value of the calculated determinant factor ( $R^2$ ) = 0.999 determines the performance rate of the relationships among variables of interest. The value of  $R^2$  calculated was between 0 and 1 and the nearness of  $R^2$  value according to this study shows that the coefficient of  $-1 \leq R \leq 1$  can be substantiated.

It shows that the FFN Network effectively learned the relationship between the input variables and expected target output with the value of  $R = 0.99487 = 99.49\%$ . The relationship among variables of interest towards the effective predictions of students' performance using Artificial Intelligent Techniques.

Conclusively, It was deduced that FFNN training, testing, and validation with those variables of interest on the Regression value  $R = 0.99487$  falls within the stipulated range  $-1 \leq R \leq 1$ .

Hence, the performance rate of FFNN with the backpropagation technique is better in training those variables of interest to be used in determining/predicting the students' academic performance which was spell-out in the conclusion.

By this, we can conclude that FFNN is 98.3% efficient at modeling data relating to the student's academic performance.

#### 4. CONCLUSION

Conclusively, this work has helped to analyze the capabilities of an Artificial Neural Network in the accurate prediction of students' academic performance using Regression and feed-forward neural network (FFNN) as evaluation metrics.

Pieces of evidence emanating from the evaluation of the metrics show that results from neural networks are precise, concise, and promising at modeling data relating to the student's academic performance. In this research, the feed-forward neural network (FFNN) is 98.3% efficient at predicting students' academic performance, and as a result, it becomes a tool much needed in academic management.

Further research can be tailored toward the research instrument (Questionnaire) that was used in this study, which can be subjected to achievement tests on students over some time as this better reflects their academic performance based on the identified variables.

#### 5. REFERENCES

1. Adekitan AI & Salau O. The impact of engineering students' performance in the first three years on their graduation result using educational data mining. Article e01250, 10.1016/j. heliyon, 5(2) (2019)
2. Ahmad Z & Shahzadi E. Prediction of students' academic performance using artificial neural network. Bulletin of Education and Research, 40 (3) (2018), pp. 157-164
3. Alyahyan E & Düşteğör D. Predicting academic success in higher education: Literature review and best practices. International Journal of Educational Technology in Higher Education, 17 (2020), pp. 1-21, 10.1186/s41239-020-0177-7
4. Asif R, Merceron A., Abbas S & Ghani N. Analyzing undergraduate students' performance using educational data mining. Computers & Education, 113 (2017), pp. 177-194, 10.1016/j.comedu.2017.05.007
5. Fagbola TM, Adeyanju IB, Olaniyan OM & Oloyede AO. Development of Multi-Factored Predictive Models for Students' Performance in Programming Courses in a Nigerian Tertiary Institution. Transylvanian Review Scopus Journal (2018) pp. 6808-6820
6. Fagbola TM, Adeyanju IB, Oloyede AO, Sijuađe, A, Olaniyan OM & Omodunbi BA. Experimental Comparison of Complexities of Artificial Neural Network and Kolmogorov Distance for News Article Classification. Journal of Engineering and Applied Sciences (2018) 14(7); pp. 2276 – 2291. ISSN:1816-949X
7. Group RP. Predicting Student Outcomes Using Discriminant Function Analysis. RP Group Proceedings (2011), publication. Retrieved December 7, 2012, from <http://www.rpgroup.org/>.
8. Kanakana G & Olanrewaju A. Predicting student performance in engineering education using an artificial neural network at Tshwane University of Technology, Proceedings of the ISEM, Stellenbosch, South Africa. (2018)
9. Lykourantzou I., Giannoukos I, Mpardis G, Nikolopoulos V & Loumos V. Early and dynamic student achievement prediction in e-learning courses using neural networks. (2019).J. Am. Soc.

- Inf. Sci., 60: 372–380. DOI: 10.1002/asi.20970
10. Oladokun VO, Charles-Owaba OE & Adebajo OE. Predicting Student's Academic performance using artificial neural network: A case study of an engineering course. *The Pacific Journal of Science and Technology*, (2018), 9 (1) pp. 72-79.
  11. Sajadin S, Zarlis M, Hartama D, Ramliana S & Elvi W. Prediction of student academic performance by an application of Data Mining Techniques. *International Conference on Management and Artificial Intelligence* (2011)
  12. Usman OL & Adenubi AO. Artificial Neural Network (ANN) Model for Predicting Students' Academic Performance. *Journal of Science and Information Technology* October 2013, Vol. 1 No.2, pp. 23-37
  13. Zacharis NZ. Predicting Student Academic Performance in Blended Learning Using Artificial Neural Networks, *Int. J. Artif. Intell. Appl.*, vol. 7, no. 5, pp. 17–29, 2016. [accessed Aug 02, 2021].